Python for Finance: Modern Portfolio Theory

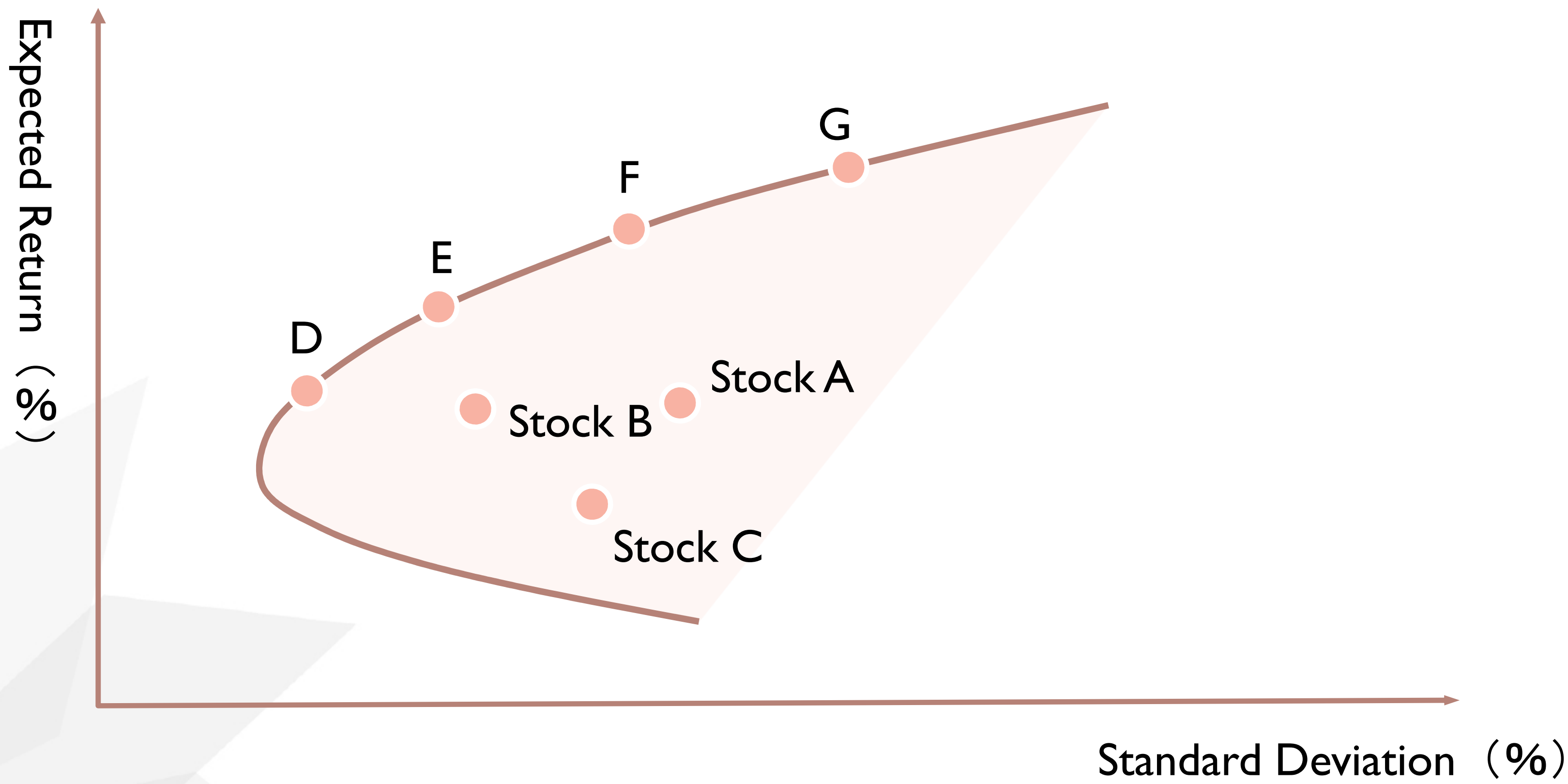
Mamoru Nagano

COPYRIGHT © 2024 MAMORU NAGANO

1. Mean and Variance Analysis

- ✓ The mean-variance approach originates from the idea that the average return and its variance (volatility) of the assets used to derive the optimal asset composition.
- ✓ This theory shows how, when holding multiple assets, the composition that provides the optimal return and risk is calculated.
- ✓ In short, average (expected) return is the mean (expected) value of the past (future) rate of return in asset prices.
- ✓ It is expected that this past rate of return distribution will be similarly scattered in the future under a certain probability distribution. The expected return is expressed as the "expected return."

The shaded areas are combinations of expected return and risk that investors can choose from.



2. Expected Return (Average Return)

- ✓ Modern finance theory assumes that expected returns as a random variable follow a certain probability distribution.

- ✓ Therefore, under this probability, it is highly likely that various returns based on a certain probability distribution will appear in the future.

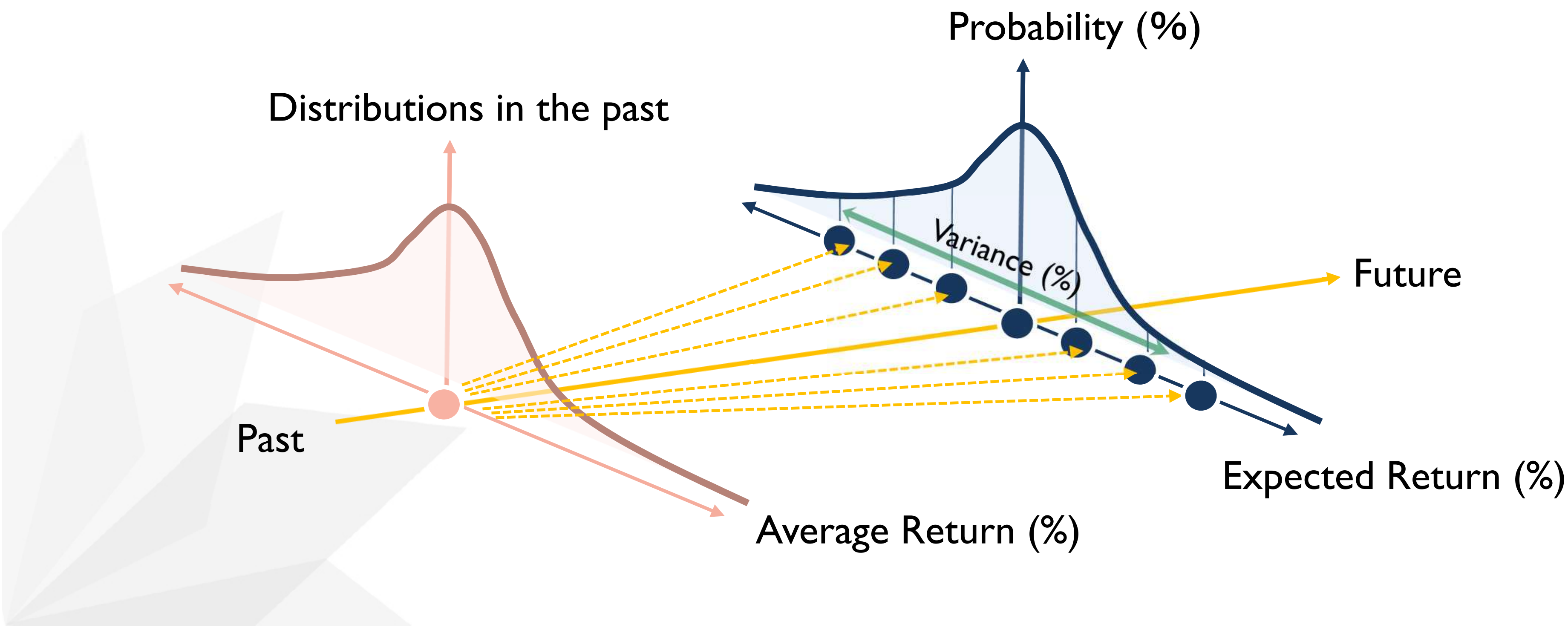
- ✓ Generally, the average return of a portfolio of assets is defined by the weighted average of the following formula.

$$E(R_p) = \sum_{i=1}^{N} \gamma_i E(R_i)$$

- ✓ In other words, if the two assets are 50%:50%

$$E(R_p) = 0.5\, E(R_A) + 0.5\, E(R_B)$$

✓ Modern portfolio theory assumes that the distribution of portfolio returns over a certain period of time in the past will be closely approximated with the future.



3 Risk ≒ Variance and Covariance

- ✓ For portfolios of two or more assets (ex. *N* assets), the formula for calculating the total variance includes the covariance Cov(R_i, R_j) in the $N \times (N-1)/2$ pattern.
- ✓ By applying this covariance, it is possible to minimize the total variance (total standard deviation) of the entire portfolio.

$$Var(R_p) = \sum_{i=1}^{N} \gamma_i^2 \, Var(R_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} 2\gamma_i \, \gamma_j Cov(R_i, R_j)$$

If you invest 50% each in two financial assets, and let R_A be the return of stock A and R_B be the return of stock B, the variance of this entire portfolio R_p is as follows.

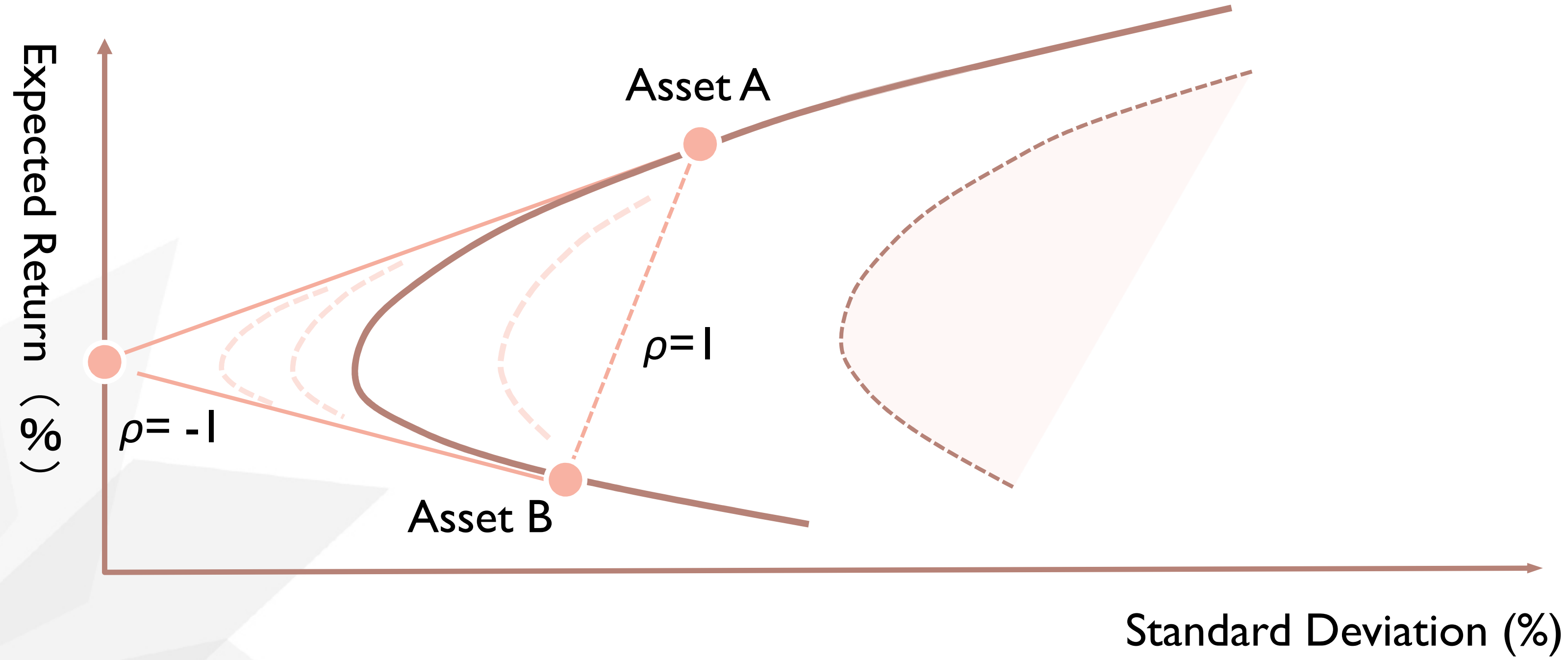
$$Var(R_p) = 0.5^2 \mathrm{Var}(R_A) + 0.5^2 \mathrm{Var}(R_B) + 2 \times 0.5 \times 0.5 \times Cov(R_A, R_B)$$

Incidentally, this covariance is calculated by the following formula.

$$Cov(R_A, R_B) = \frac{1}{N}\{R_A - E(R_A)\}\{R_B - E(R_B)\}$$

4. Correlation Coefficient and Covariance

✓ Choosing with low (high) correlation assets is nothing but shifting the efficient frontier to the left (right).



- ✓ Covariance is defined as the value obtained by subtracting the respective average returns $E(R_A)$ and $E(R_B)$ from each return R_A and R_B and dividing the result by the number of samples N.

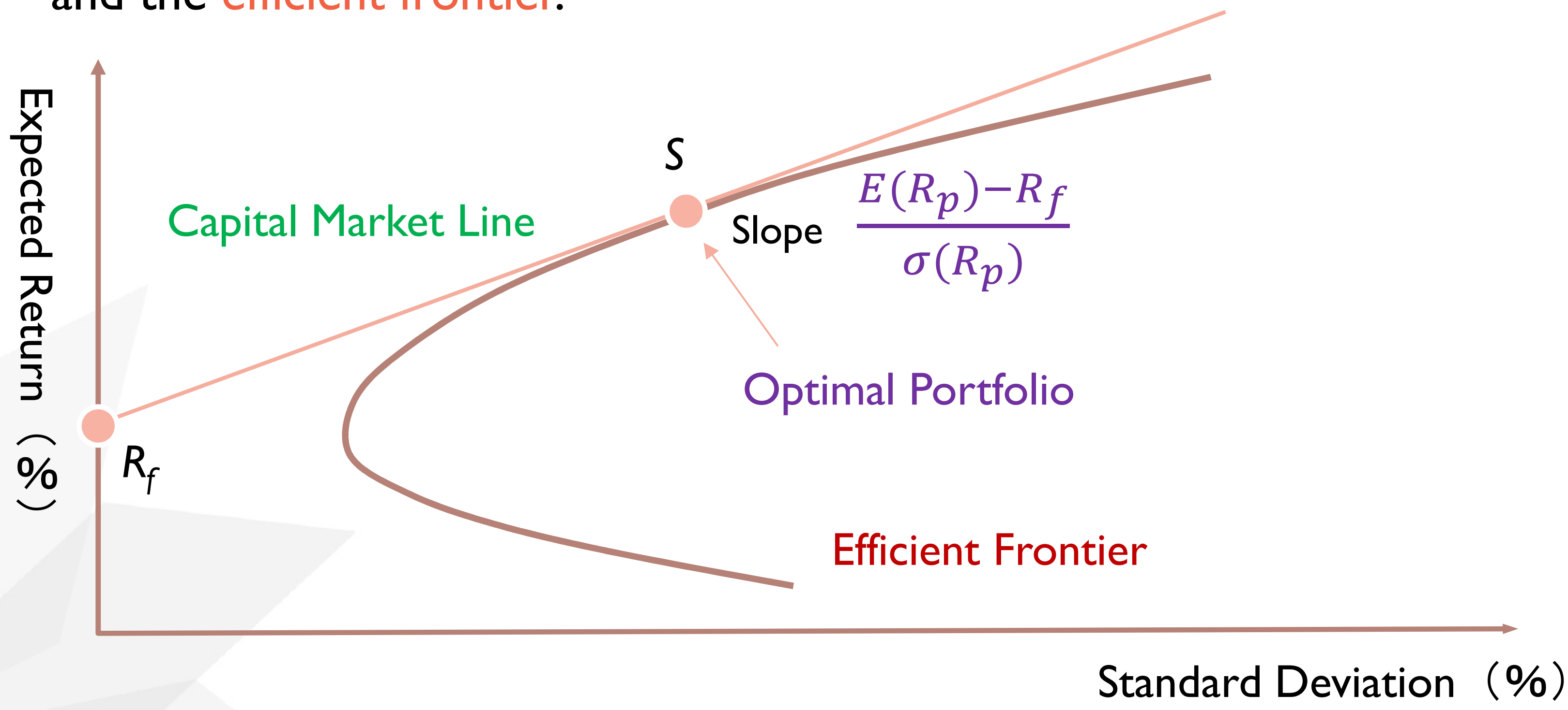
$$Cov(R_A, R_B) = \frac{1}{N}\{R_A - E(R_A)\}\{R_B - E(R_B)\}$$

- ✓ The correlation coefficient ρ between R_A and R_B is calculated by dividing the covariance Cov(R_A, R_B) by the product of the standard deviations σ(R_A) and σ(R_B) of R_A and R_B.

$$\rho(R_A, R_B) = \frac{Cov(R_A, R_B)}{\sigma(R_A)\sigma(R_B)}$$

4 Capital Market Line (CML)

✓ The optimal portfolio is the intersection between the capital market line and the efficient frontier.



✓ Assuming that the expected return of the portfolio is E(R_p), this risk (standard deviation) is σ(R_p), and the expected return of risk-free assets is R_f, the Sharpe ratio $S(R_p)$ is defined as follows.

$$S(R_p) = \frac{E(R_p) - R_f}{\sigma(R_p)}$$

【注】Sharpe, William F. (1966), "Mutual Fund Performance," *The Journal of Business*, Vol. 39, Issue 1: pp. 119–138.

5 Separation Theorem and Sharpe Ratio

- ✓ When we define the expected return on the capital market line is $E(R_M)$, the composition ratio of safe assets is 1-β, the expected return is R_f, the allocation ratio of risky assets is β, and the expected return is $E(R_p)$, then the following expression is obtained.

$$E(R_M) = (1-\beta)R_f + \beta E(R_p)$$

- ✓ The risk (standard deviation) $\sigma(R_M)$ of a portfolio on the capital market line is equal to $\beta\sigma(R_p)$, which is the risk (standard deviation) of risky assets multiplied by the composition ratio.

$$\sigma(R_M) = \beta\sigma(R_p)$$

- ✓ From two equations on the previous page, we get the following equation.

$$E(R_M) = \frac{E(R_p) - R_f}{\sigma(R_p)} \sigma(R_M) + R_f$$

- ✓ Consequently, the slope, which is the coefficient of this σ(R_M), is consistent with the definition of Sharpe ratio, and the constant term is equal to the yield of risk-free assets such as government bonds.

6 Efficient Frontier : Libraries

✓ To derive the efficient frontier of two assets, we will use the following four libraries that we learned last time.

```
#[1] Import Libraries

from pandas_datareader import data
import datetime
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

7 Efficient Frontier: Data

- ✓ In this seminar, we acquire daily stock price data of Hong Kong H-share companies Tencent (700.HK) and Zhejiang Geely Automobile Holdings (175.HK).

```
#[2]Retrieval of Daily Stock Price Data

tickers=['700.HK','175.HK’]
start=datetime.date(2019,1,1)
end=datetime.date(2021,12,31)
data=data.DataReader(tickers,'stooq',start=start,end=end).sort
_values(by='Date',ascending=True)
```

8 Efficient Frontier: Pre-Processing

- ✓ Of the acquired data, only the closing price (Close) is stored in a new data frame df.
- ✓ In the second line, the original data are converted to the returns of the closing prices stored in the dataframe df and the missing values are removed.

```
#[3]Data Pre-Processing

df=data.Close
df=df.pct_change().dropna()
```

9 Efficient Frontier: Mean and Variance

- ✓ The first line calculates the average returns of the stock price returns.
- ✓ The second line calculates the variances of the stock price returns.
- ✓ The third line calculates the covariance of the two stock price returns of 700.HK and 175.HK.

```
#[4] Calculation of mean, variance, and covariance

Rp=df.mean()*250
VAR_Rp=df.var()*250
Cov_Rp=df['700.HK'].cov(df['175.HK'])*250
```

10. Efficient Frontier: Returns

- ✓ Using np.arange(A, B, C) to generate a numerical value from the minimum A% to the maximum C% in steps of B%, and name this data as *weight*.
- ✓ In the second line, use this *weight* to increase the value by 1% from 0% (100%) to 100% to calculate the 101 patterns of expected return combinations.

```
#[5] Calculation of Average Returns by Composition

weights=np.arange(0,1.01,0.01)
for i in np.arange(0,1.01,0.01):
  E_Rp=weights*Rp["700.HK"]+(1-weights)*Rp["175.HK"]
```

11. Efficient Frontier: Risks

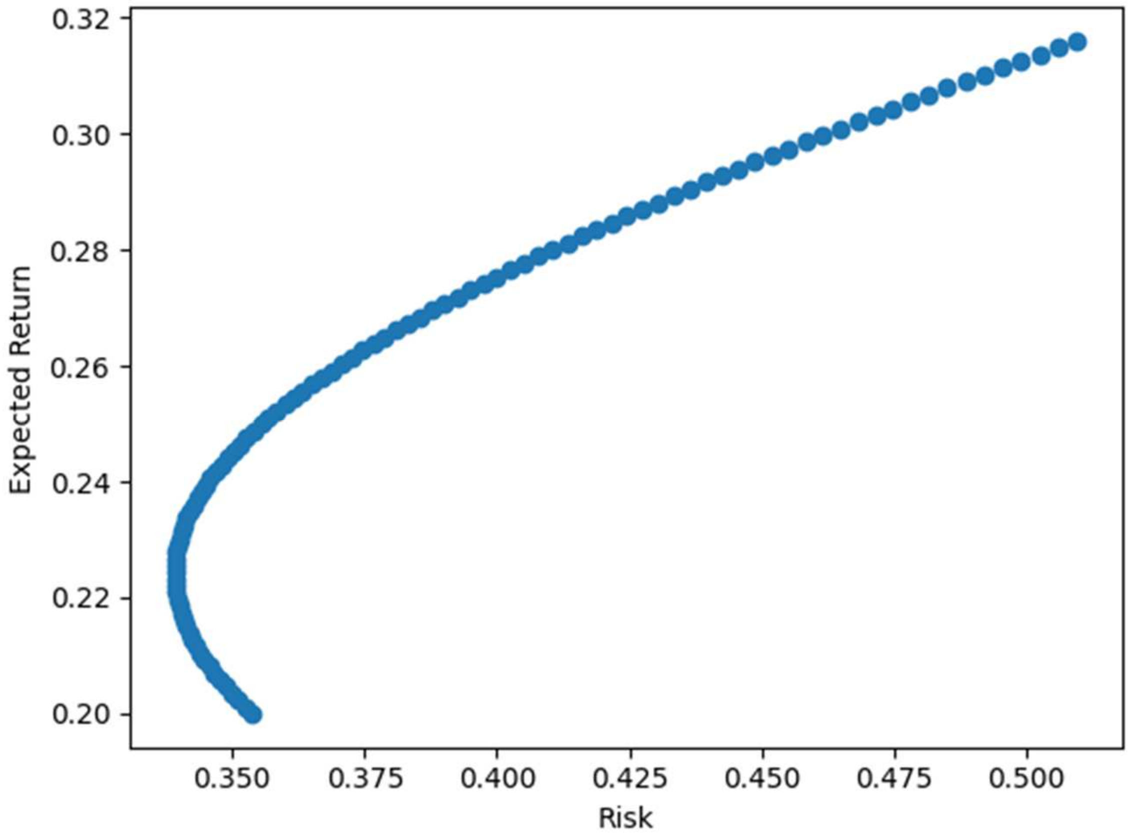
- ✓ Using varaible *weight* generated on the previous page, increase the ratio by 1% from 0% (100%) to 100% and calculate the variance by combining 101 patterns of two as sets.
- ✓ In the second line, convert this variance to standard deviation using np.sqrt.

```
#[6] Calculation of Risks by Composition

for i in np.arange(0,1.01,0.01):
  Var_Rp2=weights*weights*VAR_Rp["700.HK"]+(1-weights)*(1-weigh
ts)*VAR_Rp["175.HK"]+2*weights*(1-weights)*Cov_Rp

Sigma_Rp=np.sqrt(Var_Rp2)
```

12 Efficient Frontier: Visualization

✓ Finally, draw the scatter plot using matplotlib.

```
#[7]Vidualization

plt.scatter(Sigma_Rp,E_Rp)
plt.xlabel("Risk")
plt.ylabel("Expected Return")
```


0.32
0.30
0.28
0.26
0.24
0.22
0.20
Expected Return
0.350
0.375
0.400
0.425
0.450
0.475
0.500
Risk

13 Libraries in this Class

pandas_dataframe
datetime

Pandas	.pct_change()	.dropna()	.describe()
	.mean()	.std()	.var()
	.value()	.cov()	
NumPy	.mean()	.std()	.var()
	.arange()		
matplotlib	.plt.scatter()		

Contact:

Mamoru Nagano,

Professor of Finance,
Seikei University, Japan
mnagano@econ.seikei.ac.jp