Mathematics in Finance

Steven E. Shreve
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213 USA
shreve@andrew.cmu.edu

A Talk in the Series
Probability in Science and Industry

Carnegie Mellon University
November 20, 2007

Outline

1. Mean-Variance Analysis

Asking the right question

Question before 1952: How do I choose a good stock?

Asking the right question

Question before 1952: How do I choose a good stock?

Question after 1952: How do I choose a good portfolio?

Return of a stock

If we invest \$100 and leave it for a year, reinvesting any dividends, at the end the year the value of our investment will be some random amount Y which could be either more than \$100 or less than \$100.

Return of a stock

If we invest \$100 and leave it for a year, reinvesting any dividends, at the end the year the value of our investment will be some random amount Y which could be either more than \$100 or less than \$100.

Define the return on the investment to be

$$X = \frac{Y - 100}{100}.$$

This is a a random variable that could be either positive nor negative.

Return of a stock

If we invest \$100 and leave it for a year, reinvesting any dividends, at the end the year the value of our investment will be some random amount Y which could be either more than \$100 or less than \$100.

Define the return on the investment to be

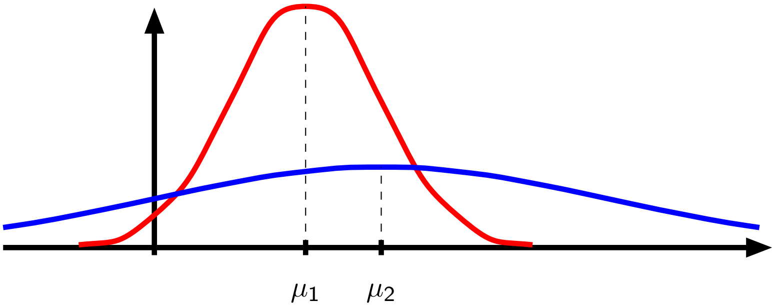
$$X = \frac{Y - 100}{100}.$$

This is a a random variable that could be either positive nor negative.

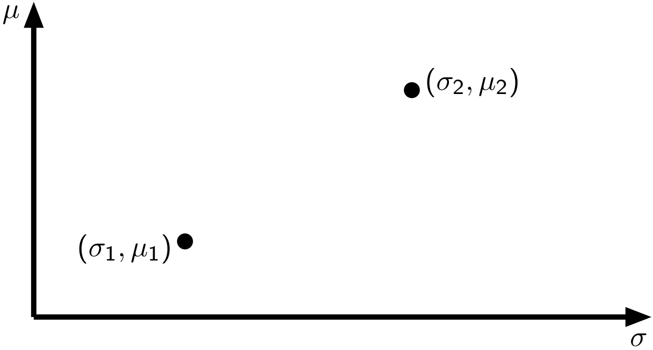
In his 1952 Ph.D. dissertation, Harry Markowitz assumed that X is a normal random variable with some mean μ and standard deviation σ.

Return on two stocks

- Two stocks with returns X_1 and X_2.
- $\mu_1 = \mathbb{E}X_1$ and $\mu_2 = \mathbb{E}X_2$.
- $\sigma_1 = \sqrt{\mathbb{E}\left[(X_1 - \mu_1)^2\right]}$ and $\sigma_2 = \sqrt{\mathbb{E}\left[(X_2 - \mu_2)^2\right]}$
- $\mu_2 > \mu_1$ and $\sigma_2 > \sigma_1$



Return on two stocks (continued)



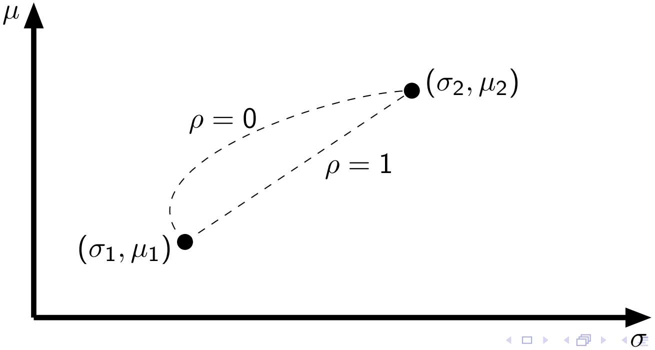
A portfolio of two stocks

Suppose the returns are jointly normal with correlation ρ. Put a fraction α of your capital in the first stock and the remaining fraction $1-\alpha$ in the second stock. Return on the portfolio is

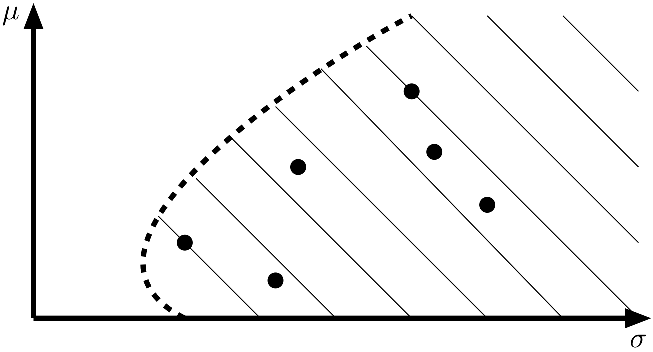
$$\mu(\alpha) = \alpha\mu_1 + (1-\alpha)\mu_2$$

Standard deviation of the portfolio is

$$\sigma(\alpha) = \sqrt{\alpha^2\sigma_1^2 + 2\rho\alpha(1-\alpha)\sigma_1\sigma_2 + (1-\alpha)^2\sigma_2^2}.$$



Efficient frontier



Optimal portfolio

Theorem Let $(X_1, X_2, \ldots, X_n)$ be an n-dimensional vector of random returns, and assume that this vector is jointly normal with a positive definite covariance matrix Γ. This can be written as $\Gamma = \sigma\sigma'$, where σ is a non-singular matrix and σ' denotes its transpose. Let μ be the vector of expected returns. Let m be a desired rate of return for a portfolio. The portfolio that achieves this rate of return with minimal standard deviation is found by solving the optimization problem

$$\begin{aligned} \text{Minimize} \quad & \sqrt{\alpha' \Gamma \alpha} \\ \text{Subject to} \quad & \alpha' e = 1, \qquad (1) \\ & \alpha' \mu = m, \qquad (2) \end{aligned}$$

where e is the n-dimensional vector whose very component is 1.

Optimal portfolio (continued)

The solution to this problem is

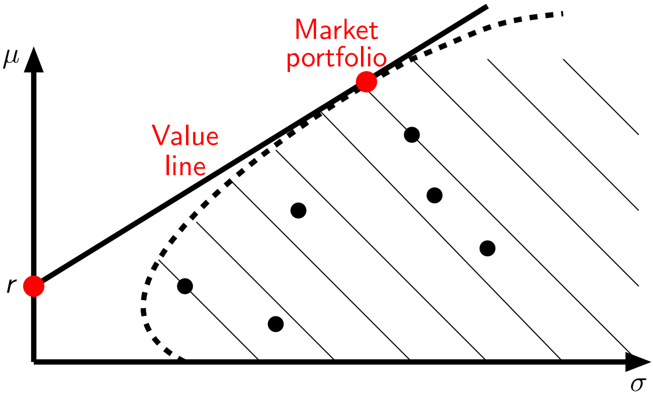
$$\alpha = ma - b,$$

where

$$\begin{aligned} a &= \frac{1}{\Delta}(\sigma')^{-1}\left[(e'\Gamma^{-1}e)\sigma^{-1}\mu - (e'\Gamma^{-1}\mu)\sigma^{-1}e\right], \\ b &= \frac{1}{\Delta}(\sigma')^{-1}\left[(e'\Gamma^{-1}\mu)\sigma^{-1}\mu - (\mu'\Gamma^{-1}\mu)\sigma^{-1}e\right], \\ \Delta &= \|\sigma^{-1}e\|^2\|\sigma^{-1}\mu\|^2 - \left((\sigma^{-1}e)'(\sigma^{-1}\mu)\right)^2. \end{aligned}$$

To guarantee that $\Delta \neq 0$, we need to assume that μ and e are linearly independent.

The bank



Separation Theorem

Theorem. Assume that Γ is positive definite, μ and e are linearly independent, and $r < (e'\Gamma^{-1}\mu)/(e'\Gamma^{-1}e)$. Then the problem of choosing the best portfolio for a particular agent separates into two parts.

- The agent should determine the market portfolio, a portfolio whose composition depends only on the asset parameters, not the attitude of the agent toward risk. This is the market portfolio.
- Secondly, the agent should allocate her wealth between the market portfolio and the bank according to her attitude toward risk. This locates the agent somewhere on the value line.

Separation Theorem

Theorem. Assume that Γ is positive definite, μ and e are linearly independent, and $r < (e'\Gamma^{-1}\mu)/(e'\Gamma^{-1}e)$. Then the problem of choosing the best portfolio for a particular agent separates into two parts.

- The agent should determine the market portfolio, a portfolio whose composition depends only on the asset parameters, not the attitude of the agent toward risk. This is the market portfolio.
- Secondly, the agent should allocate her wealth between the market portfolio and the bank according to her attitude toward risk. This locates the agent somewhere on the value line.

Remark. In practice, the "market portfolio" is an index fund and the "bank" is usually replaced by a money market account.

References

1990 Nobel Prize in Economics

1. H. Markowitz, Portfolio selection, *J. Finance* **8** (1952), 77–91.
2. W. F. Sharpe, Capital asset prices: a theory of market equilibrium under conditions of risk, *J. Finance* **19** *Econometrica* **34** (1966), 768–783.
3. M. Miller and F. Modigliani, Dividend Policy, Growth and the Valuation of Shares, *J. Business* **4** (1961), 425–442.

2. Risk Measurement

Short-comings of variance

Variance measures risk by measuring departures from mean:

$$\mathsf{Var}(X) = \mathbb{E}\left[(X - \mu)^2\right],$$

where

$$\mu = \mathbb{E}X.$$

Standard deviation, which we seek to minimize in mean-variance analysis, is $\sqrt{\mathsf{Var}(X)}$.

Short-comings of variance

Variance measures risk by measuring departures from mean:

$$\mathrm{Var}(X) = \mathbb{E}\left[(X-\mu)^2\right],$$

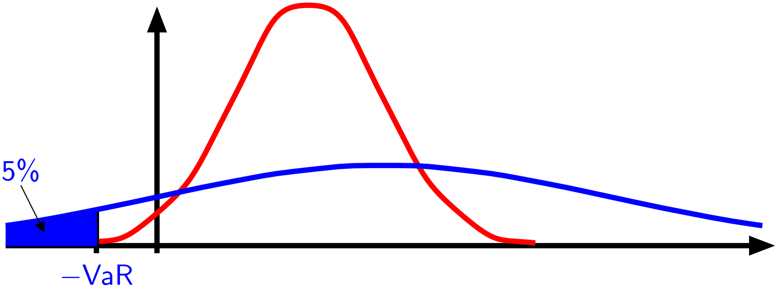
where

$$\mu = \mathbb{E}X.$$

Standard deviation, which we seek to minimize in mean-variance analysis, is $\sqrt{\mathrm{Var}(X)}$.

But a random variable can have a large variance because there is some probability that it takes a value well above its mean, i.e., that the return on an asset is well above average. This is not a risk of loss, but rather a "risk" of outstanding performance.
Mean-variance analysis seeks to avoid the "risk" of outstanding performance as well as the risk of loss.

Value-at-Risk (VaR)



VaR for the blue distribution is shown. VaR for the red distribution is zero.

Proposed by J. P. Morgan Bank:

> *RiskMetrics Technical Manual, New York, J. P. Morgan Bank, 1995.*

Adopted by the Bank of International Settlements and then by regulators in many countries, including the United States.

VaR in fixed-income markets

Example

Consider 10 firms, each of which issues identical zero-coupon bonds. The bonds of each firm have the following properties.

- The bonds sell for \$100 each.
- There is a 99% chance that the firm is solvent at the end of the year.
- After one year, if the firm is solvent, the bonds issued by the firm pay \$108 each.
- After one year, if the firm is not solvent, the bonds issued by the firm defaults, in which case they pay \$0.

We assume that the firms are independent of one another.

VaR in fixed-income markets (continued)

An agent has \$1000 to invest.

First portfolio:

- Buy 10 bonds from a single firm.
- After one year, there is a 99% probability that the portfolio is worth \$1080. In this case, the profit is

$$Y = 1080 - 1000 = 80.$$

- After one year, there is a 1% probability that the portfolio is worth \$0. In this case, the profit is

$$Y = 0 - 1000 = -1000.$$

- VaR $= 0$.

VaR in fixed-income markets (continued)

Second portfolio:

- Buy 1 bond from each firm.
- Probability that 10 firms are solvent at the end of the year is

$$(0.99)^{10} = 90.44\%.$$

  In this case, the profit is

$$Y = 1080 - 1000 = 80.$$

- Probability that at exactly one firm is insolvent is

$$10(0.99)^9(0.01)^1 = 9.14\%$$

  In this case, the profit is

$$Y = 9 \times 108 - 1000 = -28.$$

- VaR $\geq$ 28.

VaR in fixed-income markets (conclusion)

- The diversified portfolio in the example has a higher Value-at-Risk.
- Minimizing VaR encourages concentration of risk. It makes the probability of a loss small by making the consequences of the loss catastrophic.
- VaR does not add up properly. If a bank consists of 10 desks, and each desk buys one bond from a different firm, then each desk will report zero VaR. However, the VaR for the bank is at least 28, not the sum of the VaRs reported by the individual desks.

Coherent measures of risk

P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, Coherent measures of risk, *Math. Finance* **9** (1999), 203–228.

Definition

A coherent risk measure is a mapping ρ from random variables Y to the set of real numbers that has the four properties listed on the next page.

Think of $\rho(Y)$ as the amount of cash reserve you should have in order to hold a portfolio that will give you profit Y at the end of one year.

Coherent measures of risk (continued)

Defining properties of coherent risk measures:

- Monotonicity: If $Y_1 \leq Y_2$ with probability one, then $\rho(Y_1) \geq \rho(Y_2)$.
- Positive homogeneity: For all $\lambda \geq 0$, $\rho(\lambda Y) = \lambda \rho(Y)$.
- Subadditivity: $\rho(Y_1 + Y_2) \leq \rho(Y_1) + \rho(Y_2)$.
- Translation invariance: For every positive constant C,

$$\rho\big(Y + (1+r)C\big) = \rho(Y) - C,$$

where r is the annual interest rate in the economy.

Representation theorem

Theorem

Assume that the random variables under consideration are defined on a finite probability space. Every coherent risk measure ρ is given by the formula

$$\rho(Y) = \max\left\{ \mathbb{E}_{\mathbb{P}}\left[-\frac{Y}{1+r}\right] \middle| \mathbb{P} \in \mathcal{P} \right\},$$

where $\mathcal{P}$ is an arbitrary set of probability measures.

TailVaR

TailVaR is the expected loss, conditional on a "bad" event occurring. It is a coherent risk measure.

In the example in which we invest \$1000 in the bonds of a single firm, the expected loss is $0.01 \times 1000 = 10$, and the expected conditional loss at the 5% level is

$$TailVaR = \frac{10}{0.05} = 200.$$

With the diversified portfolio, we buy one bond from each firm. The expected loss at the 5% level is

$$\begin{aligned} & 28 \times (\text{Part of probability one firm defaults}) \\ & +136 \times (\text{Probability two firms default}) \\ & +244 \times (\text{Probability three firms default}) + \ldots \quad = \quad 1.87, \end{aligned}$$

and the conditional expected loss at the 5% level is

$$TailVaR = \frac{1.87}{0.05} = 37.$$

3. Controlling Risk by Hedging

Binomial model for stock price

$$
\begin{tikzcd}
 & & \boxed{S_2(HH) = 16} \\
 & \boxed{S_1(H) = 8} \arrow[ur] \arrow[dr] & \\
\boxed{S_0 = 4} \arrow[ur] \arrow[dr] & & \boxed{S_2(HT) = S_2(TH) = 4} \\
 & \boxed{S_1(T) = 2} \arrow[ur] \arrow[dr] & \\
 & & \boxed{S_2(TT) = 1}
\end{tikzcd}
$$

Call option on the stock

A call option on the stock with strike price 10 and expiration date 2 is a contract that confers the right to buy the stock at time 2 for \$10, regardless of the market price of the stock at that time.

Value of the call option at time 2:

Coin tosses	Stock price	Option value
HH	16	10
HT	4	0
TH	4	0
TT	1	0

Suppose you sell the call at time 0. Then you might need to pay \$10 at time 2. How do you manage this risk?

Replicating the call

Two trading instruments:

- Stock, given by the binomial model.
- Bank account, with interest rate r per period.

Set up a portfolio:

- X_k – Total value of portfolio at time k.
- Δ_k – Position (number of shares of stock) taken at time k.
- $(X_k - \Delta_k S_k)$ – Amount of cash at time k after position in stock is taken.
- Evolution of portfolio value:

$$X_{k+1} = \Delta_k S_{k+1} + (1+r)\big(X_k - \Delta_k S_k\big)$$

Replicating the call (continued)

$$
\begin{tikzcd}
 & & \boxed{\begin{array}{c} S_2(HH) = 16 \\ X_2(HH) = 6 \end{array}} \\
 & \boxed{\begin{array}{c} S_1(H) = 8 \\ X_1(H) = ? \end{array}} \arrow[ur] \arrow[dr] & \\
\boxed{\begin{array}{c} S_0 = 4 \\ X_0 = ? \end{array}} \arrow[ur] \arrow[dr] & & \boxed{\begin{array}{c} S_2(HT) = S_2(TH) = 4 \\ X_2(HT) = X_2(TH) = 0 \end{array}} \\
 & \boxed{\begin{array}{c} S_1(T) = 2 \\ X_1(T) = ? \end{array}} \arrow[ur] \arrow[dr] & \\
 & & \boxed{\begin{array}{c} S_2(TT) = 1 \\ X_2(TT) = 0 \end{array}}
\end{tikzcd}
$$

Six unknowns:

$$X_0, \; X_1(H), \; X_1(T), \; \Delta_0, \; \Delta_1(H), \; \Delta_1(T).$$

Six equations in six unknowns

$$
\begin{aligned}
X_1(H) &= 8\Delta_0 + \frac{5}{4}(X_0 - 4\Delta_0) \\
X_1(T) &= 2\Delta_0 + \frac{5}{4}(X_0 - 4\Delta_0)
\end{aligned}
$$

$$
\begin{aligned}
0 = X_2(TH) &= 4\Delta_1(T) + \frac{5}{4}(X_1(T) - 2\Delta_1(T)) \\
0 = X_2(TT) &= 1\Delta_1(T) + \frac{5}{4}(X_1(T) - 2\Delta_1(T)) \\
6 = X_2(HH) &= 16\Delta_1(H) + \frac{5}{4}(X_1(H) - 8\Delta_1(H)) \\
0 = X_2(HT) &= 4\Delta_1(H) + \frac{5}{4}(X_1(H) - 8\Delta_1(H))
\end{aligned}
$$

Solving by averaging

$$
\begin{aligned}
\frac{1}{2}\cdot 6+\frac{1}{2}\cdot 0 &= \left(\frac{1}{2}\cdot 16+\frac{1}{2}\cdot 4\right)\Delta_1(H)+\frac{5}{4}(X_1(H)-8\Delta_1(H)) \\
&= \frac{5}{4}X_1(H) \\
\Rightarrow \qquad X_1(H) &= \frac{4}{5}\left(\frac{1}{2}\cdot 6+\frac{1}{2}\cdot 0\right) = 2.40.
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 0 &= \left(\frac{1}{2}\cdot 4+\frac{1}{2}\cdot 1\right)\Delta_1(T)+\frac{5}{4}(X_1(T)-2\Delta_1(T)) \\
&= \frac{5}{4}X_1(T) \\
\Rightarrow \qquad X_1(T) &= \frac{4}{5}\left(\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 0\right) = 0.
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{2}\cdot 2.40+\frac{1}{2}\cdot 0 &= \frac{1}{2}\cdot X_1(H)+\frac{1}{2}\cdot X_1(T) \\
&= \left(\frac{1}{2}\cdot 8+\frac{1}{2}\cdot 2\right)\Delta_0+\frac{5}{4}(X_0-4\Delta_0) \\
&= \frac{5}{4}X_0 \\
\Rightarrow \qquad X_0 &= \frac{4}{5}\left(\frac{1}{2}\cdot 2.40+\frac{1}{2}\cdot 0\right) = 0.96.
\end{aligned}
$$

Solving for the Δ_k

$$
\begin{aligned}
X_1(H) &= 8\Delta_0 + \frac{5}{4}(X_0 - 4\Delta_0) \\
X_1(T) &= 2\Delta_0 + \frac{5}{4}(X_0 - 4\Delta_0)
\end{aligned}
$$

Subtract:

$$
\begin{aligned}
X_1(H) - X_1(T) &= (8-2)\Delta_0, \\
\Delta_0 &= \frac{X_1(H) - X_1(T)}{8-2} = \frac{2.40 - 0}{6} = 0.40.
\end{aligned}
$$

Similar calculations result in

$$\Delta_1(H) = 0.50, \quad \Delta_1(T) = 0.$$

Hedging: A revolutionary idea

- Find a buyer who will purchase the call option at time 0 for \$1.00. (Point out that the stock is tending to go up, so he has a better than 25% change of making \$6.)
- Keep \$0.04 for yourself.
- With the remaining \$0.96, set up the replicating portfolio.
- At time 2, the value of the portfolio will agree with the value of the call option, regardless of how the coin tossing turns out. Use the portfolio to pay off the option.
- You incur no risk from the coin tossing.

Ways to make the model more realistic

- The stock should evolve continuously in time, not in discrete steps.
- At each time after the initial time, there should be infinitely many possible stock prices.

Ways to make the model more realistic

- The stock should evolve continuously in time, not in discrete steps.
- At each time after the initial time, there should be infinitely many possible stock prices.

To build a model with these features, we must combine calculus with probability. This combination is called stochastic calculus. Using stochastic calculus, one can obtain the

Black-Scholes-(Merton) option pricing formula

from a replication argument.

Ways to make the model more realistic

- The stock should evolve continuously in time, not in discrete steps.
- At each time after the initial time, there should be infinitely many possible stock prices.

To build a model with these features, we must combine calculus with probability. This combination is called stochastic calculus. Using stochastic calculus, one can obtain the

Black-Scholes-(Merton) option pricing formula

from a replication argument.

1997 Nobel Prize in Economics.
F. Black and M. Scholes, The pricing of options and corporate liabilities, *Journal of Political Economy* **81** (1973), 637–659.

R. C. Merton, Theory of rational option pricing, *Bell Journal of Economics and Management Science* **4** (1973), 141–183.

4. What is Going on Today?

Mean-Variance Analysis

1. Development of optimization methods for mean-variance analysis that take into account the fact that the parameters are not known with certainty.
2. Extension of the ideas behind mean-variance analysis to multiple periods.

Risk measurement

1. Measure risk associated with a random process, evolving over time, rather than a random variable. Portfolios are dynamic and random variables are not.
2. Determine how to implement coherent risk measures. TailVaR has been implemented, but the computational requirements are greater than for simple VaR.

Controlling risk by hedging

1. Build more realistic models for prices. Build models for interest-rate dependent assets, foreign exchange, electricity, and other commodities.
2. Figure how to price and partially hedge in situations where perfect replication is not possible.

Firms that use these models

1. Derivative securities businesses at investment banks.
2. Hedge funds.
3. Energy companies.
4. Quantitative asset management firms.
5. Insurance companies.