DATA SCIENCE

Introduction to portfolio theory

Antonio Violi

antonio.violi@unirc.it

Outline

- Introduce the portfolio theory
- Portfolio is defined starting from an universe of assets (stocks)
- A portfolio is simply a specific combination of assets, usually defined by portfolio weights that sum up to 1
- How Do We Construct a Good Portfolio?
- What characteristics do we care about for a given portfolio?
  - Reward (Investors like higher expected returns)
  - Risk (Investors dislike risk)
- We describe the portfolio performance in terms of expected return and risk

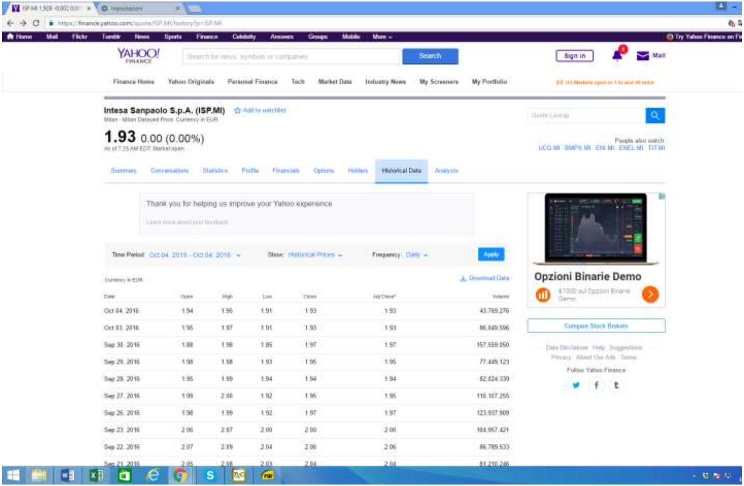
Introduction

- Typically, when making an investment, the initial outlay of capital is known, but the amount to be returned when selling an asset is uncertain
- We know the "'today price"' of a given stock, but we do not know what the "'tomorrow price"' will be
- Which criteria can guide our selection process?
- We can consider the expected return and the risk
- Typically the higher the return the riskier is the investment
- Find the proper trade-off between the two elements according to the risk profile of the investor

Introduction (2)

- Return is unknown since it will computed on the basis of the future prices
- Mathematically, the future price (return) is a random variable defined on a given probability space.
- We shall assume that the random variable is **discrete** so it can take a finite number of specific values (realizations)
- How realizations can be determined ?
  1. Assume that the future evolves as the past - use historical data as future realizations
  2. Generate realizations by using some simulation techniques (e.g. MonteCarlo)

Get Data



YAHOO! FINANCE
Search
Sign in
Mail
Finance Home
Yahoo Originals
Personal Finance
Tech
Market Data
Industry News
My Screeners
My Portfolio
Intesa Sanpaolo S.p.A. (ISP.MI)
Add to watchlist
1.93 0.00 (0.00%)
People also watch
UCG.MI
BMPS.MI
ENI.MI
ENEL.MI
TIT.MI
Summary
Conversations
Statistics
Profile
Financials
Options
Holders
Historical Data
Analysts
Thank you for helping us improve your Yahoo experience
Time Period: Oct 04, 2015 - Oct 04, 2016
Show: Historical Prices
Frequency: Daily
Apply
Currency in EUR
Download Data
Date
Open
High
Low
Close
Adj Close*
Volume
Oct 04, 2016 1.94 1.96 1.91 1.93 1.93 43,769,276
Oct 03, 2016 1.95 1.97 1.91 1.93 1.93 86,849,596
Sep 30, 2016 1.88 1.98 1.85 1.97 1.97 157,559,050
Sep 29, 2016 1.98 1.98 1.93 1.95 1.95 77,449,123
Sep 28, 2016 1.95 1.99 1.94 1.94 1.94 82,624,339
Sep 27, 2016 1.99 2.00 1.92 1.95 1.95 110,107,255
Sep 26, 2016 1.98 1.99 1.92 1.97 1.97 123,937,909
Sep 23, 2016 2.06 2.07 2.00 2.00 2.00 104,867,421
Sep 22, 2016 2.07 2.09 2.04 2.06 2.06 86,789,533
Opzioni Binarie Demo
Compare Stock Brokers
Follow Yahoo Finance

Assume the future evolves as the past . . .

- Starting from historical data it is possible to estimate what the expected return will be.
- Let us denote by P_{it} the t-th observation of the price of asset i.
- We denote by r_{it} the t-th return computed as

$$r_{it} = ln(\frac{P_{it}}{P_{it-1}})$$

- The average can be computed as

$$\bar{r}_i = \frac{1}{T}\sum_{t=1}^{T} r_{it}$$

- Reward is typically measured by the mean return
- Higher returns are better than lower returns.

Variance

The variance of the return is

$$\sigma_i^2 = \frac{1}{T-1}\sum_{t=1}^{T}(r_{it} - \overline{r}_i)^2$$

We frequently use the square root of the variance, σ, named **standard deviation**

Remind that since variance measures the variability (volatility) from the mean, it is used as a measure of risk

Get Data

	A	B	C	D	E	F	G	H	I	J	K
1	**Intesa Sanpaolo (ISP.MI)**										
2											
3	Date	Open	High	Low	Close	Volume	Adj Close	Return			
4	03/10/2016	1,95	1,974	1,915	1,928	8,6E+07	1,928	-0,023579	AVERAGE	-0,00624	-0,62%
5	30/09/2016	1,879	1,977	1,852	1,974	1,6E+08	1,974	0,014286	VARIANCE	0,001	0,10%
6	29/09/2016	1,976	1,983	1,928	1,946	7,7E+07	1,946	0,0015428	STANDARD DEVIATION		3,17%
7	28/09/2016	1,952	1,993	1,943	1,943	8,2E+07	1,943	-0,00257			
8	27/09/2016	1,992	2	1,922	1,948	1,1E+08	1,948	-0,012752			
9	26/09/2016	1,98	1,987	1,916	1,973	1,2E+08	1,973	-0,01159			
10	23/09/2016	2,062	2,066	1,996	1,996	1E+08	1,996	-0,032531			
11	22/09/2016	2,068	2,086	2,044	2,062	8,7E+07	2,062	0,0126893			
12	21/09/2016	2,054	2,08	2,034	2,036	8,1E+07	2,036	0,0138479			
13	20/09/2016	2,066	2,078	1,993	2,008	1E+08	2,008	-0,027507			
14	19/09/2016	2,04	2,086	2,028	2,064	8,9E+07	2,064	0,0245231			
15	16/09/2016	2,06	2,076	2,012	2,014	1,6E+08	2,014	-0,032245			
16	15/09/2016	2,062	2,104	2,044	2,08	1E+08	2,08	0			
17	14/09/2016	2,094	2,1	2,05	2,08	1,1E+08	2,08	-0,000961			
18	13/09/2016	2,154	2,154	2,082	2,082	9E+07	2,082	-0,01903			
19	12/09/2016	2,114	2,132	2,094	2,122	1,2E+08	2,122	-0,020523			

Generate future realizations by simulation techniques . . .

- Let us consider a single asset i and let us denote by $\widetilde{r}_i$ its random return
- We denote by $|S|$ the number of realizations (scenarios)
- For each scenario s we denote by
  - r_{is} the s-th realization
  - p_s the corresponding probability
- The expected value (or mean) $\overline{r_i}$ is

$$\mathbb{E}[\widetilde{r}_i] = \overline{r}_i = \sum_{s=1}^{|S|} p_s r_{is}$$

- Reward is typically measured by the mean return
- Higher returns are better than lower returns.

Variance

The variance of the return is measured by the average squared deviation from the mean:

$$VAR[\widetilde{r}_i] = \sigma_i^2 = \mathbb{E}[(\widetilde{r}_i - \overline{r}_i)^2] = \sum_{s=1}^{|S|} p_s (r_{is} - \overline{r}_i)^2$$

We frequently use the square root of the variance, σ, named **standard deviation**

Higher variance suggests less predictable returns and therefore a more risky investment.

Dominance criteria

By using the expected return and the standard deviation, it is possible to "'compare"' two investment alternatives

Definition

An investment A dominates an investment B if the following conditions are satisfied and one strictly holds

$$\bar{r}_A \geq \bar{r}_B$$
$$\sigma_A \leq \sigma_B$$

1. For investments of the same expected value, choose the one with the lowest standard deviation
2. For investments with same standard deviation, choose the one with the greatest expected value

Covariance

- When considering two or more random variables, their mutual dependence should be taken into account
- This is summarized conveniently by their covariance
- Let $\widetilde{r}_1$ and $\widetilde{r}_2$ two random variables denoting the uncertain return of asset 1 and 2
- The **covariance** is defined as

$$\sigma_{12} = \mathbb{E}[(\widetilde{r}_1 - \bar{r}_1)(\widetilde{r}_2 - \bar{r}_2)]$$

Covariance

- If two random variables have the property that $\sigma_{12} = 0$, then they are said to be **uncorrelated**
- In this case, the knowledge of the value of a variable gives no information about the value of the other
- If two random variables are independent, then they are uncorrelated
- If $\sigma_{12} > 0$ the variables are said **positively correlated**
  If one variable is above its mean, the other is likely to be above its mean as well
- If $\sigma_{12} < 0$, the two variables are said to **negatively correlated**

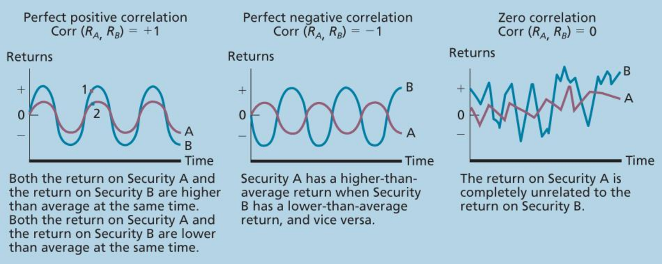
Correlation

Another useful construct is the **correlation coefficient** defined as

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$$

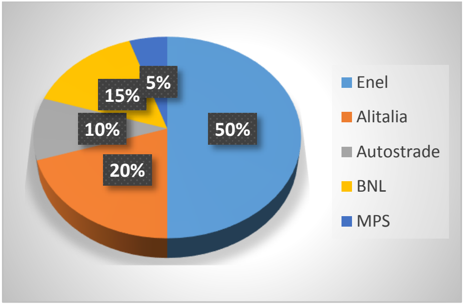
Note that $|\rho_{12}| \leq 1$

Correlation



From single asset to portfolio

- How Do We Construct a Good Portfolio?
- What characteristics do we care about for a given portfolio?
  - Risk and reward
  - Investors like higher expected returns
  - Investors dislike risk



Some properties

Given two random variables x and y the following properties hold

$$\mathbb{E}[x+y] = \mathbb{E}[x] + \mathbb{E}[y]$$
$$VAR[x+y] = \sigma^2_{x+y} = \sigma^2_x + \sigma^2_y + 2\sigma_{x,y}$$

Portfolio mean

- Suppose that there are N assets with random rates of return $\widetilde{r}_1, \ldots, \widetilde{r}_N$
- Suppose that we form a portfolio, by investing a given fraction of capital x_i in each asset i with $i = 1, \ldots, N$

$$\sum_{i=1}^{N} x_i = 1$$

- The rate of return of the portfolio in terms of the individual returns is

$$\widetilde{r}_P = \sum_{i=1}^{N} \widetilde{r}_i x_i$$

$\widetilde{r}_P$ is a random variable and we may compute its expected value and variance

Mean return of a portfolio

By applying the properties of the expected value

$$\mathbb{E}[\widetilde{r}_P] = \overline{r}_P = \sum_{i=1}^{N} \overline{r}_i x_i$$

The expect rate of return of a portfolio can be simply determined by the weighted sum of the individual expected return of the single assets

Variance of portfolio return

By applying the properties of the variance of the sum of random variables

$$\begin{aligned}
\sigma_P^2 &= \mathbb{E}[(\tilde{r}_P - \bar{r}_P)^2] = \\
&\mathbb{E}[(\sum_{i=1}^{N} \tilde{r}_i x_i - \sum_{i=1}^{N} \bar{r}_i x_i)^2] = \\
&\mathbb{E}[(\sum_{i=1}^{N} x_i(\tilde{r}_i - \bar{r}_i))(\sum_{j=1}^{N} x_j(\tilde{r}_j - \bar{r}_j))] = \\
&\mathbb{E}[\sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j(\tilde{r}_i - \bar{r}_i)(\tilde{r}_j - \bar{r}_j)] = \\
&\sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j \sigma_{ij} = \sum_{i=1}^{N} \sigma_i^2 x_i^2 + \sum_{i=1}^{N} \sum_{j=1 \;\; j \neq i}^{N} x_i x_j \sigma_{ij}
\end{aligned}$$

Diversification

- Portfolio with only a few assets may be subject to a high degree of risk, represented by a relatively large variance
- As a general rule, variance can be reduced by including additional assets in the portfolio
- This process is called **diversification**
- **Don't put all your eggs in one basket !!**

Diversification: uncorrelated case

- First we analyze the case of uncorrelated assets
- Let us assume to invest the same fraction of capital in every asset $x_i = \frac{1}{N}$

$$\sigma_P^2 = \sum_{i=1}^{N} \sigma_i^2 (\frac{1}{N})^2 = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \frac{1}{N} = \frac{1}{N} \overline{\sigma_i}^2$$

- By increasing the number of asset in the portfolio, the variance $\rightarrow 0$

Diversification: correlated case

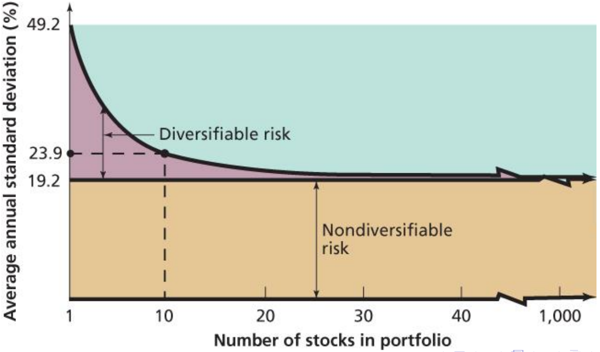
- The situation is different if the returns of the available assets are correlated

$$\sigma_P^2 = \frac{1}{N}\overline{\sigma_i}^2 + \frac{N-1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}{}_{j\neq i}\frac{\sigma_{ij}}{N(N-1)}$$

- By including more assets in the portfolio, the term related to variance tends to 0, whereas the covariance still remains

Diversification

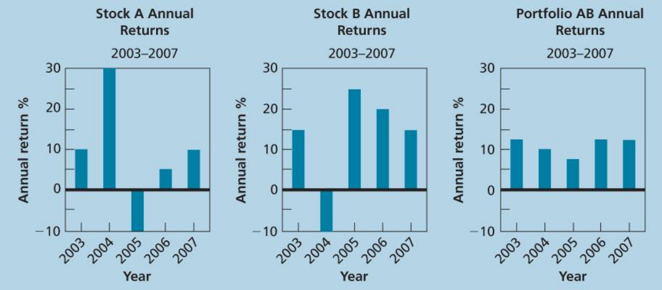
- Imperfect correlation is the key reason why diversification reduces portfolio risk
- Nondiversifiable risk is also known as **systematic** risk and it is due to common factors that cannot be diversified
- Diversifiable risk is known as **specific** risk and depends on the specific features of a firm



Why diversification works

Annual Returns on Stocks A and B			
Year	**Stock A**	**Stock B**	**PortfolioAB**
2003	10,00%	15,00%	12,50%
2004	30,00%	-10,00%	10,00%
2005	-10,00%	25,00%	7,50%
2006	5,00%	20,00%	12,50%
2007	10,00%	15,00%	12,50%
Average returns	9,00%	13,00%	11,00%
Standard deviations	11,69%	12,08%	3,07%

Why diversification works



Perfect positive correlation

- We consider two assets A and B that have a perfect positive correlation
- The mean portfolio return is

$$\bar{r}_P = \bar{r}_A x_A + \bar{r}_B x_B$$

- x_A and x_B denote the fraction of capital invested in A and B, respectively
- Assuming that $\rho_{AB} = 1$, the variance of portfolio return

$$\sigma_P^2 = \sigma_A^2 x_A^2 + \sigma_B^2 x_B^2 + 2\sigma_A \sigma_B x_A x_B = (\sigma_A x_A + \sigma_B x_B)^2$$

- Standard deviation

$$\sigma_P = \sigma_A x_A + \sigma_B x_B$$

Perfect negative correlation

- We consider two assets A and B that have a perfect negative correlation
- The mean portfolio return is

$$\bar{r}_P = \bar{r}_A x_A + \bar{r}_B x_B$$

- x_A and x_B denote the fraction of capital invested in A and B, respectively
- $x_A + x_B = 1$
- Assuming that $\rho_{AB} = -1$, the variance of portfolio return

$$\sigma_P^2 = \sigma_A^2 x_A^2 + \sigma_B^2 x_B^2 - 2\sigma_A \sigma_B x_A x_B = (\sigma_A x_A - \sigma_B x_B)^2$$

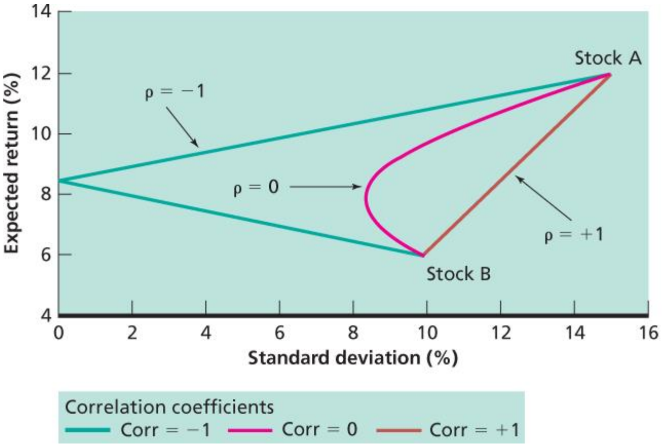
- Standard deviation

$$\sigma_P = |\sigma_A x_A - \sigma_B x_B|$$

Correlation effect

- For $-1 < \rho_{AB} < 1$ we have an hyperbolic function with a point of minimum

Effect of correlation

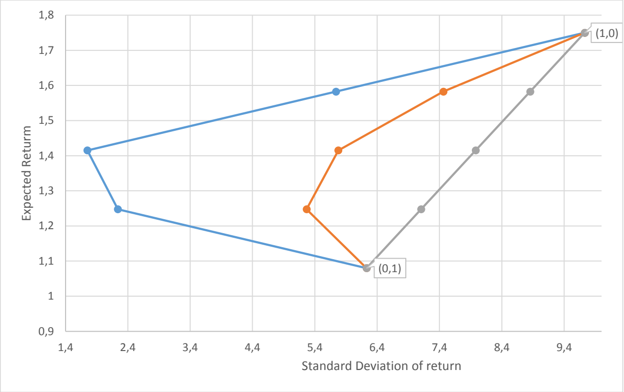


An example

- Let us consider two assets A and B with the following values of expected returns and variances
  $\bar{r}_A = 1,75$, $\bar{r}_B = 1,08$
  $\sigma_A = 9,73$, $\sigma_B = 6,23$
- Suppose that the correlation between A and B may change
- Consider the case $\rho_{AB} = -1$, $\rho_{AB} = 0$, $\rho_{AB} = 1$
- Let us analyze the correlation effect

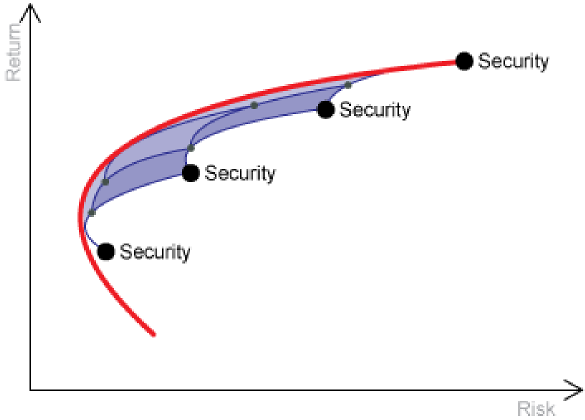
Portfolio Composition		Average Portfolio Return	Standard Deviation		
			corr = -1	corr = 0	corr = 1
0	1	1,08	6,23	6,23	6,23
0,25	0,75	1,25	2,24	5,27	7,11
0,5	0,5	1,42	1,75	5,78	7,98
0,75	0,25	1,58	5,74	7,46	8,86
1	0	1,75	9,73	9,73	9,73

Graphically



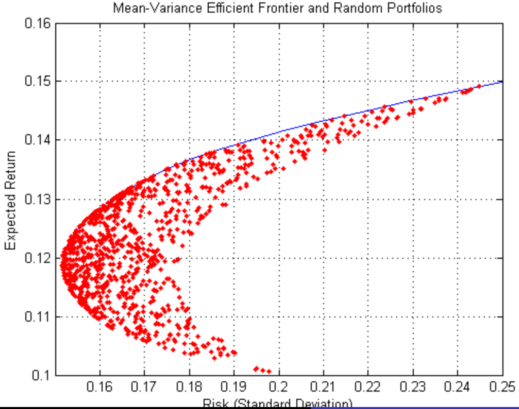
The feasible set

Combining more assets



The feasible set

- The left boundary of a feasible set is called the **minimum-variance** set, since for any value of the mean rate of return the boundary points have the smallest variance
- On this set there is a special point termed the **minimum-variance point**



Efficient frontier

The upper portion of the minimum variance set is termed **efficient frontier**

The points of the efficient frontier represent portfolio that are not dominated.

- For a given mean return, they have the minimum risk
- For a given level of risk, they guarantee the maximum reward