FIN3080 Assignment 3 Report

Data processing

Downloading data

First, I downloaded 3 different csv files, they separately contain:

1. All Daily Market Index
2. Weekly Return without Dividends for All the Stocks in the Market, the Market Type of the Stocks
3. Weekly Risk Free Rate

Problem 1

(a)

1. Filter out all the data whose 'Indexcd' is Index000300, which means it is the csi300 index.
2. Transfer the date of each data into the corresponding month and calculate the average csi300 index of the month. Save the data as another document.
3. Construct a new column and shift the average csi300 index into next month and use this month as t, last month as t-1 to calculate the monthly return. The result will be like this form.

$$R_{k,t} = I_{k,t} \,/\, I_{k.t-1} \; - 1$$

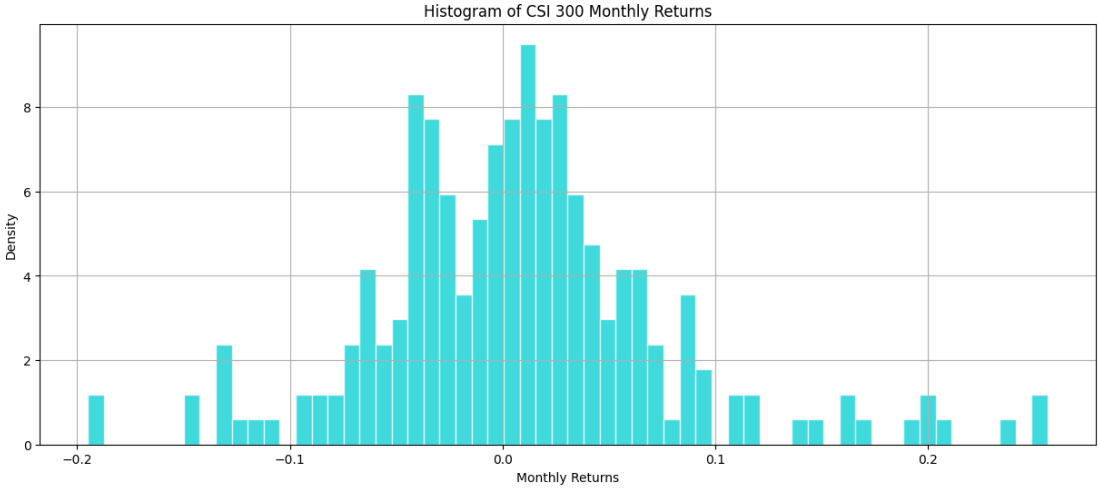
	month	mean_index	index last month	return
0	2005_05	877.821882	960.709500	-0.086278
1	2005_06	879.925045	877.821882	0.002396
2	2005_07	853.139952	879.925045	-0.030440
3	2005_08	927.240348	853.139952	0.086856
4	2005_09	942.550091	927.240348	0.016511
..	...	...	...	...
219	2023_08	3853.924435	3878.075143	-0.006227
220	2023_09	3742.056200	3853.924435	-0.029027
221	2023_10	3587.914294	3742.056200	-0.041192
222	2023_11	3568.071182	3587.914294	-0.005531
223	2023_12	3377.154857	3568.071182	-0.053507

4. Then calculate the required data.

```
Mean: 0.008022105965103086
Standard Deviation: 0.0702539576947071
Skewness: 0.6024057067617048
Kurtosis: 2.015243944338839
```

(b)

Use the monthly return calculated in (a) to plot the histogram.



(c)

Use the Shapiro–Wilk test to test whether it is a normal distribution, and as the result tells it is absolutely not a normal distribution. We can also find out the result by looking at the shape of the figure we obtained in (b).

```
Statistics=0.955, p-Value=0.000
Sample does not look Gaussian
```

Problem2

Data Processing

1. Csmar cuts the data into two files, so first merge them together.
2. Use the market type to filter out the mainboard stocks.
3. Use time to divide the data into three period.
4. Convert the date of the risk free rate into the corresponding week.

trading_date_yw	risk_free_return	Trdwnt
2017-01-01	0.000571	2017-01

5. Calculate the mean of the return of all the stocks and use it as the return of the market portfolio.
6. Merge the risk free return, the market return and the stock return together.
7. Calculate the Ri – Rf and the Rm – Rf for each stock.
8. Divide all the data into three time periods.

P1 Calculating Individual Stock βs

Run the regression for Ri – Rf and Rm – Rf of each stock, and obtain their beta. Sample data are shown below.

stock_code	Beta
1	0.337687
2	0.639046
4	0.58931
5	1.027914
6	1.196454
7	0.678335
……	……

P2 Construct Stock Portfolios and calculate portfolio βs

Process

1. Group the individual stocks into 10 portfolios according to the size of their βs.

Calculate the monthly (return of the portfolio – Rf). Sample data are shown below.

divided_by_betas	year_week	ri - rf	rm - rf
0.0	2019-01	-0.003608	0.016610
0.0	2019-02	0.012713	0.029264
0.0	2019-03	0.004021	0.001243
0.0	2019-04	-0.006002	-0.005899

2. Run the regression for the data and calculate the β for each portfolio.

Result

	0.0	1.0	2.0	3.0	4.0
alpha	-0.001297	-0.000008	-0.000062	0.000616	-0.000255
alpha_t_value	-1.992031	-0.015011	-0.148650	1.477141	-0.710558
beta	0.772549	0.864526	0.909146	0.935968	1.009280
beta_t_value	36.965487	49.004038	68.372222	69.913480	87.590838
R-squared	0.931808	0.960022	0.979057	0.979951	0.987134
Observations	102.000000	102.000000	102.000000	102.000000	102.000000

	5.0	6.0	7.0	8.0	9.0
alpha	0.000014	-0.000290	-0.000002	0.000451	-0.000854
alpha_t_value	0.041649	-0.832427	-0.005330	0.974178	-1.501781
beta	1.042325	1.053822	1.100428	1.105400	1.159132
beta_t_value	94.376043	94.077061	73.592000	74.438786	63.475120
R-squared	0.988897	0.988827	0.981870	0.982273	0.975782
Observations	102.000000	102.000000	102.000000	102.000000	102.000000

Findings

1. Betas are all around 1 and they are all very significant, it is similar with the findings in the original paper.
2. Different from the paper the alphas of my portfolios mostly are not significant, which means we fail to reject the null hypothesis.
3. R-squared seems uncorrelated with betas, it also indicates the original conclusion.

P3 CAPM Cross-sectional Regression

Process

1. Calculate the (average total return for each portfolio – Rf)
2. Run the regression. Sample data are shown below.

divided_by_betas	ri - rf	Beta
0.0	0.001354	0.772549
1.0	0.001271	0.864526
2.0	0.001187	0.909146
3.0	0.001058	0.935968
4.0	0.001752	1.009280
5.0	0.001415	1.042325
6.0	0.001941	1.053822
7.0	0.001539	1.100428
8.0	0.001900	1.105400
9.0	0.002151	1.159132

Result

	Coefficient	t-value
gamma0	-0.000649	-0.919304
gamma1	0.002216	3.146075
R-squared	0.553017	nan
F-statistics	9.897789	nan
P-value	0.013679	nan

Findings

1. Different from the paper, I fail to reject that gamma0 is equal to zero, so my conclusion will be there is no significant evidence showing that there are factors other than the systematic risk. It may because that our portfolio is formed with all the stocks so its individual risk is well-diversified.
2. Similar with the paper, my gamma1 is significantly different from 0, which indicates that as risk increases, return will also increase. It is consistent with the CAPM model.
3. My R-squared is around 0.55, it is larger than the R-squared in the paper, but it is also not a very perfect fit.