FINITE ORDER UNIVERSAL PORTFOLIO GENERATED BY RECURSIVE CALCULATION OF CONTINUOUS RANDOM VARIABLE'S MOMENT GENERATING FUNCTION

CHONG YOU CHUEN

UNIVERSITI TUNKU ABDUL RAHMAN

FINITE ORDER UNIVERSAL PORTFOLIO GENERATED BY RECURSIVE CALCULATION OF CONTINUOUS RANDOM VARIABLE'S MOMENT GENERATING FUNCTION

CHONG YOU CHUEN

A project report submitted in partial fulfilment of the requirements for the award of Bachelor of Science (Honours) Actuarial Science

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

September 2024

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature : 

Name : CHONG YOU CHUEN

ID No. : 2003732

Date : 3/9/2024

APPROVAL FOR SUBMISSION

I certify that this project report entitled **"A STUDY ON FINITE ORDER UNIVERSAL PORTFOLIO GENERATED BY RECURSIVE CALCULATION OF CONTINUOUS RANDOM VARIABLE'S MOMENT GENERATING FUNCTION"** was prepared by **CHONG YOU CHUEN** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Science (Honours) Actuarial Science at Universiti Tunku Abdul Rahman.

Approved by,

Signature : 

Supervisor : Mr. Phoon Sheong Wei

Date : 06/09/2024

ACKNOWLEDGEMENTS

I would like to thank everyone who contributed to the successful completion of this project. I would also like to express my gratitude to my project supervisor, Mr. Phoon Sheong Wei, for his invaluable advice and guidance and his enormous patience throughout the project's development.

Additionally, I would like to thank Universiti Tunku Abdul Rahman for giving me the opportunity to conduct this project project. Special thanks for Yahoo Finance and Refinitiv service providers that offer valuable dataset to conduct this project.

ABSTRACT

This project studied the finite order universal portfolio (UP) generated by five different continuous random variables' (CRV) moment generating functions (MGF), the gamma, lognormal, logistic, Weibull, and generalized Pareto distribution. five machine learning models were adopted to perform stock selection to filter the top-performing companies based on selected dataset, the ridge regression, least absolute shrinkage selection operator, elastic net, boosted regression tree and long-term short-term memory models. All machine learning models formed portfolios based on predicted performance, and all outperformed the KLSE benchmark return. The selected portfolios were used to calculate the universal portfolio by a recursive study of CRV's MGF on orders 1, 2, and 3. The terminal wealth of all portfolios did not outperform the best constant rebalance portfolio (BCRP) as a benchmark comparison but did outperform Buy and Hold (BH), Cover UP (CUP), and Successive Constant Rebalance Portfolio (SCRP) in all portfolios, and Constant Rebalance Portfolio (CRP) in portfolios B and C. As the universal portfolio will adjust the allocation weight based on past observed performance, a transaction cost of 1% was added into consideration. The Bayesian Optimization technique was used in stock selection and universal portfolio construction processes. In stock selection, each model's best parameters will be determined, which will minimize the means squared error of prediction. In contrast, in the universal portfolio, each distribution's best parameters were determined to maximize the terminal wealth generated by the distribution. Parameter sensitivity testing was conducted to study the allocation preferences and relationships between parameter value with the terminal wealth, maximum allocation, and range of allocation. Scenario testing was conducted during the COVID-19 period to study the performance of the universal portfolio during a market downturn. Additionally, scenarios of reducing the trading periods to study the performance of the universal portfolio in the short term were conducted as well. Portfolio A and B outperformed the BCRP benchmark return in all shorter trading periods, while Portfolio C underperformed the BCRP in all periods.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

q	the number of stocks in the portfolio p
p	the number of portfolios in the universal portfolio
$\boldsymbol{b}$	proportion of wealth invested in the universal portfolio
$\boldsymbol{x}$	the stock market vector
$S_n(\boldsymbol{b})$	wealth generated in n^{th} day by strategy $\boldsymbol{b}$
r	the r^{th} stock in the portfolio p
ψ	order of universal portfolio
w	the w^{th} order of the MGF
π	The denominator of $b_{n+1,r}$
Y	the continuous random variable
τ	number of iterations in BRT
v	predictor variables in ML models.

BRT	Boosted Regression Tree
ML	Machine Learning
LASSO	Least Absolute Shrinkage and Selection Operator
ENET	Elastic Net
UP	Universal Portfolio
MGF	Moment Generating Function
MSE	Mean Squared Error
RSS	Residuals Sum Squared
CRV	Continuous Random Variable
w.r.t.	With Respect To
BH	Buy-and-Hold
CUP	Cover’s Universal Portfolio
SCUP	Successive Constant Rebalance Portfolio
CRP	Constant Rebalance Portfolio
BCRP	Best Constant Rebalance Portfolio

CHAPTER 1

INTRODUCTION

1.1 Background of the study

There are various investment strategies in the financial market could be adapted by investors following their preferences or needs. This project will study the universal portfolio investment strategy introduced by Cover (1991) to improve investors' wealth. This strategy has a combination of stocks in the portfolio and the distributions of allocation in each stock will be determined by recursive calculation which will provides the possible maximum wealth generated in the portfolio. Additionally, the stock selection will be carried out by machine learning models to ensure the stock quality for the universal portfolio. The machine learning models used were Ridge regression, Least Absolute Shrinkage Selection Operator (LASSO), Elastic Net (ENET), Boosted Regression Tree (BRT) and Long-Term Short-Term memory (LSTM) which will take stock characteristic as predictor's variables and classify the stock into high and bottom performing companies. Portfolio construction will be made on the top 3, 6 and 9 stocks possessed with the highest performing companies throughout the data collected.

The proportion of wealth invested in each of q stocks can be denoted as the formula, $\boldsymbol{b} = (b_1, b_2, \dots, b_q)^t$, where each of the element in the $\boldsymbol{b}$ must be greater or equal to zero, and summation of all the elements will be equals to 1, denoted as $b_i \geq 0 \; and \; \sum_{i=0}^{q} b_i = 1$. The allocation of wealth in i stock on the n^{th} trading days can be denoted as $b_{\boldsymbol{ni}}$. For the stock market vector, it can be denoted as $\boldsymbol{x} = (x_1, x_2, \dots, x_q)^t$, where $x \geq 0$. The price relative vector in i stock on the n^{th} trading days can be denoted as x_{ni}, which is calculated by taking the closing price divide the opening price for stocks i. Then, S_n will be denoted as the wealth achieved universal portfolio strategy $\boldsymbol{b}$ with $S_n(\boldsymbol{b}) = \prod_{l=1}^{n} \boldsymbol{b_l^t} \, \boldsymbol{x_l}$,where $\boldsymbol{b_l^t x_l} = \sum_{p=1}^{q} b_{lp} x_{lp}$ and the we assume that the initial wealth $S_0(\boldsymbol{b}) = 1$. To achieve the maximum wealth, the condition $S_n^*(\boldsymbol{b}) = \max[S_n(\boldsymbol{b})]$ must be fulfilled.

The universal portfolio proposed by Cover (1991) requires a significant upfront investment of time and resources to implement the algorithm update on the weightage allocation. Therefore, the finite-order universal portfolio is introduced. Unlike traditional portfolio allocation methods that rely on static asset weightings or optimization techniques, the finite order universal portfolio employs a recursive algorithm based on the moment function of random variables. The continuous probability distributions such as the lognormal, logistic, gamma, generalized pareto and weibull distributions were used to generate the moment generating functions for the recursive calculation in this project. Next, comparison of the best universal portfolio constructed with others investment strategies such as the buy-and-hold and best constant rebalance portfolio will be conducted.

1.2 Importance of the Study

This project will study the finite order 1, 2 and 3 of universal portfolio to determine if they reduce the time and resources required for algorithm calculations. Additionally, the performance of universal portfolio against other strategies will be carried out. Besides, evaluation on the machine learning models at the use of stock selection will be conducted.

1.3 Problem Statement

According to Cover (1991), the universal portfolio considers the entire set of available assets in the market and dynamically allocates capital among them based on their historical performance and will generate positive financial returns in the long run. Nevertheless, stock selection for the universal portfolio was not mentioned and thus some high-performing stocks will be chosen by the machine learning models to construct the universal portfolio. The models were trained to select some of the high-performing stocks in Malaysia according to their weighted average of high-performing characteristic possessed. The performance of the universal portfolio built will be tested against others investment strategies.

1.4 Aim and Objectives

The aims and objectives of this project are:

1. To construct Boosted Regression Tree (BRT), Ridge regression, LASSO, ENET, and LSTM machine learning models to select high-performing stock in selected dataset.

2. To construct a universal portfolio by recursive calculation of continuous random variables' moment-generating function.
3. To investigate the optimal order for the universal portfolio generated.
4. To discuss the performance of the universal portfolio.

1.5 Scope and Limitation of the Study

The low order universal portfolio such as order 1, 2 and 3 were studied as the universal portfolio proposed by Thamas M. Cover (1991) required a large amount of computational power. Besides, the number of companies in the universal portfolio is limited to three companies only. The companies studied in this project only consist of 27 companies where were the results of data preprocessing steps.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction to Machine Learning

Machine learning at its core is to learn from data, which they will automatically learn the patterns and make predictions or decisions without having to specify how it is to be done. Machine learning can handle large datasets and can continuously improve their prediction accuracy by hyper-tuning the parameters and perform cross-validation. As the stock market mostly exhibits non-linear relationships between variables, machine learning can learn non-linear relationships with a large amount of trained data as stock market data are available at each trading day. Machine learning approaches might improve stock selection compared to linear methods as most machine learning models tend to do better on classification problems rather than on regression problems (Wolff and Echterling, 2023).

2.1.1 Boosted Regression Trees (BRT)

In this project, the boosted regression trees machine learning model is used as one of the models for the stock selection. According to Li and Rossi (2020), regression trees can handle large dimensional data sets compared to other machine learning methods. Additionally, the regression trees perform better interpretability than other predictive performances, known as the "black boxes." For example, the relative influence measures and the partial dependence plots can be used to visualize the significance of covariates and the relation between the expected return and each characteristic, respectively. The "black box" means that the internal workings of the model prediction are not transparent or easily interpretable, although they perform predictions or conclusions as a result.

Boosting algorithms are employed to perform more accurate forecasts than the original model. The boosting algorithm will combine the weak learners, such as the decision trees, and create a stronger overall model by weighted training. For instance, the algorithm will assign greater weights to the instances misclassified in previous iterations and lower weights to correctly classified instances. This algorithm will emphasize difficult instances, help the model learn from its mistakes, and improve its performance over time (Li and Rossi, 2020).

2.1.2 Ridge Regression, LASSO, and ENET

This project also adapted the ridge regression model, the Least Absolute Shrinkage and Selection Operator (LASSO), and the Elastic Net (ENET).

Ridge regression is used in linear regression to mitigate multicollinearity and prevent overfitting by adding a penalty term to the model's cost function. This penalty term, controlled by a tuning parameter (λ), helps prevent overfitting by discouraging large coefficient values. In classification tasks, ridge regression can be adapted with logistic regression. By applying ridge regularization to logistic regression, the model's coefficients are penalized to avoid extreme values, reducing the risk of overfitting and improving the model's ability to generalize to new data. Ridge logistic regression is particularly useful when dealing with high-dimensional datasets or datasets with multicollinear features, where traditional logistic regression may struggle (Hoerl and Kennard, 1970).

However, ridge regression will not penalize the coefficient values to be exactly zero; hence, it does not perform variable selection as some of the characteristics might not be contributing values of the company. In contrast, LASSO imposes a penalty on the absolute magnitude of the coefficients of the independent variable, encouraging sparsity in the model by shrinking some coefficients to zero. This helps select the most relevant features and avoid overfitting by reducing model complexity (Tibshirani, 1996).

ENET is a regularization technique combining LASSO and Ridge regression penalties to overcome their limitations. According to Zou and Hastie (2005), ENET adds a penalty term to the loss function, a mixture of the LASSO and Ridge penalties controlled by a mixing parameter (α). Like LASSO, it performs variable selection by setting some coefficients exactly to zero. However, ENET can handle situations with correlations among predictor variables, which can cause LASSO to select only one variable from a group of correlated variables. This makes ENET more robust in high-dimensional datasets. The mixing parameter alpha controls the balance between the ridge and LASSO penalties. A value of alpha closer to 1 emphasizes LASSO-like behavior, while a value closer to 0 emphasizes Ridge-like behavior. By tuning alpha, ENET provides flexibility in regularization, allowing it to adapt to different datasets and modeling goals.

2.1.3 Long-Term Short-Term Memory

The long-term, short-term memory, pioneered by Hochreiter and Schmidhuber (1997), is a Recurrent Neural Network (RNN) type that tackles the struggle of capturing long-term dependencies in sequence due to vanishing gradient problems. LSTM comprises interconnected memory cells that control the memory to preserve it in the model. The forget gate decides which information needs to be kept or discarded. The LSTM model is powerful in predicting time-dependent data due to its long-term dependencies.

2.2 Universal Portfolio

The Universal Portfolio, pioneered by Thomas Cover (1991), is an investment strategy. Unlike traditional portfolio strategies, which allocate funds to a fixed set of assets, the universal portfolio dynamically allocates capital among a large universe of assets.

The universal portfolio employs an algorithm that continuously updates the portfolio weights based on the observed performance of assets over time. This dynamic adaptation allows the strategy to capitalize on changing market conditions and exploit profitable trading opportunities. Ultimately, the optimal rebalance portfolio will generate the maximum achievable wealth.

Further studies on the universal portfolio were carried out by Cover and Ordentlich (1996), and they proposed the moving-order universal portfolio, also known as the Cover-Ordentlich universal portfolio. It is a type of portfolio that adjusts its order based on the positive moments of the Dirichlet distribution. However, implementing this strategy is impractical as the number of stocks in the portfolio increases, and the implementation time and computer requirements grow exponentially. The difference between a universal portfolio and a moving-order universal portfolio is that the universal portfolio wealth allocation is not dependent on predicting future price movements.

To address the computational challenge, further studies by Tan and Pang (2013) have proposed a finite and moving order multivariate normal universal portfolio, which reduces the number of historical stock data stored in the calculations and still provides a comparable result with moving-order Dirichlet universal portfolio, and some cases that will outperform the wealth generated (Tan, 2013). Tan and Kuang (2014) studied the orders 1, 2, and 3 dominant-diagonal-matrix-generated universal portfolios with two parameters. They showed that the order had no superiority in the performance of one order over another.

The performance of the finite order universal portfolio was further discussed by Pang, Liew, and Chang (2017), and the wealth achieved outperformed the Dirichlet universal portfolio with only order one. A study on universal portfolio with two and three parameters distribution by Ling, Phoon, and Seoh (2022) showed that the wealth generated was comparable with the best constant rebalance portfolio (BCRP) after finding the optimum parameters value for each distribution, and the Pareto distribution could generate a high-performance universal portfolio. This project will discuss the finite order universal portfolio, which reduces the number of historical stock data stored in the calculations, added with the recursive calculation of continuous random variable's moment generating function, aimed to improve the terminal wealth in the universal portfolio.

2.3 Recursive Calculation

The wealth achieved by the universal portfolio is computed using the constant rebalance strategy, which involves a recursive calculation. The recursive calculation is used so that the universal portfolio algorithm will recursively update the portfolio weights based on past performance. In short, the recursive calculation repeats or uses its previous term to calculate subsequent terms.

2.4 Lognormal and Gamma Distributions

As the finite-order universal portfolio will adjust its order based on the positive moments of generating probability distributions, a continuous random variable is used instead of discrete probability distributions, as the latter would not adequately represent the smooth, continuous changes in stock prices.

The log-normal distribution is used in the finance field as it exhibits non-negativity behaviours which align with the fact that the stock prices could not fall below zero. The logarithm of stock prices would transform the multiplicative effects into changes which will then stabilize the volatility of the stock movements (Cizeau et al., 1997). Additionally, the application of log-normal distribution can be further discovered into the Black-Scholes model that fundamentally assumes that the stock prices follow a log-normal distribution for option pricing (Black and Scholes, 1973). The Black-Scholes model has been used broadly by financial institutions, traders, and regulators to calculate the fair price of a European call or put option on a stock. Therefore, log-normal distribution is used in this study to generate universal portfolio.

Besides from the log-normal distribution, the gamma distribution also only assigns probability to positive values. However, the gamma distribution is often used in finding the rate of occurrence in the waiting times which can be implied as the time between the significant price jumps in a specific direction. The gamma distribution is used instead of exponential distribution as they are two parameters (α and β). The α will control the skewness of the distribution while the β will control the spread of the distribution. Due to its flexibility, the order of α and β can be controlled to fit most of the distribution of stocks.

2.5 Weibull and Generalized Pareto Distributions

As the log-normal and gamma distributions may underestimate the frequency of extreme events, hence the heavier-tailed distributions would be considered.

The focus of Weibull distribution is to detect extreme events such as fatality rate in actuarial industry. However, it can be adapted into focusing financial tails event such as modelling for extreme price movements in markets. Weibull distribution tends to have heavier tails than the normal distribution but with better control over the shape of distribution with parameters β and N. The parameter β will be controlling over the shape of distribution, if the $\beta < 1$, then exhibit heavy tails, if the $\beta = 1$, it becomes an exponential distribution but with heavier tails, and if $\beta > 1$, the larger the value, the more it behaves like normal distribution. The parameter N will be controlling over the spread of distribution over the axis, a larger N makes the distribution more spread out with a lower peak (Weibull, 1951).

The generalized pareto distribution is considered in this project as it portrays heavier tails, providing the ability to determine the extreme events and hence it evaluates how stock prices might perform during extreme conditions (Pickands, 1975). Besides, the distribution of generalized pareto can be controlled by adjusting the shape parameter k, location parameter α ,and the scale parameter λ. The larger the shape parameter k, the heavier the tails of the distribution, and the smaller the scale parameter λ, the narrower the distribution. As both of weibull and generalized pareto distribution's parameters are flexible, a careful estimation on choosing the values of parameters is necessary for accurate estimation.

2.6 Logistic Distribution

The logistic distribution is another valuable tool used in financial modeling due to its flexibility and symmetry around the mean. Unlike the normal distribution, the logistic distribution exhibits heavier tails, which makes it more effective in capturing the likelihood of extreme events while still maintaining simplicity in its application. The logistic distribution is often employed to model growth processes, which can be analogized to stock prices' movement over time, especially when considering the saturation point or maximum potential of growth. The key parameters, mean and scale, enable the adjustment of the distribution's central tendency and the spread of potential outcomes, respectively. This flexibility allows for the fitting of data that may not align perfectly with the assumptions of normality, particularly in markets that experience frequent, significant fluctuations (Gray et. al, 1990). Due to these characteristics, the logistic distribution is employed in this study as it effectively captures both the central tendencies and the tails of stock return distributions, providing a more comprehensive view of potential market outcomes.

2.7 Bayesian Optimization Technique

The Bayesian optimization technique start with defining an objective function $x = argmax\ f(x)_{x \in X}$, where to maximize the value generated by $f(x)$, the Bayesian optimization will construct a probabilistic surrogate model for this objective function and use this model to make decision on the next sample set. The surrogate model used in this project will be the gaussian density function, and the acquisition function used will be the expected improvement (EI) function. The surrogate model will provide a posterior distribution over the possible function and is updated continuously as new data point is added. The location of new data point added will be determined by the acquisition function, which aimed to determine the trade balance between exploration and exploitation (Shahriari et al., 2015).

CHAPTER 3

METHODOLOGY AND WORK PLAN

3.1 Introduction

Machine learning models were adopted to select high-performing stocks to build universal portfolio. To determine high-performing stocks, data collection, modelling for machine learning models and modelling for universal portfolio will be discussed.

3.2 Data collection and preprocessing

The historical stock prices were collected through the Yahoo! Finance API in the python. Next, the characteristic of company was collected through the Refinitiv website. The data preprocessing will be carried out to remove missing values and present the data into clean and presentable formats.

3.3 Machine Learning models algorithms

The classification of top-9 and bottom-9 performing companies was carried out at each test period, where the models will fit to predict the prices. After predicting the prices, the predicted price at the start of the test period will be regarded as the purchasing price, and the predicted price at the end of the test period will be the selling price. The return is calculated by taking the (selling price – purchasing price) / purchasing price. Then, the classification of top-9 and bottom-9 predicted returns will be carried out, and the classification start, and end dates will be recorded to build a portfolio. The top 9 companies will be labelled as 1, the bottom 9 as –1, and others as 0. A momentum score was added to ensure the stock performance was consistent in three test period. Additionally, the classification metrics will be computed to study the performance of different models.

The window size of time series splits for all the models will be 55. Hence, the return calculated will be in 55 periods, approximately 32 trading days. Besides, the Bayesian Optimization technique will be adopted to select the best parameters to minimize the Mean Squared Error (MSE) prediction for the BRT model. The Ridge, Lasso, and Enet will use their respective cross-validation methods to select the best parameters that will provide the least MSE of prediction.

The portfolio construction will use the predicted label during the test period, where top-performing stocks will be regarded in one portfolio. Then, the portfolio's return will be calculated by taking the mean of the return generated by companies in the portfolio as we regard the companies in the portfolio as equally weighted. Then, the overall returns generated by different models can be compared to select the best portfolio. The risk-adjusted return will be the benchmark for selecting the best portfolio.

After determining the best portfolio, the companies in the portfolio will be used to construct the universal portfolio by taking the mean value of predicted labels. The closer the company's mean predicted labels to 1, the more it is selected as top-performing and will be used in the universal portfolio.

3.4 Machine Learning and Universal Portfolio Modelling start date and end date.

The data will be collected from 2010, 1st January until 2023, 31st December. The time frame for the machine learning model to select stocks will be from 2010, 1st January, until 2016, 31st December, while the remaining will be used for modeling Universal Portfolio (UP). In machine learning, the time series cross-validation method will carry out the train test split. The expanding window validation method will be used, in which initially the training period will be determined according to the window size and the testing period subsequently. Then, the window will expand to include the testing period, forming an expended window as the training period and the testing period subsequently. The illustration can be shown in Figure 3.1.

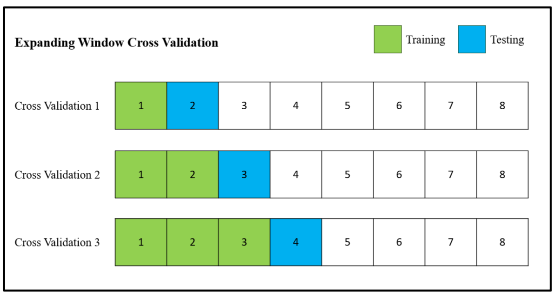


Figure 3.1: Illustration of expanding window cross-validation.

3.5 Boosted Regression Tree

Let X be the predictor variables, and Y be the response variables, τ be the number of iterations of the model. The formulas of the BRT model can be expressed as:

$$F_\tau(X) = F_0(X) + \sum_{\tau=1}^{\tau} \alpha_\tau P_\tau(X) \tag{3.5.1}$$

$$\epsilon_\tau(X) = Y - F_{\tau-1}(X) \tag{3.5.2}$$

where,
$F_\tau(X)$ is the final prediction after τ iterations,
$F_0(X)$ is the initial prediction,
α_τ is the shrinkage parameter (learning rate),
$P_\tau(X)$ is the prediction of τ regression tree,
$\epsilon_\tau(X)$ is the residuals after τ iterations.

During the training process, the BRT model aims to minimize the desired loss function by adding regression trees to the ensemble and adjusting their weight by the shrinkage parameters. The higher the α_τ, the faster the training process, as it will put more weight on each regression tree in the ensembles, leading to a faster convergence. Additionally, a higher α_τ will also increase the risk of overfitting. The loss function specified in the BRT model in our study will be the Mean Squared Error (MSE) loss function.

3.6 RIGDE, LASSO, ENET and LSTM

The main objective of RIDGE, LASSO, and ENET machine learning models is to minimize the cost function. The lower the value of the cost function, the better the performance of these models in terms of both fitting the training data and controlling overfitting issues. By minimizing the cost functions, the coefficients of these models can be calculated.

3.6.1 Ridge Regression

Let RSS be the Residuals Sum Squares. The cost function of ridge regression can be expressed as:

$$Cost(W) = RSS(W) + \lambda \cdot (sum\ of\ square\ of\ weights)$$

$$= \sum_{i=1}^{N} \left\{ y_i - \sum_{v=0}^{M} w_v x_{iv} \right\}^2 + \lambda \sum_{v=0}^{M} w_v^2 \qquad (3.6.1)$$

where *y* is the observed value, *x* is the input value and *w* are the coefficient of *v*, *M* is the total number of predictor variables, and λ as the regularization term.

3.6.2 LASSO

The cost function of LASSO can be expressed as:

$$Cost(W) = RSS(W) + \lambda \cdot (sum\ of\ absolute\ square\ of\ weights)$$

$$= \sum_{i=1}^{N} \left\{ y_i - \sum_{v=0}^{M} w_v x_{iv} \right\}^2 + \lambda \sum_{v=0}^{M} |w_v| \qquad (3.6.2)$$

where *y* is the observed value, *x* is the input value and *w* are the coefficient of *v*, *M* is the total number of predictor variables, and λ as the regularization term.

3.6.3 ENET

The cost function of ENET can be expressed as:

$$Cost(W) = \sum_{i=1}^{N} \left\{ y_i - \sum_{v=0}^{M} w_v x_{iv} \right\}^2 + \lambda_1 \sum_{v=0}^{M} |w_v| + \lambda_2 \sum_{v=0}^{M} w_v^2 \qquad (3.6.3)$$

where *y* is the observed value, *x* is the input value and *w* are the coefficient of *v*, *M* is the total number of predictor variables, and λ as the regularization term.

The λ_1 is the penalty term for lasso regression and λ_2 is the penalty term for ridge regression.

3.6.4 LSTM

The LSTM model consists of three types of gates: the forget gate, the input gate, and the output gate. The forget gate decides what to forget, the input gate decides what to add to the model, and the output gate decides the information to be outputted. Each gate is controlled by a sigmoid function, which will take values from 0 to 1. A value close to 0 will allow no information to pass through, while a value close to 1 allows information to flow freely. The tanh function acts as a cell state update and output activation to prevent the information in the model from growing exponentially. An illustration of the LSTM model can be shown in Figure 3.2.

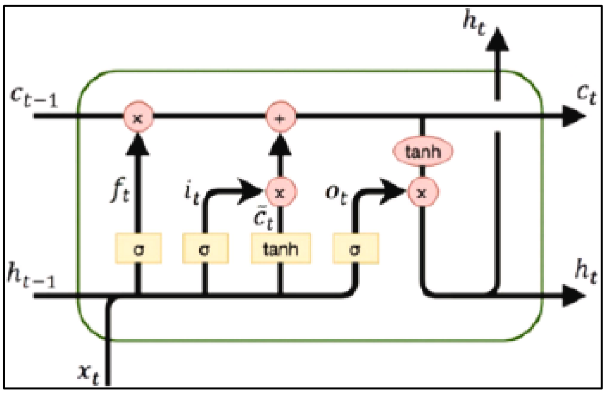


Figure 3.2: Illustration of LSTM model.

3.7 Universal Portfolio Formulas and Definitions

According to Cover (1991), the proportion of wealth invested in each m stock can be denoted as $\boldsymbol{b} = (b_1, b_2, \dots, b_q)^t$, where $\boldsymbol{b_i} \geq 0$ and $\sum_{i=0}^{q} b_i = 1$. Then $\boldsymbol{b}_{ni}$ is the wealth allocation invested in i stock on the n^{th} trading days. The stock market vector is denoted as $\boldsymbol{x} = (x_1, x_2, \dots, x_q)^t$, where $\boldsymbol{x_i} \geq 0$. The price relative vector in i stock on the n^{th} trading days can be denoted as $\boldsymbol{x_{ni}}$. Then, the overall return for n^{th} trading days, S_n will be:

$$S_n(\boldsymbol{b}) = \prod_{l=1}^{n} \boldsymbol{b}_l^t \, \boldsymbol{x}_l \tag{3.7.1}$$

where $\boldsymbol{b}_l^t \boldsymbol{x}_l = \sum_{p=1}^{q} b_{lp} x_{lp}$ and the initial wealth $S_0(b) = 1$.

3.8 Low-Order Universal Portfolio generated by Recursive Calculations

According to Tan (2013), let $Y_1, Y_2, \dots Y_k$ be k mutually independent continuous random variables having a joint probability mass function $f(y_1, y_2, \dots, y_k)$ defined over the domain D such that:

$$D = \{(y_1, y_2, \dots, y_q): f(y_1, y_2, \dots, y_q) > 0\}$$

and the universal portfolio proportion at r^{th} portfolio, $\hat{b}_{n+1,r}$ of order ψ generated by $Y_1, Y_2, \dots, Y_k$ is defined as:

$$\hat{b}_{n+1,r} = \frac{\int_D y_r (y^t x_n)(y^t x_{n-1}) \dots \left(y^t x_{n-(\psi-1)}\right) f(y_1, y_2, \dots, y_r)\, dy}{\int_D (y_1 + y_2 + \cdots y_r)(y^t x_n) \dots \left(y^t x_{n-(\psi-1)}\right) f(y_1, y_2, \dots, y_r)\, dy} \tag{3.8.1}$$

for $r = 1, 2, \dots, q$.

By assuming the $E\left(Y_1^{w_1}Y_2^{w_2}\ldots Y_q^{w_q}\right)\geq 0$, for w in the range of $0\leq w_i\leq \psi+1$, where $i=1,2,\ldots,q$ and $\sum_{i=1}^{q} w_i = \psi+1$, the numerator of $\hat{b}_{n+1,r}$ can be rewritten as below:

$$\hat{b}_{n+1,r} = \pi_{n+1}\left\{\sum_{i_1=1}^{q}\cdots\sum_{i_\psi=1}^{q}\left(x_{n,i_1}\ldots x_{n-\psi+1,i_1}\right)E\left[Y_1^{w_1(r;i)}\ldots Y_q^{n_q(r;i)}\right]\right\} \tag{3.8.2}$$

for r = 1,2, …, q. where π_{n+1} is denoted as:

$$\pi_{n+1} = \left\{\sum_{i_1=1}^{q}\sum_{i_2=1}^{q}\cdots\sum_{i_\psi}^{q}\begin{matrix}\left(x_{ni_1},x_{n-1,i_2}\ldots x_{n-\psi+1,i_1}\right)\\ \times E\left[\left(Y_1+Y_2+\cdots+Y_q\right)\left(Y_1^{w_1(i)}Y_2^{w_2(i)}\ldots Y_q^{w_q(i)}\right)\right]\end{matrix}\right\}^{-1} \tag{3.8.3}$$

where,
$w_j(i)$ = number of y_j's in the sequence $y_{i_1},y_{i_2},\ldots,y_{i_\psi}$ for $j=1,2,\ldots,q$;
$0\leq w_j(i)\leq\psi$; $i=(i_1,i_2,\ldots,i_n)$; $\sum_{j=1}^{q} w_j(i)=\psi$,
$w_j(r;i)$ = number of y_j's in the sequence $y_{i_1},y_{i_2},\ldots,y_{i_\psi}$ for $j=1,2,\ldots,q$;
$0\leq w_j(r;i)\leq\psi+1$; $\sum_{j=1}^{q} w_j(r;i)=\psi+1$.

Let $X_n\left(w_1,w_2,\ldots,w_q\right)$ be the sum of all products $x_{1i_1},x_{2i_2,},\ldots,x_{ni_n}$ having the same set of counts as $\sum_{j=1}^{q} w_j(i)=n$, which can be expressed as:

$$X_n\left(w_1,w_2,\ldots,w_q\right)=\sum_{w_1(i)+w_2(i)+\cdots+w_q(i)=n}\left(x_{1i_1},x_{2i_2,},\ldots,x_{ni_n}\right) \tag{3.8.4}$$

And the quantity X_n is calculated recursively as follows:

$$X_n\left(w_1,w_2,\ldots,w_q\right)=\sum_{j=1}^{q} x_{nj}X_{n-1}\left(w_1,w_2,\ldots,w_q\right) \tag{3.8.5}$$

With the initial conditions:

$$X_n(0,0,\dots,n,\dots,0) = x_{nj}X_{n-1}(0,0,\dots,n-1,\dots,0) \qquad (3.8.6)$$

Then, define the $R_n(w_1, w_2, \dots, w_q)$ as:

$$R_n(w_1, w_2, \dots, w_q) = \prod_{r=1}^{q} E\left[Y_r^{w_r}\right] \qquad (3.8.7)$$

where $\sum_{r=1}^{q} w_r = n$.

Hence, the numerator of $\hat{b}_{n+1,r}$ can be further simplified into:

$$= \sum_{i_1=1}^{q} \dots \sum_{i_\psi=1}^{q} (x_{n,i_1} \dots x_{n-\psi+1,i_1}) E[Y_1^{w_1(r;i)} \dots Y_q^{w_q(r;i)}]$$

$$= \sum_{w_1(i)+w_2(i)+\dots+w_q(i)=n} X_n(w_1, w_2, \dots, w_q)\, E(Y_r^{w_r(i)+1}) \prod_{r=1}^{q} E\left[Y_r^{w_r}\right]$$

$$= \sum_{w_1(i)+w_2(i)+\dots+w_q(i)=n} X_n(w_1, w_2, \dots, w_q)\, R_{n+1}(w_1, w_2, \dots, w_q) \qquad (3.8.8)$$

where,

$$R_{n+1}(w_1, w_2, \dots, w_q) = \left(\frac{E\left[Y_w^{w_r+1}\right]}{E\left[Y_w^{w_r}\right]} \right) R_n(w_1, \dots, w_r, \dots w_q) \qquad (3.8.9)$$

for $w_r \geq 1, r = 1,2, \dots, q$.

The portfolio component $\hat{b}_{n+1,r}$ now can be rewritten as:

$$\frac{\sum_{w_1+\dots+w_q=n} X_n(w_1, \dots, w_q) R_{n+1}(w_1, \dots, w_r + 1, \dots, w_q)}{\sum_{j=1}^{q} \left[\sum_{w_1+\dots+w_q=n} X_n(w_1, \dots, w_q)\, R_{n+1}(w_1, \dots, w_j + 1, \dots, w_q)\right]} \qquad (3.8.10)$$

Then, the wealth function $\hat{S}_n(x^n)$ can be calculated recursively as:

$$\hat{S}_n(x^{n+1}) = \left(\hat{\boldsymbol{b}}_{n+1}\boldsymbol{x}_{n+1}\right)\hat{S}_n(x^n) \tag{3.8.11}$$

where from (3.8.1), $\left(\hat{\boldsymbol{b}}_{n+1}\boldsymbol{x}_{n+1}\right)$ can be expressed as:

$$\left(\hat{\boldsymbol{b}}_{n+1}\boldsymbol{x}_{n+1}\right) = \frac{\int_D \prod_{i=1}^{n+1} y^t x_i\, f\left(y_1, y_2, \dots, y_q\right) dy}{\int_D \left(y_1 + y_2 + \cdots y_q\right) \prod_{i=1}^{n} y^t x_i\, f\left(y_1, y_2, \dots, y_q\right) dy}$$

$$= \frac{\sum_{w_1+\cdots+w_q=n+1} X_{n+1}\left(w_1, \dots, w_q\right) R_{n+1}\left(w_1, \dots, w_q\right)}{\sum_{j=1}^{q} [\sum_{w_1+\cdots+w_q=n} X_n\left(w_1, \dots, w_q\right) R_{n+1}\left(w_1, \dots, w_j + 1, \dots, w_q\right)]} \tag{3.8.12}$$

The initial portfolio $\hat{\boldsymbol{b}}_{\boldsymbol{1}} = (b_{1r})$ can be chosen as $\hat{b}_{1r} = \frac{1}{q}\ , r = 1,2, \dots, q$.

3.9 Log-normal Distribution

Let Y_r follows log-normal distribution where:

$$Y_r \sim (\mu_r\ , \sigma_r)$$

where the probability density function $f(y_r)$:

$$f(y_r) = \frac{1}{y_r \sigma_r \sqrt{2\pi}} e^{-\frac{(\ln(y_r)-\mu_r)^2}{2\sigma_r^2}}\ , y_r > 0$$

The expected values of w^{th} order of log-normal distribution can be expressed as:

$$E\left[Y_r^{w_r}\right] = e^{w_r\mu_r + \frac{\sigma_r^2}{2}}$$

where μ and σ are the mean and standard deviation of the normal distribution respectively.

Hence, the ratio is,

$$\left(\frac{E\left[Y_r^{w_r+1}\right]}{E\left[Y_r^{w_r}\right]}\right) = \frac{e^{(w_r+1)\mu_r+\frac{(w_r+1)^2\sigma_r^2}{2}}}{e^{(w_r)\mu_r+\frac{(w_r)^2\sigma_r^2}{2}}}$$

$$= e^{\mu_r+\sigma_r^2\left(w_r+\frac{1}{2}\right)}$$

From the equation (3.8.9), the recursive calculation for log-normal distribution is:

$$R_{n+1}(w_1, \dots, w_r + 1, \dots, w_q) = \left(e^{\mu_r+\sigma_r^2\left(w_r+\frac{1}{2}\right)}\right) R_n(w_1, \dots, w_r, \dots, w_q) \quad (3.9.1)$$

for $r = 1,2, \dots, q$.

3.10 Gamma Distribution

Let Y_r follows gamma distribution where:

$$Y_r \sim (\alpha_r, \beta_r)$$

where the probability density function $f(y_r)$:

$$f(y_r) = \frac{1}{\Gamma(\alpha_r)\beta_r^{\alpha_r}} \, e^{-\frac{y_r}{\beta_r}}, y_r > 0$$

and $\alpha_r > 0, \beta_r > 0$.

The expected values of w^{th} order of gamma distribution can be expressed as:

$$E\left[Y_r^{w_r}\right] = \frac{\beta_r^{w_r}\Gamma(w_r + \alpha_r)}{\Gamma(\alpha_r)} \; ; \; \alpha_r > 0, \beta_r > 0, w_r = 1,2,3 \dots$$

Hence, the ratio is,

$$\left(\frac{E\left[Y_r^{w_r+1}\right]}{E\left[Y_r^{w_r}\right]}\right) = \frac{\beta_r^{w_r+1}\Gamma(\mathrm{w_r} + 1 + \mathrm{\alpha_r})}{\Gamma(\mathrm{\alpha_r})} \times \frac{\Gamma(\mathrm{\alpha_r})}{\beta_r^{w_r}\Gamma(\mathrm{w_r} + \mathrm{\alpha_r})}$$
$$= (\alpha_r + w_r)\beta_r$$

From the equation (3.8.9), the recursive calculation for gamma distribution is:

$$R_{n+1}(w_1, \ldots, w_r + 1, \ldots, w_q) \\ = \big((\alpha_r + w_r)\beta_r\big)R_n(w_1, \ldots, w_r, \ldots, w_q) \quad (3.10.1)$$

for $r = 1,2, \ldots, q$.

3.11 Weibull Distribution

Let Y_r follows weibull distribution where:

$$Y_r \sim (\beta_r, N_r)$$

where the probability density function $f(y_r)$:

$$f(y_r) = \frac{N_r}{\beta_r}\left(\frac{y_r}{\beta_r}\right)^{N_r-1}, y_r \geq 0$$

β_r being the shape parameter and N_r being the scale parameter. The expected values of w^{th} order of Weibull Distribution can be expressed as:

$$E\left[Y_r^{w_r}\right] = \beta_r^{w_r}\,\Gamma(1 + N_r w_r)$$

Hence, the ratio is,

$$\left(\frac{E\left[Y_r^{w_r+1}\right]}{E\left[Y_r^{w_r}\right]}\right) = \frac{\beta_r^{w_r+1}\,\Gamma(1 + N_r(w_r + 1))}{\beta_r^{w_r}\,\Gamma(1 + N_r w_r)}$$

$$= \beta_r \prod_{i=1}^{N_r}(N_r w_r + i)$$

From the equation (3.8.9), the recursive calculation for Weibull distribution is:

$$R_{n+1}(w_1, \dots, w_r + 1, \dots, w_q) = (\beta_r \prod_{i=1}^{N_r} (N_r w_r + i))\, R_n(w_1, \dots, w_r, \dots, w_q) \tag{3.11.1}$$

for $r = 1,2, \dots, q$.

3.12 Generalized Pareto Distribution

Let Y_r follows generalized pareto distribution (GPD) where:

$$Y_r \sim (\alpha_r, \lambda_r, k_r)$$

where the probability density function $f(x_r)$:

$$f(y_r) = \frac{\Gamma(\alpha_r + k_r)\, \lambda_r^{\alpha_r}\, y_r^{k_r - 1}}{\Gamma(\alpha_r)\Gamma(k_r)(\lambda_r + y_r)^{\alpha_r + k_r}}\ , y_r > 0$$

α_r will be the location parameter, λ_r will be the scale parameter whereas k_r will be the shape parameter. The expected values of w^{th} order of generalized pareto distribution can be expressed as:

$$E[Y_r^{w_r}] = \frac{\lambda_r^{w_r} \prod_{i=0}^{w_r - 1}(k_r + i)}{\prod_{j=1}^{w}(\alpha_r - j)}$$

Hence, the ratio is,

$$\left(\frac{E[Y_r^{w_r+1}]}{E[Y_r^{w_r}]}\right) = \left(\frac{\lambda_r^{w_r+1} \prod_{i=0}^{w_r}(k_r + i)}{\prod_{j=1}^{w_r+1}(\alpha_r - j)} \times \frac{\prod_{j=1}^{w_r}(\alpha_r - j)}{\lambda_r^{w_r} \prod_{i=0}^{w_r - 1}(k_r + i)}\right)$$

$$= \frac{\lambda_r^{w_r}(k_r + w_r)}{(\alpha_r - w_r)}$$

From the equation (3.8.9), the recursive calculation for generalized pareto distribution is:

$$\begin{aligned} &R_{n+1}(w_1, \dots, w_r + 1, \dots, w_q) \\ &= \left(\frac{\lambda_r^{w_r}(k_r + w_r)}{(\alpha_r - w_r)}\right) R_n(w_1, \dots, w_r, \dots, w_q) \end{aligned} \quad (3.12.1)$$

for $r = 1,2, \dots, q$.

3.13 Logistic Distribution

Let Y_r follows logistic distribution where:

$$Y_r \sim (\mu_r, s_r)$$

where the probability density function $f(y_r)$:

$$f(y_r) = \frac{\exp\left(-\frac{y_r - \mu_r}{s_r}\right)}{\mathrm{s_r}\left(1 + \exp\left(-\frac{y_r - \mu_r}{s_r}\right)\right)}$$

μ_r will be the location parameter, s_r will be the scale parameter. The expected values of w^{th} order of Generalized Pareto distribution can be expressed as:

$$E\left[Y_r^{w_r}\right] = \frac{\exp(\mu_r w_r)\,\Gamma(1 - s_r w_r)\Gamma(1 + s_r w_r)}{\Gamma(2)}, \quad |w_r| < \frac{1}{s_r}$$

Hence, the ratio is,

$$\left(\frac{E\left[Y_r^{w_r+1}\right]}{E\left[Y_r^{w_r}\right]}\right) = \left(\frac{\exp(\mu_r(w_r + 1))\,\Gamma\big(1 - s_r(w_r + 1)\big)\Gamma(1 + s_r(w_r + 1))}{\Gamma(2)} \times \frac{\Gamma(2)}{\exp(\mu_r w_r)\,\Gamma(1 - s_r w_r)\Gamma(1 + s_r w_r)}\right)$$

$$= \frac{\exp(\mu_r)\Gamma\big(1 - s_r(w_r + 1)\big)\Gamma(1 + s_r(w_r + 1))}{\Gamma(1 - s_r w_r)\Gamma(1 + s_r w_r)}$$

From the equation (3.8.9), the recursive calculation for generalized pareto distribution is:

$$R_{n+1}(w_1, \dots, w_r + 1, \dots, w_q) = \left(\frac{\exp(\mu_r)\Gamma(1 - s_r(w_r + 1))\Gamma(1 + s_r(w_r + 1))}{\Gamma(1 - s_r w_r)\Gamma(1 + s_r w_r)}\right) \times R_n(w_1, \dots, w_r, \dots, w_q) \quad (3.13.1)$$

3.14 Transaction cost in Universal Portfolio

Blum and Kalai (1999) had proposed the concept of transaction costs into the universal portfolio, where assuming the costs is borne equally between seller and buyer can be shown in below:

$$\frac{\theta}{2} \times \sum_{q} |\boldsymbol{b}_{n+1}(q) - \boldsymbol{b}_n(q)| \quad (3.14.1)$$

where $0 < \theta < 1$ is the transaction rate in percentage, and the transaction costs only occurred when there exists a change in allocation b_n.

3.15 Low order Universal Portfolio

Consider first 3 orders of Universal Portfolio in this study and assume there are 3 mutually independent stock portfolios. From equation (3.7.7), the recursive formula can be expressed in:

$$R_n(w_1, w_2, w_3) = \prod_{r=1}^{3} E\left[Y_r^{w_r}\right] = E\left[Y_1^{w_1}\right] \times E\left[Y_2^{w_2}\right] \times E\left[Y_3^{w_3}\right] \qquad (3.15.1)$$

and from equation (3.8.9), the $R_n(w_1, w_2, w_3)$ can be calculated recursively as:

$$R_{n+1}(w_1, w_2, w_3) = \left(\frac{E\left[Y_r^{w_r+1}\right]}{E\left[Y_r^{w_r}\right]}\right) R_n(w_1, w_2, w_3) \qquad (3.15.2)$$

3.15.1 Order 1 Universal Portfolio

From equation (3.8.2) and equation (3.8.3), the order 1 universal portfolio will have the values of

$\psi = 1, m = 3,$

$\sum_{j=1}^{3} w_r(i) = \psi = 1,$

$\sum_{j=1}^{q} w_j(r; i) = \psi + 1 = 2,$

and can be expressed as:

$$\hat{b}_{n+1,r} = \pi_{n+1} \left\{ \sum_{i_1=1}^{3} \left(x_{n,i_1}\right) E[Y_1^{w_1(r;i)} Y_2^{w_2(r;i)} Y_3^{w_3(r;i)}] \right\} \qquad (3.15.3)$$

$$\pi_{n+1} = \left\{ \sum_{i_1=1}^{3} \left(x_{n,i_1}\right) E[(Y_1 + Y_2 + Y_3)(Y_1^{w_1(i)} Y_2^{w_2(i)} Y_3^{w_3(i)})] \right\}^{-1} \qquad (3.15.4)$$

$$= \left[b_{n+1,1} + b_{n+1,2} + b_{n+1,3}\right]^{-1}.$$

From equation (3.8.10), the $\hat{b}_{n+1,r}$ can be expressed in:

$$\hat{b}_{n+1,r} = \frac{\sum_{i_1=1}^{3}\left(x_{n,i_1}\right)R_{n+1}(w_1,w_2,w_3)}{\sum_{j=1}^{3}\left[\sum_{i_1=1}^{3}\left(x_{n,i_1}\right) R_{n+1}(w_1,w_2,w_3)\right]} \quad (3.15.5)$$

for $r = 1,2,3$.

3.15.2 Order 2 Universal Portfolio

From equation (3.8.2) and equation (3.8.3), the order 2 universal portfolio will have the values of

$\psi = 2, m = 3,$

$\sum_{j=1}^{3} w_r(i) = \psi = 2,$

$\sum_{j=1}^{q} w_j(r;i) = \psi + 1 = 3,$

and can be expressed as:

$$\hat{b}_{n+1,r} = \pi_{n+1}\left\{\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\left(x_{n,i_1}\right)(x_{n-1,i_2})E[Y_1^{w_1(r;i)}Y_2^{w_2(r;i)}Y_3^{w_3(r;i)}]\right\} \quad (3.15.6)$$

$$\pi_{n+1} = \left\{\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\left(x_{n,i_1}\right)(x_{n-1,i_2})E[(Y_1 + Y_2 + Y_3)(Y_1^{w_1(i)}Y_2^{w_2(i)}Y_3^{w_3(i)})]\right\}^{-1} \quad (3.15.7)$$

$$= \left[b_{n+1,1} + b_{n+1,2} + b_{n+1,3}\right]^{-1}.$$

From equation (3.8.10), the $\hat{b}_{n+1,r}$ can be expressed in:

$$\hat{b}_{n+1,r} = \frac{\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\left(x_{n,i_1}\right)(x_{n-1,i_2})R_{n+1}(w_1,w_2,w_3)}{\sum_{j=1}^{3}\left[\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\left(x_{n,i_1}\right)(x_{n-1,i_2})R_{n+1}(w_1,w_2,w_3)\right]} \quad (3.15.8)$$

for $r = 1,2,3$.

3.15.3 Order 3 Universal Portfolio

From equation (3.8.2) and equation (3.8.3), the order 3 universal portfolio will have the values of

$\psi = 3, m = 3,$

$\sum_{j=1}^{3} w_r(i) = \psi = 3,$

$\sum_{j=1}^{q} w_j(r;i) = \psi + 1 = 4,$

and can be expressed as:

$$\hat{b}_{n+1,r} = \pi_{n+1}\left\{\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\sum_{i_3=1}^{3}(x_{n,i_1})(x_{n-1,i_2})(x_{n-2,i_3}) \times E[Y_1^{w_1(r;i)}Y_2^{w_2(r;i)}Y_3^{w_3(r;i)}\right\} \quad (3.15.9)$$

$$\pi_{n+1} = \left\{\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\sum_{i_3=1}^{3} \begin{matrix}(x_{n,i_1})(x_{n-1,i_2})(x_{n-2,i_3}) \\ \times E[(Y_1+Y_2+Y_3)(Y_1^{w_1(i)}Y_2^{w_2(i)}Y_3^{w_3(i)})]\end{matrix}\right\}^{-1} \quad (3.15.10)$$

$$= \left[b_{n+1,1} + b_{n+1,2} + b_{n+1,3}\right]^{-1}$$

From equation (3.8.10), the $\hat{b}_{n+1,r}$ can be expressed in:

$$\hat{b}_{n+1,r} = \frac{\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\sum_{i_3=1}^{3}(x_{n,i_1})(x_{n-1,i_2})(x_{n-2,i_3})R_{n+1}(w_1,w_2,w_3)}{\sum_{j=1}^{3}\left[\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\sum_{i_3=1}^{3}(x_{n,i_1})\,(x_{n-1,i_2})(x_{n-2,i_3})R_{n+1}(w_1,w_2,w_3)\right]} \quad (3.15.11)$$

for $r = 1,2,3$.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Data Collection and Preprocessing

Data collection for stock price movement from Yahoo! Finance API and the quarterly financial metrics from Refinitiv. The data preprocessing procedure was conducted and discussed in the following section.

4.1.1 Historical Stock Prices

The symbol of companies listed in Malaysia was obtained through the Bursa website, including 1022 companies. Then, data filtering was done to filter the company in the Bursa but not in Yahoo! Finance API, leaving 744 companies. Next, the stock price data were obtained from the Yahoo! Finance API from January 1st, 2010, to December 31st, 2023, including 3448 trading days. Data cleaning on the 744 companies was done by excluding missing stock data during the 3448 trading days. Hence, the cleaned dataset will have no missing historical stock data and consist of 253 companies.

4.1.2 Characteristics of Companies

The characteristics of the 253 companies were obtained through Refinitiv by matching the ticker symbol in Yahoo! Finance and Refinitiv, as they used a different ticket symbol representing the company. There are 35 companies left after matching the names, the characteristics of the companies were collected through the quarterly financial summary section in Refinitiv. The financial summary consisted of the important metrics in the quarter. There were 125 unique characteristics in each of the companies, and as they are from different industries, common characteristics were filtered out. The companies in the banking industry were removed as they consist of 12 individual metrics for their industry only. The characteristics with missing values in the starting date, January 1st, 2010, were excluded too. Hence, the final cleaned dataset comprised 27 companies with 26 common characteristics. As the characteristics collected were in a quarterly format, and the date they published their financial statements differed, the matching of dates for characteristics was carried out. Forward

forward-filling technique was applied to handle the missing values. Forward filling was applied as the company's characteristics available at the moment were the last published financial statements. The characteristics of the company will be shown in Table 4.1. Nevertheless, the time frame of characteristics was consistent with the historical stock prices, with 3448 trading days, 27 companies and 26 characteristics.

Table 4.1: Quarterly Common Characteristics and Details in Financial Summary collected from Refinitiv.

Characteristics	Details
STLR	Revenue from Business Activities - Total
SOPR	Operating Profit before Non-Recurring Income/Expense
SEBITA	Earnings before Interest, Taxes, Depreciation & Amortization (EBITDA)
SICO	Income before Discontinued Operations & Extraordinary Items
SCSI	Cash & Short-Term Investments
ATOT	Total Assets
SLSD	Debt – Total
SQCM	Common Equity - Total
STLO	Net Cash Flow from Operating Activities
SNCC	Net Change in Cash - Total
DivYield_TTC	Dividend Yield - Common Stock - Net - Issue Specific - %, TTM
SDCOC	EPS - Diluted - excluding Extraordinary Items Applicable to Common - Total
SDWSC	Shares used to calculate Diluted EPS - Total
REBITDAM	EBITDA Margin - %
ROPMAR	Operating Margin - %
RIBTM	Income before Tax Margin - %
RINTR	Income Tax Rate - %
RNIMAR	Net Margin - %

RetOnSTCSE_TTM	Return on Average Common Equity - % (Income available to Common excluding Extraordinary Items), TTM
RetOnTotAst_TTM	Return on Average Total Assets - % (Income before Discontinued Operations & Extraordinary Items), TTM
STCOC	Common Shares - Outstanding - Total
RTDTA	Total Debt Percentage of Total Assets
RTDTC	Total Debt Percentage of Total Capital
RTDTE	Total Debt Percentage of Total Equity
EarnRetenRate	Earnings Retention Rate
PayoutRatio	Dividend Payout Ratio - %

A summary of the data preprocessing steps can be shown in Figure 4.1 below:

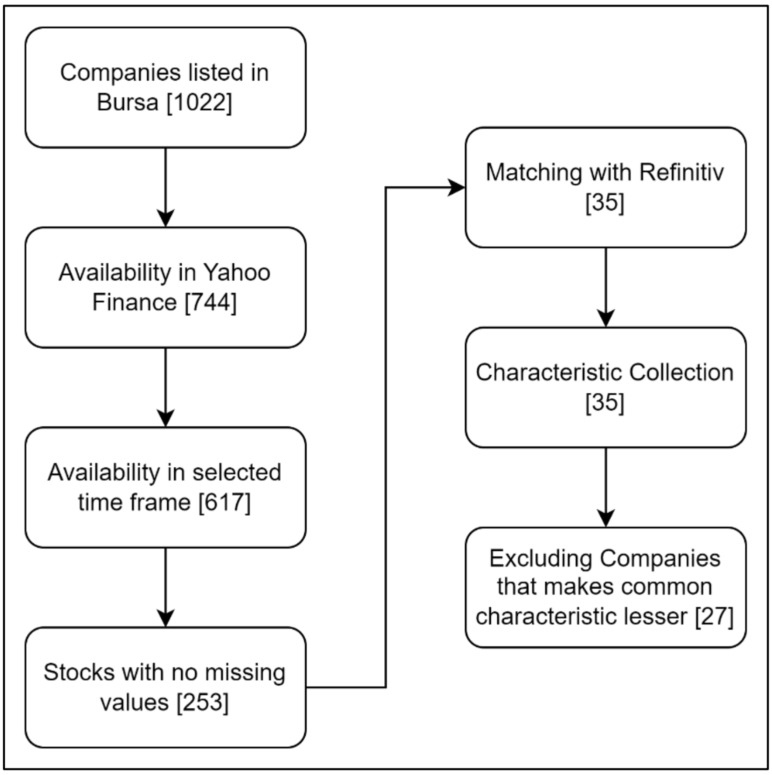


Figure 4.1: Flow and Summary of data preprocessing steps. Squared Bracket indicates number of companies left.

4.2 Stock Selection

This section will discuss the stock selection made by BRT, Ridge, LASSO, ENET and LSTM models. The random seed has been set to 42 to ensure reproductivity. The best parameters that will minimize the MSE of prediction will be determined using appropriate techniques. The BRT model will be using the Bayesian optimization technique and LSTM model with ADAM optimization, where ridge, lasso and enet will be using the ridgecv, lassocv and enetcv respectively. The range of parameters searched were shown in the Table 4.2.

Table 4.2: Range of parameters tuned in machine learning models to perform stock selection.

ML models	Parameters	Range
BRT	Learning rate	(0.01, 0.3)
	n estimators	(50, 300)
	Max depth	(3, 10)
Ridge	Alpha	Logspace(-1, 1, 50)
Lasso	Alpha	Logspace(-2, 1, 50)
Enet	Alpha	Logspace(-2, 1, 50)
	L1 ratio	Logspace(0.01, 1, 50)
LSTM	ADAM Optimization	

Table 4.3: Machine Learning Models Classification Metrics.

ML models	Accuracy	Precision	Recall	F1 Score
BRT	31.04%	38.04%	34.01%	30.04%
Ridge	34.28%	38.36%	34.28%	34.45%
Lasso	33.67%	38.32%	33.67%	31.86%
Enet	34.28%	38.87%	34.28%	34.36%
LSTM	**35.42%**	**39.50%**	**35.42%**	**35.97%**

Table 4.3 showed that the accuracy provided by all ML models consistently falls around 31% to 36%, with the highest accuracy, precision, recall, and F1 score achieved by the LSTM model. This indicates that the LSTM

model effectively captures both the positive and negative labels and highest accuracy as well. The portfolio construction is carried out with the top-9 predicted labelled companies and comparison of benchmark return as KLSE to determine the suitable portfolio constructed by different machine learning models.

Table 4.4: Top-9 portfolio constructed by machine learning models.

ML models	**Top-9 Portfolio**			**KLSE**	
	Terminal Return	***Volatility***	***Sharpe Ratio***	***Terminal Return***	***Volatility***
BRT	1.454%	0.0160	90.98		
Ridge	**1.586%**	0.0106	**149.83**		
Lasso	0.763%	**0.0101**	75.96	0.407%	0.0082
Enet	1.467%	0.0147	99.67		
LSTM	1.386%	0.0133	104.53		

All portfolios constructed by different ML models outperform the KLSE benchmark return at a higher risk rate. The Lasso model had the lowest return and Sharpe ratio among all portfolios generated. The other models have similar returns gained, ranging from 1.38% to 1.56%, with the best Sharpe ratio achieved by the Ridge model. This indicates that the model's portfolio is ideal for maximizing returns with a lower risk tolerance.

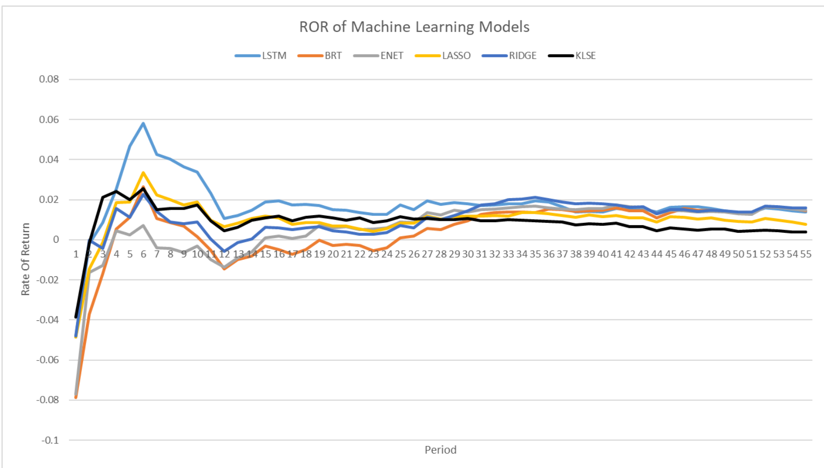


Figure 4.2: Rate of return of different machine learning models.

The rate of return of the LSTM model always outperforms the KLSE after the fourth trading period, and all the models have a decrease in the rate of return from the 6th period to the 12th period, with the Ridge, Enet, and BRT models decreasing beyond zero. However, the rate of return for BRT, ridge, Lasso, and ENET only outperform the KLSE return after the 30th period. From Figure 4.2, the LSTM and Lasso models performed better than KLSE during a market downturn, but the LSTM model underperformed the ridge model after the 52nd period. Nevertheless, the Ridge model will be selected to build a portfolio as it possesses the highest sharpe ratio. The company selected by the Ridge model can be shown in Table 4.5.

Table 4.5: Selected companies from the Ridge Model.

Portfolio	Companies
A	PDZH.KL, PBAH.KL, MISC.KL
B	KENH.KL, TAFI.KL, GENT.KL
C	HAPS.KL, KPJH.KL, TROP.KL

Table 4.6: Ranking details from the Ridge Model.

Company	Ranking Score	Company	Ranking Score
PDZH	0.4000	MESB	-0.0182
PBAH	0.3091	STAR	-0.0182
MISC	0.2545	UMSH	-0.0182
KENH	0.1818	MBMR	-0.0727
TAFI	0.1818	PETR	-0.1091
GENT	0.1636	CWGH	-0.1455
HAPS	0.1636	ENRA	-0.1455
KPJH	0.1636	YLHI	-0.1455
TROP	0.1455	KRIB	-0.1636
PPHB	0.1091	AYER	-0.2000
GAMU	0.0909	CHHB	-0.2727
GIIB	0.0364	ACME	-0.3818
DKLS	0.0182	DBMS	-0.3818
KYMH	-0.0182		

From Table 4.6, the ranking of companies ranges from 0.4 to -0.4, indicating no strongly high-performing or underperforming companies, with the top-ranked companies shown in Table 4.5 that will be used for universal portfolio computation.

4.3 Universal Portfolio

Table 4.7 shows the top three portfolios considered in the UP, along with the companies' symbols, industries, and full names for portfolios A, B and C.

Table 4.7: Three Portfolios generated by the Ridge Regression model for UP.

Portfolio A

Symbol	Full Name	Industry
PDZH.KL	PDZ Holdings Berhad	Freight & Logistics Services
PBAH.KL	PBA Holdings Berhad	Water & Related Utilities
MISC.KL	MISC Berhad	Transport Infrastructure

Three companies.

Portfolio B

Symbol	Full Name	Industry
KENH.KL	KEN Holdings Berhad	Real Estate Operations
TAFI.KL	TAFI Industries Berhad	Professional & Commercial Services
GENT.KL	Genting Berhad	Hotels & Entertainment Services

Three companies.

Portfolio C

Symbol	Full Name	Industry
HAPS.KL	Hap Seng Consolidated Berhad	Consumer Goods Conglomerates
KPJH.KL	KPJ Healthcare Berhad	Healthcare Providers & Services
TROP.KL	Tropicana Corporation Berhad	Real Estate Operations

Three companies.

The best parameters for each distribution will be determined using the Bayesian optimization technique targeted to maximize the wealth generated. The range of parameters searched can be shown in Table 4.8.

Table 4.8: Range of parameters tuned to find the best terminal wealth generated in Universal Portfolio Calculation.

Distribution	Parameters	Type	Range
Gamma	α_r, β_r	Real	(0.1, 120)
Lognormal	μ_r , σ_r	Real	(0, 3)
Weibull	β_r	Real	(0.1, 70)
	N_r	Integer	(1, 11)
GPD	α_r	Integer	(3, 150) [a]
	λ_r, k_r	Integer	(1, 150) [a]
Logistic	μ_r	Real	(0, 5)
	s_r	Real	(0.001, 0.24)

[a.] The maximum range will be 120 for order 3 as the value is close to infinity.

The range of gamma distribution starts from 0.1 to 120, as both the α and β can never be zero. This property is also consistent with the Weibull distribution, where the *Nr* must be an integer as it determines the number of sequences. The lognormal takes values from 0 to 3 as the exponent of a large value will cause the value to close to infinity. The range of the GPD searched is between 3 and 150, as the α, λ and k can never be zero. Additionally, the α value for GPD, order 1 is 3, order 2 is 4, and order 3 is 5 as the denominator will be invalid. The s_r of logistic distribution maxed at 0.24 to fulfill the condition of $|w_r| < \frac{1}{s_r}$, given w_r is 3.

The universal portfolio is calculated at a transaction cost of 1%. Table 4.9 shows the return for different portfolios, distributions, and orders, with the Best Constant Rebalance Portfolio (BCRP) as a benchmark. The BCRP was determined using the olpsR package in RStudio, with 10,000 random samplings.

Table 4.9: Terminal Wealth for Universal Portfolio.

Portfolio	**Order**	**Distributions**				
		Gamma	**Lognormal**	**Weibull**	**Logistic**	**GPD**
A	1	1.84018	1.29297	1.29113	2.00895	1.99981
	2	2.00366	1.29113	1.29569	**2.01060**	1.99605
	3	2.00471	1.29113	1.32018	2.01003	2.00157
BCRP		2.15021				
B	1	4.65403	4.65340	4.65400	4.64665	4.65684
	2	**4.65738**	4.65340	4.65400	4.64922	4.65613
	3	4.65719	4.65340	4.65400	4.65327	4.65680
BCRP		4.91469				
C	1	**1.53122**	1.49166	1.41354	1.52781	1.52459
	2	1.53107	1.52753	1.40024	1.52781	1.52893
	3	1.53107	1.40135	1.40135	1.52687	1.53004
BCRP		1.54327				

All the portfolio terminal wealth underperformed the benchmark BCRP returns, with the highest return generated by both gamma distributions and logistic distributions. However, portfolios B and C provide relatively close returns to BCRP, but portfolio A has the highest difference of 0.86 by the lognormal distribution. In three of the portfolios, the Weibull and Lognormal distributions were the worst performers among the distribution. Additionally, the gamma distribution was the best performer in all portfolios except portfolio A, with GPD as the best performer. The logistic distribution provided the middle performance among all portfolios.

All distributions provide similar results across portfolios and orders except for gamma distribution, order 1. This similarity suggests that these distributions effectively capture the universal portfolio algorithm, maximizing the returns. Order 1 of the gamma distribution shows a lower return in all portfolios except Portfolio C, suggesting orders 2 and 3 to maximize return in the universal portfolios A and B. Gamma shows slightly higher returns in portfolios A and B, indicating it is doing better in accommodating the portfolio’s weight based on the past price vector of two or three days.

In portfolio A, the order increment will lead to improvement in terminal wealth, except for lognormal and logistic distributions. In portfolios B and C, the order increment leads to an uncertain movement. The gamma distribution was the best performer, followed by GPD and logistic distribution. The Weibull and Lognormal distributions were the worst performers in the selected portfolios. As the finite order universal portfolio will decrease computational resource requirements and running time, the time taken to generate results in Table 4.9 can be shown in Table 4.10.

Table 4.10: Time taken to generate universal portfolio results.

Order	**Distributions**					**Single Iteration**
	Gamma	***Lognormal***	***Weibull***	***GPD***	**Logistic**	
1	7m	8m	8m	8m	7m	4s
2	32m	36m	27m	27m	29m	15s
3	86m	100m	105m	74m	94m	65s

The *m* is regarded as minutes, and *s* is regarded as seconds. All of the distributions were run through 75 iterations on Bayesian optimizations with a total of 3 portfolios, and the time taken to compute the weight allocation given certain parameters is shown in the single iterations column. From the table, order 3 will take around 12 times the time for order 1 except for the GPD distribution. The computer specification is attached for reference to the run time in Table 4.11.

Table 4.11: Specification of Laptop used in this Project.

Description	**Specification**
Model	ASUS Predator Helios 300
Processor	Intel Core i5-10500H
Operating System	Windows 11
Graphic	NVIDIA GeForce GTX 3060 6GB DDR6
Memory	16GB DDR4 RAM
Storage	1TB SATA HDD + 512 GB NVMe SSD

4.4 Parameter Sensitivity Testing among Portfolios

This section will discuss the effect of parameter changes generated by the best and worst performer distributions in order 1. The price vector summary and performance of stocks in three portfolios can be shown from Table 4.12 to Table 4.14, and Figure 4.3 to Figure 4.5. The terminal wealth of portfolios in each company was computed to consist only of their own stock movement, portraying the sole performance achieved by the company. The price vector counts, and graphical representation of each stock were displayed to better understand the stocks selected in each portfolio.

Table 4.12: Price vectors summary for portfolio A.

Portfolio A	**Price Vector**			**Terminal Wealth**
	1	**<1**	**>1**	
PDZH.KL	999	398	324	0.03353
PBAH.KL	702	531	488	1.11304
MISC.KL	194	764	763	1.26261

Table 4.13: Price vectors summary for portfolio B.

Portfolio B	**Price Vector**			**Terminal Wealth**
	1	**<1**	**>1**	
KENH.KL	1034	358	329	0.74118
TAFI.KL	866	459	396	4.53248
GENT.KL	124	848	749	0.67631

Table 4.14: Price vectors summary for portfolio C.

Portfolio C	**Price Vector**			**Terminal Wealth**
	1	**<1**	**>1**	
HAPS.KL	268	735	718	0.51580
KPJH.KL	454	657	610	1.38753
TROP.KL	485	578	658	1.39426

In both portfolios A and C, the second and third companies had relatively close terminal wealth, except for portfolio B where the second company (TAFI.KL) performed the best with a maximum difference of 3.86.

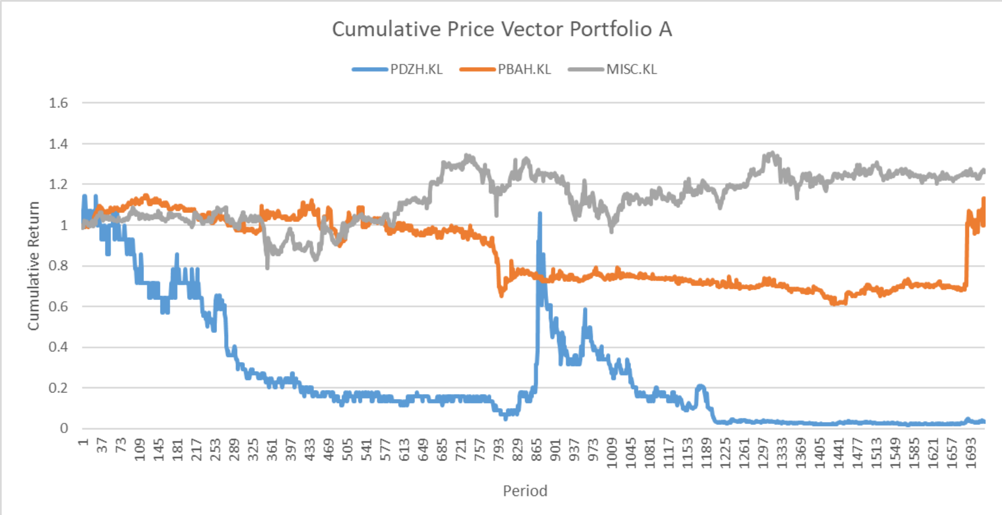


Figure 4.3: Portfolio A, Cumulative Price Vector Movement graph.

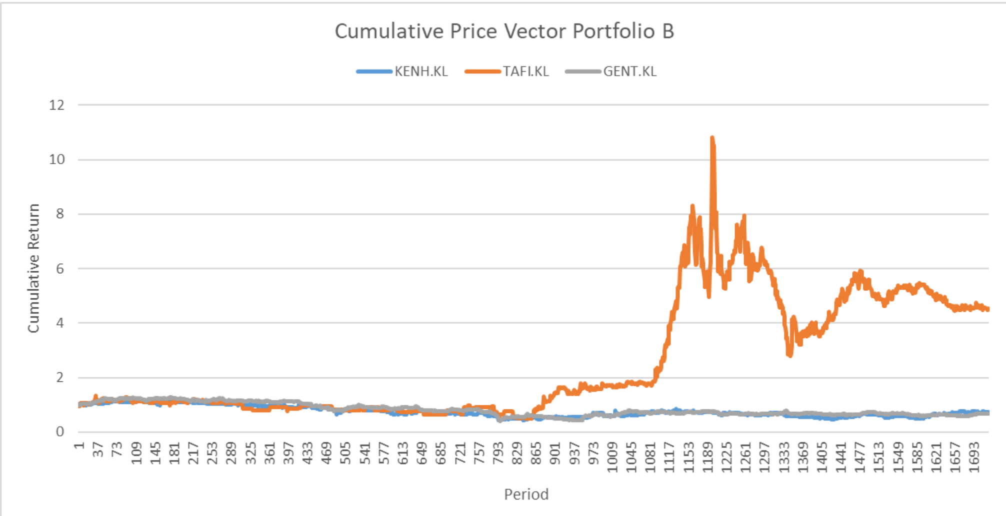


Figure 4.4: Portfolio B, Cumulative Price Vector Movement graph.

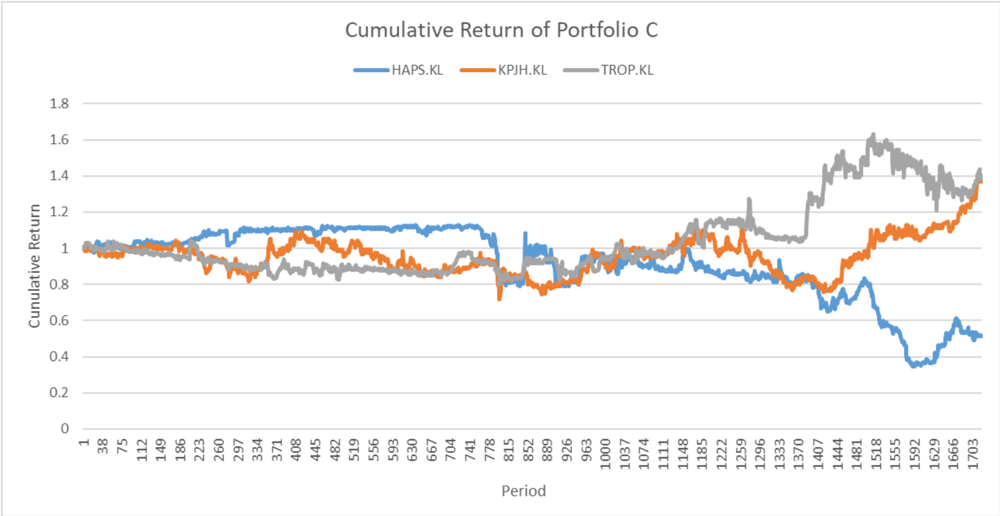


Figure 4.5: Portfolio C, Cumulative Price Vector Movement graph.

4.4.1 Portfolio A, Weibull Distribution

The Weibull distribution of order 1 had the lowest terminal wealth generated. The parameter sensitivity graphs of the Weibull distribution can be shown in Figure 4.6.

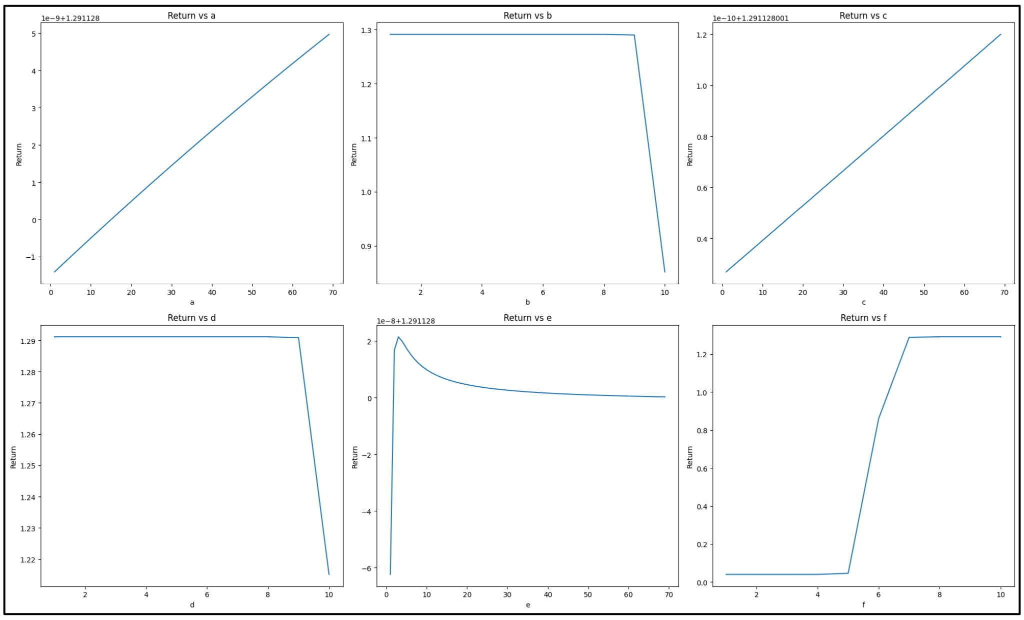


Figure 4.6: Parameter sensitivity graphs for portfolio A, Weibull distribution and Order 1.

The top left graph will be β_1 followed by N_1 to the right and, β_2 and so on. The β_1 and β_2 shows the same increasing trend between parameter value and the terminal wealth generated, suggesting a higher β value leads to higher terminal wealth. The N_1 and N_2 also shows a similar pattern whereby the terminal wealth drops drastically beyond value 9. The highest wealth can be achieved by higher N_3 value (beyond value 7), and the β_3 was highest at value 5, with terminal wealth at 2. Hence, by combining the parameters of different stocks, the allocation details can be shown in Table 4.15.

Table 4.15: Range of allocation for Portfolio A, Weibull distribution and Order 1.

Portfolio A	**Parameters**	**Allocation**			**BCRP**
	$[\boldsymbol{\beta_r}, \boldsymbol{N_r}]$	**Maximum**	**Minimum**	**Range**	
PDZH.KL	[26.47, 6]	0.000000	0.000000	0.000000	0.241
PBAH.KL	[61.19, 5]	0.000000	0.000000	0.000000	0.431
MISC.KL	[48.12, 10]	1.000000	1.000000	0.000000	0.328

The universal portfolio generated for portfolio A, Weibull distribution Order 1 only consists of the MISC.KL company, where the PDZH.KL and PBAH.KL has allocations close to zero, although both companies have similar terminal wealth. Nevertheless, BCRP asserts a significant allocation at the lowest terminal wealth of PDZH.KL, which exploits profitable opportunities from it during certain periods, as shown in Figure 4.3. From Table 4.12, the Weibull distribution had allocated all weight on MISC.KL was more vigorous than the other distributions with a price vector movement of more than 1 or less than 1. Figure 4.7 illustrates that the price vector movement for all companies in Portfolio A was quite similar, and the highest volatility on price vector changes is the PDZH.KL. Weibull distribution emphasized more on the MISC.KL as it could not capture the high volatility of PDZH.KL. However, the BCRP allocation suggests allocation in all companies with the highest allocation on the PBAH.KL company. Hence, the Weibull distribution did not capture the return of all companies in the portfolio effectively.

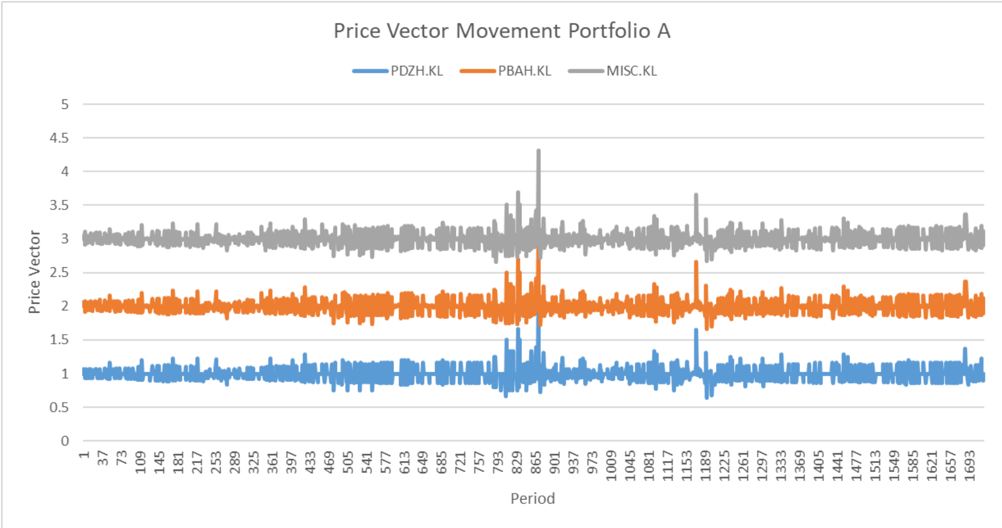


Figure 4.7: Price vector movement for Portfolio A.

4.4.2 Portfolio A, Logistic Distribution

Logistic distribution generated the highest terminal wealth in portfolio A, and Figure 4.8 shows the parameter sensitivity graphs of its parameters when other parameters are held at their best values.

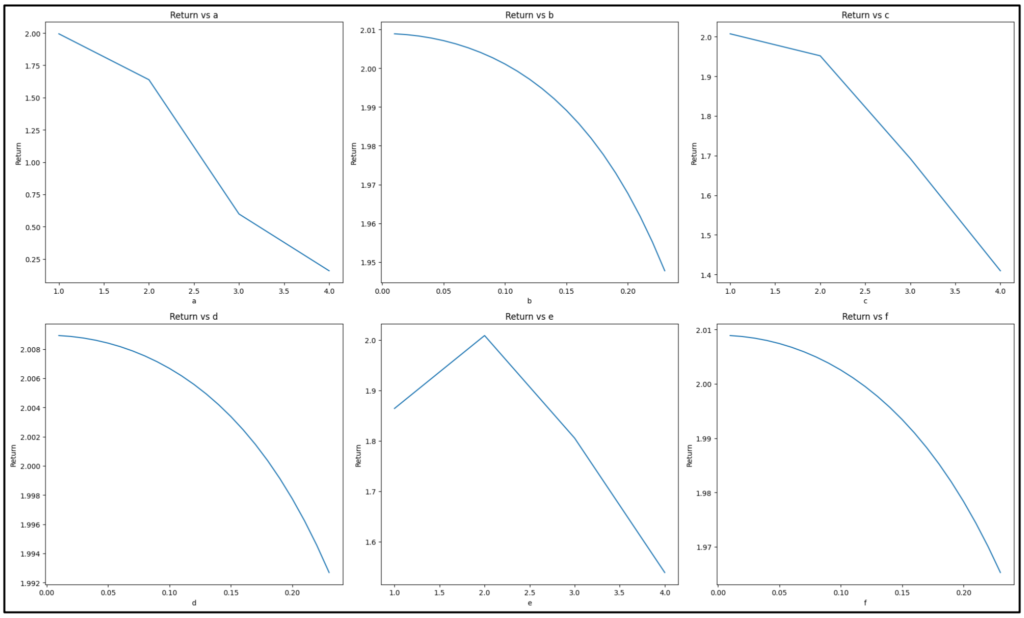


Figure 4.8: Parameter sensitivity graphs for portfolio A, Logistic distribution and Order 1.

The top left graph will be μ_1 followed by s_1 to the right, and μ_2 and so on. Both the μ_r and s_r of all stocks 1, 2, and 3 showed a similar curve, where the best parameters with the highest terminal wealth achieved at the lowest possible value of the parameters, except for μ_3 at 1.99. The allocation details given the best parameters can be shown in Table 4.16.

Table 4.16: Range of allocation for Portfolio A, Logistic distribution and Order 1.

Portfolio A	**Parameters**	**Allocation**			**BCRP**
	$[\mu_r, s_r]$	**Maximum**	**Minimum**	**Range**	
PDZH.KL	[1.14, 0.001]	0.220305	0.220305	0.000000	0.241
PBAH.KL	[1.32, 0.001]	0.263258	0.263258	0.000000	0.431
MISC.KL	[1.99, 0.001]	0.516438	0.516438	0.000000	0.328

From Table 4.16, the allocation of the logistic distribution universal portfolio was close to zero at a given weight determined by the Bayesian Optimization technique. Logistic distribution asserts more weight on MISC.KL which aligned with Order 1, Weibull distribution Portfolio A, but differs from it by allowing weights on PDZH.KL and PBAH.KL.

4.4.3 Portfolio B, Generalized Pareto Distribution

The GPD distribution had the highest return in order 1 for portfolio B. The change in parameters with respect to terminal wealth was plotted by holding the other parameters constant at their best rate, which can be shown in Figure 4.9.

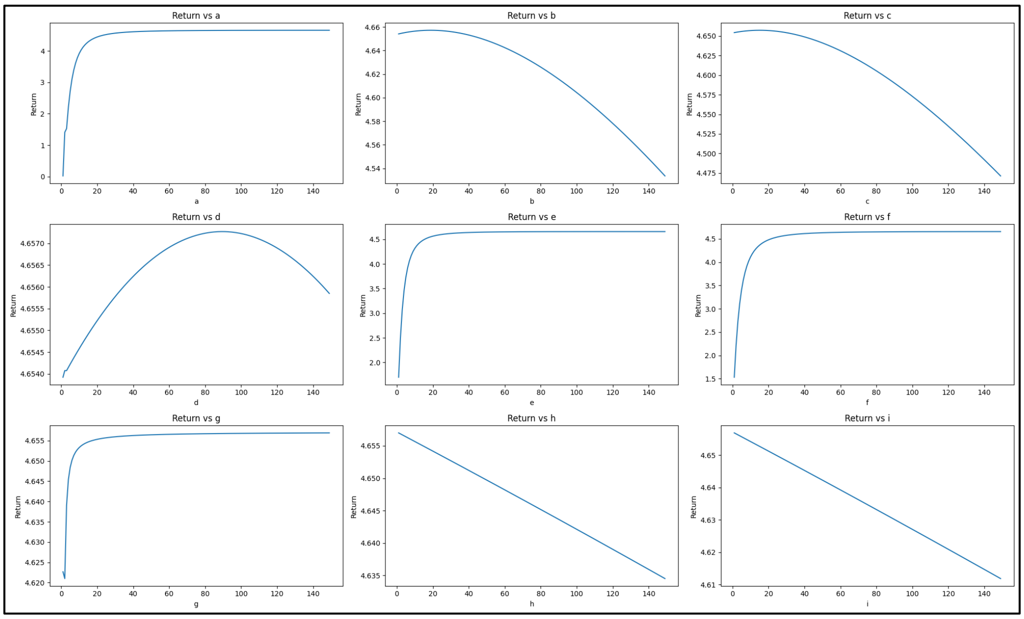


Figure 4.9: Parameter sensitivity graphs for portfolio B, GPD and Order 1.

The top left corner graph will be α_1, followed by $\lambda_1$to the right of the graph, and $k_1$to the right and so on. From the graphs, the α_1, λ_2, k_2 and α_3 exhibit the same pattern as the value of these parameters increases, the return generated will increase as well but will remain constant when it reaches the highest wealth achievable. Besides, the parameters λ_2 and λ_3 shows a decreasing trend as the parameter value increases. Hence, the parameters that will maximize the portfolio's returns were determined, and they are shown in Table 4.17.

Table 4.17: Range of allocation for Portfolio B, GPD and Order 1.

Portfolio B	**Parameters**	**Allocation**			**BCRP**
	$[\alpha_r, \lambda_r, k_r]$	***Maximum***	***Minimum***	***Range***	
KENH.KL	[150,25,20]	0.025794	0.025764	0.000030	0.009
TAFI.KL	[122,103,147]	0.974105	0.974075	0.000030	0.990
GENT.KL	[118,2,1]	0.000131	0.000131	0.000000	0.001

The generalized Pareto distribution asserted more weight than BCRP on the KENH.KL that have the highest price vector count equal to 1 has similar weight with BCRP where GENT.KL at minimal allocation. Additionally, the range in change of allocation is relatively small as well.

4.4.4 Portfolio B, Lognormal Distribution

Lognormal distribution generates the lowest terminal wealth in portfolio B. The graphs of parameter sensitivity can be shown in Figure 4.10.

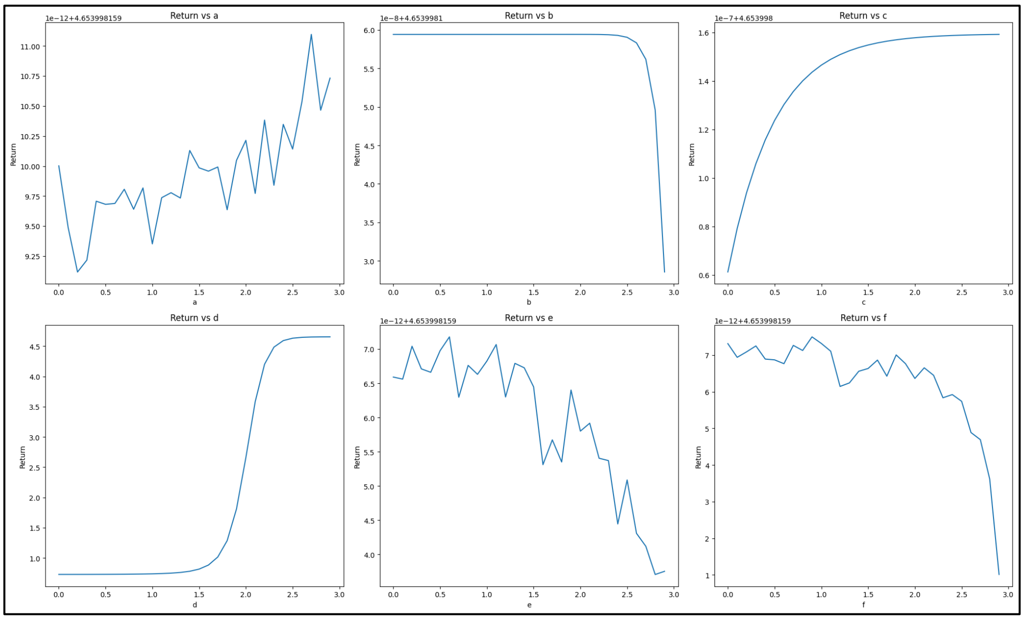


Figure 4.10:Parameter sensitivity graphs for portfolio B, Lognormal distribution and Order 1.

The leftmost graph shows μ_1 followed by σ_1 to the right, and μ_2 and so on. None of them possessed similar patterns among each other and fluctuations on μ_1, μ_3, σ_3 with respect to terminal wealth can be shown. The best parameters were determined and can be shown in Table 4.18.

Table 4.18: Range of allocation for Portfolio B, Lognormal distribution and Order 1.

Portfolio B	Parameters $[\mu_r, \sigma_r]$	Allocation			BCRP
		Maximum	Minimum	Range	
KENH.KL	[4.86, 0]	0.000000	0.000000	0.000000	0.009
TAFI.KL	[4.90, 0]	1.000000	1.000000	0.000000	0.990
GENT.KL	[0, 0.36]	0.000000	0.000000	0.000000	0.001

The allocation preference was the same as the Portfolio A, Order 1 Weibull distribution, where the universal portfolio consists of only one stock. The lognormal distribution asserted all of its weight on TAFI.KL has the highest terminal wealth generated.

4.4.5 Portfolio C, Gamma Distribution

The parameter sensitivity graphs of the Gamma distribution in portfolio C can be shown in Figure 4.11.

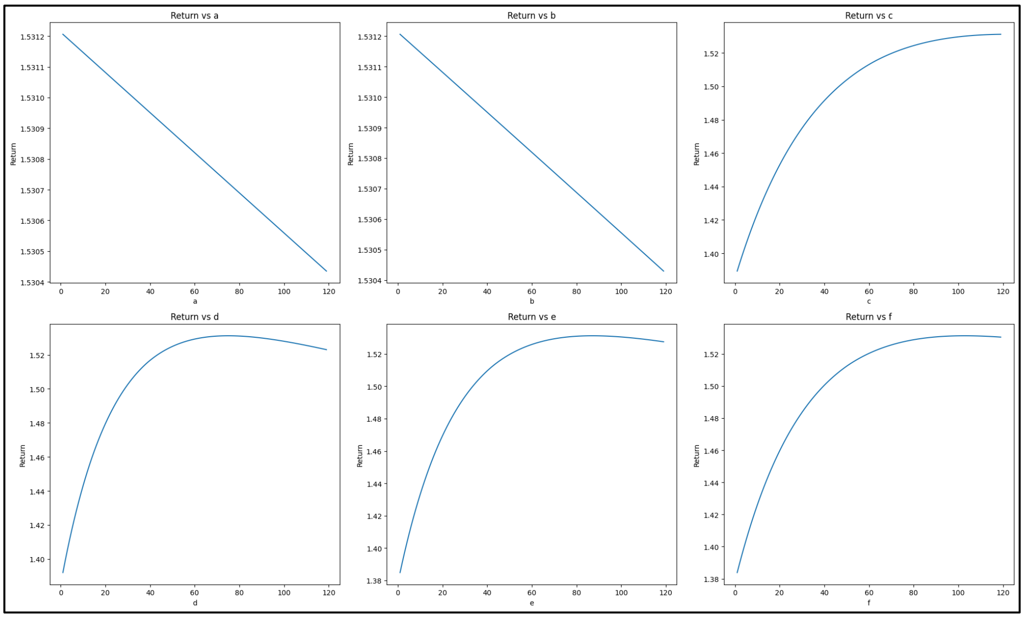


Figure 4.11:Parameter sensitivity graphs for portfolio C, Gamma Distribution and Order 1.

The graphs started from α_1, followed by β_1 and α_2 and so on to the right. Similar linear decreasing pattern can be found in α_1 and β_1, where the as the parameter value increase, the terminal wealth decreases. Additionally, $\alpha_2, \beta_2, \alpha_3,$ and β_3 have similar pattern as well. As the parameter values increase, the terminal wealth increases to a certain level and is held constant afterward. The best parameters that provide the highest terminal wealth can be shown in in Table 4.19.

Table 4.19: Range of allocation for Portfolio C, Gamma distribution and Order 1.

Portfolio C	**Parameters**	**Allocation**			**BCRP**
	$[\boldsymbol{\alpha_r}, \boldsymbol{\beta_r}]$	**Maximum**	**Minimum**	**Range**	
HAPS.KL	[0.1, 0.1]	0.000001	0.000001	0.000000	0.000
KPJH.KL	[120, 73.86]	0.490264	0.490011	0.000253	0.487
TROP.KL	[88.15, 104.41]	0.509988	0.509736	0.000253	0.513

The weight allocation shows that the HAPS.KL has the lowest weight, with approximately zero allocation and no change in weight. Hence, this universal portfolio will focus on the KPJH.KL and TROP.KL stocks only. Gamma distribution weights are aligned with the BCRP's weights but with slight changes in allocation for KPJH.KL and TROP.KL.

4.4.6 Portfolio C, Lognormal Distribution:

The lognormal distribution had the second lowest terminal wealth gain in portfolio C. Figure 4.12 shows the graph of parameter sensitivity testing.

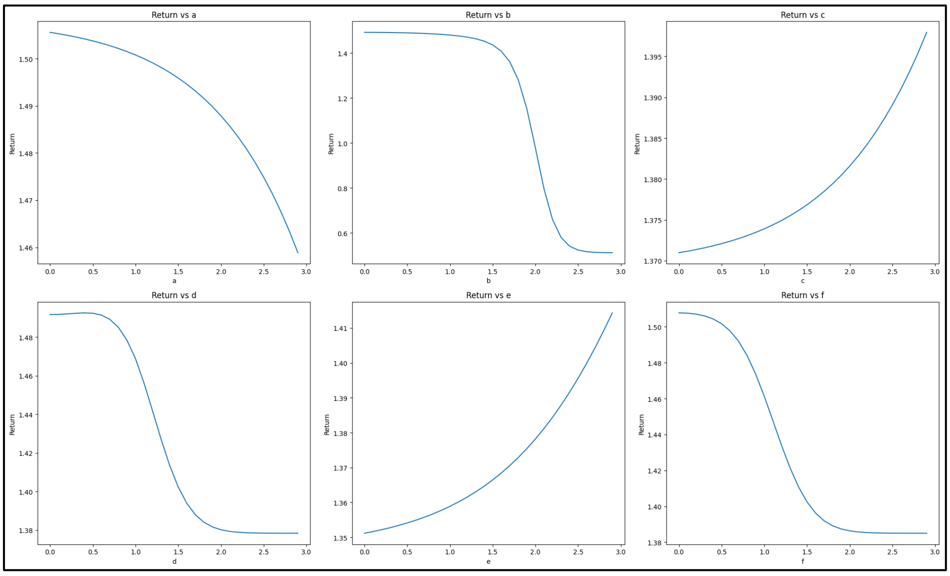


Figure 4.12: Parameter sensitivity graphs for portfolio C, lognormal Distribution and Order 1.

From Figure 4.12, the μ_2 and μ_3, σ_2 and σ_3 had a similar pattern in parameter value and terminal wealth, respectively. This distribution's highest terminal wealth will have a minimum value at $\mu1$, $\sigma1$, $\sigma2$, and $\sigma3$ and a maximum value at μ_2 and μ_3. The best parameters and allocation details are shown in Table 4.20.

Table 4.20: Range of allocation for Portfolio C, Lognormal distribution and Order 1.

Portfolio C	**Parameters**	**Allocation**			**BCRP**
	$[\mu_r, \sigma_r]$	**Maximum**	**Minimum**	**Range**	
HAPS.KL	[1.79, 0.20]	0.015018	0.014794	0.000224	0.000
KPJH.KL	[5.00, 0.00]	0.363417	0.358015	0.005402	0.487
TROP.KL	[5.00, 0.70]	0.627191	0.621565	0.005626	0.513

The lognormal distribution does not assert weight to be exactly zero for HAPS.KL as shown in Portfolio B, but it still possesses dominance over the

KPJH.KL and TROP.KL stocks in the universal portfolio. Lognormal distribution had a higher allocation range than the gamma distribution as well.

4.5 Parameter Sensitivity Testing among Portfolio C

This section will discuss the parameter sensitivity testing in portfolio C for gamma, lognormal, logistic, and weibull distributions in order 1.

4.5.1 Gamma Distributions

Gamma distribution in portfolio C, order 1, has the highest terminal wealth obtained. The change of parameters of the first stock, α_1 and β_1 was studied to determine terminal wealth, and its effect on other stock allocations can be shown in Figure 4.13. The best terminal wealth was acquired when both the α_1 and β_1 were around 0 and the terminal wealth had an inverse distribution with the maximum allocation of the first stock, indicating that increasing allocation will lead to decrement in terminal wealth. The terminal wealth, the maximum allocation of the second and third stock shows the same distribution, inferring that both contribute to the terminal wealth. Additionally, adjusting the parameter of α_1 and β_1 from 0 to 120 enables the stock to maximize its allocation until 0.5.

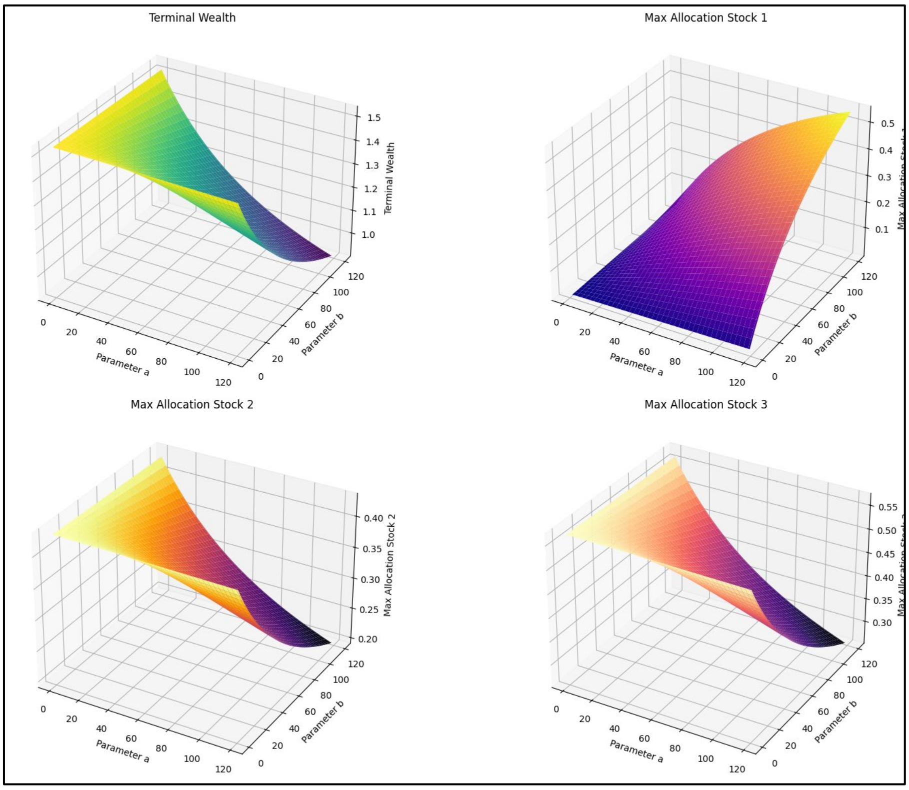


Figure 4.13: Portfolio C, Gamma distribution Parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\alpha_2, \beta_2, \alpha_3, \beta_3$ at constant.

Next, the parameter changes of α_2 and β_2 were studied and can be shown in Figure 4.14. The parameter sensitivity graph shows that maximum terminal wealth was obtained when the α_2 and β_2 fall around the yellow shaded area which is (x, y) = (80,120) or (120,80), maximizing both parameters will lead to a decrease in terminal wealth. Besides, the maximum allocation of the first and second stock had the same distribution but with a different magnitude, the first stock allocation z-axis is applicable for multiplier 10^{-6} as the best terminal wealth suggests the first stock to have minimum allocation.

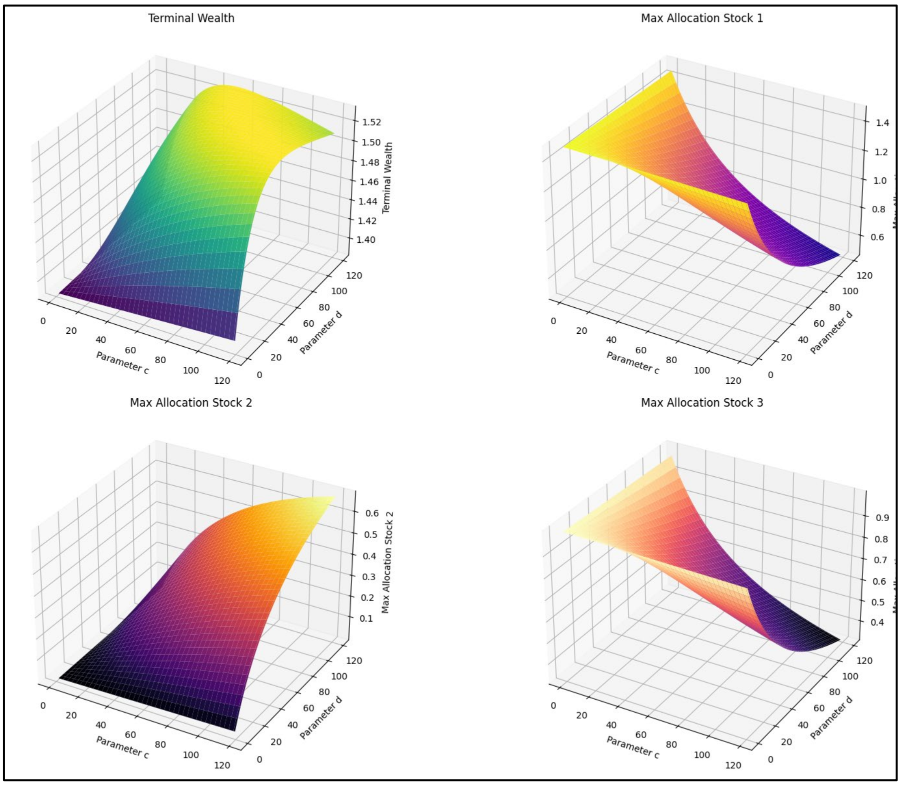


Figure 4.14: Portfolio C, Gamma distribution Parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\alpha_1, \beta_1, \alpha_3, \beta_3$ at constant.

Figure 4.15 shows the effects of the maximum allocation and terminal wealth of change in the parameter α_3 and β_3. A similar distribution and connection can be found in Figure 4.14, where the parameters that will provide the maximum terminal wealth were around (x, y) = (70,110) or (110,70). Besides, over-emphasizing the third stock will also lead to a decrement in terminal wealth, and the allocation of the first stock will be applied to multipliers 10^{-6} as well. Both Figures 4.14 and 4.15 suggested that the change in parameters in the second and third stock will not greatly affect the allocation of the first stock as its best parameter value was too small.

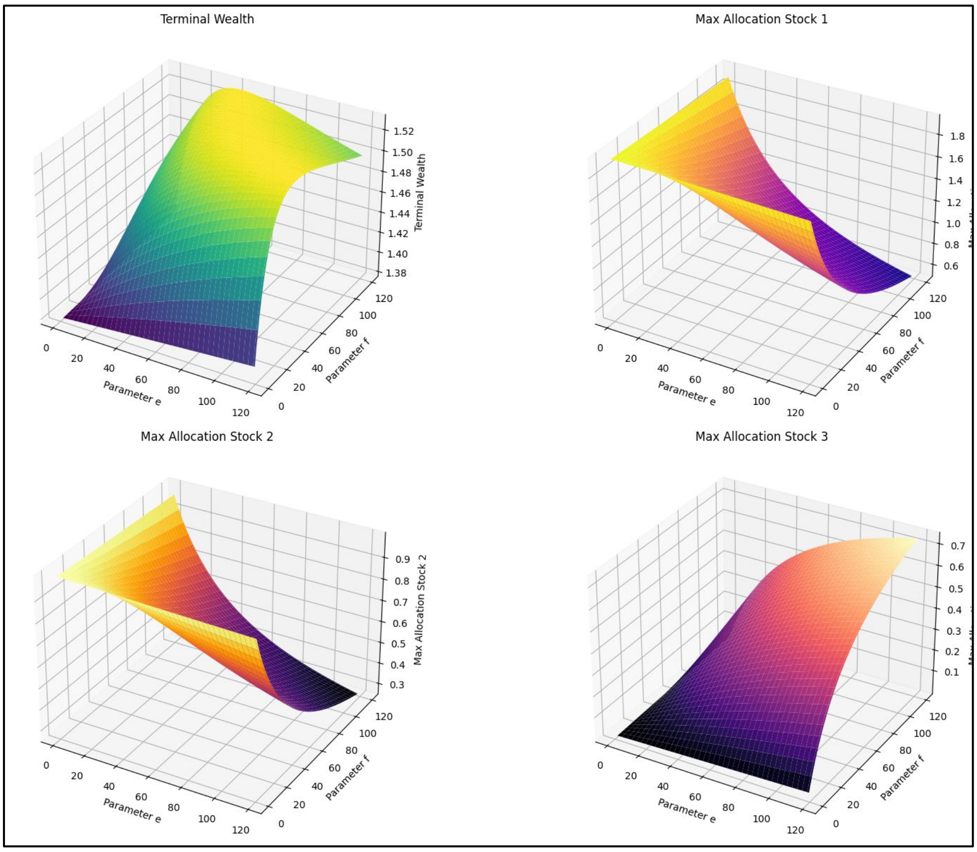


Figure 4.15: Portfolio C, Gamma distribution Parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\alpha_1, \beta_1, \alpha_2, \beta_2$ at constant.

The universal portfolio generated by recursive calculation of moment generating function of continuous random variable could change the allocation of stock with parameters. Hence, the parameter changes of gamma distribution regarding the range of allocation can be shown in Figure 4.16. The range of allocations for the second and third companies had similar distribution in the three stock parameter changes. In the first graph where $\alpha_2, \beta_2, \alpha_3, \beta_3$ were constant, the increase in the range of allocation in the first stock will have an inverse relationship with both the range of allocation for the second and third stock.

The maximum range of allocation of the second and third companies in last two graphs were when the α_2 and α_3 at around 20 to 40 and $\beta_2$and $\beta_3$at 110 to 120. However, both graphs suggest that the highest allocation range will not bring the highest terminal wealth, as transaction costs were also one of the considerations when determining the best parameters. Additionally, the range of allocation of the first stock in the last two graphs has an uncertain distribution but has the same peak distribution at lower alpha and higher beta, the same as the other stocks. The same pattern for second and third stock, suggesting they will possess a similar range of allocation at any parameter, but not the first as it was applicable for multipliers 10^{-8} and 10^{-10} respectively. In short, the combination of both α_r and β_r in all stocks was determinant in portfolio weightage allocation, higher values on both parameters will have a higher allocation amount. The lower the value of α_r and the higher the value of β_r will provide a higher range of allocation.

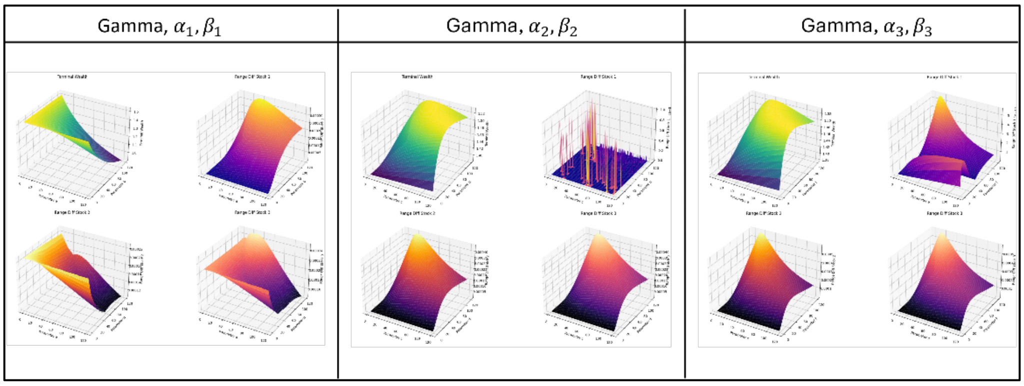


Figure 4.16: Portfolio C, Gamma distribution parameter testing w.r.t range of allocation and terminal wealth.

4.5.2 Lognormal Distribution

The graphs of parameter changes regarding the maximum allocation and the terminal wealth can be shown in Figures 4.17 to 4.19, and lognormal distribution was the second worst performer in portfolio C, order 1.

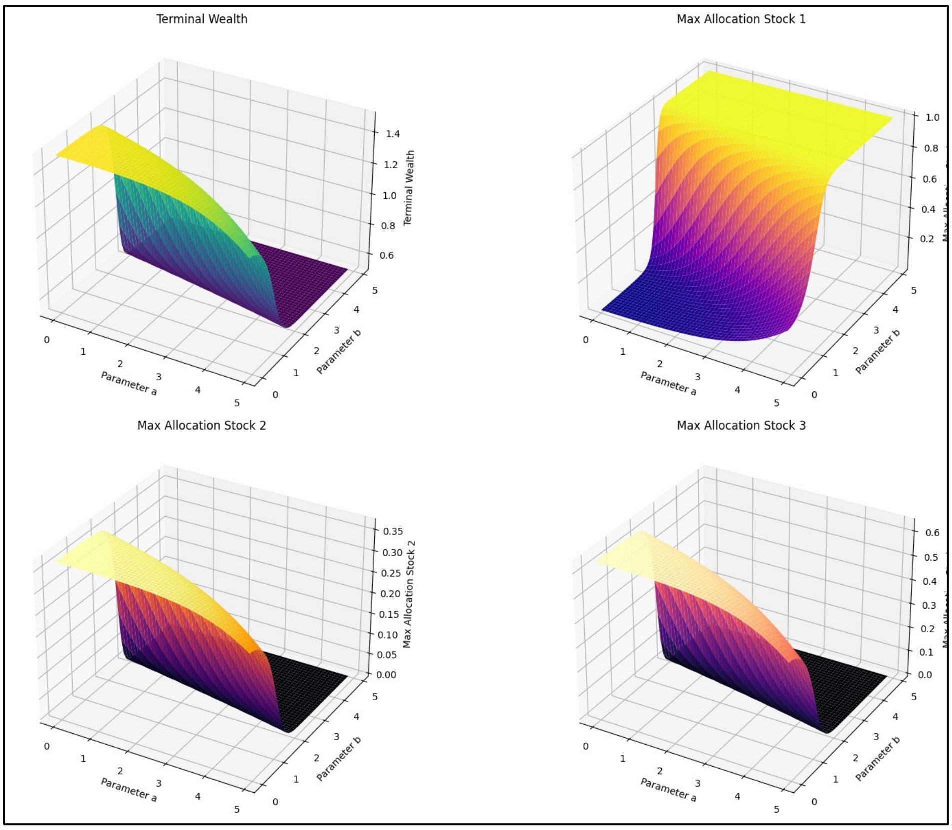


Figure 4.17: Portfolio C, Lognormal distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\mu_2, \sigma_2, \mu_3, \sigma_3$ at constant.

Similar to the gamma distribution where the terminal wealth, second stock and third stock in the portfolio exist a similar pattern when the μ_1 and σ_1 changed, and the highest terminal wealth was obtained at both μ_1 and σ_1 close to zero. However, ranging the parameters from 0 to 5 enables the first stock to have full allocation at 1.0 in the portfolio, whereby differ from the gamma distribution that only allows allocation up to 0.5. The σ_1 value contribute more to the increment in allocation than the μ_1 value for lognormal distribution as the allocation is near 1.0 after the σ_1 reach 3 and beyond.

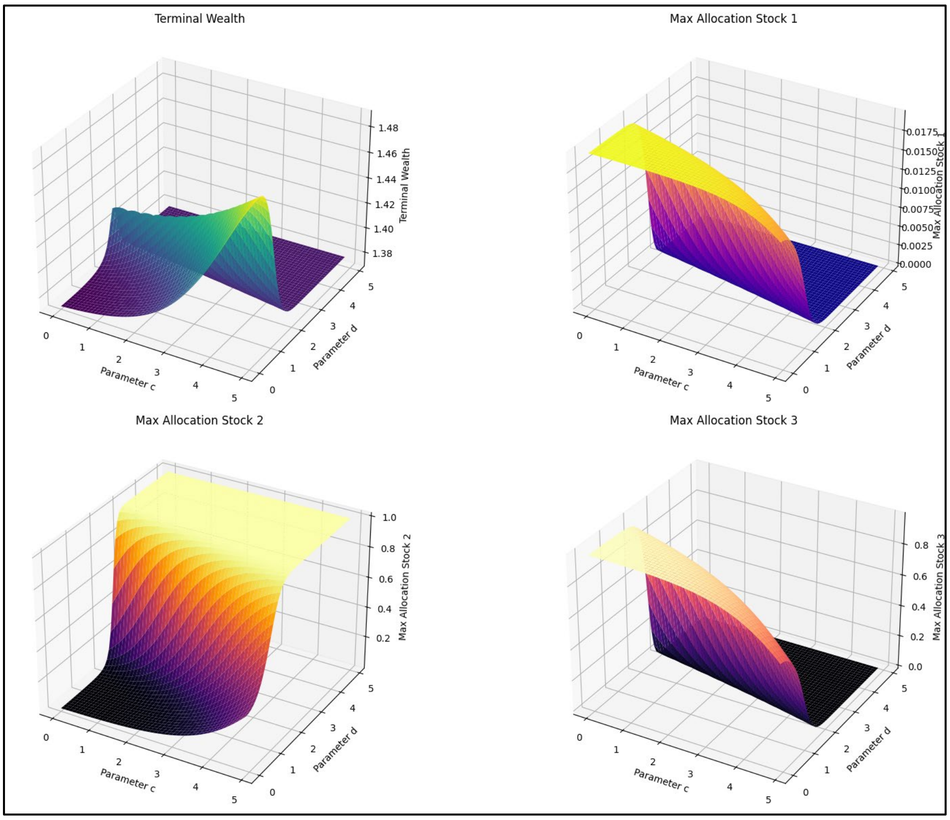


Figure 4.18: Portfolio C, Lognormal distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\mu_1, \sigma_1, \mu_3, \sigma_3$ at constant.

The terminal wealth in Figure 4.18 showed that there was a sudden increment when the σ_2 value is 2, and the terminal wealth stayed constant at 1.38 after value 3. The increment in μ_2 value will increase the terminal wealth given that the σ_2 less than 2.5. Additionally, the maximum allocation for first and second stock exhibit similar pattern with the previous section, but first stock at maximum allocation less than 0.0175. The decrease in allocation of second stock will be proportionally increase the allocation of third stock.

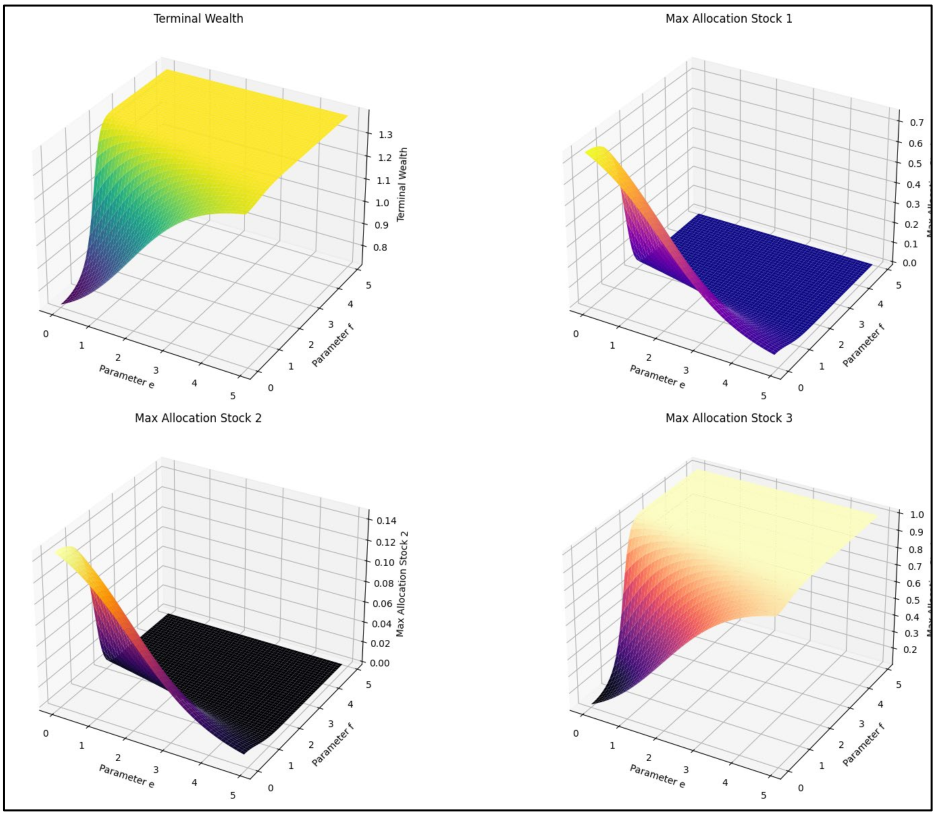


Figure 4.19: Portfolio C, Lognormal distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\mu_1, \sigma_1, \mu_2, \sigma_2$ at constant.

When the third stock's parameter changed, the decrease in allocation was greatly compensated by the increase in allocation of first stock. Besides, the σ_3 value at any point will produce the highest allocation especially beyond the value 2.5. The controlling variable on the allocation was highly dependent in the μ_3 value whereby varying it could change the allocation of third stock from 0.2 to 1.0. From Figure 4.19, the area of the highest terminal wealth was widely spread at 1.3, expect for σ_3 value less than 2 that will decrease the terminal wealth drastically. The range of allocation of lognormal distribution by varying the μ_r and σ_r parameters can be shown in Figure 4.20. All stocks in the portfolio shows similar distribution in range of allocation varying parameters where the highest range was obtainable at lowest μ_r value, but at a different σ_r value. Decreasing the third stock's μ_r will have a smooth decrement in range of allocation, but in first and second stock will have a wave of decrement, where the allocation range was still decreasing, but will be increased at certain μ_r

value. In short, the μ_r have strong control over the range of allocation and σ_r have strong control over the allocation amount.

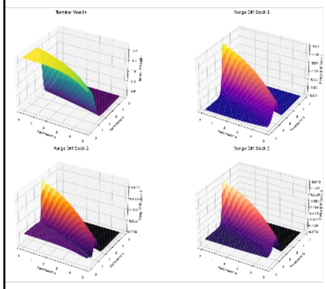
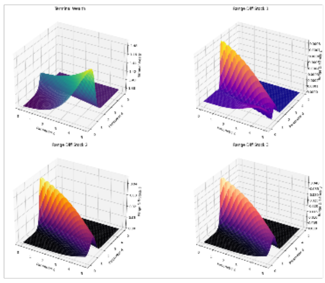
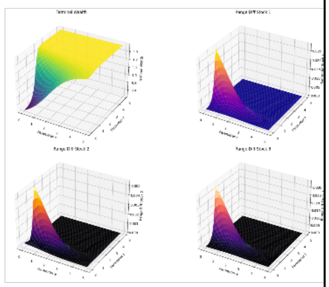
Lognormal, μ_1, σ_1	Lognormal, μ_2, σ_2	Lognormal, μ_3, σ_3
		

Figure 4.20: Portfolio C, Lognormal distribution parameter testing w.r.t range of allocation and terminal wealth.

4.5.3 Logistic Distribution

The second-best performer in portfolio C, order 1 was the logistic distribution, and the change in terminal wealth and maximum allocation of each stock can be shown in Figure 4.21 To 4.23.

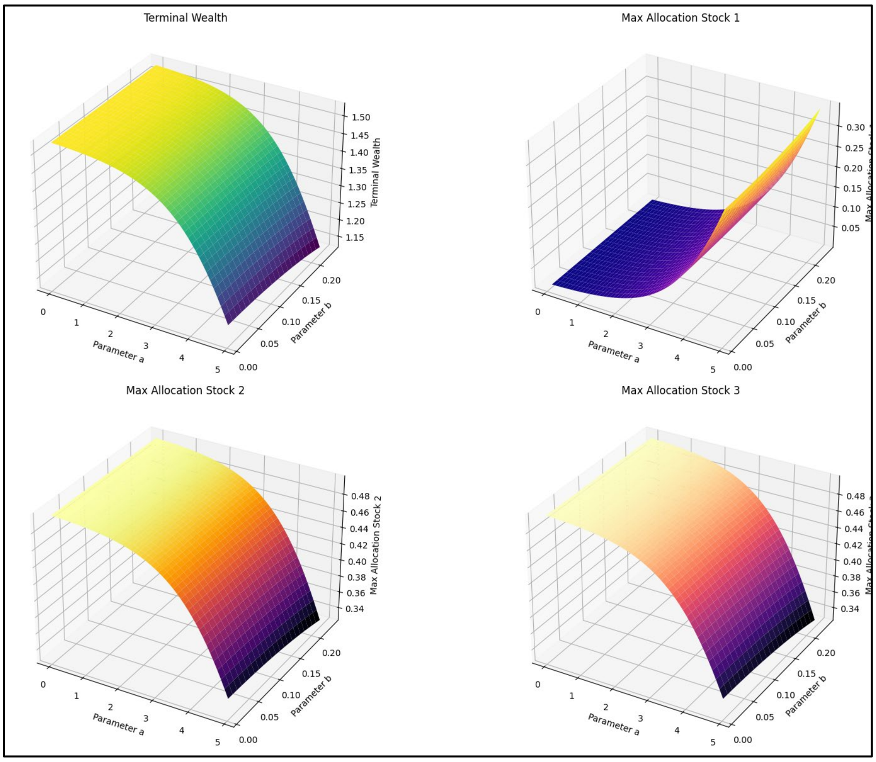


Figure 4.21: Portfolio C, Logistic distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding μ_2, s_2, μ_3, s_3 at constant.

Varying the first stock's parameters in the portfolio showed that the best terminal wealth was obtained when the μ_1 value was close to zero and varying the s_1 value will not change the terminal wealth obtained. The maximum allocation distribution of first stock will increase as the μ_1 value increases, regardless of the s_1 value. From Figure 4.21, the increase in allocation of first stock will evenly decrease the allocation of second and third stock.

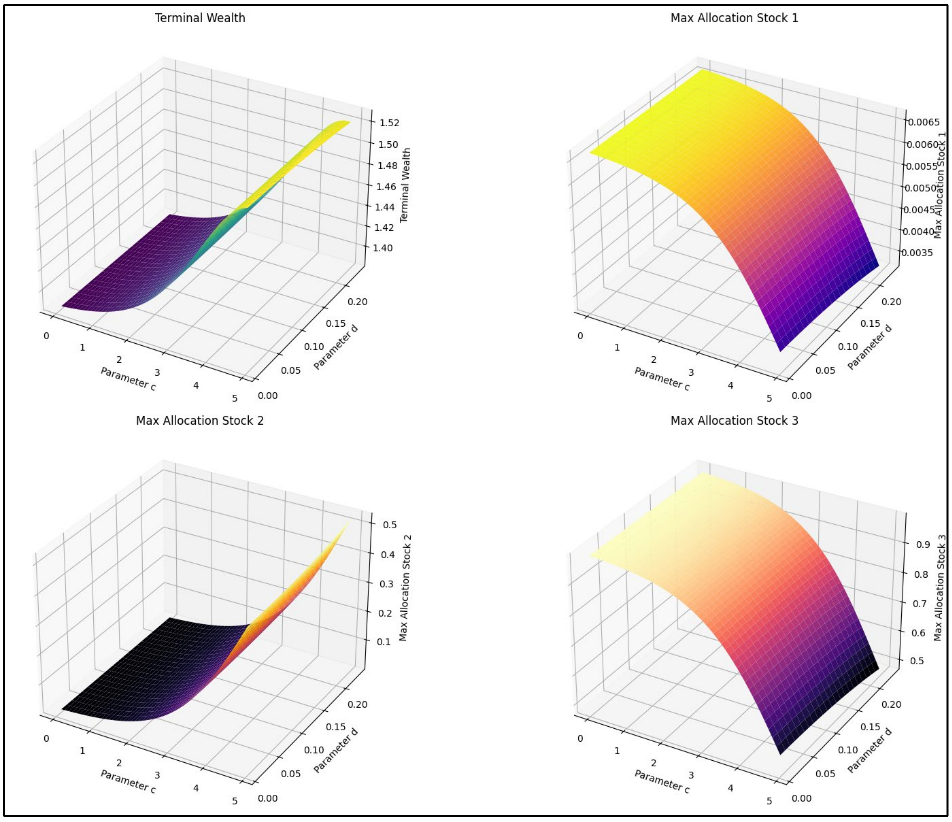


Figure 4.22: Portfolio C, Logistic distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding μ_1, s_1, μ_3, s_3 at constant.

In the second sets of parameters, the increase in allocation of second stock was greatly compensated by the decrease in allocation of third stock as the best μ_1 value was close to zero. Besides, the terminal wealth had positive relationship with the μ_2 value but varying the s_2 value will not affect the terminal wealth.

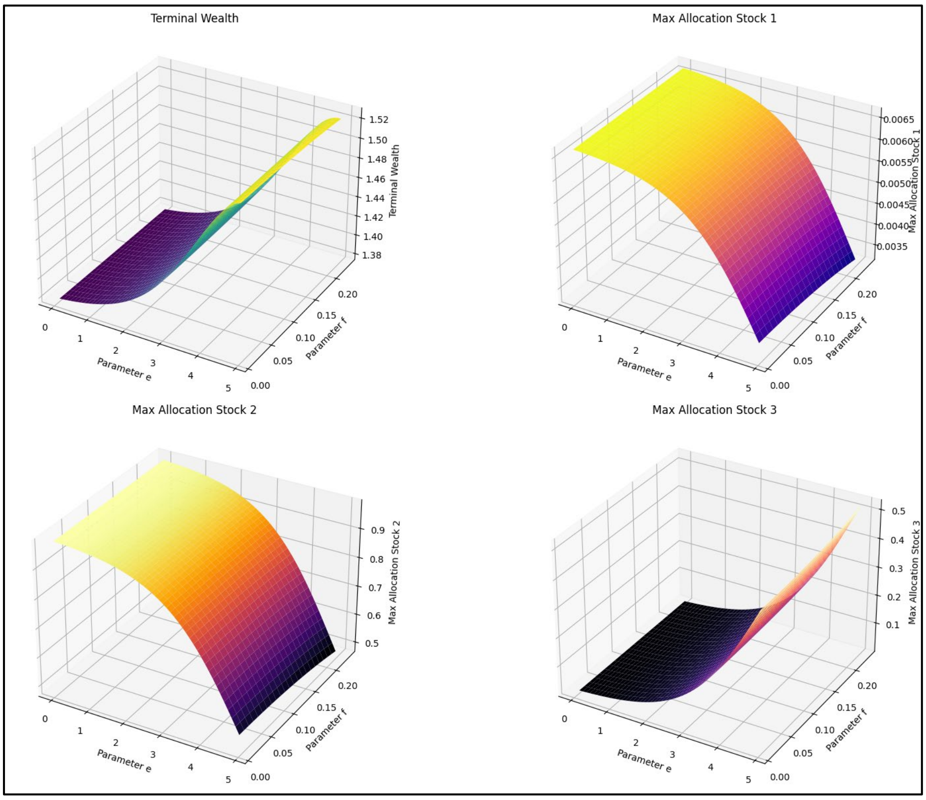


Figure 4.23: Portfolio C, Logistic distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding μ_1, s_1, μ_2, s_2 at constant.

In the third set of parameters, it exhibits similar patterns with the Figure 4.22, but with changes in allocation of third stock. The highest terminal wealth was determined when the μ_3 value was the highest, and regardless of the s_3 value too. Hence, from Figure 4.21 to 4.23, the higher the μ_r value the higher the allocation amount of the stock but it still limited to a certain amount as the μ_r value of other stock will affect the maximum obtainable allocation. Besides, the s_r value contribute no effect to the terminal wealth as well as the maximum allocation amount.

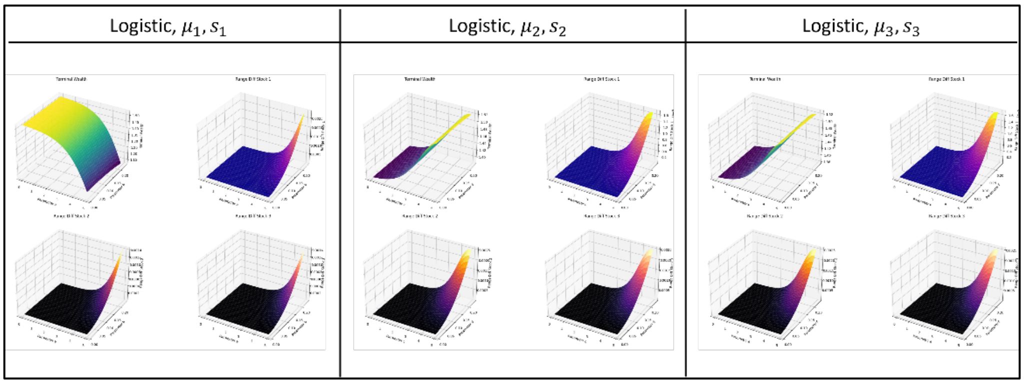


Figure 4.24: Portfolio C, Logistic distribution parameter testing w.r.t range of allocation and terminal wealth.

Figure 4.24 shows the range of allocation with different parameter values, and the change in the range of allocation was dependent on the s_r value. The higher the s_r value, the higher the range of allocation, but with criteria that the μ_r value must be greater than 4.9 for the first set of parameters, and 3.5 for the second and third set of parameters. In short, the μ_r value of logistic distribution will control the amount of allocation of the company and was also dependent on the other's stock μ_r value. Additionally, s_r value will be controlled over the range of allocation, but only applicable when the μ_r value is higher than a certain value.

4.5.4 Weibull Distribution

Weibull distribution has the lowest terminal wealth obtained in portfolio C, order1. The parameters of distribution included were β_r, N_r and were searched in the range from 0.1 to 70.0 for β_r, and 1 to 11 for N_r.

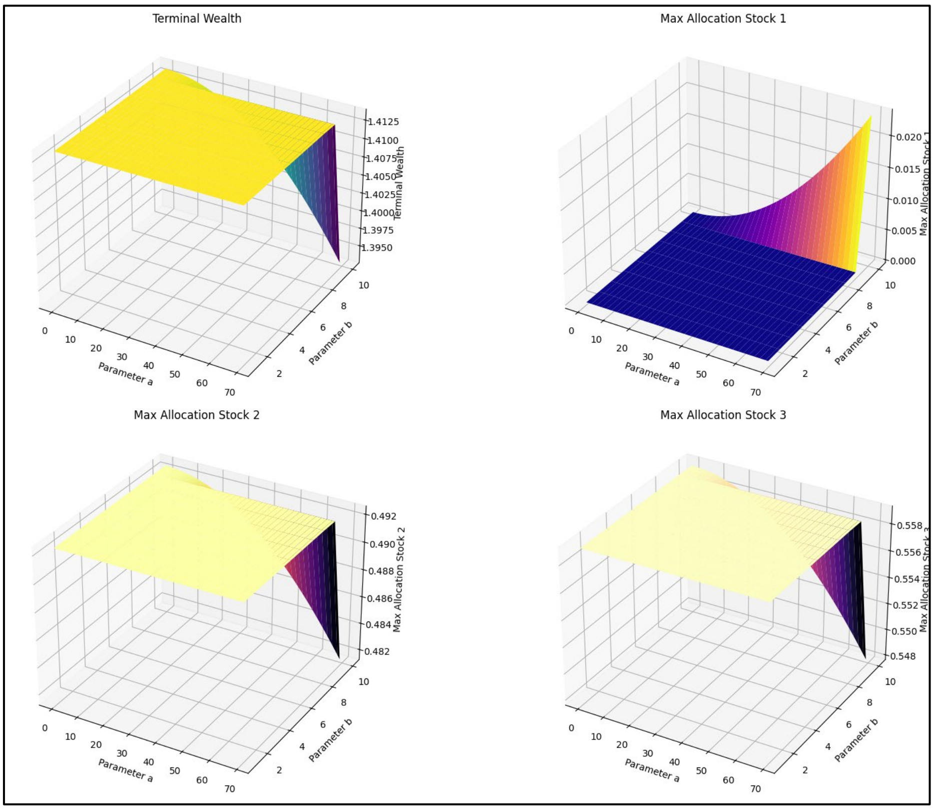


Figure 4.25: Portfolio C, Weibull distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\beta_2, N_2, \beta_3, N_3$ at constant.

Figure 4.25 showed that there was a wide range of highest terminal wealth of could be obtained by Weibull distribution except for the case where N_1 value was greater than 9. In overall, changing the parameters did not bring effect to the maximum allocation amount before the N_1 value of 9. After N_1 value reach 9, the highest allocation obtainable by first stock was only 0.020 and it will drop the terminal wealth to 1.3950. Besides, the increase in allocation of first stock will evenly decrease both allocation in second and third stock.

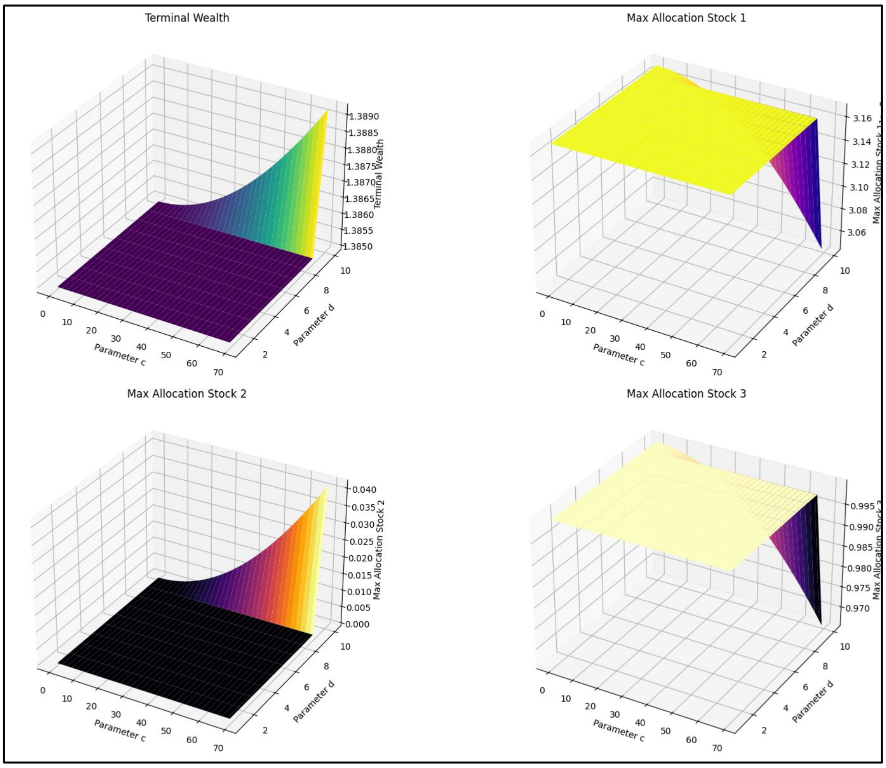


Figure 4.26: Portfolio C, Weibull distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\beta_1, N_1, \beta_3, N_3$ at constant.

Similar pattern of allocation could be obtained from Figure 4.26 that studied the parameters value of the second stock in the portfolio. The highest obtainable allocation was at 0.04. The change in allocation also only works after the N_1 value of 9. The wide area of non-changing allocation in all stocks of the portfolio suggests that the combination of parameters of all stocks was important. The increase in allocation in second stock will greatly affect the allocation on third stock, but not the first stock as it was applicable to multiplier of 10^{-6}.

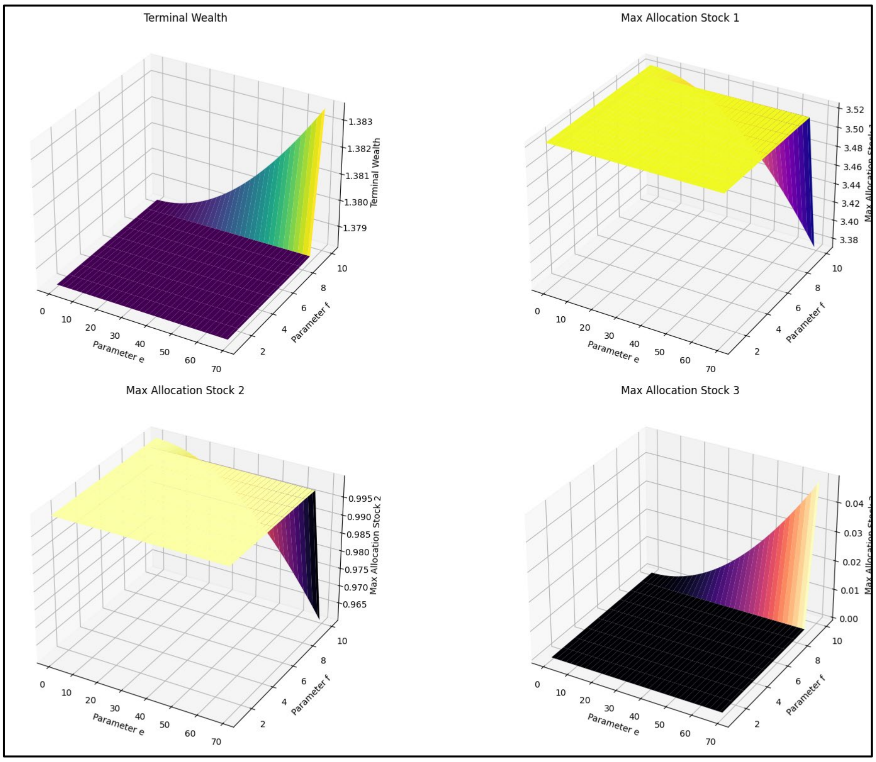


Figure 4.27: Portfolio C, Weibull distribution parameter sensitivity testing w.r.t allocation and terminal wealth by holding $\beta_1, N_1, \beta_2, N_2$ at constant.

Figure 4.27 shows a similar pattern of distribution in allocation and terminal wealth with Figure 4.26, where the highest changeable allocation only at 0.04, with N_3 value greater than 9. Besides, the increase in β_r value after N_r value greater than 9 will lead to increment in stock allocation exponentially. Three of the sets of parameters suggest that the combination of best parameters, especially the N_r value will greatly affect the allocation weightage of stock.

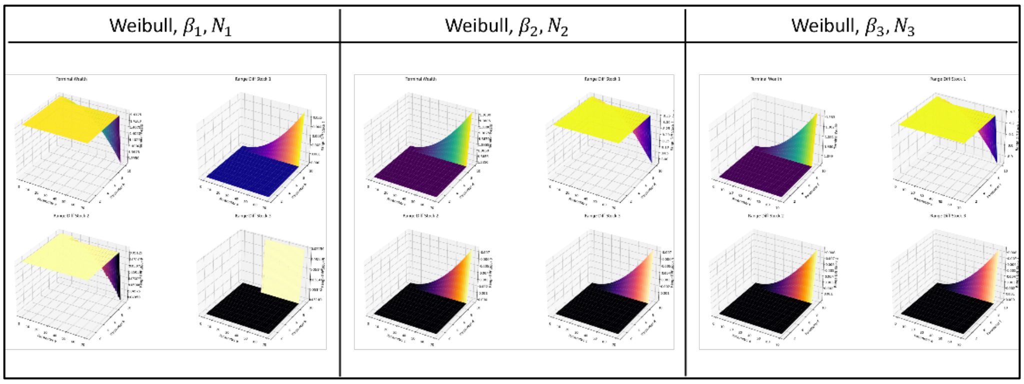


Figure 4.28: Portfolio C, Weibull distribution parameter testing w.r.t range of allocation and terminal wealth.

Figure 4.28 shows similar results from Figure 4.25 to 4.26, where the increment in stock allocation will also leads to increment in range allocation, but still only available after N_r value greater than 9. In short, the combination of sets of parameters for all stocks in the portfolio needs careful consideration as if a very high N_r value exists, the higher the other stock's N_r value is needed to vary the allocation amount.

4.5.5 Short Summary

This section will provide the changes in parameter value concerning the terminal wealth and allocation with a summary. In portfolio C, the first stock had the best allocation at 0 weight.

Table 4.21: Summary on allocation preferences with parameter value in Portfolio C.

Distribution	**Parameter VS Max. Allocation**	**Parameter VS Range of Allocation**	**Maximum Achievable Allocation**
Gamma	Direct Positive for both α_r and β_r, but β_r with higher determinant role.	The lower the value of α_r and the higher the value of β_r will provide a higher range of allocation.	Depends on combination of parameters.
Lognormal	σ_r have strong positive control over the allocation amount.	μ_r have strong positive control over the range of allocation.	Full control
Logistic	μ_r weak positive control over the allocation amount.	s_r value weak positive control over range of allocation	Depends on combination of parameters.

Weibull	N_r strong positive control over the allocation amount but very dependent on other N_r.	Same with maximum allocation preference.	Depends on combination of parameters.

4.6 Risk and Return among Portfolios

The Annual Percentage Yield (APY) and Annualized Standard Deviation (ASTDV) have been adopted to understand the risk and return associated with the universal portfolio generated by different orders and distributions. APY provides a comprehensive measure of the portfolio's return over a year, taking into account the effects of compounding interest. It serves as a critical benchmark for evaluating the profitability of the investment strategy, allowing investors to compare it against alternative strategies or market indices. On the other hand, ASTDV offers a detailed assessment of the volatility associated with the portfolio. By annualizing the standard deviation of returns, ASTDV provides insights into the risk inherent in the strategy, reflecting how much the portfolio's returns can be expected to fluctuate over time. Together, APY and ASTDV form a robust framework for analyzing the risk and return performance of the universal portfolio, balancing the potential rewards against the associated risks. This dual perspective enables investors to make more informed decisions, optimizing their portfolio for both growth and stability while considering the broader market conditions and their individual risk tolerance.

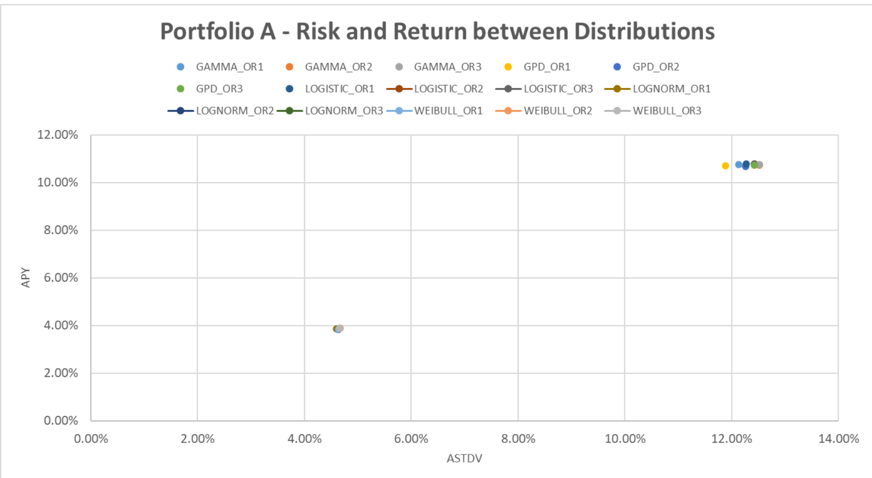


Figure 4.29: Performance comparisons between distributions for portfolio A.

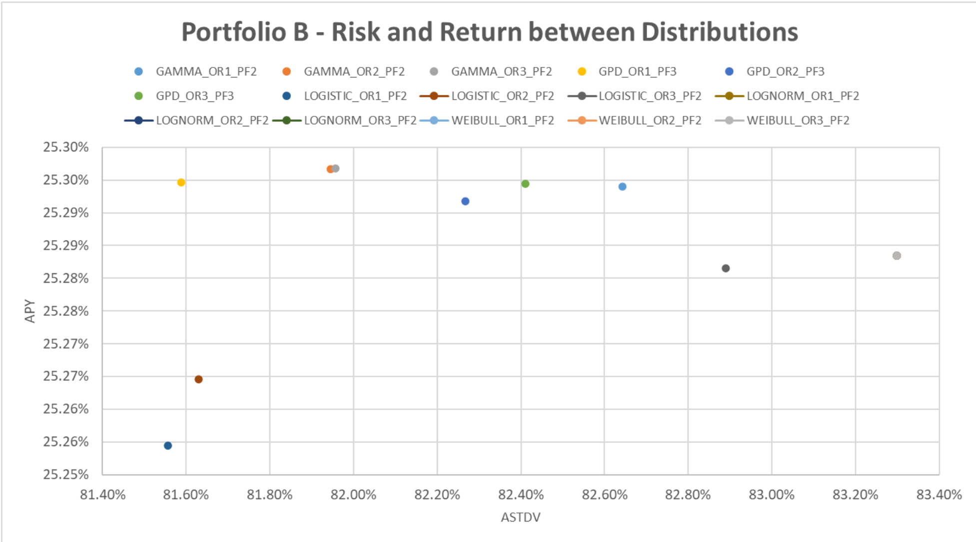


Figure 4.30: Performance comparisons between distributions for portfolio B.

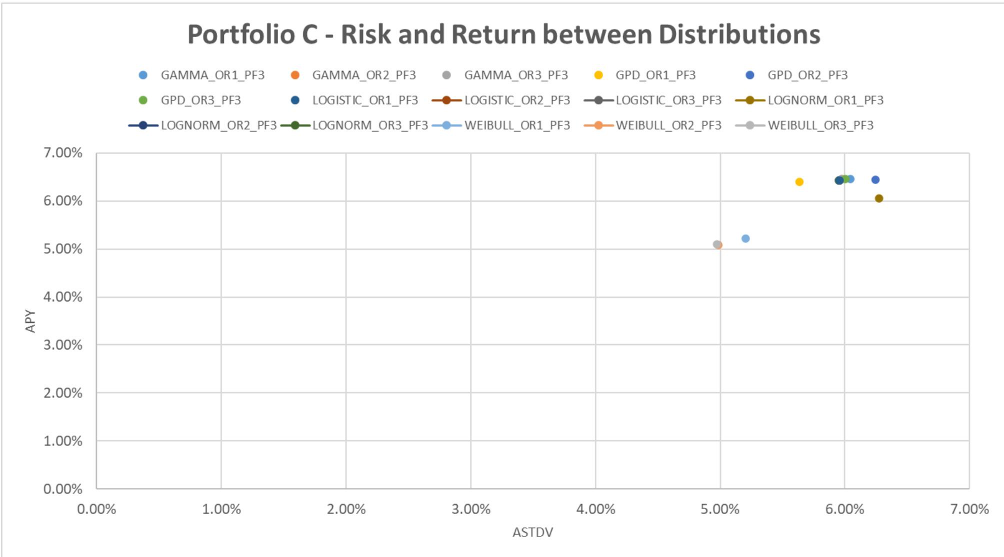


Figure 4.31: Performance comparisons between distributions for portfolio C.

From Figure 4.29 to Figure 4.31, the portfolio A and C have clusters in the top right corner, indicating a higher risk level compensated with a higher return. However, there are some distribution clusters at lower return and risk as well, such as the Weibull distribution in both of the portfolios and the lognormal distribution for portfolio A. Given the same level of return, the GPD of order 1 has the lowest level of risk, indicating better risk compensation. In portfolio A, the second lowest level of risk given the same level of return is the gamma distribution followed by logistic distribution, while portfolio C has the reverse sequence. Hence, GPD, gamma, and logistic distribution are ideal for investors who seek to maximize their profit given a lower risk tolerance level. In portfolio B, the highest level of return was achieved by gamma orders 2 and 3, at the second and third lowest level of risk among the distribution, respectively. The Weibull distribution will have the highest risk level at a relatively lower return. Additionally, the increase in orders increases the risk level in portfolio A but leads to uncertain movement in portfolios B and C. Nevertheless, the higher risk is usually tied with a higher return, careful consideration in selecting the distributions to build the universal portfolio based on the investor's risk tolerance was required.

4.7 Performance Comparisons

The best and worst performers on each portfolio will be compared against the Buy-and-Hold (BH), BCRP, Constant Rebalanced Portfolio (CRP), Cover's Universal Portfolio (CUP), and Successive Constant Rebalanced Portfolio (SCRP). The initial weight was taken 1/N, where N is the number of companies in the portfolio. The universal portfolio was constructed with a 1% transaction cost, while the others did not include any transaction cost.

Table 4.22: Performance Comparison Terminal Wealth for Portfolios.

Portfolio	**Terminal Wealth**					
	UP	**BCRP**	**BH**	**CUP**	**CRP**	**SCRP**
A (Best)	2.0106	**2.1502**	0.8559	1.5441	2.0190	0.5463
A (Worst)	1.2911					
B (Best)	4.6574	**4.9134**	2.1163	2.2188	2.2068	0.6028
B (Worst)	4.6467					
C (Best)	1.5311	**1.5433**	1.1025	1.1381	1.1476	0.4872
C (Worst)	1.4014					

All portfolios underperform the BCRP but were relatively close with it except portfolio B. The best and worst performers in portfolios B and C outperformed all other investment strategies, such as BH, CUP, CRP, and SCRP. For portfolio A, the best performer only outperformed the BH and SCRP strategies. The worst and best performers had a relatively close terminal wealth for portfolios B and C, but A with differences at 0.72.

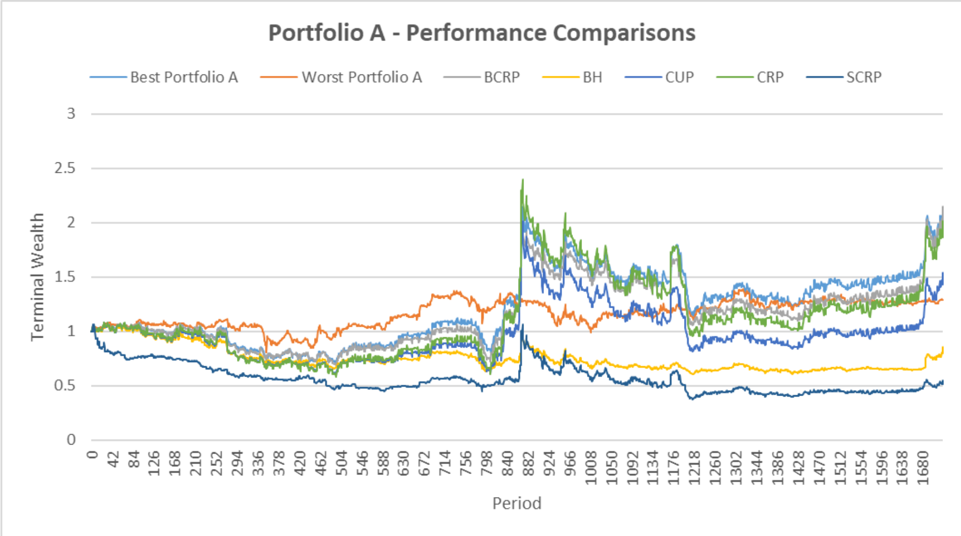


Figure 4.32: Performance comparisons between investment strategies for portfolio A.

From the time series graph, the best performer outperformed the BCRP return for the 1462 trading period and underperformed it on the last trading days. The worst performer in portfolio A was the best performer among other investment strategies before trading period 840th, as it does not capture the sudden increment in portfolio value at trading period 868th like others. The lowest volatility rate in portfolio A was BH, followed by the worst performer and SCRP. Next, the performance comparisons of portfolio B can be shown in Figure 4.33, where the best and worst performers overlapped with the BCRP return, and the best performers exceeded the BCRP return for 1516 trading days.

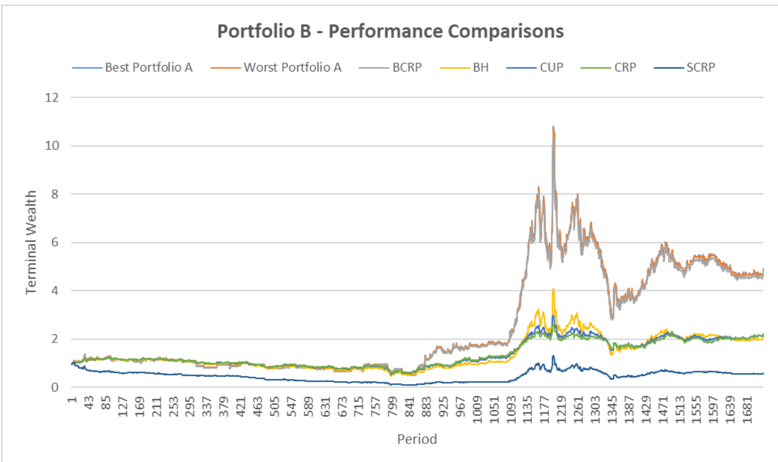


Figure 4.33: Performance comparisons between investment strategies for portfolio B.

Nevertheless, the BCRP, best, and worst performers all had allocations close to 1 for the TAFI.KL stock, so they exhibited similar movement in portfolio value. The volatility rate for BCRP is the lowest among the best, worst, and BCRP, followed by the worst performer and best performer.

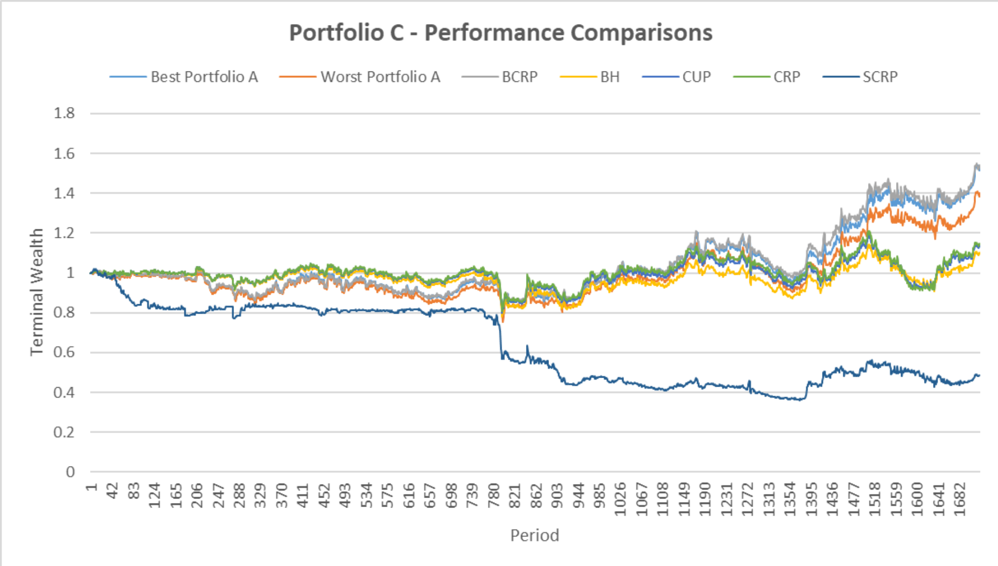


Figure 4.34: Performance comparisons between investment for portfolio C.

From Figure 4.34, the BCRP consistently outperformed both the best and worst performers in the portfolio. The best performer in this portfolio surpassed the BCRP only in 458 trading days. However, the worst performer had the lowest volatility, followed by the best performer and BCRP.

4.8 Risk and Return among Investment Strategies

The Annual Percentage Yield (APY) and Annualized Standard Deviation (ASTDV) have been adopted to understand the risk and return associated with the universal portfolio generated together with other investment strategies.

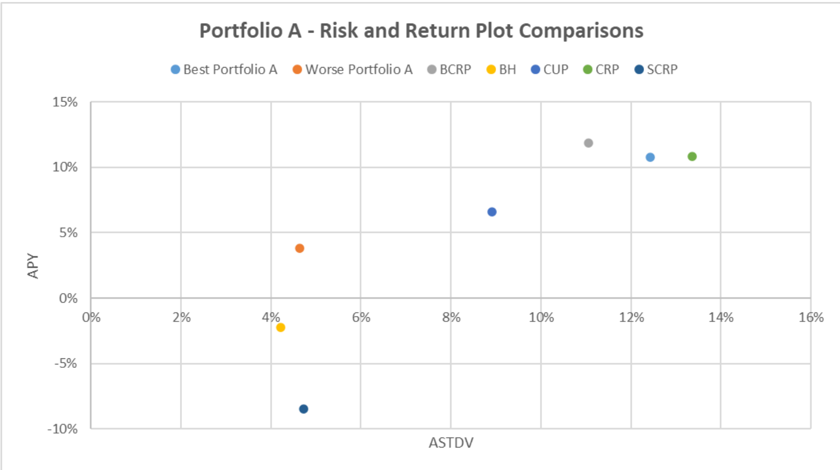


Figure 4.35: Risk and Return comparison for Portfolio A.

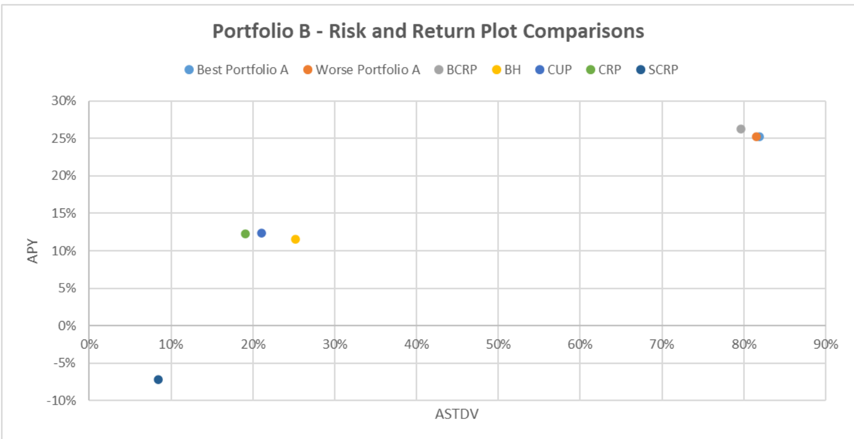


Figure 4.36: Risk and Return comparison for Portfolio B.

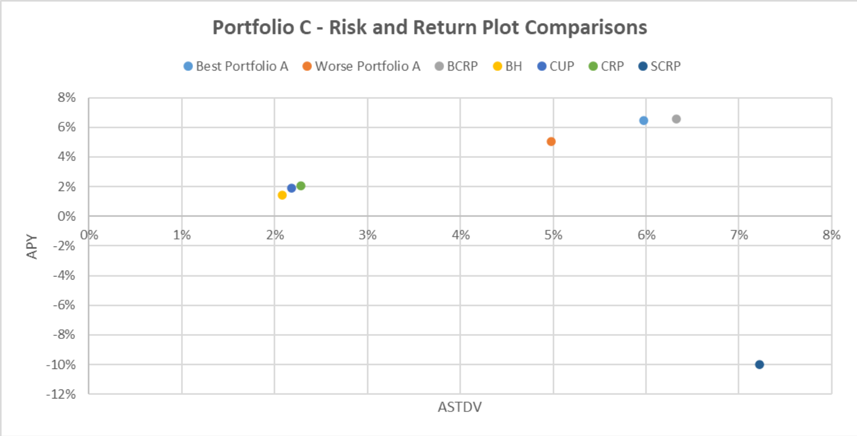


Figure 4.37: Risk and Return comparison for Portfolio C.

In portfolio A, the BCRP's risk is relatively lower than the universal portfolio generated by the recursive calculation of continuous random variables given the same level of APY. Similarly, portfolio B is also relatively close to BCRP risk and return and at a higher risk level given the same level of return. However, in portfolio C, the best performer has a relatively lower risk rate than the BCRP given the same level of return, providing a better risk compensation. In short, the universal portfolio generated by recursive calculation of continuous random variables provides a relatively close with BCRP's risk and return in portfolios A and B, and the best performer in portfolio C outperformed the risk compensation of BCRP.

4.9 Scenario Testing

4.9.1 Covid-19 Period

The universal portfolio was recalculated to study its performance generated by MGF of CRV during the COVID-19 period from 1 October 2019 to 1 July 2022. Table 4.22 shows the terminal wealth by different orders and distributions during the COVID-19 period.

Table 4.23: Terminal Wealth for Universal Portfolio.

Portfolio	**Order**	**Distributions**				
		Gamma	**Lognormal**	**Weibull**	**Logistic**	**GPD**
A	1	1.41789	1.06117	1.05689	**1.41822**	1.41350
	2	1.41728	1.05682	1.05704	1.41609	1.41526
	3	1.41798	*1.05681*	1.05688	1.41609	1.40139
BCRP		**1.4096**				
B	1	6.01892	**6.01893**	6.01893	*5.97242*	6.01872
	2	6.01892	6.01893	6.01893	5.98596	6.01892
	3	6.01892	6.01893	6.01893	6.01103	6.01888
BCRP		**6.2355**				
C	1	1.19313	**1.19313**	1.19313	*1.19071*	1.19311
	2	1.19313	1.19313	1.19313	1.19141	1.19312
	3	1.19313	1.19313	1.19313	1.19271	1.19296
BCRP		**1.2115**				

From Table 4.23, the gamma, logistic, and generalized Pareto distributions managed to surpass the BCRP terminal wealth in portfolio A, while the other portfolios, B and C, still did not surpass the BCRP. Besides, the terminal wealth by different distributions in portfolios B and C had a consistent value. Additionally, the order increment also leads to uncertain movement in terminal wealth generated in all distributions. The best and worst performers in each portfolio were the Lognormal and Logistic distributions, respectively, in portfolios B and C, but reserved in Portfolio A. However, by only focusing on the COVID-19 period, portfolio B surpassed the whole trading period shown in Table 4.9 by 1.32. Next, the best and worst performers in the portfolios were studied against the other investment strategies and can be shown in Figure 4.38. All the best performers have closely aligned with BCRP and surpassed the other investment strategies.

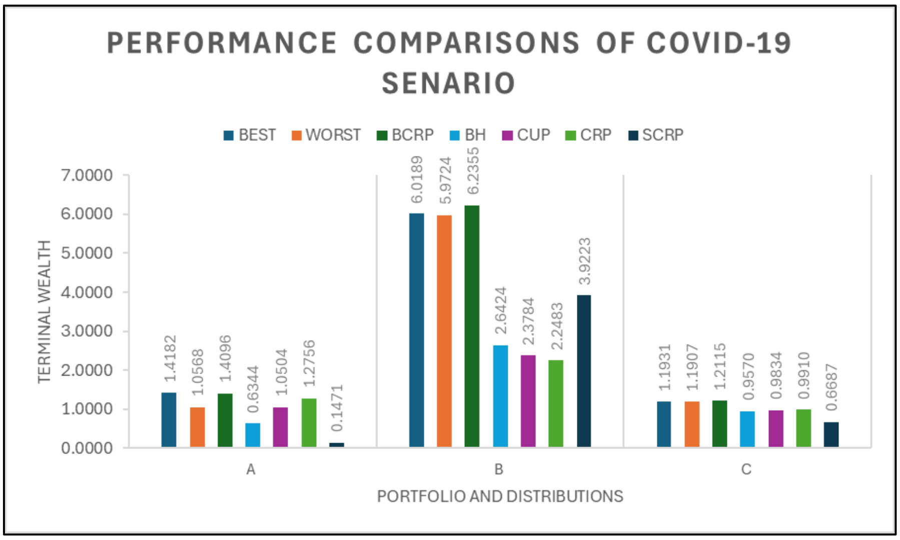


Figure 4.38: Best and Worst performer in each portfolio during the COVID-19 period's performance comparisons with other investment strategies.

4.9.2 Short-Term and Long-Term Trading

The universal portfolio introduced by Cover (1991) demonstrated that it would provide profitable returns in the long term regardless of the performance of the stocks chosen. Hence, studies on adding each 250-trading day to the universal portfolio were carried out.

Table 4.24: Terminal Wealth for Universal Portfolio A.

Period	**Order**	**Distributions**				
		Gamma	**Lognormal**	**Weibull**	**Logistic**	**GPD**
250	1	1.04817	1.04762	1.04762	1.04663	1.04798
	2	**1.04826**	1.04762	1.04762	1.04692	1.04793
	3	1.04826	1.04762	1.04762	1.04740	1.04817
BCRP		**1.04440**				
500	1	**1.03172**	1.03163	1.03163	1.02670	1.03174
	2	1.03172	1.03163	1.03164	1.02807	1.03180
	3	1.03172	1.03163	1.03164	1.03072	1.03163
BCRP		**1.01045**				
	1	**1.32729**	1.32729	1.32729	1.32519	1.32717

Period	Order	Distributions				
		Gamma	Lognormal	Weibull	Logistic	GPD
750	2	1.32729	1.32729	1.32729	1.32581	1.32728
	3	1.32729	1.32729	1.32729	1.32694	1.32725
BCRP		**1.30322**				
1000	1	1.87014	1.10794	1.12723	**1.87369**	1.85930
	2	1.87141	1.10661	1.10659	1.87252	1.85673
	3	1.87194	1.10659	1.10659	1.87252	1.86514
BCRP		**1.86623**				
1250	1	1.43206	1.21785	1.21796	1.43187	1.42828
	2	1.43243	1.21780	1.21801	1.43105	1.42863
	3	**1.43265**	1.21780	1.21786	1.43106	1.42726
BCRP		**1.44929**				
1500	1	1.64176	1.29150	1.29143	1.64280	1.63656
	2	1.64247	1.29113	1.29149	1.64099	1.62421
	3	**1.64282**	1.29113	1.29123	1.64099	1.63706
BCRP		**1.62031**				

Table 4.25: Terminal Wealth for Universal Portfolio B.

Period	Order	Distributions				
		Gamma	Lognormal	Weibull	Logistic	GPD
250	1	**1.19346**	1.18086	1.16669	1.19309	1.19332
	2	1.19346	1.16557	1.17597	1.19314	1.19300
	3	1.19345	1.16557	1.17587	1.19314	1.19313
BCRP		**1.17788**				
500	1	0.87359	0.86110	0.86042	0.87443	0.87391
	2	0.87368	0.86241	0.86030	0.87445	0.87418
	3	0.87366	0.84197	0.86030	**0.87445**	0.87382
BCRP		**0.86402**				
750	1	**0.96712**	0.94015	0.94017	0.96672	0.96588
	2	0.96705	0.91581	0.92653	0.96682	0.96540
	3	0.96699	0.91580	0.91602	0.96682	0.96479
BCRP		**0.95842**				

Period	Order	Distributions				
		Gamma	Lognormal	Weibull	Logistic	GPD
1000	1	1.74845	1.74835	1.74835	1.74740	**1.74882**
	2	1.74867	1.74835	1.74835	1.74770	1.74870
	3	1.74866	1.74835	1.74835	1.74823	1.74876
BCRP		**1.72920**				
1250	1	7.53917	**7.53918**	7.53918	7.46495	7.53917
	2	7.53917	7.53918	7.53918	7.48653	7.53910
	3	7.53917	7.53918	7.53918	7.52654	7.53911
BCRP		**6.92344**				
1500	1	5.39342	5.39342	**5.39342**	5.37043	5.39339
	2	5.39342	5.39342	5.39342	5.37721	5.39341
	3	5.39342	5.39342	5.39342	5.38957	5.39342
BCRP		**5.20079**				

Table 4.26: Terminal Wealth for Universal Portfolio C.

Period	Order	Distributions				
		Gamma	Lognormal	Weibull	Logistic	GPD
250	1	1.07631	1.07631	1.07631	1.07509	1.07631
	2	**1.07631**	1.07631	1.07631	1.07544	1.07631
	3	1.07631	1.07631	1.07631	1.07610	1.07631
BCRP		**1.08481**				
500	1	1.09761	1.09762	1.09762	1.09652	1.09761
	2	1.09761	**1.09762**	1.09762	1.09683	1.09761
	3	1.09761	1.09762	1.09762	1.09743	1.09761
BCRP		**1.09953**				
750	1	**1.10995**	1.10995	1.10995	1.10879	1.10994
	2	1.10995	1.10995	1.10995	1.10912	1.10995
	3	1.10995	1.10995	1.10995	1.10975	1.10995
BCRP		**1.11419**				
1000	1	1.00250	1.00259	0.97313	1.00260	1.00234
	2	1.00255	1.00232	0.99335	**1.00262**	1.00240
	3	1.00253	1.00232	0.99096	1.00262	1.00225

Period	Order	Distributions				
		Gamma	Lognormal	Weibull	Logistic	GPD
BCRP		**1.00472**				
1250	1	1.14668	1.13612	1.13612	1.14622	**1.14468**
	2	1.14667	1.13612	1.13629	1.14621	1.14644
	3	1.14666	1.13612	1.13619	1.14621	1.14340
BCRP		**1.15104**				
1500	1	1.46073	**1.46073**	1.46073	1.45781	1.46072
	2	1.46073	1.46073	1.46073	1.45866	1.46072
	3	1.46073	1.46073	1.46073	1.46023	1.46072
BCRP		**1.46813**				

Tables 4.24 to 4.26 show that the terminal wealth of the universal portfolio generated by MGF of different CRV provides comparable results with the BCRP, and portfolio B managed to surpass the BCRP terminal wealth in all trading periods. In contrast, portfolio C constantly underperforms the BCRP terminal wealth, and portfolio A underperforms the BCRP on the 1250 trading days period. An illustration of different distribution terminal wealth against trading days can be shown in Figure 4.39, where all distributions in portfolios B and C were moving in the same trend, but portfolio A with lognormal and Weibull distribution constantly underperformed after 750 trading days.

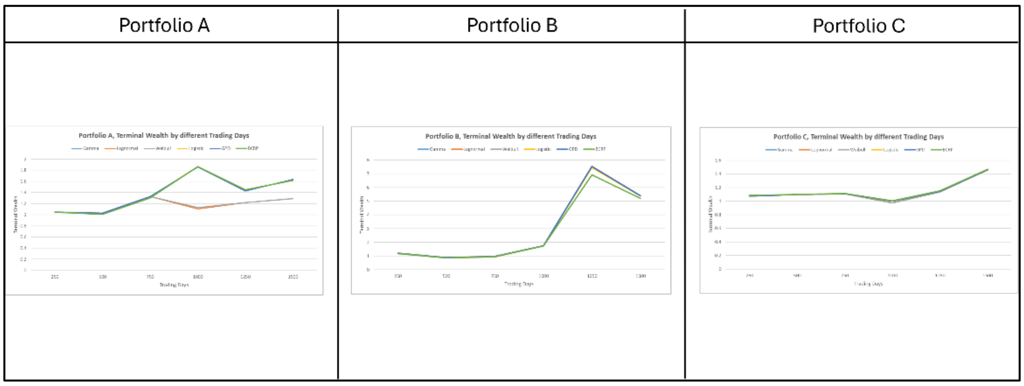


Figure 4.39: Terminal Wealth in each trading periods for different distributions of order 1, and for portfolio A, B and C.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This project empirically studies the universal portfolio generated by recursive calculation of the continuous random variable's moment function, with portfolios created by machine learning models by using the data from Refinitiv and Yahoo Finance. The portfolio formed through various machine learning models outperformed the KLSE return, with the highest risk-adjusted ratio obtained by the ridge regression model. Portfolios A, B and C each consist of three stocks that were formed with the best model to generate the universal portfolio. Universal portfolio A had the highest difference in terminal wealth among orders and distributions, while B and C had similar terminal wealth regardless of orders and distribution.

Parameter sensitivity testing was carried out to study the changes in allocation and terminal wealth concerning the change in parameter values. Portfolio A had different allocation preferences among different distributions hence generating a high difference between distributions. In Portfolio B, all distributions assert all allocation to one single stock, and evenly on the last two stocks in Portfolio C. Three-dimensional graphs were plotted on Portfolio C to study the changes in parameters value in detail where the lognormal distribution of a single stock parameter values have full control over the maximum allocation amount, while other distributions were dependent on the combination of parameters. From performance comparisons, all portfolios underperformed the BCRP terminal wealth but did outperform the BH, CUP, and SCRP in all portfolios.

Risk and return plots were conducted to give investors better insight into each distribution, order, and portfolio. Scenario testing showed that the UP generated by MGF of CRV could outperform the BCRP return in shorter trading periods in portfolios A and B and a relatively close terminal wealth in portfolio C. During the COVID-19 period, all portfolios still underperformed the BCRP return but remained comparable to it.

5.2 Recommendations for Future Work

The universal portfolio generated by recursive calculation of moment generating function of continuous random variables could provide a comparable terminal wealth with BCRP, as it would change the allocation over time due to observed performance in the past. Further studies on other distributions would be recommended to study the performance and allocation preferences. Besides, the best parameters could be determined by other strategies, such as the maximum likelihood function. Varying parameters in different trading periods were recommended, and a combination of distributions in one universal portfolio calculation was recommended as well.

REFERENCES

Black, F. and Scholes, M. (1973). ‘The pricing of options and corporate liabilities’, *Journal of Political Economy*, 81(3), pp. 637–654. https://doi.org/10.1086/260062

Blum, A. and Kalai, A. (1999). Universal Portfolio With and Without Transaction Costs. Machine Learning, 35, 193 205.

Cizeau, P., Liu, Y., Meyer, M., Peng, C.K. and Stanley, H.E., (1997). Volatility distribution in the S&P500 stock index. *Physica A: Statistical Mechanics and its Applications*, *245*(3-4), pp.441-445.

Cover, T.M. (1991). ‘Universal portfolios’, Mathematical Finance, 1(1), pp. 1–29. https://doi.org/10.1111/j.1467-9965.1991.tb00002.x

Cover, T.M. and Ordentlich, E. (1996). ‘Universal portfolios with Side Information’, IEEE Transactions on Information Theory, 42(2), pp. 348–363. https://doi.org/10.1109/18.485708

Gray, J. B., & French, D. W. (1990). EMPIRICAL COMPARISONS OF DISTRIBUTIONAL MODELS FOR STOCK INDEX RETURNS. *Journal of Business Finance & Accounting*, *17*(3).

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), pp.1735-1780.

Hoerl, A.E. and Kennard, R.W. (1970) ‘Ridge regression: Biased estimation for nonorthogonal problems’, *Technometrics*, 12(1), pp. 55–67. doi:10.1080/00401706.1970.10488634.

Li, B. and Rossi, A.G. (2020). Selecting mutual funds from the stocks they hold: A machine learning approach, SSRN. Available at: https://projects.ssrn.com/sol3/projects.cfm?abstract_id=3737667

Ling, H.C., Phoon, S.W. and Seoh, Y.K., (2022). Analysis of Low Order Universal Portfolios Generated by Two and Three Parameters Distributions on Malaysia's Stocks During the Covid-19 Pandemic. *Thai Journal of Mathematics*, pp.1-16.

Pang, S.T., Liew, H.H. and Chang, Y.F., (2017), April. Performance of finite order distribution-generated universal portfolios. In *AIP Conference Proceedings* (Vol. 1830, No. 1). AIP Publishing.

Pickands, J. (1975). Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics*, 3(1), 119–131. http://www.jstor.org/stable/2958083

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & De Freitas, N. (2015). Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1), 148-175.

Tan, C.P. (2013). Performance bounds for the distribution-generated universal portfolio. In *Proceedings of the 59th World Statistics Congress of the International Statistical Institute* (pp. 5327-5332).

Tan, C.P. and Kuang, K.S., (2014). Performance of Universal Portfolios Generated by Dominant-Diagonal Matrices and Probability Distributions. *STATISTICS, AND FINANCIAL MATHEMATICS 2014 WITH IASC-ARS SESSIONS*

Tan, C.P. and Pang, S.T., (2013), April. The finite and moving order multinomial universal portfolio. In *Journal of Physics: Conference Series* (Vol. 435, No. 1, p. 012039). IOP Publishing.

Tibshirani, R. (1996). 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), pp. 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

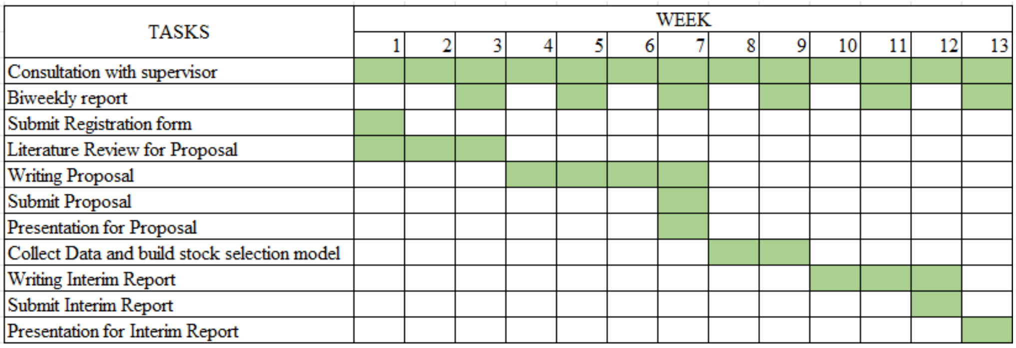
Weibull, W., (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*.

Wolff, D. and Echterling, F. (2023). 'Stock picking with machine learning', *Journal of Forecasting*, 43(1), pp. 81–102. https://doi.org/10.1002/for.3021

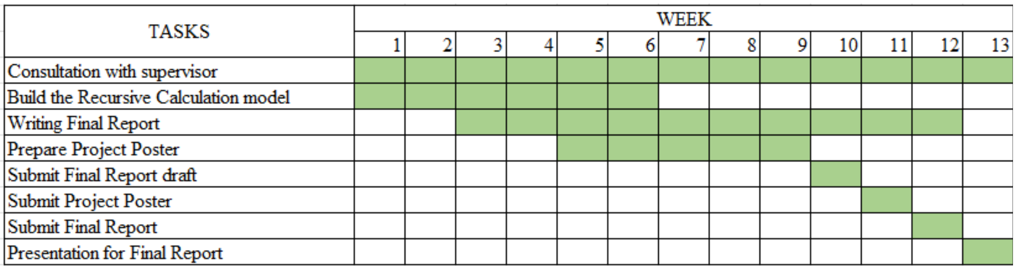
Zou, H. and Hastie, T. (2005). 'Regularization and variable selection via the elastic net', *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), pp. 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.

APPENDICES

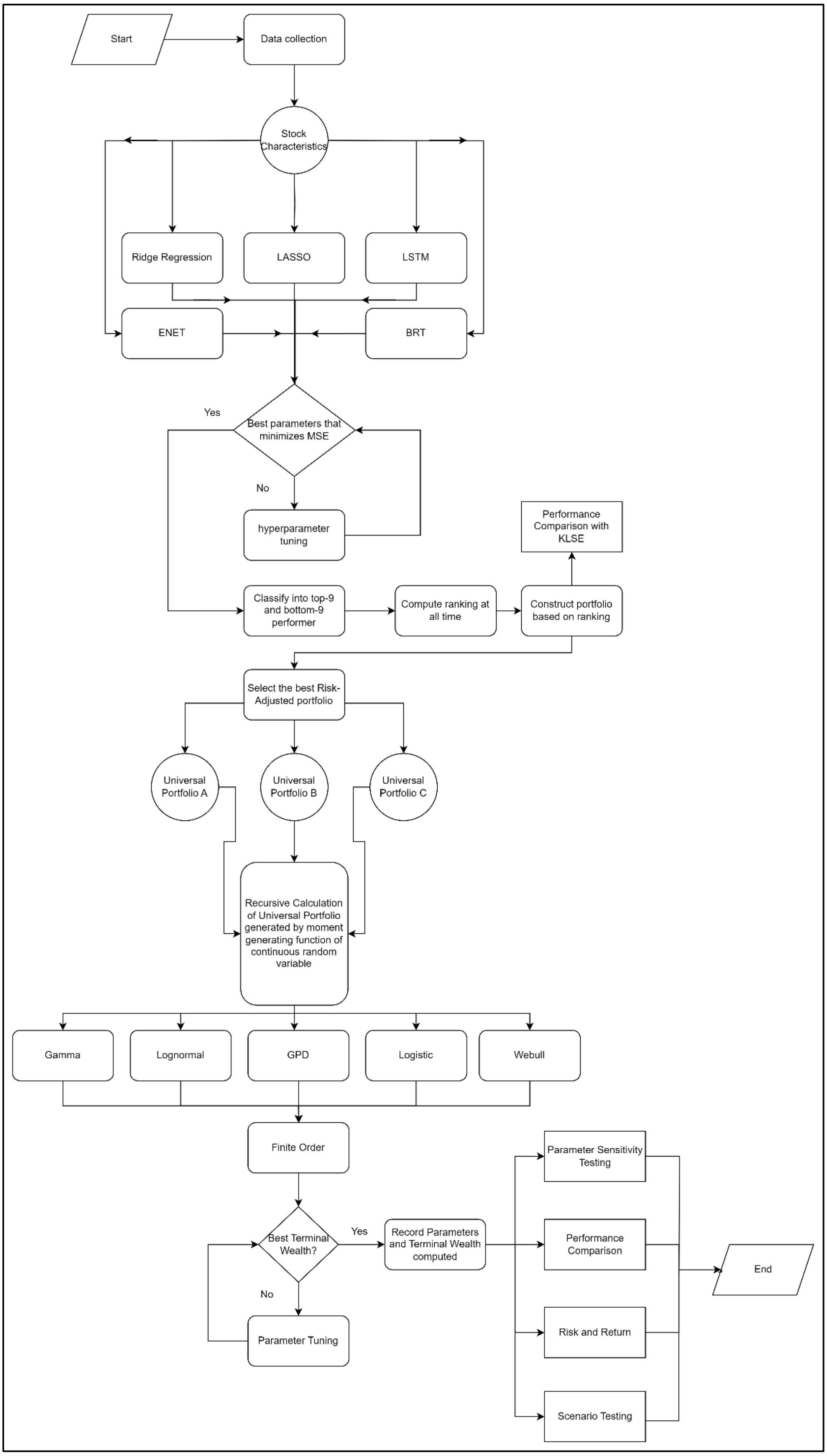
Appendix A: Figures



Appendix A-1: Gannt Chart for Project I.



Appendix A-2: Gannt Chart for Project II.



Appendix A-3: Flow Chart for Project I and II