FE 535: Introduction for Financial Risk Management
Project 1

Due Date 11:59 PM, October 18th, 2020

Instructions

1. This is a teamwork project - please check your assigned team in the course website.
2. There are 5 main questions in this project.
3. Each team member should contribute equally to the project.
4. Feel free to use the handouts and other published codes to do your project.
5. You are welcome to use any programming language or statistical software. Also, you are welcome to use any library/package unless stated otherwise.
6. You will need to download data on your own - unless provided otherwise.
7. The final report should be written using a special document editor, e.g. Word, Latex, Markdown, etc. **Any form of document with a handwriting will not be accepted**.
8. Please submit a pdf copy of your final report - include your codes in the appendix. Note that Canvas will not accept any other formats than pdf.
9. Utilize the length of your report carefully. Not including the code appendix, the maximum report's length **should not be more than 15 pages** - with font 11 point size, 1.5 line space, and 1in margin (just like this document).
10. Please **avoid taking picture snapshots**. You should report your results in an organized table. The same applies to plots and other visualizations - **do not paste any low resolution figures**.
11. Please use a special equation editor to write any math, in case needed.

Data

In this project, you will be working with the same data for all questions. In particular, you need to download market prices for 27 stocks. These stocks correspond to the Dow Jones 30 Constituents. The symbol of each is given below

AAPL	CSCO	HON	KO	NKE	WBA
AMGN	CVX	IBM	MCD	PG	WMT
AXP	DIS	INTC	MMM	TRV	
BA	GS	JNJ	MRK	UNH	
CAT	HD	JPM	MSFT	VZ	

As a market portfolio, you will also need to download data for the SPY ETF. Merged altogether, the data should range between May 1999 and Sep 2020 (included). The rest of the project will be referring to this dataset.

Note: I recommend using the **R quantmod** package to download the data. You may also refer to this application to download the data manually. The end result of the data downloading procedure should result in a $T \times 27$ matrix of returns, where the columns correspond to the stocks and the rows are the daily returns. Overall, you should have around $T \approx 5,400$ daily returns for each individual stock.

1 Performance Summary (20 Points)

After downloading the data, refer to the adjusted price of each stock to compute the daily return and merge altogether. This should result in a $T \times 27$ data matrix, representing T periods and 27 assets. Given this data matrix, address the following:

1. For each stock, compute an absolute performance summary table. In particular, this should include mean return, volatility, and SR (all reported in annual terms). Rather than reporting the results in 3×27 table - where the rows correspond to the performance measure and the columns correspond to the stock - report a summary for each measure. Specifically, for each performance metric, report the minimum, mean, and maximum, such that, in total, your summary consists of a 3×3 table. (6 Points)

2. Plot the asset mean returns against their volatilities. Provide a couple insights. (2 Points)

3. Let's consider the SPY ETF as the market risk factor to price the relative performance of each asset. To do so, merge the SPY daily returns with the stock data and compute the following measures for each stock:

   (a) Jensen's α

   (b) Market β

   (c) Treynor Ratio (TR)

   (d) Tracking error (ω)

   (e) Information Ratio (IR)

   - As in Task 1 above, report the relative performance measures in a 5×3 summary table. (6 Points)
   - Which are the worst and best performing stocks? Explain the rationale behind your assessment (2 Points)

4. Theoretically, the CAPM states that there is a linear relationship between the asset mean return and the market beta. To test this relationship, address the following:

   (a) Plot the mean return of each asset against its beta. (2 Points)

   (b) Does the CAPM hold true? How is that different for financial stocks, e.g. JPM, GS, and AXP? Provide a couple of insights. (2 Points)

Note: To perform the above computations, assume that the risk-free rate is zero over time for simplicity.

2 Back-Testing (20 Points)

Let's recall Portfolio 1 and Portfolio 2 from Session 1. The weight allocated to asset i in Portfolio 1 is given by

$$w_i^{\sigma} = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^{d} \frac{1}{\sigma_j^2}} \tag{1}$$

where σ_i is the volatility of asset i $\forall i = 1, ..., 27$ and $\sum_{i=1}^{27} w_i^{\sigma} = 1$. On the other hand, the weight allocated to asset i in Portfolio 2 is

$$w_i^{SR} = \frac{SR_i}{\sum_{j=1}^{d} SR_j} \tag{2}$$

where SR_i denotes the SR of stock i $\forall i = 1, .., 27$ and $\sum_{i=1}^{27} w_i^{SR} = 1$.

Let's now introduce Portfolio 3 that allocates equal weights to each asset, such that the weight allocated to asset i is

$$w_i^{N} = \frac{1}{27} \tag{3}$$

This portfolio is known as the *naive* portfolio (hence the N). This is mainly due to the fact that it does not incorporate any evaluations in the portfolio formation and, hence, always allocates equal weights to each asset regardless. Obviously, this portfolio is less susceptible to **model risk**. But, at the same time, it also does not take into account the performance of each asset into the allocation decision, unlike the case for Portfolio 1 and 2.

Your Task is to evaluate the performance for each of the three portfolios. To do so, you need to perform back-testing in the following manner (this mainly applies to Portfolios 1 and 2):

1. Split the data into in-sample (IN) and out-of-sample (OUT). For the IN sample, consider the data from 2017 and 2018 included. For the OUT sample, consider the data from 2019 and 2020. The latter corresponds to the back-testing period. To confirm, report the date range for each sample (2 Points)

2. Use the IN sample to compute the portfolio weights and the OUT sample to compute the return of each portfolio. In a 27×3 table, report the weight (in percentage) allocated to each asset. Note that the sum of the weights for each portfolio should equal 1. (6 Points)

3. Using the realized return for each portfolio, **you need to**

- Provide a plot showing the cumulative return of each portfolio with respect to the SPY (4 Points)
- Provide a summary table showing the SR, the beta, and the Jensen's alpha for each portfolio. (4 Points)
- Discuss the absolute/relative performance of each. Which portfolio would you pick and why? What do these result say about portfolio selection compared to a passive fund as the SPY (4 Points)

3 Random Numbers and Monte Carlo Simulation

Core Questions (30 Points)

In this question, you will need to address different tasks related to random number generators and Monte Carlo (MC) simulations

1. **Game 2 Phase 2**: Let's consider a game in which a dice is rolled 6 times. The game pays you \$1, if the difference between the maximum number and the minimum number is less than 3. Otherwise, it pays 0. However, to participate, you need to pay $\$p$. What is the fair price for this game, i.e. what's the maximum dollar amount you are willing to pay in order to break even? (6 Points)

   To answer this, you need to consider two approaches. In the first one, you need to consider an MC approach, in which you randomly generate numbers from a uniform distribution. In the second case, you need to compute the exact answer to this problem (not necessarily analytical). Note that the exact solution of this problem should consider all possible permutations in this experiment, i.e. $6^6 = 46,656$. Plot the \$ price as a function N (total number of iterations) and illustrate the convergence of the MC price to the exact one.

---

2. **Breaking Even**: Consider the following game in which a fair coin is repeatedly tossed. The rules of the game are straightforward. You win a \$1 when you get two heads in a row for the first time. However, you have to pay $\$k$ for each toss. For instance, if it takes you 10 tosses to get two heads in a row for the first time, then your profit and loss (PL) is $\$1 - \$10k$. However, the number of tosses that will take you to win a \$1 is uncertain, which is denoted by a random variable X. Hence, in reality your PL is stochastic and can be described as

$$PL = \$1 - X \times \$k \tag{4}$$

   for some fixed cost k and random variable X. By expectation, what is the maximum price you would be willing to pay per a coin toss? In other words, what is the value of k that makes the game break even? Use a MC simulation to answer this. (7 Points)

   Hint: Breaking even means that the expected value of the PL is zero. Your task, hence, is to find the value k that leads to $\mathbb{E}[PL] = 0$, which requires finding $\mathbb{E}[X]$.

---

3. **Pi**: In this task, you need to demonstrate a way to numerically compute the irrational number π. You will need to provide a plot demonstrating the approximation as a function of N. Make sure to add a line on the plot of the number π that your machine provides you as a reference point, e.g. you can treat 3.14159265358979 as the "true" number. (7 Points)

   Hint: while π is considered an irrational number, it still has an important property when it comes to computing the circle area. Recall that the area of a given circle with radius r is given by πr^2. In order to compute π, you need to numerically compute the area of a given circle. To do so, consider a circle of radius r being inscribed inside a square with side of length $2r$. In order to compute the circle area, you need to generate random numbers for (x, y) coordinates from a uniform distribution between $-r$ and r. The proportion of these numbers that fall inside the circle will be proportional to the area of the circle. To determine whether a point falls inside the circle, recall that the circle equation is given by

$$(x-a)^2 + (y-b)^2 = r^2 \tag{5}$$

with the center of the circle being at point (a, b). For simplicity, assume that $a = 0$ and $b = 0$. At the same time, let $r = 1$, such that the circle area is given by π while the square area is $2^2 = 4$. Finally, let I_n be an index variable that takes 1 if a random point (x_n, y_n) falls inside the circle, where $x_n \sim U(-1, 1)$ and $y_n \sim U(-1, 1)$, and zero otherwise. Note that the ratio between the circle area and the square's area should satisfy

$$\frac{\pi}{4} = \mathbb{E}[I_n] \rightarrow \pi = 4 \times \mathbb{E}[I_n] \tag{6}$$

where $\mathbb{E}[I_n] = \mathbb{P}[I_n = 1]$, i.e. the probability that the random point falls inside the circle.

---

4. Consider the case of a continuously distributed uniform random variable $X \sim U(0, 1)$. At the same time, consider $Y_k = X^k$, where $k \in \mathbb{N}$ is a natural number. Your task is the following:

   (a) Find a closed-form expression for the expectation and variance of Y_k, i.e. $\mathbb{E}[Y_k]$ and $\mathbb{V}[Y_k]$. Show the steps needed to derive each expression. (4 Points)

   (b) In terms of k, find the condition for which $\mathbb{E}[Y_k]$ and $\mathbb{V}[Y_k]$ are finite while at the same time satisfying the condition $\mathbb{V}[Y_k] > 0$. (2 Points)

   (c) Using MC simulation, create a function that estimates $\mathbb{E}[Y_k]$ and $\mathbb{V}[Y_k]$ for a given k based on $N = 10^5$ samples. Plot each estimate versus the true value from step 1, for k = -10,..,-1,0,1,..,10. What do the conditions from the previous step tell us? (3 Points)

   (d) Finally, impose the conditions for k from step (b) and repeat the same plot from step (c). Elaborate. (1 Points)

Bonus Question (5 Points)

Breaking Even II: In fact, the Breaking Even question above has a nice closed form solution. This means that we can find the solution without relying on a MC simulation at all. Can you find the solution analytically? To qualify for a full bonus credit, you have to detail the steps needed to derive the final answer. An answer without explanation will result in a zero score.

Hint: Given the law of total expectation, try conditioning on the outcome of the first two tosses. Also, note that the problem resets whenever you get a tail.

4 Value at Risk and Stress Testing (20 Points)

Task 1 (16 Points)

As a risk manager, you need to evaluate the downside risk of each portfolio from Question 2 over the next year (2021). In particular, you need to compute the Value at Risk for Portfolio 1, 2, and 3 over a one year period. To do so, you need to do the following steps:

1. Using the back-testing results, you need to calibrate the price path for each portfolio, i.e. compute $\hat{\mu}_p$ and $\hat{\sigma}_p$ for $p = 1, 2, 3$ using the daily returns in the OUT period. Report the results in a 2×3 table. (4 Points)

2. Starting with $F_{0,p} = 100$, the value of Portfolio p for all $p = 1, 2, 3$ at time t obeys to a Geometric Brownian Motion (GMB), i.e.

$$F_{t,p} = F_{0,p} \times \exp\left((\hat{\mu}_p - \frac{\hat{\sigma}_p^2}{2})t + \hat{\sigma}_p Z_t \right) \tag{7}$$

   where Z_t is a standard Brownian Motion (BM). For each portfolio, you need to simulate $N = 1000$ paths. Given the simulated paths, provide a distribution plot, e.g. boxplot or histogram, for each portfolio. Provide a couple of insights (4 Points)

3. What's the expected value of each portfolio one year from now? (4 Points)

4. With 95% level of confidence, what is the Value-at-Risk, i.e. $VaR(0.05)$, for each portfolio? (4 Points)

Note: Ideally, you should report the final numbers in a single table where columns refer to portfolios and rows refer to the computed statistics.

Task 2 (4 Points)

You need to assess the performance of each portfolio under different scenarios. In particular, you need to evaluate the VaR of each portfolio with respect to market risk. To do so, you need to estimate the market β_p for each portfolio and the market volatility σ_M during the OUT period. Refer to the SPY ETF as the market, similar to Question 2. After doing so, consider the scenario in which the market volatility increases by $a = 10\%$ and generate 1000 paths for each portfolio. Given these simulations, compute the $VaR(0.05)$ for each portfolio and summarize the results in a single table as you did before.

Hint: All else equal, according to the CAPM, an increase of a in the market volatility should increase the portfolio volatility by $\beta_p \times \sigma_M \times a$.

5 Mean-Variance Efficient Frontier (20 Points)

As a financial risk manager (FRM) working for an asset management firm, you are facing the task of constructing a set of optimal portfolios, which eventually will be proposed to the clients. Ideally, you should deliver a summary of different strategies stating the risk and return of each. In doing so, you would like to make sure that each strategy is delivering the best risk-return trade-off.

Your **task** is to construct a Mean-Variance Efficient Frontier (MVEF) given the universe of the 27 stocks. To do so, you need to solve the following optimization problem for a given m:

$$\min_{\boldsymbol{w}} \sigma_p^2 = \boldsymbol{w}'\Sigma\boldsymbol{w} \tag{8}$$

subject to

$$\boldsymbol{w}'\mathbf{1} = 1 \tag{9}$$

$$\boldsymbol{w}'\boldsymbol{\mu} = m \tag{10}$$

where

- $\boldsymbol{w}$ denotes a $d \times 1$ vector of portfolio weights
- $\boldsymbol{\Sigma}$ is the covariance matrix of the asset returns
- $\boldsymbol{\mu}$ is the vector of mean returns, while m denotes the mean target

Note that for each given m, there is an optimal portfolio $\boldsymbol{w}(m)$ that has a mean return of $\mu_p(m) = m$ and volatility of $\sigma_p(m)$. Your final goal is to provide a list of $(m, \sigma_p(m))$ for a set of m values. To do so, address the following:

1. Refer to the proposed solution below and the full sample period to construct the MVEF. As a summary plot the MVEF and make sure that the m (y-axis) and $\sigma_p(m)$ (x-axis) values are reported in an annual basis. (8 Points)

2. Given the MVEF, highlight two specific points on the plot. **One** corresponds to the maximum Sharpe-ratio (SR) portfolio and the **other** is the global minimum variance (GMV) portfolio. (2 Points)

3. The two-funds separation theorem states that the MV optimal portfolio choice problem can be written as a linear combination of two funds, one is low risk and the other is the SR portfolio. In fact, when the risk-free asset is absent, the theorem states that the MV portfolio is given by

$$\boldsymbol{w} = \theta\boldsymbol{w}_0 + (1-\theta)\boldsymbol{w}_{SR} \tag{11}$$

for constant $\theta \in (0,1)$, where $\boldsymbol{w}_0$ and $\boldsymbol{w}_{SR}$ denote, respectively, the GMV and the SR portfolios. Since you know $\boldsymbol{w}_0$ and $\boldsymbol{w}_{SR}$ from part 2, derive the MVEF using Equation (11). As a summary,

(a) highlight this frontier using a red-dashed line on the previous plot you have. (5 Points)

(b) What does it mean to have a $\theta > 1$? Provide an economic rationale. (5 Points)

Proposed Solution

1. We covered this problem in both the class and the handouts. In particular, there is a closed form solution for the above optimization problem that can be represented as a function of A_m, which represents the risk aversion of the client. According to Session 1 and Handouts, it follows that the optimal portfolio is a combination of two funds:

$$\boldsymbol{w} = \boldsymbol{w}_0 + \frac{1}{A_m}\boldsymbol{w}_1 \tag{12}$$

where $\boldsymbol{w}_0$ is the global minimum variance portfolio (GMV) and $\boldsymbol{w}_1 = \boldsymbol{B}\boldsymbol{\mu}$ is the more "aggressive" portfolio. The weight allocated to the latter is determined by A_m, which is proxies the risk-aversion of the investor. In particular, under certain assumptions, it follows that

$$A_m = \left[\frac{m - \boldsymbol{w}_0'\boldsymbol{\mu}}{\boldsymbol{w}_1'\boldsymbol{\mu}}\right]^{-1} \tag{13}$$

2. Using the stock returns data, estimate the vector of mean returns $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. This should result in one mean vector and one covariance matrix. Using both estimates, perform the following computations:

(a) Given $\boldsymbol{\Sigma}$, compute the GMV portfolio, i.e. $\boldsymbol{w}_0$ from Equation (7) from the slides of Session 1

(b) Given $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, compute $\boldsymbol{w}_1 = \boldsymbol{B}\boldsymbol{\mu}$ - see Equation (6) from the slides of Session 1 for the definition of the $\boldsymbol{B}$ matrix.

(c) Compute the mean return on each fund, i.e. $\mu_0 = \boldsymbol{w}_0'\boldsymbol{\mu}$ and $\mu_1 = \boldsymbol{w}_1'\boldsymbol{\mu}$

(d) For a given m, there is a unique value A_m as described in Equation (13). In particular, set the values of m to range between μ_0 and $2{\times}\max(\mu_i)$ $\forall i = 1, .., 27$. **You need to do this for at least 20 unique values of m in this range**.

(e) Ideally, you should write a function that takes m, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$ and returns the optimal portfolio $\boldsymbol{w}(m)$ along with m and its optimal volatility $\sigma_p(m)$ where

$$\sigma_p(m) = \sqrt{\boldsymbol{w}(m)'\boldsymbol{\Sigma}\boldsymbol{w}(m)} \tag{14}$$

Note: The above task requires matrix multiplication only, such that you do not need to perform any numerical optimization. I will not accept numerical solutions that rely on already established package. You need to perform the above steps and provide details on how the MVEF was derived.

Last Bonus Question (5 Points)

Note that the optimal portfolio from Equation (12) is consistent with the one from Equation (11). This implies that the scalar θ can written as a function of A_m. Can you find the value of this θ?

Hint: note that the SR portfolio has the following closed-form expression:

$$\boldsymbol{w}_{SR} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \tag{15}$$

with $\mathbf{1}$ is a column vector of ones. The key to this proof is to keep in mind that there is a unique level of A_m for which $\boldsymbol{w} = \boldsymbol{w}_{SR}$. Refer to this level as A_S. Based on which express θ as a function of A_m and A_S. Note that, after all, the value of θ should depend on m, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$.