# Introduction to Financial Risk Management

Fall 2022

Project 1

## Instructions

1. This is a teamwork project - please check your assigned team in the course website.
2. There are four main questions in this project.
3. All team members are expected to contribute to the project. Please **avoid working** on each question **independently**. This defeats the whole purpose of the assignment.
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 <u>not</u> be accepted**.
8. Please submit a **pdf copy of your final report**. Also, please submit the source file of the conducted computations. Note that the main evaluation of the project will be conducted based on the pdf file.
9. Utilize the length of your report carefully. 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 pictures/snapshots**. You should report your results in an organized table. The same applies to plots and other visualizations - **do <u>not</u> 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 following ETF data for all questions. In particular, you need to download adjusted closing market prices for the following 30 exchange-traded funds (ETFs). The ETF information comes from ishares.com and the raw data of the following table is available on canvas under the data module:

| Symbol | Name | Net Assets | Expense Ratio | ESG Quality |
|---|---|---|---|---|
| IVV | iShares Core S&P 500 | 339934.00 | 0.03 | 6.6 |
| IJH | iShares Core S&P Mid-Cap | 70755.00 | 0.05 | 6 |
| IWF | iShares Russell 1000 Growth | 69161.00 | 0.19 | 6.64 |
| IJR | iShares Core S&P Small-Cap | 66490.00 | 0.06 | 5.39 |
| IWM | iShares Russell 2000 | 52127.00 | 0.19 | 5.25 |
| IWD | iShares Russell 1000 Value | 48008.00 | 0.19 | 6.37 |
| IVW | iShares S&P 500 Growth | 33788.00 | 0.18 | 6.77 |
| IWB | iShares Russell 1000 | 29122.00 | 0.15 | 6.52 |
| IVE | iShares S&P 500 Value | 23998.00 | 0.18 | 6.39 |
| IUSG | iShares Core S&P U.S. Growth | 13434.00 | 0.04 | 6.74 |
| IUSV | iShares Core S&P U.S. Value | 13277.00 | 0.04 | 6.36 |
| IWV | iShares Russell 3000 | 11019.00 | 0.20 | 6.45 |
| IYW | iShares U.S. Technology | 10719.00 | 0.40 | 6.9 |
| IWN | iShares Russell 2000 Value | 10350.00 | 0.24 | 5.09 |
| IWO | iShares Russell 2000 Growth | 9422.00 | 0.24 | 5.42 |
| IJK | iShares S&P Mid-Cap 400 Growth | 7385.00 | 0.17 | 6.12 |
| IJJ | iShares S&P Mid-Cap 400 Value | 6613.00 | 0.18 | 5.87 |
| IJS | iShares S&P Small-Cap 600 Value | 6306.00 | 0.18 | 5.53 |
| IJT | iShares S&P Small-Cap 600 Growth | 4856.00 | 0.18 | 5.25 |
| IYH | iShares U.S. Healthcare | 3107.00 | 0.40 | 6.93 |
| IYF | iShares U.S. Financials | 1781.00 | 0.40 | 5.42 |
| IYY | iShares Dow Jones U.S. | 1577.00 | 0.20 | 6.52 |
| IYK | iShares U.S. Consumer Staples | 1409.00 | 0.40 | 7.21 |
| IYE | iShares U.S. Energy | 1356.00 | 0.40 | 6.45 |
| IYG | iShares U.S. Financial Services | 1151.00 | 0.40 | 6.25 |
| IYJ | iShares U.S. Industrials | 1100.00 | 0.40 | 6.96 |
| IDU | iShares U.S. Utilities | 757.00 | 0.40 | 6.78 |
| IYC | iShares U.S. Consumer Discretionary | 733.00 | 0.40 | 5.84 |
| IYM | iShares U.S. Basic Materials | 675.00 | 0.40 | 6.72 |
| IYZ | iShares U.S. Telecommunications | 265.00 | 0.40 | 6.12 |

For the market portfolio, consider the IVV ETF to represent the stock market. Merged altogether, the data should range between Jan 2010 and Sept 2023 (included). The rest of the project will be referring to this dataset. In all cases, assume that the risk-free rate is zero.

**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 $T \times 30$ matrix of returns, where the columns correspond to the 30 ETFs and the rows to the daily log-returns. Overall, you should have around $T \approx 3,500$ daily returns for each individual ETF.

## 1 Performance Summary (50 Points)

Given this data matrix of daily returns, address the following:

1. For each ETF, 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 a $30 \times 3$ table - where columns correspond to the performance measure and rows correspond to the ETF - report a summary for each measure. Specifically, for each performance metric, report the mean, first quartile (Q1), median, and third quartile (Q3) such that, in total, your summary consists of a $4 \times 3$ table. (10 Points)

2. Plot the asset means returns against their volatilities. Provide a couple of insights. (5 Points)

3. Let us consider the IVV ETF as the market risk factor to price the relative performance of each asset and compute the following measures for each ETF:

   - Annual Jensen's $\alpha$
   - Market $\beta$
   - Treynor Ratio (TR)
   - Tracking error ($\omega$)
   - Information Ratio (IR)

   (a) As in Task 1 above, report the relative performance measures in a $4 \times 5$ summary table, where each column represents a relative performance measure, and rows correspond to the four statistics, i.e., the mean, first quartile (Q1), median, and third quartile (Q3). (10 Points)

   (b) Which are the worst and best-performing ETFs? Explain the rationale behind your assessment (5 Points)

   (c) How does the expense ratio of these ETFs relate to the fund's performance? (10 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. (5 Points)

   (b) Does the CAPM hold? Why? (5 Points)

## 2 Mean-Variance Efficient Frontier (65 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 strategy. In doing so, you would like to ensure that each strategy delivers the best risk-return trade-off. Your **task** is to construct a mean-variance efficient frontier (MVEF) given the universe of the 30 assets. Specifically, out of the 30 assets, you choose the following three ETFs: IVV, IYW, and IYF. To construct, solve the following optimization problem for a given $m$:

$$\min_{\boldsymbol{w}} \sigma_p^2 = \boldsymbol{w}'\Sigma\boldsymbol{w} \tag{1}$$

subject to

$$\boldsymbol{w}'\mathbf{1} = 1 \tag{2}$$

$$\boldsymbol{w}'\boldsymbol{\mu} = m \tag{3}$$

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
- $\mathbf{1}$ denotes a $d \times 1$ vector of ones

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 **Solution 1** below and the full sample period to construct the MVEF. As a summary plot the MVEF. (10 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. (5 Points)

3. The two-funds separation theorem states that the MV optimal portfolio choice problem can be written as a convex 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

$$\boldsymbol{w}(\lambda) = \lambda\boldsymbol{w}_0 + (1 - \lambda)\boldsymbol{w}_{SR} \tag{4}$$

for constant $\lambda \in (-1, 1)$, 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 (4). As a summary,

(a) highlight this frontier using a red-dashed line on the previous plot you have. (10 Points)

(b) What does it mean to have a $\lambda < 0$? Provide an economic rationale. (5 Points)

4. Repeat the previous part but assume that you have access to a risk-free asset with return $R_F = 0$. Note that the same two-fund separation theorem from Equation (4) holds true in this case; however, the main difference is that the low risk fund becomes the risk-free asset. In other words, the vector of weights allocated to the risky assets becomes

$$\boldsymbol{w}_T = (1 - \lambda)\boldsymbol{w}_{SR} \tag{5}$$

whereas the weight allocated to the risk-free asset is

$$w_F = 1 - (1 - \lambda)\boldsymbol{w}_{SR}^\top \mathbf{1} = \lambda \tag{6}$$

Note that when $\lambda = 1$ you allocate 100% to the risk-free asset and when $\lambda = 0$ you allocate 0% to the risk-free asset. As a summary,

(a) Highlight this frontier using a blue-dotted line on the previous plot you had. (10 Points)

(b) Given this frontier, regress the portfolio mean returns on the portfolio volatility. Report the intercept and slope of this regression. (10 Points)

(c) What is meaning of the intercept and slope of this regression? (5 Points)

5. The last part of this question requires constructing the MVEF numerically using **Solution 2** described below. Following the instructions of Solution 2, consider a $10 \times 10$ grid of $m$ and $w$ values, which corresponds to 100 portfolios. Given these 100 portfolios, compute their mean returns and volatilities. Note that their mean returns should be equal to $m$, by construction. As a summary,

- Plot the mean returns of these portfolios against their volatilities
- The above should result in a "cloud" of 100 points.
- Given this cloud, highlight the upper envelope of this set.
- The highlighted line should correspond to the MVEF constructed based on Solution 1.
- For full credit, you need to add the new results to the MVEF from Solution 1 and compare both solutions. (10 Points)

**Solution 1**

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:

$$\boldsymbol{w}(A_m) = \boldsymbol{w}_0 + \frac{1}{A_m}\boldsymbol{w}_1 \tag{7}$$

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

$$A_m = \left[\frac{m - \boldsymbol{w}_0'\boldsymbol{\mu}}{\boldsymbol{w}_1'\boldsymbol{\mu}}\right]^{-1} \tag{8}$$

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 (8). In particular, set the values of $m$ to range between $\mu_0$ and $2\times\max(\mu_i)$ $\forall i = 1, .., d$. **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

$$\sigma_p(m) = \sqrt{\boldsymbol{w}(m)'\boldsymbol{\Sigma}\boldsymbol{w}(m)} \tag{9}$$

   **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.

**Solution 2**

Note that since we have three assets, the two constraints of the optimization problem are not binding. For this reason, the portfolio weights that satisfy the two constraints are not necessarily optimal. Therefore, we form the optimization problem to determine the optimal weights - as Solution 1 does. Specifically, Solution 1 indicates that for each mean target, there is a unique portfolio $\boldsymbol{w}(m)$, that satisfies $\boldsymbol{w}(m)'\boldsymbol{\mu} = m$ while attaining the lowest possible volatility, $\sigma_p(m)$.

Solution 1 is analytical and corresponds to a single portfolio for each $m$ value. In this challenge, you need determine the solution using numerically by generating different portfolios that satisfy the two constraints, while searching the portfolios that yield the lowest variance for each value $m$. In order to do so, note that the two constraints denote that

$$w_1 + w_2 + w_3 = 1 \tag{10}$$

$$w_1\mu_1 + w_2\mu_2 + w_3\mu_3 = m \tag{11}$$

For a given value of $w \in (-1, 1)$ and arbitrary value of $m$, you can define the values of $w_2$ and $w_3$, accordingly. Hence, your task here to generate a grid based on arbitrary values of $m$ and $w$.

## 3 Random Numbers and Monte Carlo Simulation (50 Points)

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

> **Very Important**: Make sure to explain the steps taken to conduct each experiment. An answer based solely on a code only will result in zero credit.

1. **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 **three** 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 three heads in a row for the first time, then your profit and loss ($PNL$) 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 $PNL$ is stochastic and can be described as

$$PNL = \$1 - X \times \$k \tag{12}$$

   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. (15 Points)

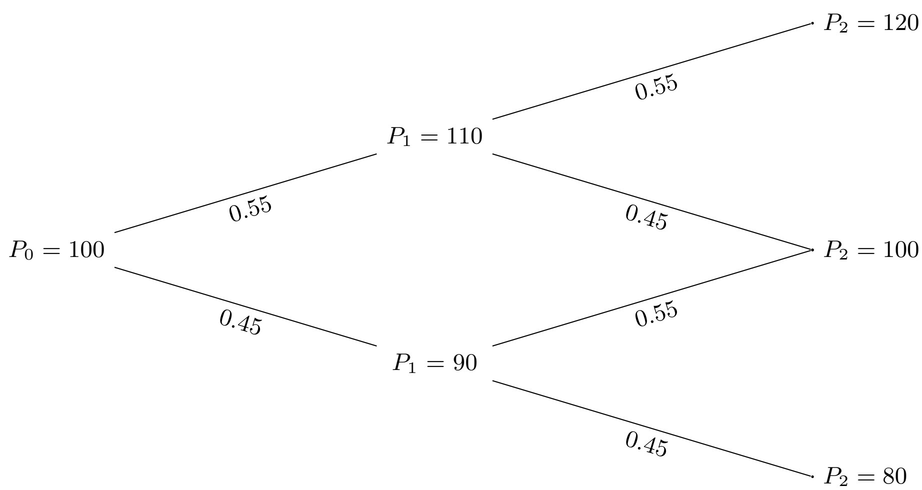
   **Hint:** Breaking even means that the expected value of the $PNL$ is zero. Your task, hence, is to find the value $k$ that leads to $\mathbb{E}[PNL] = 0$, which requires finding $\mathbb{E}[X]$.

---

2. **Turtles** On a one-way street, there are $n$ turtles crawling in the same direction. Their initial velocity is different, and can be assumed to be 1 (cm/s), 2 (cm/s), respectively... , $n$ (cm/s), and the initial position is completely random. Because it is a one-way street, when the rear turtle catches up with the front one, it will have to slow down and then crawl behind the front one to form a group. After enough time, how many groups will the turtles divide into? You will need to create $N > 10^3$ experiments and compute the average number of groups to approximate the correct answer. (15 Points)

---

3. **Trees** Consider a simple binomial model in which the stock price either goes up or down at the end of each period (step). Suppose that the probability of going up is 55%, whereas that of going down is 45%. Additionally, assume that the current price is $P_0 = 100$. At each step, the price can either go up by \$10 or drop by \$10. For instance, the potential price outcomes over two steps can be described using the following tree:




(a) Your first task is to find $\mathbb{E}[P_2]$ and $\mathbb{V}[P_2]$ analytically. This requires taking expectations for first and second moments of the price at step 2. Note that should result in the true rather than estimated values. (10 Points)

(b) Your second task is to use MC to simulate the price process for $M = 10$ steps, i.e., simulate $P_{10}(i)$ for $i = 1, ..., N$ with $N > 10^3$ denoting the total number of experiments. Note that the above tree corresponds to two steps ($M = 2$) and considers all potential outcomes. Your task does not require identifying all outcomes for the ten steps. Instead, your goal is to simulate different price outcomes and learn about the price distribution using the simulated prices. Based on the simulation, approximate $\mathbb{E}[P_{10}]$ and $\mathbb{V}[P_{10}]$, i.e., the expectation and variance at the tenth step given that the initial price is known $P_0 = 100$. (10 Points)

**Hint**: Since in the first task you find the true values for $\mathbb{E}[P_2]$ and $\mathbb{V}[P_2]$, run the simulation for $M = 2$ and check whether the results are consistent. Only after confirming so, you should do the ten steps simulation.

## 4 Risk Modeling (75 Points)

As a risk manager, your task is to evaluate the downside risk of a given asset. In particular, you need to compute the Value at Risk (VaR) for the IVV ETF based on the full period. To do so, you need to perform the following tasks:

(a) To get started, you need to work with **monthly** log-returns. The total time series should correspond to 165 months. Based on this data, calibrate the $\mu$ and $\sigma$, assuming that the asset price follows a geometric Brownian motion. Report the estimates in a table. (5 Points)

(b) Given the calibrated parameters. Assume that the initial price is given by the last closing price from the week of Jan 2010. This price should be around \$80. Based on this information and Monte Carlo (MC) simulation, you need to simulate the asset price path 1000 times over 165 months. To ensure your results are correct, compute the simulated conditional expectation and variance of $S_T$ given $S_0$ and compare them to the true values (recall lab 2). You should report all four values in a single table. (10 Points)
**Hint**: recall Lab 2.

(c) In reality, there was a single trajectory of how the IVV price played out during this sample period. On the other hand, the MC simulation results in 1000 paths. Clearly, out of the 1000 paths, there is one simulation that deviates the most and one that deviates least from the true trajectory. Can you spot the latter? As a summary, plot this simulated path versus the actual one. (5 Points)
**Hint**: There is a zero chance that a simulated path will be exactly the same as the true one. However, one can measure the distance between the simulated path and the true one over time. One potential candidate to measure such distance is the second norm of the differential between the two paths. This is similar to the sum of squared errors. The goal is to identify the path with the lowest discrepancy with the actual data.

(d) Suppose you purchased 100 IVV ETFs at the end of Jan 2010. Based on the simulated paths, compute the 99% VaR of this portfolio position at the end of Sep 2023. (10 Points)

(e) Clearly, the above computed VaR depends on the calibrated $\mu$ and $\sigma$. Instead of computing the VaR based on the calibrated $\hat{\sigma}$, consider a sequence of values for $\sigma \in \{0.10, 0.11, ..., 0.49, 0.50\}$, while keeping $\mu$ constant and equal to the calibrated value from Part (a). For each value, you need to rerun the 1000 simulations and compute the 99% VaR from the last part as a function of $\sigma$. As a summary, address the following

   i. Plot the 99% VaR as a function of $\sigma$. (10 Points)

   ii. What does the last part tells us in terms of *model risk*, i.e., the importance of getting the suitable model or calibrated parameters? (5 Points)

   iii. How do you justify the shape of the function from Part (e)? Is it linear or nonlinear? For instance, how does it compare with the case of an IID Gaussian process? (5 Points)
   **Hint**: Recall Lab 3.

(f) Instead of computing the 99% VaR of the portfolio position at the end of Sept 2023, you are interested in computing the monthly 99% VaR of the portfolio return using two approaches:

   i. **Historical**: In this case, assume that the 165 monthly returns represent the return on the portfolio position. Based on these realized returns, compute the 99% VaR? (5 Points)
   **Note**: In this case, the 99% VaR corresponds to the average historical returns minus the 99% historical percentile.

   ii. **Parametric**: This approach relies on the MC simulated paths from Part (b). Given the simulated paths, you can compute the monthly log-returns. This will result in a large simulated matrix of returns $165 \times 1000$. Based on this "large" data, compute the 99% VaR as you did in the previous part (historical approach). (5 Points)

   iii. How do both risk metrics compare? (5 Points)

   iv. As a final perspective, plot the density of the historical returns versus the simulated returns. What do you observe? (5 Points)

   v. What did you learn from this exercise in terms of model risk? (5 Points)

---

### Bonus Question 2 (10 Points)

**VaR for GBM**: In fact, the VaR of the portfolio position from Part (d) has a closed-form solution. The fact that we can identify the closed-from solution makes the MC analysis redundant! Can you derive such a closed-form solution? To qualify for full credit, you need to demonstrate the derivations/steps carefully. Additionally, you need to demonstrate the consistency of this closed-form solution with the simulated one from Part (e).