

Covariance and correlation

Let random variables X, Y with means μ_X, μ_Y respectively. The covariance, denoted with $cov(X, Y)$, is a measure of the association between X and Y .

Definition:

$$cov(X, Y) = E(X - \mu_X)(Y - \mu_Y)$$

This can be simplified as follows:

$$cov(X, Y) = E(X - \mu_X)(Y - \mu_Y) = E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_X \mu_Y$$

Therefore,

$$cov(X, Y) = E(XY) - (EX)(EY)$$

Note: If X, Y are independent then $E(XY) = (EX)E(Y)$ Therefore $cov(X, Y) = 0$.

Let W, X, Y, Z random variables, and a, b, c, d constants:

- Find $cov(a + X, Y)$

$$cov(a + X, Y) = E(a + X - \mu_{a+X})(Y - \mu_Y) = E(a + X - \mu_X - a)(Y - \mu_Y)$$

Therefore, $cov(a + X, Y) = cov(X, Y)$.

- Find $cov(aX, bY)$

$$cov(aX, bY) = E(aX - \mu_{aX})(bY - \mu_{bY}) = E(aX - a\mu_X)(bY - b\mu_Y)$$

Therefore, $cov(aX, bY) = abE(X - \mu_X)(Y - \mu_Y) = ab cov(X, Y)$

- Find $cov(X, Y + Z)$

$$cov(X, Y + Z) = E(X - \mu_X)(Y - \mu_{Y+Z}) = E(X - \mu_X)(Y + Z - \mu_Y - \mu_Z)$$

Or

$$cov(X, Y + Z) = E(X - \mu_X)(Y - \mu_Y + Z - \mu_Z) = E(X - \mu_X)(Y - \mu_Y) + E(X - \mu_X)(Z - \mu_Z)$$

Therefore, $cov(X, Y + Z) = cov(X, Y) + cov(X, Z)$

- Using the results above we can find $cov(aW + bX, cY + dZ)$.

$$cov(aW + bX, cY + dZ) = ab cov(W, Y) + ad cov(W, Z) + bc cov(X, Y) + bd cov(X, Z)$$

- Correlation:

However, the covariance depends on the scale of measurement and so it is not easy to say whether a particular covariance is small or large. The problem is solved by standardize the value of covariance (divide it by $\sigma_X \sigma_Y$), to get the so called coefficient of correlation ρ_{XY} .

$$\rho = \frac{cov(X, Y)}{\sigma_X \sigma_Y}, \quad \text{Always, } -1 \leq \rho \leq 1$$

$$cov(X, Y) = \rho \sigma_X \sigma_Y$$

If X, Y are independent then $\rho_{XY} = 0$

- Important:

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y)$$

Proof:

- Find $\text{var}(aX + bY)$

- In general: Let X_1, X_2, \dots, X_n be random variables, and a_1, a_2, \dots, a_n be constants. Find the variance of the linear combination $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n$.

Portfolio risk and return

An investor has a certain amount of dollars to invest into two stocks (*IBM* and *TEXACO*). A portion of the available funds will be invested into *IBM* (denote this portion of the funds with a) and the remaining funds into *TEXACO* (denote it with b) - so $a + b = 1$. The resulting portfolio will be $aX + bY$, where X is the monthly return of *IBM* and Y is the monthly return of *TEXACO*. The goal here is to find the most efficient portfolios given a certain amount of risk. Using market data from January 1980 until February 2001 we compute that $E(X) = 0.010$, $E(Y) = 0.013$, $Var(X) = 0.0061$, $Var(Y) = 0.0046$, and $Cov(X, Y) = 0.00062$.

We first want to minimize the variance of the portfolio. This will be:

$$\begin{aligned} &\text{Minimize } \text{Var}(aX + bY) \\ &\text{subject to } a + b = 1 \end{aligned}$$

Or

$$\begin{aligned} &\text{Minimize } a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y) \\ &\text{subject to } a + b = 1 \end{aligned}$$

Therefore our goal is to find a and b , the percentage of the available funds that will be invested in each stock. Substituting $b = 1 - a$ into the equation of the variance we get

$$a^2\text{Var}(X) + (1 - a)^2\text{Var}(Y) + 2a(1 - a)\text{Cov}(X, Y)$$

To minimize the above expression we take the derivative with respect to a , set it equal to zero and solve for a . The result is:

$$a = \frac{\text{Var}(Y) - \text{Cov}(X, Y)}{\text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)}$$

and therefore

$$b = \frac{\text{Var}(X) - \text{Cov}(X, Y)}{\text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)}$$

The values of a and b are:

$$a = \frac{0.0046 - 0.0062}{0.0061 + 0.0046 - 2(0.00062)} \Rightarrow a = 0.42.$$

and $b = 1 - a = 1 - 0.42 \Rightarrow b = 0.58$. Therefore if the investor invests 42% of the available funds into *IBM* and the remaining 58% into *TEXACO* the variance of the portfolio will be minimum and equal to:

$$\text{Var}(0.42X + 0.58Y) = 0.42^2(0.0061) + 0.58^2(0.0046) + 2(0.42)(0.58)(0.00062) = 0.002926.$$

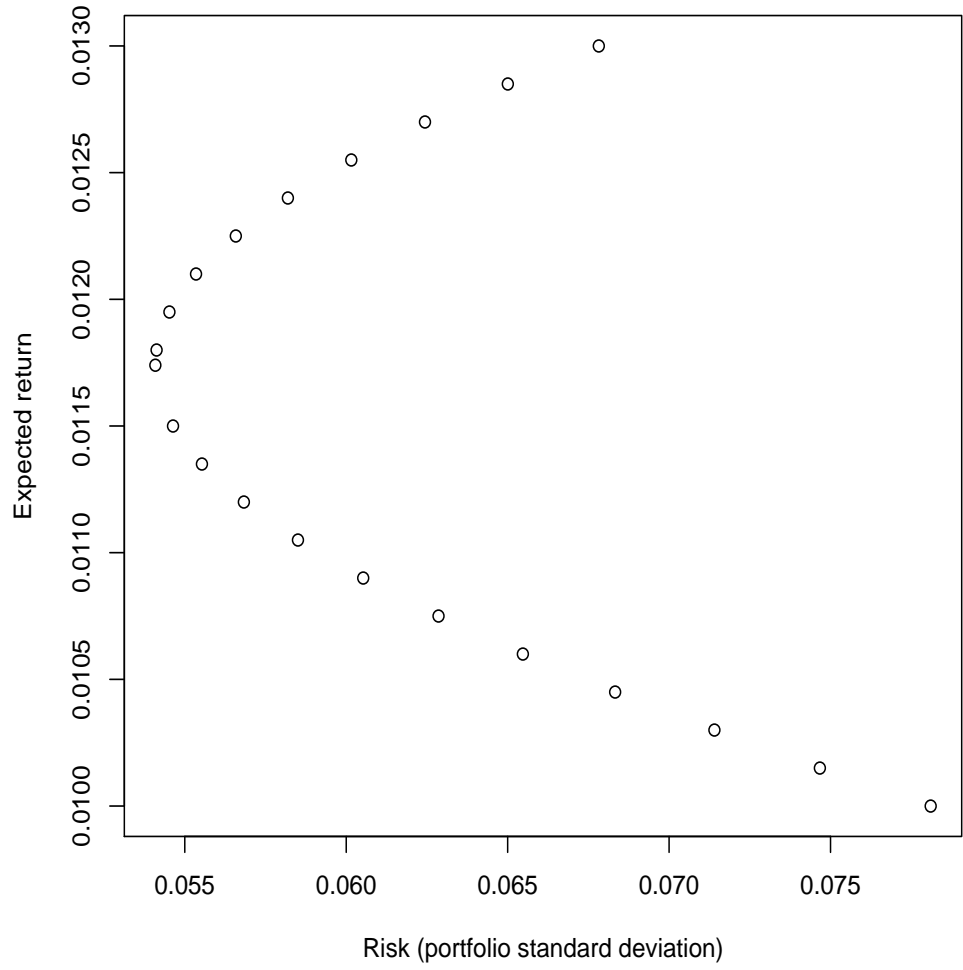
The corresponding expected return of this portfolio will be:

$$E(0.42X + 0.58Y) = 0.42(0.010) + 0.58(0.013) = 0.01174.$$

We can try many other combinations of a and b (but always $a + b = 1$) and compute the risk and return for each resulting portfolio. This is shown in the table below and the graph of return against risk on the other side.

| a | b | Variance | Std | Return |
|------|------|----------|----------|---------|
| 1.00 | 0.00 | 0.006100 | 0.078102 | 0.01000 |
| 0.95 | 0.05 | 0.005576 | 0.074673 | 0.01015 |
| 0.90 | 0.10 | 0.005099 | 0.071407 | 0.01030 |
| 0.85 | 0.15 | 0.004669 | 0.068330 | 0.01045 |
| 0.80 | 0.20 | 0.004286 | 0.065468 | 0.01060 |
| 0.75 | 0.25 | 0.003951 | 0.062857 | 0.01075 |
| 0.70 | 0.30 | 0.003663 | 0.060523 | 0.01090 |
| 0.65 | 0.35 | 0.003423 | 0.058506 | 0.01105 |
| 0.60 | 0.40 | 0.003230 | 0.056833 | 0.01120 |
| 0.55 | 0.45 | 0.003084 | 0.055534 | 0.01135 |
| 0.50 | 0.50 | 0.002985 | 0.054635 | 0.01150 |
| 0.42 | 0.58 | 0.002926 | 0.054093 | 0.01174 |
| 0.40 | 0.60 | 0.002930 | 0.054129 | 0.01180 |
| 0.35 | 0.65 | 0.002973 | 0.054525 | 0.01195 |
| 0.30 | 0.70 | 0.003063 | 0.055344 | 0.01210 |
| 0.25 | 0.75 | 0.003201 | 0.056577 | 0.01225 |
| 0.20 | 0.80 | 0.003386 | 0.058189 | 0.01240 |
| 0.15 | 0.85 | 0.003619 | 0.060158 | 0.01255 |
| 0.10 | 0.90 | 0.003899 | 0.062442 | 0.01270 |
| 0.05 | 0.95 | 0.004226 | 0.065008 | 0.01285 |
| 0.00 | 1.00 | 0.004600 | 0.067823 | 0.01300 |

Portfolio possibilities curve



Efficient frontier with three stocks

```
> summary(returns)
      ribm          rxom          rboeing
Min.   :-0.2264526  Min.   :-0.5219233  Min.   :-0.34570
1st Qu.:-0.0515524  1st Qu.:-0.0172273  1st Qu.:-0.04308
Median :-0.0089916  Median : 0.0007013  Median : 0.01843
Mean   : 0.0003073  Mean    :-0.0011666  Mean    : 0.01079
3rd Qu.: 0.0462550  3rd Qu.: 0.0337488  3rd Qu.: 0.07357
Max.   : 0.3537987  Max.    : 0.2269380  Max.    : 0.17483
```

```
> cov(returns)
      ribm          rxom          rboeing
ribm   9.930174e-03  0.001798962  3.020685e-05
rxom   1.798962e-03  0.006743820  1.781462e-03
rboeing 3.020685e-05  0.001781462  8.282167e-03
```

Portfolio possibilities curve with 3 stocks

