Problem Chosen	2022 MCM/ICM Summary Sheet	Team Control Number
C		2227906

A Short-Term Arbitrage Model of Two Risk Portfolio

Summary

Keywords: keyword1; keyword2

A Short-Term Arbitrage Model of Two Risk Portfolio

LATEX Studio

February 22, 2022

Summary

Keywords: keyword1; keyword2

Contents

1 Introduction

1.1 Background

which means , to develop a model that uses only the past stream of daily prices to date to determine each day if the trader should buy, hold, or sell their assets in their portfolio.

1.2 Problem Statement

1.
2.
3. ,
4.
5.
6. Determine how sensitive the strategy is to transaction costs

1.3 Problem Analysis

ARIMA ARIMAARIMA

2 Assumption

3 Data Processing

3.1 Data Screening

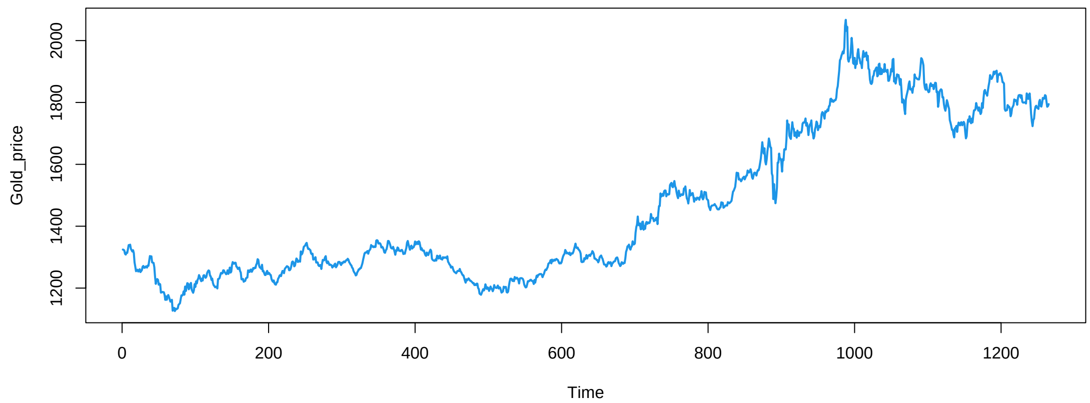
We analyzed the raw data in the LBMA-GOLD.csv and BCHAIN-MKPRU.csv files,final data status is as follows:

For the missing data in LBMA-GOLD.csv, we fill in the date according to the average of the day before and the day after.

	gold	bitcoin
count	1255	1826
mean	1464.54	12206.06
std	249.29	14043.89
min	1125.7	594.08
max	2067.15	11084.73
missing data	10	0

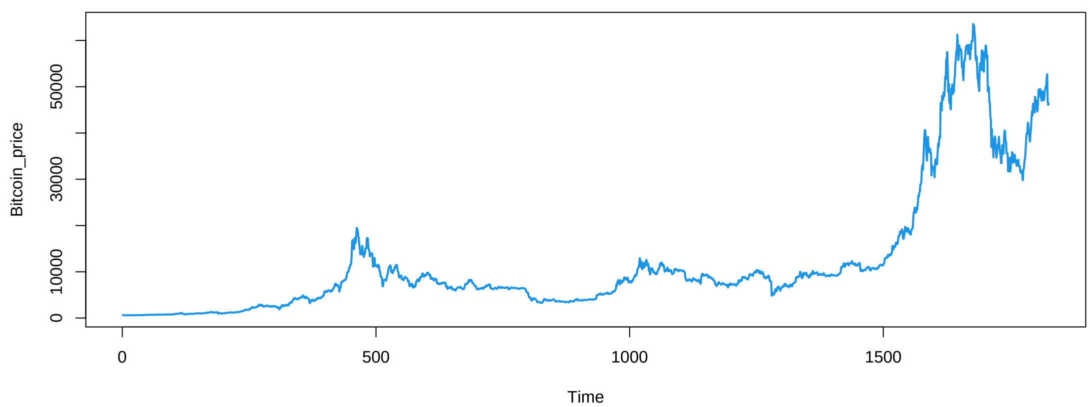
3.2 Data Visualization

To observe the price trends of gold and bitcoin more visually, we visualize the given data and draw figure??and??



(a) 1

Figure 1: Gold price tendency



(a) 2

Figure 2: Bitcoin price tendency

3.3 Mining Time Series

For subsequent data prediction using the time series model ARIMA, We perform stability test and white noise test on the raw data and processed data as a way to mine meaningful time series.

3.3.1 Stability Test

First,we test the stability of the original data by comparing two methods, the image observation and the unit root test. Testing unit rootand result is shown below: GoldDickey-Fuller = -2.4368,

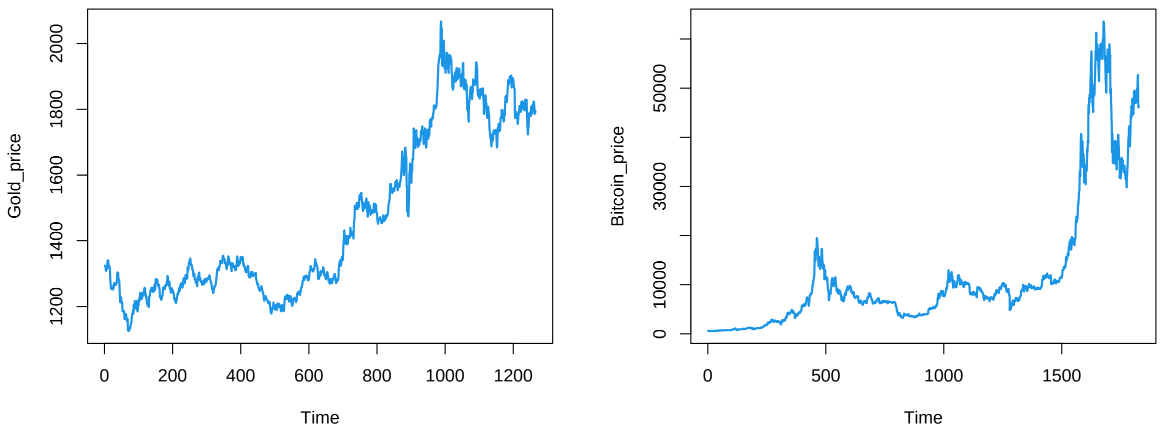


Figure 3: Raw data visualization

Lag order = 10, p-value = 0.3934<0.1 BitcoinDickey-Fuller = -1.4395, Lag order = 12, p-value = 0.8156<0.1

The raw data, the image mean varies with time and the unit root test p>0.05, so it is an unstable time series. We then perform first order differencing on the original data to obtain updated data.

Secondly,the first-order difference data is obtained according to the first-order difference of the original data, and the two methods abrove are also used to test.The result is as follows.**??** Visualizing the first order difference data is illustrated in Figure 4

Testing unit rootand result is shown below:

GoldDickey-Fuller = -11.357, Lag order = 10, p-value <0.01 BitcoinDickey-Fuller = -11.633, Lag order = 12, p-value <0.01

It can be concluded that the first-order difference data, with image mean essentially zero and unit root test p<0.05, is a stable time series.

Thirdly,utilizing second order difference we obtained second order difference data with two methods testing The result is shown in**??**.

Testing unit rootand result is shown below:

GoldDickey-Fuller = -20.351, Lag order = 10, p-value <0.01 BitcoinDickey-Fuller = -18.999, Lag order = 12, p-value <0.01

It can be seen that the image mean of the second order difference data is basically 0 and the unit root test p<0.05, which means it is a stable time series.

Final conclusion: we cannot use the original data directly for time series modeling because it is unstable, and need to use its first-order difference or second-order difference data for time series model.

3.3.2 White Noise Test

We need to evaluate whether the data is white noise or not, and will discard the one that is white noise because it has no research significance. So We chose Ljung-Box test to meet the demands.

The first step is to examine the raw data,The test yielded the following graph

GoldX-squared = 7495.2, df = 6, p-value < 2.2e-16 X-squared = 14834, df = 12, p-value < 2.2e-16 X-squared = 22018, df = 18, p-value < 2.2e-16 BitcoinX-squared = 10716, df = 6, p-value < 2.2e-16 X-squared = 20966, df = 12, p-value < 2.2e-16 X-squared = 30765, df = 18, p-value < 2.2e-16

Raw data p<0.05, not white noise

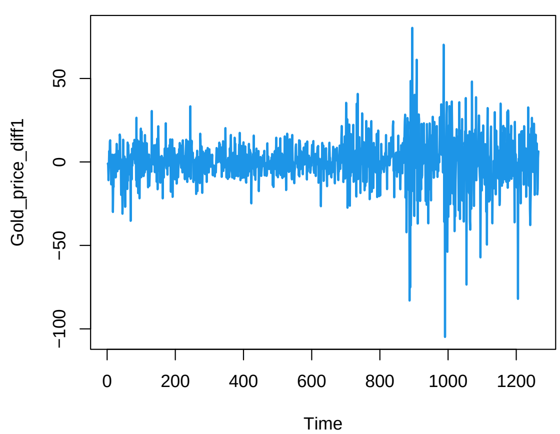
The second step,we test first order difference data,result can be seen below:

GoldX-squared = 35.268, df = 6, p-value = 3.824e-06 X-squared = 47.324, df = 12, p-value = 4.097e-06 X-squared = 56.106, df = 18, p-value = 8.576e-06 BitcoinX-squared = 21.896, df = 6, p-value = 0.001265 X-squared = 63.942, df = 12, p-value = 4.275e-09 X-squared = 71.685, df = 18, p-value = 2.339e-08

First order differential data p<0.05, not white noise

Third,we test second order difference data,result is shown below

Second-order differential data p<0.05, not white noise **Final conclusion:** The data come from a professional data statistics center and should not be white noise, and the test results prove it. They are not white noise.



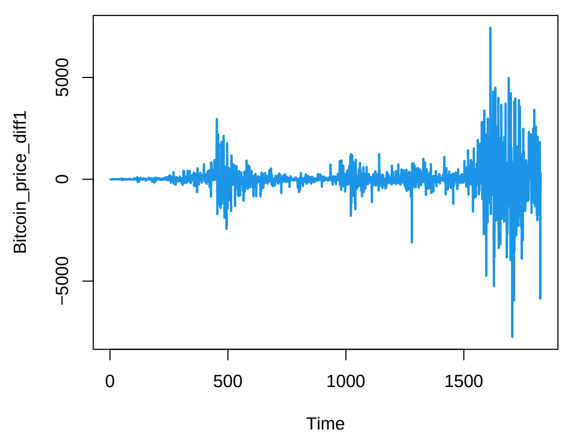


Figure 4: first order difference data

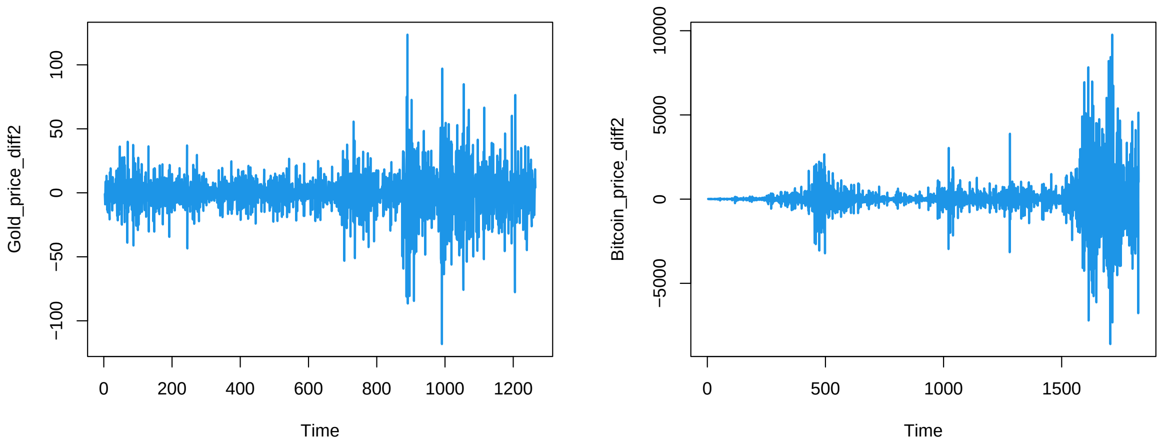


Figure 5: second order difference data

4 PartModel Development

4.1 Time Series Model ARIMA - Data Forecasting

4.1.1 Model Theory

AutoregressiveăIntegratedăMovingăAverage model is the differential integrated moving average autoregressive model, also known as the integrated moving(or slilding)average autoregressive model, is one of the time series forecasting analysis methods. In ARIMA(p, d, q),AR is "autoregressive", p is the number of autoregressive terms; MA is "sliding average ",q is the number of sliding average terms, and d is the number of differences (order) made to make it a smooth series. Although the word "difference" does not appear in the English name of ARIMA, it is a key step to analyse time series.

«««< HEAD «««< HEAD «««< HEAD $\left(1-\sum_{i=1}^{p} L^{i}\right)(1-L)^{d} X_{t}=\left(1+\sum_{i=1}^{q} \theta_{i} L^{i}\right) \varepsilon_{t}$

======= »»»> e0d6d6aa972c268925f53c491cc6a08ca18daf0e ======= »»»> e0d6d6aa972c268925f53c491
======= »»»> e0d6d6aa972c268925f53c491cc6a08ca18daf0e

4.1.2 Determining the parameters p, q

We take advantage of the autocorrelation and partial autocorrelation plots to find out the parameters p, q. The following figures show the the format of autocorrelation and partial autocorrelation plots.

In theory Tail-dragging: always have non-zero values, not constant equal to zero after k is greater than some constant (or fluctuate randomly around 0).

Truncated tail: After greater than a constant k, it quickly tends to 0 as a k-order truncated tail when both autocorrelation and partial.

By figure ??and ??, it can be seen that the first order difference data and the second order

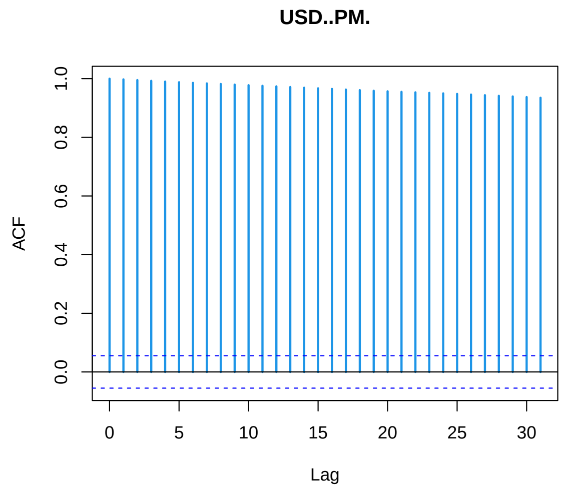


Figure 6: Autocorrelation diagram

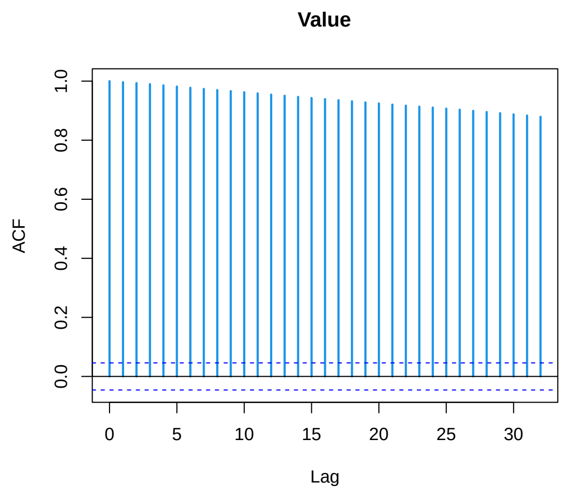


Figure 7: Partial autocorrelation diagram

difference data are meaningful time series. Therefore,we use the same methods in the subsequent section.The analysis charts are as follows.

4.1.3 R Language Determines the Optimal Parameters p, d, q

Given that only the price data as of the day can be used each day, i.e., the training data used each day are inconsistent, it is not practical to determine the optimal parameters for the model through autocorrelation and partial autocorrelation plots, so we use the auto.arima function in R language to automate the parameter determination.

The best model information was obtained after using the auto.arima function with all given data. And the model is as follows: Gold:p=4d=1q=5;ARIMA(4,1,5) Bitcoin:p=2d=1q=1;ARIMA(2,1,1)

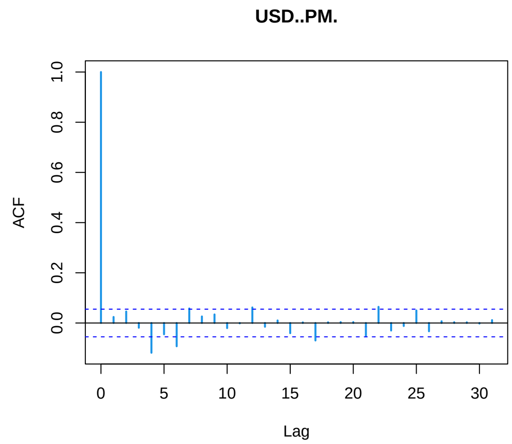


Figure 8: First order differential autocorrelation diagram-gold

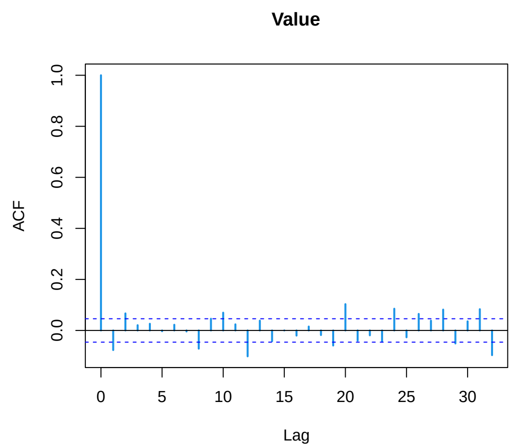


Figure 9: First order differential partial autocorrelation diagram-bitcoin

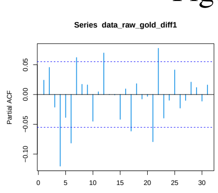


Figure 10: First order differential autocorrelation diagram-gold

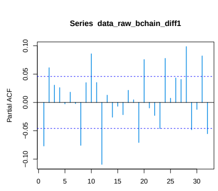


Figure 11: First order differential partial autocorrelation diagram-bitcoin

4.1.4 White noise test for model residuals

It is usually assumed that the model residuals of a reasonable model should be white noise. logically we conducted a white noise test on the residuals of the resulting model. The results are as follows.

GoldX-squared = 4.9084, df = 7.1428, p-value = 0.6862 BitcoinX-squared = 1.3484, df = 7.5099, p-value = 0.9919 We can see model residuals p>0.05, is white noise.Our model is valid.

4.1.5 Model Prediction and Visualization

To make the results more intuitive,we use the model to calculate and predict the historical data. And the visualization results are shown in the figure

The reason for the large overlap of lines in Figure 6 is that the data sample is too large. So we choose 100 of these samples and make graph7. The model fits well apparently from the images.

4.1.6 Batch prediction of data

Based on the price data as of the day,we predicted gold and bitcoin price for the next 7 days and the same automated arima modeling is performed using the auto.arima function to obtain the price data for the next 7 days. By analyzing the forecast data, we observed a roughly linear variation. Then,we integrated these data using linear regression to clarity the future price trend and fitted slope quantifies the trend to make the better investment decisions.

4.2 Trading Strategy Model - Dynamic Programming

Notation: w_G:Gold holding ratio w_B:Bitcoin holding ratio r_G:Gold expected return r_B:Bitcoin expected return α_G:Commission ratio of gold transaction α_B:commission ratio of bitcoin transaction σ_G:Gold lower semi-variance of historical yield σ_B:Bitcoin lower semi-variance of historical yield β:Risk aversion coefficient of trader T:The average number of days held after each purchase of assets

4.2.1 Expected return on assets

Traders pay a percentage of commission when they buy and sell assets, in other words, there is hidden cost to holding assets every day. We use$\alpha \div T$ to represent such costs.

We determine trading behavior by comparing the expected benefit with the size of that cost. When expected revenue is large enough to offset this cost,model decide to buy; When the expected return is less than the negative cost, it represents maintaining; When the asset is about to lose more than commission cost,model needs to sold Holding share immediately to stop loss.

If$r > (1 + \beta)\alpha \div T$it means that the asset will appreciation in the future,so trader can be buy in.

If$r < -(1 - \beta)\alpha \div T$it represents that the asset will depreciation in upcoming period,so trader must sell them out .

If$-(1 - \beta)\alpha \div T < r < (1 + \beta)\alpha \div T$,it indicates that recent prices are stable,trader can either buy or sell.

moreover,due to the difference in investors and investment products, We introduceβ to characterize the diffrent degree of risk aversion.

The larger the β,refers to the more conservative the investors is,stricter restrictions on buying and more lenient restrictions on selling. On the contrary,thr smaller β indicates that the more aggressive the investor is, who has more lenient buying criteria and more stringent selling criteria.

Additionally,we have to consider the following two situations when purchasing: ,

1. If only one of gold and bitcoin meets the upside condition, then we simply buy all of our currently available funds into that asset.

2. Whereas if gold and bitcoin rise at the same time,there is a need to consider how to allocate the available funds. In this case, we use the Sharpe ratio to measure the different proportional investment groups

$$Sharpe\ Ratio = \frac{w_G \times r_G + w_B \times r_B}{\sqrt{w_G^2 \sigma_G^2 + w_B^2 \sigma_B^2 + 2Cov_{w_B w_G}}}$$

We use the lower semi-variance as a quantitative index of risk, as the portion of the standard deviation that represents fluctuations less than the mean. which is more indicative of the risk of asset losses. We divide the expected return by the following half standard deviation yields the Sharpe ratio,which implies the magnitude of the return per unit of risk of the current portfolio. When the Sharpe ratio is maximum, it undoubtedly means that the current proportion of the portfolio is optimal

Based on the assumptions that "All cash is consumed at each purchase" and "The price fluctuations of gold and bitcoin are independent of each other", We simplify the problem of solving the optimal investment ratio as an optimization problem. And we can use the computer to find its numerical solution.

$$\begin{aligned} max \quad & \frac{w_G \times r_G + w_B \times r_B}{\sqrt{w_G^2 \sigma_G^2 + w_B^2 \sigma_B^2}} \\ s.t. \quad & w_G + w_B = 1 \\ & 0 \leq w_G \leq 1 \end{aligned}$$

4.2.2 Current Asset Holding

After determining the return on the assets, we also consider the current asset allocation in order to determine what and how much assets we will eventually buy or sell. With the assumption that "All cash is consumed at each purchase",it is impossible to hold cash, gold, bitcoin or no assets at the same time. So there are a total of $2^3 - 2 = 6$ possible scenarios, to be specific: (solid circles indicate that the asset is held, hollow circles indicate that the asset is not held)

cash	gold	bitcoin
●	○	○
●	○	●
●	●	○
○	○	●
○	●	○
○	●	●

4.2.3 Determination of the final act of transaction

Daily expected returns for gold and bitcoin are up, down, and stable. 9So,there are 9 cases when they are combined together. 6And the initial state of the daily assets, as pointed out in the previous subsection, has six profiles. 54Thus, the final trading behavior totals 54 scenarios. 546Fortunately, we can summarize this into the following six situations:

1. **one maintain,the other appreciate:**In this case, if cash is available trader uses it all to buy appreciating assets. If no cash is available do not make the purchase.
2. **one maintain,the other depreciate:**Under such circumstances, trader sells all the devalued asset if they hold.
3. **one appreciate,the other depreciate:**In this case, the trader first sells the depreciating asset in full and then buys the appreciating asset with all the cash gained.
4. **both maintain:**Trader does not make any transactions
5. **both appreciate:** 4.2.1 As mentioned in 4.2.1, when the expected returns of two assets rise simultaneously, we need to determine the optimal asset mix based on the Sharpe ratio. If we only hold cash as an asset, then we can simply buy two assets based on the optimal ratio. However, if we hold either gold, bitcoin or both, we have to decide if we still need to adjust the ratio to the optimal ratio.After all, each adjustment requires a significant commission. 10To simplify the model, we define the criteria for transferring positions: If the proportion of simultaneous changes in both assets exceeds 10 percent when the position is transferred, it will be adjusted to the optimal proportion, otherwise it will remain unchanged.
6. **both depreciate:**After selling two assets at the same time, we only have one asset in cash

4.2.4 Additional Explanation

-

5 Part:Strategy Evaluation

5.1 Set Perturbation Terms

To demonstrate strategy optimality,firstly we add small perturbations to the buy and sell criteria.

Then,we make minor adjustments to the buy and sell criteria, invoking the model (at this point the risk appetite indicator is 0.2) to calculate the final asset size,.

Finallyobserve the asset stability.We found In the range of -0.25 to 0.25, the final asset is basically smooth and there is no sudden drop, which shows that the model stability is relatively reliable. result shown in the figure.

6 Part:Sensitivity Analysis

6.1 Assuming Changes In Commission

We set the commission percentage of gold as a, the commission percentage of bitcoin as b, and the final asset as f USD. After we kept adjusting the commission rate, we obtained the final assets under different commission rates, and the results are shown in the figure.

Final conclusion Both the gold commission ratio a and the bitcoin commission ratio b will significantly affect the final asset f dollars when they changes. According to the analysis,f is essentially negatively correlated with b and f will be more sensitive to changes in b compared to a. Our model is not very good at grasping the price of gold because it changes so frequently. So it is not true that if the transaction cost is low, trader will benefit more. But for bitcoin, the changes are large and infrequent, the model is easier to make a correct judgment, and when the transaction cost is very low or even 0, we can arbitrage and protect the value easily.

In summary, it can be seen that the model is sensitive to trading commissions.

7 Evaluate of the Model

7.1 Strengths and weaknesses

1. lie1

2. lie2

7.2 Sensitivity Analysis

8 A Memo

References

[1] D. E. KNUTH The TEXbook the American Mathematical Society and Addison-Wesley Publishing Company , 1984-1986.

[2] Lamport, Leslie, LATEX: " A Document Preparation System ", Addison-Wesley Publishing Company, 1986.

[3] https://www.latexstudio.net/

Appendices

Appendix A First appendix

In addition, your report must include a letter to the Chief Financial Officer (CFO) of the Goodgrant Foundation, Mr. Alpha Chiang, that describes the optimal investment strategy, your modeling approach and major results, and a brief discussion of your proposed concept of a return-on-investment (ROI). This letter should be no more than two pages in length.

Dear, Mr. Alpha Chiang

Sincerely yours,

Your friends

Here are simulation programmes we used in our model as follow.

Input matlab source:

```
function [t,seat,aisle]=OI6Sim(n,target,seated)
pab=rand(1,n);
for i=1:n
    if pab(i)<0.4
       aisleTime(i)=0;
    else
        aisleTime(i)=trirnd(3.2,7.1,38.7);
    end
end
```

Appendix B Second appendix

some more text **Input C++ source:**

```
//===========================================================================
// Name        : Sudoku.cpp
// Author      : wzlf11
// Version     : a.0
// Copyright   : Your copyright notice
// Description : Sudoku in C++.
//===========================================================================

#include <iostream>
#include <cstdlib>
#include <ctime>

using namespace std;

int table[9][9];

int main() {
```

```
    for(int i = 0; i < 9; i++){
        table[0][i] = i + 1;
    }

    srand((unsigned int)time(NULL));

    shuffle((int *)&table[0], 9);

    while(!put_line(1))
    {
        shuffle((int *)&table[0], 9);
    }

    for(int x = 0; x < 9; x++){
        for(int y = 0; y < 9; y++){
            cout << table[x][y] << " ";
        }

        cout << endl;
    }

    return 0;
}
```

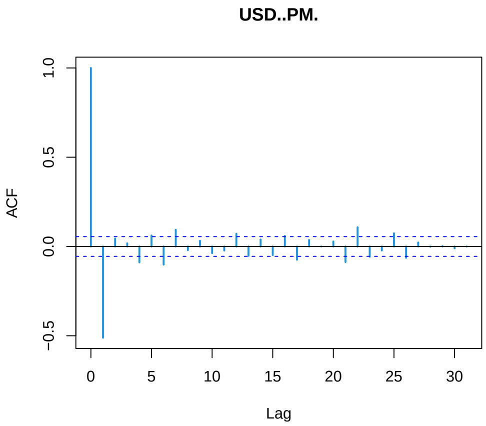


Figure 12: Second order differential autocorrelation diagram-gold

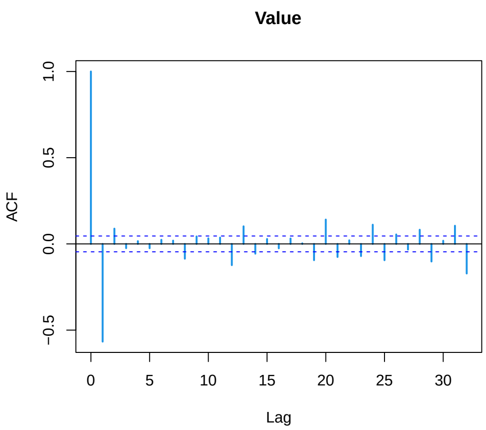


Figure 13: Second order differential partial autocorrelation diagram-bitcoin

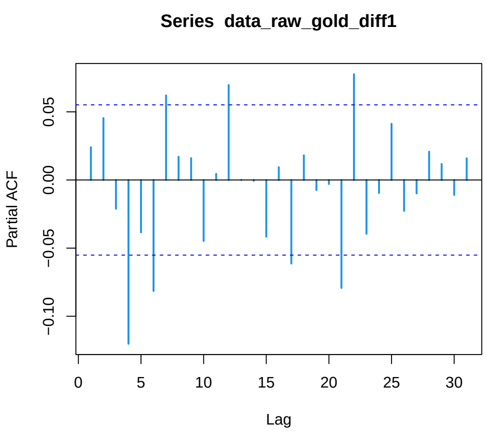


Figure 14: Second order differential autocorrelation diagram-gold