Financial Engineering & Risk Management

Review of Basic Probability

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Discrete Random Variables

Definition. The cumulative distribution function (CDF), $F(\cdot)$, of a random variable, X, is defined by

$$\mathsf{F}(x) := \mathsf{P}(X \leq x).$$

Definition. A discrete random variable, X, has probability mass function (PMF), $p(\cdot)$, if $p(x) \geq 0$ and for all events A we have

$$\mathsf{P}(X \in A) = \sum_{x \in A} p(x).$$

Definition. The expected value of a discrete random variable, X, is given by

$$\mathsf{E}[X] := \sum_i x_i \, p(x_i).$$

Definition. The variance of any random variable, X, is defined as

$$\begin{aligned}\mathsf{Var}(X) &:= \mathsf{E}\left[(X - \mathsf{E}[X])^2\right] \\ &= \mathsf{E}[X^2] - \mathsf{E}[X]^2.\end{aligned}$$

The Binomial Distribution

We say X has a binomial distribution, or $X \sim \mathsf{Bin}(n, p)$, if

$$\mathsf{P}(X = r) = \binom{n}{r} p^r (1-p)^{n-r}.$$

For example, X might represent the number of heads in n independent coin tosses, where $p = \mathsf{P}(\text{head})$. The mean and variance of the binomial distribution satisfy

$$\begin{aligned}\mathsf{E}[X] &= np \\ \mathsf{Var}(X) &= np(1-p).\end{aligned}$$

A Financial Application

- Suppose a fund manager outperforms the market in a given year with probability p and that she underperforms the market with probability $1 - p$.
- She has a track record of 10 years and has outperformed the market in 8 of the 10 years.
- Moreover, performance in any one year is independent of performance in other years.

Question: How likely is a track record as good as this if the fund manager had no skill so that $p = 1/2$?

Answer: Let X be the number of outperforming years. Since the fund manager has no skill, $X \sim \mathsf{Bin}(n = 10, p = 1/2)$ and

$$\mathsf{P}(X \geq 8) = \sum_{r=8}^{n} \binom{n}{r} p^r (1-p)^{n-r}$$

Question: Suppose there are M fund managers? How well should the best one do over the 10-year period if none of them had any skill?

The Poisson Distribution

We say X has a Poisson(λ) distribution if

$$\mathsf{P}(X = r) = \frac{\lambda^r e^{-\lambda}}{r!}.$$

$$\mathsf{E}[X] = \lambda \text{ and } \mathsf{Var}(X) = \lambda.$$

For example, the mean is calculated as

$$\begin{aligned}
\mathsf{E}[X] &= \sum_{r=0}^{\infty} r\,\mathsf{P}(X = r) = \sum_{r=0}^{\infty} r \frac{\lambda^r e^{-\lambda}}{r!} = \sum_{r=1}^{\infty} r \frac{\lambda^r e^{-\lambda}}{r!} \\
&= \lambda \sum_{r=1}^{\infty} \frac{\lambda^{r-1} e^{-\lambda}}{(r-1)!} \\
&= \lambda \sum_{r=0}^{\infty} \frac{\lambda^r e^{-\lambda}}{r!} = \lambda.
\end{aligned}$$

Bayes' Theorem

Let A and B be two events for which $\mathsf{P}(B) \neq 0$. Then

$$\begin{aligned}
\mathsf{P}(A \mid B) &= \frac{\mathsf{P}(A \bigcap B)}{\mathsf{P}(B)} \\
&= \frac{\mathsf{P}(B \mid A)\mathsf{P}(A)}{P(B)} \\
&= \frac{\mathsf{P}(B \mid A)\mathsf{P}(A)}{\sum_j \mathsf{P}(B \mid A_j)\mathsf{P}(A_j)}
\end{aligned}$$

where the A_j's form a partition of the sample-space.

An Example: Tossing Two Fair 6-Sided Dice

	6	7	8	9	10	11	12
	5	6	7	8	9	10	11
	4	5	6	7	8	9	10
Y_2	3	4	5	6	7	8	9
	2	3	4	5	6	7	8
	1	2	3	4	5	6	7
		1	2	3	4	5	6
					Y_1		

Table : $X = Y_1 + Y_2$

- Let Y_1 and Y_2 be the outcomes of tossing two fair dice independently of one another.
- Let $X := Y_1 + Y_2$. Question: What is $\mathsf{P}(Y_1 \geq 4 \mid X \geq 8)$?

Continuous Random Variables

Definition. A continuous random variable, X, has probability density function (PDF), $f(\cdot)$, if $f(x) \geq 0$ and for all events A

$$\mathsf{P}(X \in A) = \int_A f(y)\, dy.$$

The CDF and PDF are related by

$$\mathsf{F}(x) = \int_{-\infty}^{x} f(y)\, dy.$$

It is often convenient to observe that

$$\mathsf{P}\left(X \in \left(x - \frac{\epsilon}{2},\, x + \frac{\epsilon}{2}\right)\right) \approx \epsilon f(x)$$

The Normal Distribution

We say X has a Normal distribution, or $X \sim \mathsf{N}(\mu, \sigma^2)$, if

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

The mean and variance of the normal distribution satisfy

$$\begin{aligned} \mathsf{E}[X] &= \mu \\ \mathsf{Var}(X) &= \sigma^2. \end{aligned}$$

The Log-Normal Distribution

We say X has a log-normal distribution, or $X \sim \mathsf{LN}(\mu, \sigma^2)$, if

$$\log(X) \sim \mathsf{N}(\mu, \sigma^2).$$

The mean and variance of the log-normal distribution satisfy

$$\begin{aligned} \mathsf{E}[X] &= \exp(\mu + \sigma^2/2) \\ \mathsf{Var}(X) &= \exp(2\mu + \sigma^2)\ (\exp(\sigma^2) - 1). \end{aligned}$$

The log-normal distribution plays a very important in financial applications.

Financial Engineering & Risk Management

Review of Conditional Expectations and Variances

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Conditional Expectations and Variances

Let X and Y be two random variables.

The conditional expectation identity says

$$\mathsf{E}[X] = \mathsf{E}\left[\mathsf{E}[X|Y]\right]$$

and the conditional variance identity says

$$\mathsf{Var}(X) = \mathsf{Var}(\mathsf{E}[X|Y]) + \mathsf{E}[\mathsf{Var}(X|Y)].$$

Note that $\mathsf{E}[X|Y]$ and $\mathsf{Var}(X|Y)$ are both functions of Y and are therefore random variables themselves.

A Random Sum of Random Variables

Let $W = X_1 + X_2 + \ldots + X_N$ where the X_i's are IID with mean μ_x and variance σ_x^2, and where N is also a random variable, independent of the X_i's.

Question: What is $\mathsf{E}[W]$?

Answer: The conditional expectation identity implies

$$
\begin{aligned}
\mathsf{E}[W] &= \mathsf{E}\left[\mathsf{E}\left[\sum_{i=1}^{N} X_i \mid N\right]\right] \\
&= \mathsf{E}\left[N\mu_x\right] = \mu_x \, \mathsf{E}\left[N\right].
\end{aligned}
$$

Question: What is $\mathsf{Var}(W)$?

Answer: The conditional variance identity implies

$$
\begin{aligned}
\mathsf{Var}(W) &= \mathsf{Var}(\mathsf{E}[W|N]) + \mathsf{E}[\mathsf{Var}(W|N)] \\
&= \mathsf{Var}(\mu_x N) + \mathsf{E}[N\sigma_x^2] \\
&= \mu_x^2 \mathsf{Var}(N) + \sigma_x^2 \, \mathsf{E}[N].
\end{aligned}
$$

An Example: Chickens and Eggs

A hen lays N eggs where $N \sim \mathsf{Poisson}(\lambda)$. Each egg hatches and yields a chicken with probability p, independently of the other eggs and N. Let K be the number of chickens.

Question: What is $\mathsf{E}[K|N]$?

Answer: We can use indicator functions to answer this question.

In particular, can write $K = \sum_{i=1}^{N} 1_{H_i}$ where H_i is the event that the i^{th} egg hatches. Therefore

$$
1_{H_i} = \begin{cases} 1, & \text{if } i^{th} \text{ egg hatches;} \\ 0, & \text{otherwise.} \end{cases}
$$

Also clear that $\mathsf{E}[1_{H_i}] = 1 \times p + 0 \times (1-p) = p$ so that

$$
\mathsf{E}[K|N] = \mathsf{E}\left[\sum_{i=1}^{N} 1_{H_i} \mid N\right] = \sum_{i=1}^{N} \mathsf{E}\left[1_{H_i}\right] = Np.
$$

Conditional expectation formula then gives $\mathsf{E}[K] = \mathsf{E}[\mathsf{E}[K|N]] = \mathsf{E}[Np] = \lambda p$.

Financial Engineering & Risk Management

Review of Multivariate Distributions

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Multivariate Distributions I

Let $\mathbf{X} = (X_1 \ldots X_n)^\top$ be an n-dimensional vector of random variables.

Definition. For all $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the joint cumulative distribution function (CDF) of $\mathbf{X}$ satisfies

$$
F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}}(x_1, \ldots, x_n) = \mathsf{P}(X_1 \leq x_1, \ldots, X_n \leq x_n).
$$

Definition. For a fixed i, the marginal CDF of X_i satisfies

$$
F_{X_i}(x_i) = F_{\mathbf{X}}(\infty, \ldots, \infty, x_i, \infty, \ldots \infty).
$$

It is straightforward to generalize the previous definition to joint marginal distributions. For example, the joint marginal distribution of X_i and X_j satisfies

$$
F_{ij}(x_i, x_j) = F_{\mathbf{X}}(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty, x_j, \infty, \ldots \infty).
$$

We also say that $\mathbf{X}$ has joint PDF $f_{\mathbf{X}}(\cdot, \ldots, \cdot)$ if

$$
F_{\mathbf{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(u_1, \ldots, u_n) \, du_1 \ldots du_n.
$$

Multivariate Distributions II

Definition. If $\mathbf{X_1} = (X_1, \ldots X_k)^\top$ and $\mathbf{X_2} = (X_{k+1} \ldots X_n)^\top$ is a partition of $\mathbf{X}$ then the conditional CDF of $\mathbf{X_2}$ given $\mathbf{X_1}$ satisfies

$$F_{\mathbf{X_2}|\mathbf{X_1}}(\mathbf{x_2} \,|\, \mathbf{x_1}) \;=\; P(\mathbf{X_2} \leq \mathbf{x_2} \,|\, \mathbf{X_1} = \mathbf{x_1}).$$

If $\mathbf{X}$ has a PDF, $f_{\mathbf{X}}(\cdot)$, then the conditional PDF of $\mathbf{X_2}$ given $\mathbf{X_1}$ satisfies

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2 \,|\, \mathbf{x}_1) \;=\; \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \;=\; \frac{f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1 \,|\, \mathbf{x}_2) f_{\mathbf{X}_2}(\mathbf{x}_2)}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \tag{1}$$

and the conditional CDF is then given by

$$F_{\mathbf{X_2}|\mathbf{X_1}}(\mathbf{x_2}|\mathbf{x_1}) \;=\; \int_{-\infty}^{x_{k+1}} \cdots \int_{-\infty}^{x_n} \frac{f_{\mathbf{X}}(x_1, \ldots, x_k, u_{k+1}, \ldots, u_n)}{f_{\mathbf{X_1}}(\mathbf{x_1})} \, du_{k+1} \ldots du_n$$

where $f_{\mathbf{X_1}}(\cdot)$ is the joint marginal PDF of $\mathbf{X_1}$ which is given by

$$f_{\mathbf{X_1}}(x_1, \ldots, x_k) \;=\; \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, \ldots, x_k, u_{k+1}, \ldots, u_n) \, du_{k+1} \ldots du_n.$$

Independence

Definition. We say the collection $\mathbf{X}$ is independent if the joint CDF can be factored into the product of the marginal CDFs so that

$$F_{\mathbf{X}}(x_1, \ldots, x_n) \;=\; F_{X_1}(x_1) \ldots F_{X_n}(x_n).$$

If $\mathbf{X}$ has a PDF, $f_{\mathbf{X}}(\cdot)$ then independence implies that the PDF also factorizes into the product of marginal PDFs so that

$$f_{\mathbf{X}}(\mathbf{x}) \;=\; f_{X_1}(x_1) \ldots f_{X_n}(x_n).$$

Can also see from (1) that if $\mathbf{X_1}$ and $\mathbf{X_2}$ are independent then

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2 \,|\, \mathbf{x}_1) \;=\; \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \;=\; \frac{f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \;=\; f_{\mathbf{X}_2}(\mathbf{x}_2)$$

– so having information about $\mathbf{X}_1$ tells you nothing about $\mathbf{X}_2$.

Implications of Independence

Let X and Y be independent random variables. Then for any events, A and B,

$$\mathsf{P}(X \in A,\ Y \in B) \;=\; \mathsf{P}(X \in A)\,\mathsf{P}(Y \in B) \tag{2}$$

More generally, for any function, $f(\cdot)$ and $g(\cdot)$, independence of X and Y implies

$$\mathsf{E}[f(X)g(Y)] \;=\; \mathsf{E}[f(X)]\mathsf{E}[g(Y)]. \tag{3}$$

In fact, (2) follows from (3) since

$$\begin{aligned} \mathsf{P}(X \in A,\ Y \in B) &= \mathsf{E}\left[1_{\{X \in A\}} 1_{\{Y \in B\}}\right] \\ &= \mathsf{E}\left[1_{\{X \in A\}}\right] \mathsf{E}\left[1_{\{Y \in B\}}\right] \qquad \text{by (3)} \\ &= \mathsf{P}(X \in A)\,\mathsf{P}(Y \in B). \end{aligned}$$

Implications of Independence

More generally, if $X_1, \ldots X_n$ are independent random variables then

$$\mathsf{E}\left[f_1(X_1) f_2(X_2) \cdots f_n(X_n)\right] \;=\; \mathsf{E}[f_1(X_1)]\mathsf{E}[f_2(X_2)] \cdots \mathsf{E}[f_n(X_n)].$$

Random variables can also be conditionally independent. For example, we say X and Y are conditionally independent given Z if

$$\mathsf{E}[f(X)g(Y) \,|\, Z] \;=\; \mathsf{E}[f(X) \,|\, Z]\ \mathsf{E}[g(Y) \,|\, Z].$$

– used in the (in)famous Gaussian copula model for pricing CDOs!

In particular, let D_i be the event that the i^{th} bond in a portfolio defaults.
Not reasonable to assume that the D_i's are independent. Why?
But maybe they are conditionally independent given Z so that

$$\mathsf{P}(D_1, \ldots, D_n \,|\, Z) \;=\; \mathsf{P}(D_1 \,|\, Z) \cdots \mathsf{P}(D_n \,|\, Z)$$

– often easy to compute this.

The Mean Vector and Covariance Matrix

The mean vector of $\mathbf{X}$ is given by

$$\mathsf{E}[\mathbf{X}] \;:=\; \left(\mathsf{E}[X_1] \;\ldots\; \mathsf{E}[X_n]\right)^\top$$

and the covariance matrix of $\mathbf{X}$ satisfies

$$\mathbf{\Sigma} := \mathsf{Cov}(\mathbf{X}) \;:=\; \mathsf{E}\left[(\mathbf{X}-\mathsf{E}[\mathbf{X}])\,(\mathbf{X}-\mathsf{E}[\mathbf{X}])^\top\right]$$

so that the $(i,j)^{th}$ element of $\mathbf{\Sigma}$ is simply the covariance of X_i and X_j.

The covariance matrix is symmetric and its diagonal elements satisfy $\mathbf{\Sigma}_{i,i} \geq 0$.

It is also positive semi-definite so that $\mathbf{x}^\top \mathbf{\Sigma} \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

The correlation matrix, $\rho(\mathbf{X})$, has $(i,j)^{th}$ element $\rho_{ij} := \mathsf{Corr}(X_i, X_j)$

- it is also symmetric, positive semi-definite and has 1's along the diagonal.

Variances and Covariances

For any matrix $\mathbf{A} \in \mathbb{R}^{k \times n}$ and vector $\mathbf{a} \in \mathbb{R}^k$ we have

$$\mathsf{E}\left[\mathbf{A}\mathbf{X}+\mathbf{a}\right] \;=\; \mathbf{A}\mathsf{E}\left[\mathbf{X}\right] \;+\; \mathbf{a} \tag{4}$$

$$\mathsf{Cov}(\mathbf{A}\mathbf{X}+\mathbf{a}) \;=\; \mathbf{A}\;\mathsf{Cov}(\mathbf{X})\;\mathbf{A}^\top. \tag{5}$$

Note that (5) implies

$$\mathsf{Var}(aX+bY) = a^2\mathsf{Var}(X) + b^2\mathsf{Var}(Y) + 2ab\,\mathsf{Cov}(X,Y).$$

If X and Y independent, then $\mathsf{Cov}(X,Y) = 0$

– but converse not true in general.

Financial Engineering & Risk Management

The Multivariate Normal Distribution

M. Haugh   G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

The Multivariate Normal Distribution I

If the n-dimensional vector $\mathbf{X}$ is multivariate normal with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$ then we write

$$\mathbf{X} \sim \mathsf{MN}_n(\boldsymbol{\mu}, \mathbf{\Sigma}).$$

The PDF of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) \;=\; \frac{1}{(2\pi)^{n/2}|\mathbf{\Sigma}|^{1/2}}\; e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top\,\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

where $|\cdot|$ denotes the determinant.

Standard multivariate normal has $\boldsymbol{\mu} = \mathbf{0}$ and $\mathbf{\Sigma} = \mathbf{I_n}$, the $n \times n$ identity matrix

- in this case the X_i's are independent.

The moment generating function (MGF) of $\mathbf{X}$ satisfies

$$\phi_{\mathbf{X}}(\mathbf{s}) \;=\; \mathsf{E}\left[e^{\mathbf{s}^\top\mathbf{X}}\right] \;=\; e^{\mathbf{s}^\top\boldsymbol{\mu}+\frac{1}{2}\mathbf{s}^\top\mathbf{\Sigma}s}.$$

The Multivariate Normal Distribution II

Recall our partition of $\mathbf{X}$ into $\mathbf{X_1} = (X_1 \; \ldots \; X_k)^\top$ and $\mathbf{X_2} = (X_{k+1} \; \ldots \; X_n)^\top$.
Can extend this notation naturally so that

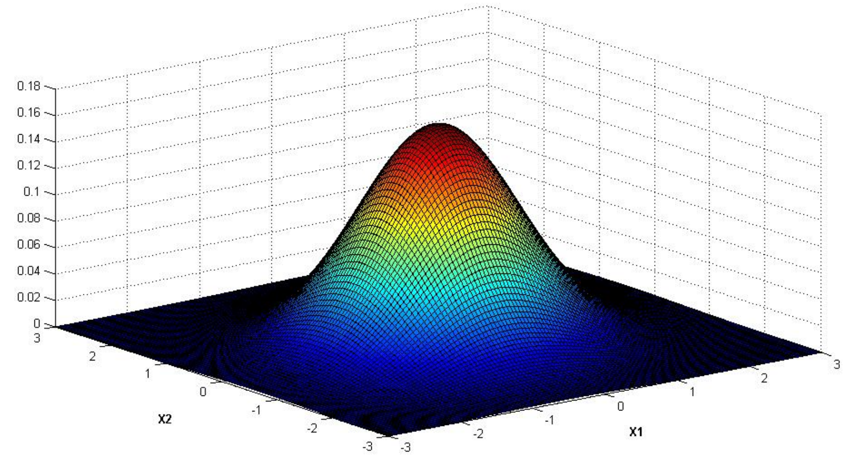
$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}.$$

are the mean vector and covariance matrix of $(\mathbf{X_1}, \mathbf{X_2})$.

Then have following results on marginal and conditional distributions of $\mathbf{X}$:

Marginal Distribution
The marginal distribution of a multivariate normal random vector is itself normal.
In particular, $\mathbf{X_i} \sim \mathsf{MN}(\boldsymbol{\mu}_i, \mathbf{\Sigma}_{ii})$, for $i = 1, 2$.

The Bivariate Normal PDF



The Bivariate Normal PDF

The Multivariate Normal Distribution III

Conditional Distribution
Assuming $\mathbf{\Sigma}$ is positive definite, the conditional distribution of a multivariate normal distribution is also a multivariate normal distribution. In particular,

$$\mathbf{X_2} \mid \mathbf{X_1} = \mathbf{x_1} \;\sim\; \mathsf{MN}(\boldsymbol{\mu}_{2.1}, \mathbf{\Sigma}_{2.1})$$

where $\boldsymbol{\mu}_{2.1} = \mu_2 + \mathbf{\Sigma}_{21}\, \mathbf{\Sigma}_{11}^{-1}\, (\mathbf{x_1} - \boldsymbol{\mu}_1)$ and $\mathbf{\Sigma}_{2.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$.

Linear Combinations
A linear combination, $\mathbf{AX} + \mathbf{a}$, of a multivariate normal random vector, $\mathbf{X}$, is normally distributed with mean vector, $\mathbf{A}\mathsf{E}[\mathbf{X}] + \mathbf{a}$, and covariance matrix, $\mathbf{A}\; \mathsf{Cov}(\mathbf{X})\; \mathbf{A}^\top$.

Financial Engineering & Risk Management

Introduction to Martingales

M. Haugh G. Iyengar
Department of Industrial Engineering and Operations Research
Columbia University

Martingales

Definition. A random process, $\{X_n \; : \; 0 \leq n \leq \infty\}$, is a martingale with respect to the information filtration, $\mathcal{F}_n$, and probability distribution, P, if

1. $\mathsf{E}^P[|X_n|] < \infty$ for all $n \geq 0$
2. $\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] = X_n$ for all $n, m \geq 0$.

Martingales are used to model fair games and have a rich history in the modeling of gambling problems.

We define a submartingale by replacing condition #2 with

$$\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] \geq X_n \quad \text{for all } n, m \geq 0.$$

And we define a supermartingale by replacing condition #2 with

$$\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] \leq X_n \quad \text{for all } n, m \geq 0.$$

A martingale is both a submartingale and a supermartingale.

Constructing a Martingale from a Random Walk

Let $S_n := \sum_{i=1}^n X_i$ be a random walk where the X_i's are IID with mean μ.

Let $M_n := S_n - n\mu$. Then M_n is a martingale because:

$$\begin{aligned}
\mathsf{E}_n[M_{n+m}] &= \mathsf{E}_n\left[\sum_{i=1}^{n+m} X_i - (n+m)\mu\right] \\
&= \mathsf{E}_n\left[\sum_{i=1}^{n+m} X_i\right] - (n+m)\mu \\
&= \sum_{i=1}^{n} X_i \; + \; \mathsf{E}_n\left[\sum_{i=n+1}^{n+m} X_i\right] - (n+m)\mu \\
&= \sum_{i=1}^{n} X_i \; + \; m\mu - (n+m)\mu \; = \; M_n.
\end{aligned}$$

A Martingale Betting Strategy

Let $X_1, X_2, \ldots$ be IID random variables with

$$\mathsf{P}(X_i = 1) \; = \; \mathsf{P}(X_i = -1) \; = \; \frac{1}{2}.$$

Can imagine X_i representing the result of coin-flipping game:

- Win $\$1$ if coin comes up heads
- Lose $\$1$ if coin comes up tails

Consider now a doubling strategy where we keep doubling the bet until we eventually win. Once we win, we stop and our initial bet is $\$1$.

First note that size of bet on n^{th} play is 2^{n-1}

– assuming we're still playing at time n.

Let W_n denote total winnings after n coin tosses assuming $W_0 = 0$.

Then W_n is a martingale!

A Martingale Betting Strategy

To see this, first note that $W_n \in \{1, -2^n + 1\}$ for all n. Why?

1. Suppose we win for first time on n^{th} bet. Then

$$\begin{aligned}
W_n &= -\left(1 + 2 + \cdots + 2^{n-2}\right) \; + \; 2^{n-1} \\
&= -\left(2^{n-1} - 1\right) \; + \; 2^{n-1} \\
&= 1
\end{aligned}$$

2. If we have not yet won after n bets then

$$\begin{aligned}
W_n &= -\left(1 + 2 + \cdots + 2^{n-1}\right) \\
&= -2^n + 1.
\end{aligned}$$

To show W_n is a martingale only need to show $\mathsf{E}[W_{n+1} \,|\, W_n] = W_n$

– then follows by iterated expectations that $\mathsf{E}[W_{n+m} \,|\, W_n] = W_n$.

A Martingale Betting Strategy

There are two cases to consider:

1: $W_n = 1$: then $\mathsf{P}(W_{n+1} = 1 | W_n = 1) = 1$ so

$$\mathsf{E}[W_{n+1} \mid W_n = 1] = 1 = W_n \tag{6}$$

2: $W_n = -2^n + 1$: bet 2^n on $(n+1)^{th}$ toss so $W_{n+1} \in \{1, -2^{n+1} + 1\}$.
Clear that

$$\begin{aligned} \mathsf{P}(W_{n+1} = 1 \mid W_n = -2^n + 1) &= 1/2 \\ \mathsf{P}(W_{n+1} = -2^{n+1} + 1 \mid W_n = -2^n + 1) &= 1/2 \end{aligned}$$

so that

$$\begin{aligned} \mathsf{E}[W_{n+1} \mid W_n = -2^n + 1] &= (1/2)1 \;+\; (1/2)(-2^{n+1} + 1) \\ &= -2^n \;+\; 1 \;=\; W_n. \end{aligned} \tag{7}$$

From (6) and (7) we see that $\mathsf{E}[W_{n+1} \mid W_n] = W_n$.

Polya's Urn

Consider an urn which contains red balls and green balls.
Initially there is just one green ball and one red ball in the urn.

At each time step a ball is chosen randomly from the urn:

1. If ball is red, then it's returned to the urn with an additional red ball.
2. If ball is green, then it's returned to the urn with an additional green ball.

Let X_n denote the number of red balls in the urn after n draws. Then

$$\begin{aligned} \mathsf{P}(X_{n+1} = k + 1 \mid X_n = k) &= \frac{k}{n+2} \\ \mathsf{P}(X_{n+1} = k \mid X_n = k) &= \frac{n+2-k}{n+2}. \end{aligned}$$

Show that $M_n := X_n/(n+2)$ is a martingale.

(These martingale examples taken from *"Introduction to Stochastic Processes"* (Chapman & Hall) by Gregory F. Lawler.)

Financial Engineering & Risk Management

Introduction to Brownian Motion

M. Haugh **G. Iyengar**

Department of Industrial Engineering and Operations Research
Columbia University

Brownian Motion

Definition. We say that a random process, $\{X_t : t \geq 0\}$, is a Brownian motion with parameters (μ, σ) if

1. For $0 < t_1 < t_2 < \ldots < t_{n-1} < t_n$
$$(X_{t_2} - X_{t_1}),\ (X_{t_3} - X_{t_2}),\ \ldots,\ \ (X_{t_n} - X_{t_{n-1}})$$
are mutually independent.
2. For $s > 0$, $X_{t+s} - X_t \;\sim\; \mathsf{N}(\mu s, \sigma^2 s)$ and
3. X_t is a continuous function of t.

We say that X_t is a $B(\mu, \sigma)$ Brownian motion with drift μ and volatility σ

Property #1 is often called the independent increments property.

Remark. Bachelier (1900) and Einstein (1905) were the first to explore Brownian motion from a mathematical viewpoint whereas Wiener (1920's) was the first to show that it actually exists as a well-defined mathematical entity.

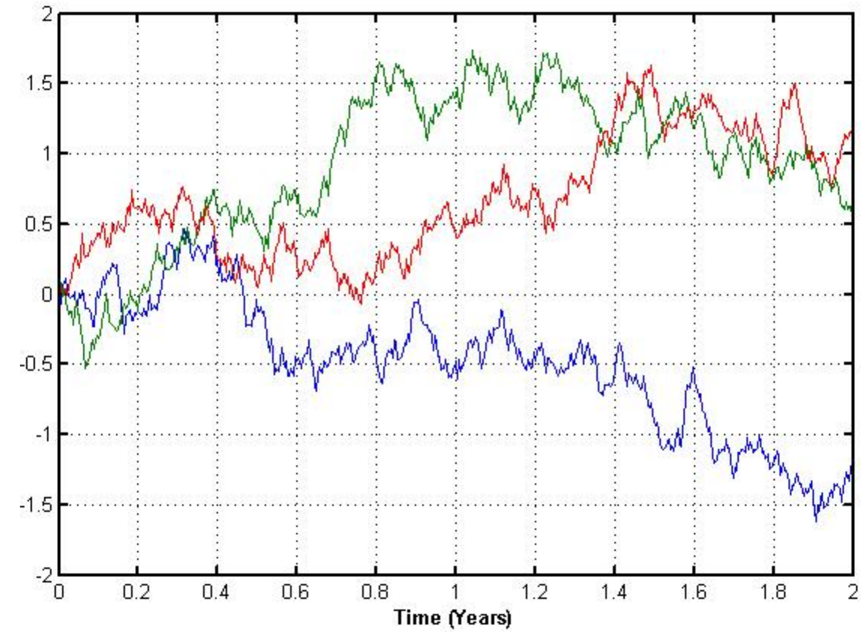
Standard Brownian Motion

- When $\mu = 0$ and $\sigma = 1$ we have a standard Brownian motion (SBM).
- We will use W_t to denote a SBM and we always assume that $W_0 = 0$.
- Note that if $X_t \sim B(\mu, \sigma)$ and $X_0 = x$ then we can write

$$X_t = x + \mu t + \sigma W_t \tag{8}$$

where W_t is an SBM. Therefore see that $X_t \sim \mathsf{N}(x + \mu t, \sigma^2 t)$.

Sample Paths of Brownian Motion



Information Filtrations

- For any random process we will use $\mathcal{F}_t$ to denote the information available at time t
  - the set $\{\mathcal{F}_t\}_{t \geq 0}$ is then the information filtration
  - so $\mathsf{E}[\cdot \,|\, \mathcal{F}_t]$ denotes an expectation conditional on time t information available.
- Will usually write $\mathsf{E}[\cdot \,|\, \mathcal{F}_t]$ as $\mathsf{E}_t[\,\cdot\,]$.

Important Fact: The independent increments property of Brownian motion implies that any function of $W_{t+s} - W_t$ is independent of $\mathcal{F}_t$ and that

$$(W_{t+s} - W_t) \sim \mathsf{N}(0,\, s).$$

A Brownian Motion Calculation

Question: What is $\mathsf{E}_0[W_{t+s} W_s]$?

Answer: We can use a version of the conditional expectation identity to obtain

$$\begin{aligned} \mathsf{E}_0[W_{t+s} W_s] &= \mathsf{E}_0[(W_{t+s} - W_s + W_s)\, W_s] \\ &= \mathsf{E}_0[(W_{t+s} - W_s)\, W_s] \;+\; \mathsf{E}_0\left[W_s^2\right]. \end{aligned} \tag{9}$$

Now we know (why?) $\mathsf{E}_0[W_s^2] = s$.
To calculate first term on r.h.s. of (9) a version of the conditional expectation identity implies

$$\begin{aligned} \mathsf{E}_0[(W_{t+s} - W_s)\, W_s] &= \mathsf{E}_0[\mathsf{E}_s[(W_{t+s} - W_s)\, W_s]] \\ &= \mathsf{E}_0[W_s\, \mathsf{E}_s[(W_{t+s} - W_s)]] \\ &= \mathsf{E}_0[W_s\, 0] \\ &= 0. \end{aligned}$$

Therefore obtain $\mathsf{E}_0[W_{t+s} W_s] = s$.

Financial Engineering & Risk Management

Geometric Brownian Motion

M. Haugh   G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Geometric Brownian Motion

Definition. We say that a random process, X_t, is a geometric Brownian motion (GBM) if for all $t \geq 0$

$$X_t = e^{\left(\mu - \frac{\sigma^2}{2}\right) t \,+\, \sigma W_t}$$

where W_t is a standard Brownian motion.

We call μ the drift, σ the volatility and write $X_t \sim \mathrm{GBM}(\mu, \sigma)$.

Note that

$$\begin{aligned} X_{t+s} &= X_0 \, e^{\left(\mu - \frac{\sigma^2}{2}\right)(t+s) \,+\, \sigma W_{t+s}} \\ &= X_0 \, e^{\left(\mu - \frac{\sigma^2}{2}\right) t \,+\, \sigma W_t \,+\, \left(\mu - \frac{\sigma^2}{2}\right) s \,+\, \sigma (W_{t+s} - W_t)} \\ &= X_t \, e^{\left(\mu - \frac{\sigma^2}{2}\right) s \,+\, \sigma (W_{t+s} - W_t)} \qquad (10) \end{aligned}$$

– a representation that is very useful for simulating security prices.

Geometric Brownian Motion

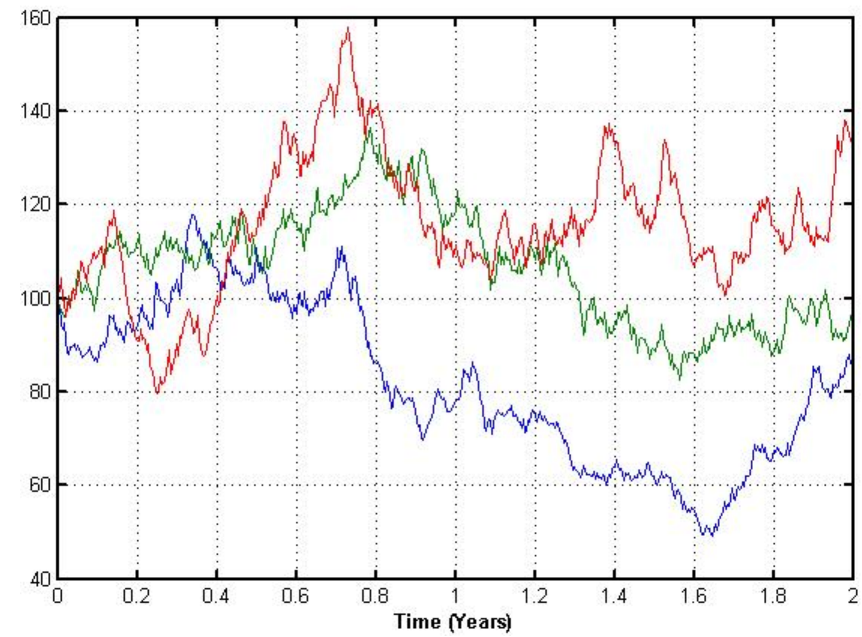
Question: Suppose $X_t \sim \mathrm{GBM}(\mu, \sigma)$. What is $\mathsf{E}_t[X_{t+s}]$?

Answer: From (10) we have

$$\begin{aligned} \mathsf{E}_t[X_{t+s}] &= \mathsf{E}_t\left[X_t \, e^{\left(\mu - \frac{\sigma^2}{2}\right) s \,+\, \sigma (W_{t+s} - W_t)} \right] \\ &= X_t \, e^{\left(\mu - \frac{\sigma^2}{2}\right) s} \, \mathsf{E}_t\left[e^{\sigma (W_{t+s} - W_t)} \right] \\ &= X_t \, e^{\left(\mu - \frac{\sigma^2}{2}\right) s} \, e^{\frac{\sigma^2}{2} s} \\ &= e^{\mu s} X_t \end{aligned}$$

– so the expected growth rate of X_t is μ.

Sample Paths of Geometric Brownian Motion



Geometric Brownian Motion

The following properties of GBM follow immediately from the definition of BM:

1. Fix $t_1, t_2, \ldots, t_n$. Then $\frac{X_{t_2}}{X_{t_1}}, \ \frac{X_{t_3}}{X_{t_2}}, \ \ldots, \frac{X_{t_n}}{X_{t_{n-1}}}$ are mutually independent.
   (For a period of time t, consider 0<t1 <t2 <t3 <t4 …..tn < t)
2. Paths of X_t are continuous as a function of t, i.e., they do not jump.
3. For $s > 0$, $\log\left(\frac{X_{t+s}}{X_t}\right) \ \sim \ \mathsf{N}\left((\mu - \frac{\sigma^2}{2})s, \ \sigma^2 s\right)$.

Modeling Stock Prices as GBM

Suppose $X_t \sim \mathsf{GBM}(\mu, \sigma)$. Then clear that:

1. If $X_t > 0$, then X_{t+s} is always positive for any $s > 0$.
   - so limited liability of stock price is not violated.
2. The distribution of X_{t+s}/X_t only depends on s and not on X_t

These properties suggest that GBM might be a reasonable model for stock prices.

Indeed it is the underlying model for the famous Black-Scholes option formula.

Financial Engineering and Risk Management

Review of vectors

Martin Haugh   Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Reals numbers and vectors

- We will denote the set of real numbers by $\mathbb{R}$
- Vectors are finite collections of real numbers
- Vectors come in two varieties
  - Row vectors: $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$
  - Column vectors $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$
  - By default, vectors are column vectors
- The set of all vectors with **n** components is denoted by $\mathbb{R}^{\mathbf{n}}$

Linear independence

- A vector $\mathbf{w}$ is linearly dependent on $\mathbf{v}_1, \mathbf{v}_2$ if

$$\mathbf{w} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \text{ for some } \alpha_1, \alpha_2 \in \mathbb{R}$$

Example:

$$\begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

- Other names: linear combination, linear span
- A set $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ are linearly independent if **no** $\mathbf{v}_i$ is linearly dependent on the others, $\{\mathbf{v}_j : j \neq i\}$

Basis

- Every $\mathbf{w} \in \mathbb{R}^n$ is a linear combination of the linearly independent set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} \qquad \mathbf{w} = w_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{e}_1} + w_2 \underbrace{\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{e}_2} + \dots + w_n \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{e}_n}$$

- Basis $\equiv$ any linearly independent set that spans the entire space
- Any basis for $\mathbb{R}^n$ has exactly n elements

Norms

- A function $\rho(\mathbf{v})$ of a vector $\mathbf{v}$ is called a norm if
  - $\rho(\mathbf{v}) \geq 0$ and $\rho(\mathbf{v}) = 0$ implies $\mathbf{v} = \mathbf{0}$
  - $\rho(\alpha \mathbf{v}) = |\alpha| \, \rho(\mathbf{v})$ for all $\alpha \in \mathbb{R}$
  - $\rho(\mathbf{v}_1 + \mathbf{v}_2) \leq \rho(\mathbf{v}_1) + \rho(\mathbf{v}_2)$ (triangle inequality)

  ρ generalizes the notion of "length"

- **Examples**:
  - ℓ_2 norm: $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ ... usual length
  - ℓ_1 norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$
  - ℓ_∞ norm: $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x|_i$
  - ℓ_p norm, $1 \leq p < \infty$: $\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x|_i^p \right)^{\frac{1}{p}}$

Inner product

- The inner-product or dot-product of two vector $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is defined as

$$\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} v_i w_i$$

- The ℓ_2 norm $\|\mathbf{v}\|_2 = \sqrt{\mathbf{v} \cdot \mathbf{v}}$
- The angle θ between two vectors $\mathbf{v}$ and $\mathbf{w}$ is given by

$$\cos(\theta) = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|_2 \, \|\mathbf{w}\|_2}$$

- Will show later: $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}$ = product of $\mathbf{v}$ transpose and $\mathbf{w}$

Financial Engineering and Risk Management

Review of matrices

Martin Haugh **Garud Iyengar**

Columbia University
Industrial Engineering and Operations Research

Matrices

- Matrices are rectangular arrays of real numbers
- Examples:
  - $\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}$: 2×3 matrix
  - $\mathbf{B} = \begin{bmatrix} 2 & 3 & 7 \end{bmatrix}$: 1×3 matrix $\equiv$ row vector
  - $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$: $\mathbf{m} \times \mathbf{n}$ matrix ... $\mathbb{R}^{\mathbf{m} \times \mathbf{n}}$
  - $\mathbf{I} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$ ... $n \times n$ **Identity** matrix
- Vectors are clearly also matrices

Matrix Operations: Transpose

- Transpose: $\mathbf{A} \in \mathbb{R}^{m \times d}$

$$\mathbf{A}^\top = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{md} \end{bmatrix}^\top = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{d2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1d} & a_{2d} & \dots & a_{md} \end{bmatrix} \in \mathbb{R}^{d \times m}$$

  - Transpose of a row vector is a column vector

- **Example**:
  - $\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}$: 2×3 matrix ... $\mathbf{A}^\top = \begin{bmatrix} 2 & 1 \\ 3 & 6 \\ 7 & 5 \end{bmatrix}$: 3×2 matrix
  - $\mathbf{v} = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix}$: column vector ... $\mathbf{v}^\top = \begin{bmatrix} 2 & 6 & 4 \end{bmatrix}$: row vector

Matrix Operations: Multiplication

- Multiplication: $\mathbf{A} \in \mathbb{R}^{\mathbf{m} \times \mathbf{d}}$, $\mathbf{B} \in \mathbb{R}^{\mathbf{d} \times \mathbf{p}}$ then $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{\mathbf{m} \times \mathbf{p}}$

$$c_{ij} = \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{id} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{dj} \end{bmatrix}$$

  - row vector $\mathbf{v} \in \mathbb{R}^{1 \times d}$ times column vector $\mathbf{w} \in \mathbb{R}^{d \times 1}$ is a scalar.
  - Identity times any matrix $\mathbf{A}\mathbf{I}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A}$

- **Examples**:
  - $\begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 2(2) + 3(6) + 7(4) \\ 1(2) + 6(6) + 5(4) \end{bmatrix} = \begin{bmatrix} 50 \\ 58 \end{bmatrix}$
  - ℓ_2 norm: $\left\| \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\|_2 = \sqrt{1^2 + (-2)^2} = \sqrt{\begin{bmatrix} 1 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix}} = \sqrt{\begin{bmatrix} 1 \\ -2 \end{bmatrix}^\top \begin{bmatrix} 1 \\ -2 \end{bmatrix}}$
  - inner product: $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}$

Linear functions

- A function $f : \mathbb{R}^d \mapsto \mathbb{R}^m$ is linear if

$$f(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}), \quad \alpha, \beta \in \mathbb{R}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^d$$

- A function f is linear if and only if $f(\mathbf{x}) = \mathbf{Ax}$ for matrix $\mathbf{A} \in \mathbb{R}^{m \times d}$

- **Examples**
  - $f(\mathbf{x}) : \mathbb{R}^3 \mapsto \mathbb{R}$: $f(\mathbf{x}) = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x_1 + 3x_2 + 4x_3$
  - $f(\mathbf{x}) : \mathbb{R}^3 \mapsto \mathbb{R}^2$: $f(\mathbf{x}) = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ x_1 + 2x_3 \end{bmatrix}$

- Linear constraints define sets of vectors that satisfy linear relationships
  - Linear equality: $\{\mathbf{x} : \mathbf{Ax} = \mathbf{b}\}$ ... line, plane, etc.
  - Linear inequality: $\{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b}\}$ ... half-space

Rank of a matrix

- column rank of $\mathbf{A} \in \mathbb{R}^{m \times d}$ = number of linearly independent columns
  - range($\mathbf{A}$) = $\{\mathbf{y} : \mathbf{y} = \mathbf{Ax}$ for some $\mathbf{x}\}$
  - column rank of $\mathbf{A}$ = size of basis for range($\mathbf{A}$)
  - column rank of $\mathbf{A} = m \Rightarrow$ range($\mathbf{A}$) $= \mathbb{R}^m$

- row rank of $\mathbf{A}$ = number of linearly independent rows

- **Fact**: row rank = column rank $\leq \min\{m, d\}$

- **Example**:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}, \quad \text{rank} = 1, \quad \text{range}(\mathbf{A}) = \left\{ \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} : \lambda \in \mathbb{R} \right\}$$

- $\mathbf{A} \in \mathbb{R}^{\mathbf{n} \times \mathbf{n}}$ and $\mathbf{rank}(\mathbf{A}) = n \Rightarrow \mathbf{A}$ invertible, i.e. $\mathbf{A}^{-1} \in \mathbb{R}^{\mathbf{n} \times \mathbf{n}}$

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

Financial Engineering and Risk Management

Review of linear optimization

Martin Haugh   Garud Iyengar
Columbia University
Industrial Engineering and Operations Research

Hedging problem

- d assets
- Prices at time $t = 0$: $\mathbf{p} \in \mathbb{R}^d$

- Market in m possible states at time $t = 1$
- Price of asset j in state $i = S_{ij}$

$$\mathbf{S}_j = \begin{bmatrix} S_{1j} \\ S_{2j} \\ \vdots \\ S_{mj} \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 & \dots & \mathbf{S}_d \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1d} \\ S_{21} & S_{22} & \dots & S_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \dots & S_{md} \end{bmatrix} \in \mathbb{R}^{m \times d}$$

- Hedge an obligation $\mathbf{X} \in \mathbb{R}^m$
  - Have to pay X_i if state i occurs
  - Buy/short sell $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)^\top$ shares to cover obligation

Hedging problem (contd)

- Position $\boldsymbol{\theta} \in \mathbb{R}^d$ purchased at time $t = 0$
  - θ_j = number of shares of asset j purchased, $j = 1, \dots, d$
  - Cost of the position $\boldsymbol{\theta} = \sum_{j=1}^d p_j \theta_j = \mathbf{p}^\top \boldsymbol{\theta}$

- Payoff from liquidating position at time $t = 1$
  - payoff y_i in state i: $y_i = \sum_{j=1}^d S_{ij}\theta_j$
  - Stacking payoffs for all states: $\mathbf{y} = \mathbf{S}\boldsymbol{\theta}$
  - Viewing the payoff vector $\mathbf{y}$: $\mathbf{y} \in \text{range}(\mathbf{S})$

$$\mathbf{y} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 & \dots & \mathbf{S}_d \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} = \sum_{j=1}^d \theta_j \mathbf{S}_j$$

- Payoff $\mathbf{y}$ hedges $\mathbf{X}$ if $\mathbf{y} \geq \mathbf{X}$.

Hedging problem (contd)

- Optimization problem:

$$\begin{array}{rll} \min & \sum_{j=1}^d p_j \theta_j & (\equiv \mathbf{p}^\top \boldsymbol{\theta}) \\ \text{subject to} & \sum_{j=1}^d S_{ij}\theta_j \geq X_i, \quad i = 1, \dots, m & (\equiv \mathbf{S}\boldsymbol{\theta} \geq \mathbf{X}) \end{array}$$

- Features of this optimization problem
  - Linear objective function: $\mathbf{p}^\top \boldsymbol{\theta}$
  - Linear inequality constraints: $\mathbf{S}\boldsymbol{\theta} \geq \mathbf{X}$

- Example of a linear program
  - Linear objective function: either a min/max
  - Linear inequality and equality constraints

$$\begin{array}{ll} \max/\min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}_{eq}\mathbf{x} = \mathbf{b}_{eq} \\ & \mathbf{A}_{in}\mathbf{x} \leq \mathbf{b}_{in} \end{array}$$

Linear programming duality

- Linear program

$$P = \begin{array}{rl} \min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{array}$$

- Dual linear program

$$D = \begin{array}{rl} \max_{\mathbf{u}} & \mathbf{b}^\top \mathbf{u} \\ \text{subject to} & \mathbf{A}^\top \mathbf{u} = \mathbf{c} \\ & \mathbf{u} \geq \mathbf{0} \end{array}$$

Theorem.

- **Weak Duality**: $P \geq D$
- **Bound**: $\mathbf{x}$ feasible for P, $\mathbf{u}$ feasible for D, $\mathbf{c}^\top \mathbf{x} \geq P \geq D \geq \mathbf{b}^\top \mathbf{u}$
- **Strong Duality**: Suppose P or D finite. Then $P = D$.
- Dual of the dual is the primal (original) problem

More duality results

- Here is another primal-dual pair

$$\begin{array}{rl} \min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} \quad = \quad \begin{array}{rl} \max_{\mathbf{u}} & \mathbf{b}^\top \mathbf{u} \\ \text{subject to} & \mathbf{A}^\top \mathbf{u} = \mathbf{c} \end{array}$$

- General idea for constructing duals

$$\begin{aligned} P &= \min\{\mathbf{c}^\top \mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} \\ &\geq \min\{\mathbf{c}^\top \mathbf{x} - \mathbf{u}^\top(\mathbf{A}\mathbf{x} - \mathbf{b}) : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} \text{ for all } \mathbf{u} \geq \mathbf{0} \\ &\geq \mathbf{b}^\top \mathbf{u} + \min\{(\mathbf{c} - \mathbf{A}^\top \mathbf{u})^\top \mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} \\ &= \begin{cases} \mathbf{b}^\top \mathbf{u} & \mathbf{A}^\top \mathbf{u} = \mathbf{c} \\ -\infty & \text{otherwise} \end{cases} \\ &\geq \max\{\mathbf{b}^\top \mathbf{u} : \mathbf{A}^\top \mathbf{u} = \mathbf{c}\} \end{aligned}$$

- Lagrangian relaxation: dualize constraints and relax them!

Financial Engineering and Risk Management

Review of nonlinear optimization

Martin Haugh Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Unconstrained nonlinear optimization

- Optimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

- Categorization of minimum points
  - $\mathbf{x}^*$ global minimum if $f(\mathbf{y}) \geq f(\mathbf{x}^*)$ for all $\mathbf{y}$
  - $\mathbf{x}^*_{loc}$ local minimum if $f(\mathbf{y}) \geq f(\mathbf{x}^*_{loc})$ for all $\mathbf{y}$ such that $\|\mathbf{y} - \mathbf{x}^*_{loc}\| \leq r$
- Sufficient condition for local min
  - gradient $\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \mathbf{0}$: local stationarity
  - Hessian $\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$ positive semidefinite
- Gradient condition is sufficient if the function $f(\mathbf{x})$ is convex.

Unconstrained nonlinear optimization

- Optimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^2} \; x_1^2 + 3x_1x_2 + x_2^3$$

- Gradient

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 + 3x_2 \\ 3x_1 + 3x_2^2 \end{bmatrix} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \mathbf{0}, \quad \begin{bmatrix} -\frac{9}{4} \\ \frac{3}{2} \end{bmatrix}$$

- Hessian at $\mathbf{x}$: $\mathbf{H} = \begin{bmatrix} 2 & 3 \\ 3 & 6x_2 \end{bmatrix}$
  - $\mathbf{x} = \mathbf{0}$: $\mathbf{H} = \begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix}$. Not positive definite. Not local minimum.
  - $\mathbf{x} = \begin{bmatrix} -\frac{9}{4} \\ \frac{3}{2} \end{bmatrix}$: $\mathbf{H} = \begin{bmatrix} 2 & 3 \\ 3 & 9 \end{bmatrix}$. Positive semidefinite. Local minimum

Lagrangian method

- Constrained optimization problem

$$\begin{aligned} \max_{\mathbf{x}\in\mathbb{R}^2} \quad & 2\ln(1+x_1) + 4\ln(1+x_2), \\ \text{s.t.} \quad & x_1 + x_2 = 12 \end{aligned}$$

- Convex problem. But constraints make the problem hard to solve.
- Form a Lagrangian function

$$\mathcal{L}(\mathbf{x}, v) = 2\ln(1+x_1) + 4\ln(1+x_2) - v(x_1 + x_2 - 12)$$

- Compute the stationary points of the Lagrangian as a function of v

$$\nabla \mathcal{L}(\mathbf{x}, v) = \begin{bmatrix} \frac{2}{1+x_1} - v \\ \frac{4}{1+x_2} - v \end{bmatrix} = \mathbf{0} \quad \Rightarrow \quad x_1 = \frac{2}{v} - 1, \quad x_2 = \frac{4}{v} - 1$$

- Substituting in the constraint $x_1 + x_2 = 12$, we get

$$\frac{6}{v} = 14 \quad \Rightarrow \quad v = \frac{3}{7} \quad \Rightarrow \quad \mathbf{x} = \frac{1}{3}\begin{bmatrix} 11 \\ 25 \end{bmatrix}$$

Portfolio Selection

- Optimization problem

$$\begin{aligned} \max_{\mathbf{x}} \quad & \boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \mathbf{V} \mathbf{x} \\ \text{s.t.} \quad & \mathbf{1}^\top \mathbf{x} = 1 \end{aligned}$$

  Constraints make the problem hard!

- Lagrangian function

$$\mathcal{L}(\mathbf{x}, v) = \boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \mathbf{V} \mathbf{x} - v(\mathbf{1}^\top \mathbf{x} - 1)$$

  - Solve for the maximum value with no constraints

$$\nabla_x \mathcal{L}(\mathbf{x}, v) = \boldsymbol{\mu} - 2\lambda \mathbf{V}\mathbf{x} - v\mathbf{1} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \frac{1}{2\lambda} \cdot \mathbf{V}^{-1}(\boldsymbol{\mu} - v\mathbf{1})$$

  - Solve for v from the constraint

$$\mathbf{1}^\top \mathbf{x} = 1 \quad \Rightarrow \quad \mathbf{1}^\top \mathbf{V}^{-1}(\boldsymbol{\mu} - v\mathbf{1}) = 2\lambda \quad \Rightarrow \quad v = \frac{\mathbf{1}^\top \mathbf{V}^{-1}\boldsymbol{\mu} - 2\lambda}{\mathbf{1}^\top \mathbf{V}^{-1}\mathbf{1}}$$

  - Substitute back in the expression for $\mathbf{x}$

Financial Engineering and Risk Management

Introduction to no-arbitrage

Martin Haugh   Garud Iyengar

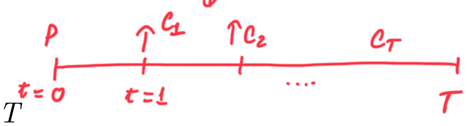
Columbia University
Industrial Engineering and Operations Research

Contracts, prices and no-arbitrage

sometime fix c_k, sometime bound it

Consider the following contract

- Pay price p at time $t = 0$
- Receive c_k at time $t = k$, $k = 1, \ldots, T$



Note that the cash flow c_k could be negative!

The no-arbitrage condition bounds the price p for this contract.

- Weak No-Arbitrage: $c_k \geq 0$ for all $k \geq 1 \Rightarrow p \geq 0$ if there's profit, priced higher bc of demand supply
- Strong No-Arbitrage: $c_k \geq 0$ for all $k \geq 1$ and $c_\ell > 0$ for some $\ell \Rightarrow p > 0$

Essentially eliminate the possibility of a **free-lunch**!

Rationale for the weak no-arbitrage condition: Suppose $p < 0$

- Since $c_k \geq 0$ for all $k \geq 1$, the buyer receives $-p > 0$ at time 0, and then does not lose money thereafter. Free lunch!
- Seller can increase price as long as $p \leq 0$, and still have buyers available.
- Buyers will be willing to pay a higher price in order to compete.

Assumptions underlying no-arbitrage

Rationale for the strong no-arbitrage condition

- Suppose $p \leq 0$.
- Recall that $c_\ell > 0$ for some $\ell \geq 1$. Therefore, a free lunch as long as $p \leq 0$.
- We can only guarantee that $p > 0$ but not the precise value!

Implicit assumptions underlying the no-arbitrage condition

- Markets are liquid: sufficient number of buyers and sellers
- Price information is available to all buyers and sellers
- Competition in supply and demand will correct any deviation from no-arbitrage prices

Pricing a simple bond

What is the price p of a contract that pays A dollars in 1 year?

Suppose one is able to borrow and lend unlimited amounts at an interest rate of r per year.

Construct the following portfolio

- **Buy** the contract at price p
- **Borrow** $A/(1+r)$ at interest rate r

Cash flows associated with this portfolio

Price of portfolio	Cashflow in 1 year
$z = p - \frac{A}{1+r}$	$A - A = 0$

Weak No-arbitrage: $c_1 \geq 0$ implies price $z \geq 0$, i.e. $p \geq \frac{A}{1+r}$.

Pricing a simple bond (contd.)

Next, construct the following portfolio

- **Sell** the contract at price p
- **Lend** $A/(1+r)$ at interest rate r

Cash flows associated with this portfolio

Price of portfolio	Cashflow in 1 year
$z = \frac{A}{1+r} - p$	$-A + A = 0$

Weak No-arbitrage: $c_1 \geq 0$ implies price $z \geq 0$, i.e. $p \leq \frac{A}{1+r}$.

Two results together imply: $p = \frac{A}{1+r}$. Surprise?

The result relied on the ability to borrow **and** lend at rate r.

- What if borrowing and lending rates are different?
- What if the borrowing and lending markets are elastic?

Financial Engineering and Risk Management

Interest rates and fixed income instruments

Martin Haugh   Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Simple and compound interest

Definition. An amount A invested for n periods at a simple interest rate of r per period is worth $A(1 + n \cdot r)$ at maturity.

Definition. An amount A invested for n periods at a compound interest rate of r per period is worth $A(1 + r)^n$ at maturity.

Interest rates are typically quoted on annual basis, even if the compounding period is less than 1 year.

- n compounding periods in each year
- rate of interest r
- A invested for y years yields $A\big(1 + \frac{r}{n}\big)^{y \cdot n}$

Definition. Continuous compounding corresponds to the situation where the length of the compounding period goes to zero. Therefore, an amount A invested for y years is worth $\lim_{n\to\infty} A(1 + r/n)^{yn} = Ae^{ry}$ at maturity.

Present value

Price p of a contract that pays $\mathbf{c} = (c_0, c_1, c_2, \ldots, c_N)$

- $c_k > 0 \equiv$ cash inflow, and $c_k < 0 \equiv$ cash outflow

Present Value (PV) assuming interest rate r per period

$$PV(\mathbf{c}; r) = c_0 + \frac{c_1}{(1+r)} + \frac{c_2}{(1+r)^2} + \ldots \frac{c_N}{(1+r)^N} = \sum_{k=0}^{N} \frac{c_k}{(1+r)^k}.$$

No-arbitrage argument: Suppose one can borrow and lend at rate r

Cash flows	$t=0$	$t=1$	$t=2$	$t=k$	$t=T$
Buy contract	$-p+c_0$	c_1	c_2	c_k	c_T
Borrow $c_1/(1+r)$ up to time 1	$c_1/(1+r)$	$-c_1$			
Borrow $c_2/(1+r)^2$ up to time 2	$c_2/(1+r)^2$		$-c_2$		
Borrow $c_k/(1+r)^k$ up to time k	$c_k/(1+r)^k$			$-c_k$	
Borrow $c_T/(1+r)^T$ up to time T	$c_T/(1+r)^T$				$-c_T$

- Portfolio cash flows $= 0$ for times $k \geq 1$
- Price of portfolio: $p - \sum_{k=0}^{T} c_k/(1+r)^k \geq 0 \Rightarrow p \geq \sum_{k=0}^{T} c_k/(1+r)^k$

Present value (contd.)

To obtain the upper bound: **reverse** the portfolio.

Cash flows	$t=0$	$t=1$	$t=2$	$t=k$	$t=T$
Sell contract	$p-c_0$	$-c_1$	$-c_2$	$-c_k$	$-c_T$
Lend $c_1/(1+r)$ up to time 1	$-c_1/(1+r)$	c_1			
Lend $c_2/(1+r)^2$ up to time 2	$-c_2/(1+r)^2$		c_2		
Lend $c_k/(1+r)^k$ up to time k	$-c_k/(1+r)^k$			c_k	
Lend $c_T/(1+r)^T$ up to time T	$-c_T/(1+r)^T$				c_T

- Portfolio cash flows $= 0$ for times $k \geq 1$
- Price of portfolio: $\sum_{k=0}^{T} c_k/(1+r)^k - p \geq 0 \Rightarrow p \leq \sum_{k=0}^{T} c_k/(1+r)^k$

The two bounds together imply: $p = PV(\mathbf{c}; r)$

Important we could both lend and borrow at rate r

- What if the lending rate is different from borrowing rate?

Different lending and borrowing rates

Can lend at rate r_L and borrow rate at rate r_B: $r_L \leq r_B$

Portfolio: buy contract, and borrow $\frac{c_k}{(1+r_B)^k}$ for k years, $k = 1, \ldots, N$

- Cash flow in year k: $c_k - \frac{c_k}{(1+r_B)^k}(1+r_B)^k = 0$ for $k \geq 1$
- No-arbitrage: price $= p - c_0 - \sum_{k=1}^{N} \frac{c_k}{(1+r_B)^k} \geq 0$
- Lower bound on price $p \geq PV(\mathbf{c}; r_B)$

Portfolio: sell contract, and lend $\frac{c_k}{(1+r_L)^k}$ for k years, $k = 1, \ldots, N$

- Cash flow in year k: $-c_k + \frac{c_k}{(1+r_L)^k}(1+r_L)^k = 0$ for $k \geq 1$
- No-arbitrage: price $= -p + c_0 + \sum_{k=1}^{N} \frac{c_k}{(1+r_L)^k} \geq 0$
- Upper bound on price $p \leq PV(\mathbf{c}; r_L)$

Bounds on the price $PV(\mathbf{c}; r_B) \leq p \leq PV(\mathbf{c}; r_L)$

How is the price set?

Fixed income securities

Fixed income securities "guarantee" a fixed cash flow. Are these risk-free?

- Default risk
- Inflation risk
- Market risk

Perpetuity: $c_k = A$ for all $k \geq 1$

$$p = \sum_{k=1}^{\infty} \frac{A}{(1+r)^k} = \frac{A}{r}$$

Annuity: $c_k = A$ for all $k = 1, \ldots, n$

$$\begin{aligned} \text{Annuity} &= \text{Perpetuity} - \text{Perpetuity starting in year } n+1 \\ \text{Price } p &= \frac{A}{r} - \frac{1}{(1+r)^n} \cdot \frac{A}{r} = \frac{A}{r}\left(1 - \frac{1}{(1+r)^n}\right) \end{aligned}$$

Bonds

Features of bonds

- Face value F: usually 100 or 1000
- Coupon rate α: pays $c = \alpha F/2$ every six months
- Maturity T: Date of the payment of the face value and the last coupon
- Price P
- Quality rating: S&P Ratings AAA, AA, BBB, BB, CCC, CC

Bonds differ in many dimensions ... hard to compare bonds

Yield to maturity λ

$$P = \sum_{k=1}^{2T} \frac{c}{(1+\lambda/2)^k} + \frac{F}{(1+\lambda/2)^{2T}}$$

Annual interest rate at which price P = present value of coupon payments

Yield to maturity

Yield to maturity λ

$$P = \sum_{k=1}^{2T} \frac{c}{(1+\lambda/2)^k} + \frac{F}{(1+\lambda/2)^{2T}}$$

Why do we think in terms of yields?

- Summarizes face value, coupon, maturity, and quality
- Relates to quality: lower quality $\rightarrow$ lower price $\rightarrow$ higher yield to maturity
- Relates to interest rate movements

But ... yield to maturity is a crude measure. Does not capture everything.

Financial Engineering and Risk Management

Floating rate bonds and term structure of interest rates

Martin Haugh   Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Linear pricing

Theorem. (Linear Pricing) Suppose there is no arbitrage. Suppose also

- Price of cash flow $\mathbf{c}_A$ is p_A
- Price of cash flow $\mathbf{c}_B$ is p_B

Then the price of cash flow that pays $\mathbf{c} = \mathbf{c}_A + \mathbf{c}_B$ must be $p_A + p_B$.

Let p denote the price of the total cash flow $\mathbf{c}$. Suppose $p < p_A + p_B$, i.e. $\mathbf{c}$ is cheap! Will create an arbitrage portfolio, i.e. a free-lunch portfolio.

- Purchase $\mathbf{c}$ at price p
- Sell cash flow $\mathbf{c}_A$ and $\mathbf{c}_B$ separately

Price of the portfolio $= p - p_A - p_B < 0$, i.e. net income at time $t = 0$.

The cash flows cancel out at all times. Future cash flows $=$ **zero**. Free lunch!

No arbitrage $\equiv$ no free lunch. Therefore, $p \geq p_A + p_B$

We can reverse the argument if $p > p_A + p_B$

- Note that we need a liquid market for buying/selling all the cash flows.

Simple example of linear pricing

Cash flow $\mathbf{c} = (c_1, \ldots, c_T)$ is a portfolio of T separate cash flows

- $\mathbf{c}^{(t)}$ pays c_t at time t and zero otherwise.

Suppose the cash flows are annual and the annual interest rate is r.

Price of cash flow $\mathbf{c}^{(t)} = \frac{c_t}{(1+r)^t}$.

Price of cash flow $\mathbf{c} = \sum_{t=1}^{T}$ Price of cash flow $\mathbf{c}^{(t)} = \sum_{t=1}^{T} \frac{c_t}{(1+r)^t}$

Floating interest rates

Interest rates are **random** quantities ... they fluctuate with time.

Let r_k denote the per period interest rate over period $[k, k+1)$

- The exact value of r_k becomes known only at time k
- 1-period loans issued in period k to be repaid in period $k+1$ are charged r_k

Cash flow of floating rate bond

- coupon payment at time k: $r_{k-1}F$
- face value at time n: F

Goal: Compute the arbitrage-free price P_f of the floating rate bond

Split up the cash flows of floating rate bond into simpler cash flows

- $p_k =$ Price of contract paying $r_{k-1}F$ at time k
- $P =$ Price of Principal F at time $n = \frac{F}{(1+r)^n}$

Price of floating rate bond $P_f = P + \sum_{k=1}^{n} p_k$

Price of contract that pays $r_{k-1}F$ at time k

Goal: Construct a portfolio that has a **deterministic** cash flow

- The price of a deterministic cash flow at time $t = 0$ is given by the NPV

	$t = 0$	$t = k-1$	$t = k$
Buy contract	$-p_k$		$r_{k-1}F$
Borrow α over $[0, k-1]$	α	$-\alpha(1+r_0)^{k-1}$	
Borrow $\alpha(1+r_0)^{k-1}$ over $[k-1, k]$		$\alpha(1+r_0)^{k-1}$	$-\alpha(1+r_0)^{k-1}(1+r_{k-1})$
Lend α from $[0, k]$	$-\alpha$		$\alpha(1+r_0)^k$

Cash flow at time k

$$\begin{aligned} c_k &= r_{k-1}F - \alpha(1+r_0)^{k-1}(1+r_{k-1}) + \alpha(1+r_0)^k \\ &= \underbrace{\left(F - \alpha(1+r_0)^{k-1}\right) r_{k-1}}_{\text{random}} + \underbrace{\alpha r_0 (1+r_0)^{k-1}}_{\text{deterministic}} \end{aligned}$$

Set $\alpha = \frac{F}{(1+r_0)^{(k-1)}}$. Then the random term is 0.

Net cash flow is now deterministic ... $c_k = \alpha r_0 (1+r_0)^{k-1} = F r_0$

Price of floating rate bond (contd)

Price of the portfolio $= p_k - \alpha + \alpha = p_k = \frac{c_k}{(1+r)^k} = \frac{F r_0}{(1+r)^k}$

Recall that

$$\begin{aligned} P_f &= \frac{F}{(1+r_0)^n} + \sum_{k=1}^{n} p_k \\ &= \frac{F}{(1+r_0)^n} + \sum_{k=1}^{n} \frac{F r_0}{(1+r_0)^k} \\ &= \frac{F}{(1+r_0)^n} + \frac{F r_0}{(1+r_0)} \sum_{k=1}^{n} \frac{1}{(1+r_0)^{k-1}} \\ &= \frac{F}{(1+r_0)^n} + \frac{F r_0}{(1+r_0)} \cdot \frac{1 - \frac{1}{(1+r_0)^n}}{1 - \frac{1}{1+r_0}} \\ &= F \end{aligned}$$

The price P_f of a floating rate bond is equal to its face value F

Term structure of interest rates

Interest rates depend on the term or duration of the loan. Why?

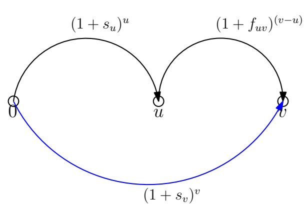
- Investors prefer their funds to be liquid rather than tied up.
- Investors have to be offered a higher rate to lock in funds for a longer period.
- Other explanations: expectation of future rates, market segmentation.

Spot rates: s_t = interest rate for a loan maturing in t years

$$A \text{ in year } t \quad \Rightarrow \quad PV = \frac{A}{(1+s_t)^t}$$

Discount rate $d(0,t) = \frac{1}{(1+s_t)^t}$. Can infer the spot rates from bond prices.

Forward rate f_{uv}: interest rate quoted today for lending from year u to v.



$$(1+s_v)^v = (1+s_u)^u (1+f_{uv})^{(v-u)} \Rightarrow f_{uv} = \left(\frac{(1+s_v)^v}{(1+s_u)^u}\right)^{\frac{1}{v-u}} - 1$$

Relation between spot and forward rates

$$(1+s_t)^t = \prod_{k=0}^{t-1} (1 + f_{k,k+1})$$

Financial Engineering and Risk Management

Forwards contracts

Martin Haugh Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Forward contract

Definition. A forward contract gives the buyer the right, and also the obligation, to purchase

- a specified amount of an asset
- at a specified time T
- at a specified price F (called the forward price) set at time $t = 0$

Example.

- Forward contract for delivery of a stock with maturity 6 months
- Forward contract for sale of gold with maturity 1 year
- Forward contract to buy 10m $ worth of Euros with maturity 3 months
- Forward contract for delivery of 9-month T-Bill with maturity 3 months.

Setting the forward price F

Goal: Set the forward price F for a forward contract at time $t = 0$ for 1 unit of an asset with

- asset price S_t at time t
- and maturity T

f_t = value/price at time t of a long position in the forward contract

Value at time T: $f_T = (S_T - F)$

- long position in forward: must purchase the asset at price F
- spot price of asset: S_T

Forward price F is set so that time $t = 0$ value/price f_0 is 0

Use no-arbitrage principle to set F

Short selling an asset

Short selling is the selling of shares in a stock that the seller doesn't own

- The seller borrows the shares from the broker
- The shares comes from the brokerage's own inventory
- The shares are sold and the proceeds are credited to the seller's account

However ... sooner or later

- the seller must "close" the short by buying back the shares (called covering)

Profit/loss associated with a short sale

- Results in a profit when the price drops
- Results in a loss when the price increases

Short positions can be very risky

- Price can only drop to zero ... potential profit is bounded
- Price can increase to arbitrarily large values ... potential loss is unbounded

No-arbitrage argument to set F

Assume asset has no intermediate cash flows, e.g. dividends, or storage costs.

Portfolio: Buy contract, **short** sell the underlying and lend S_0 up to time T

Cash flow	$t=0$	$t=T$
Buy contract	$f_0 = 0$	$f_T = S_T - F$
Short sell asset and buy back at time T	$+S_0$	$-S_T$
Lend S_0 up to T	$-S_0$	$S_0/d(0,T)$
Net cash flow	0	$S_0/d(0,T) - F$

The portfolio has a deterministic cash flow at time T and the cost = 0. Therefore,

$$0 = \left(\frac{S_0}{d(0,T)} - F\right) d(0,T) \quad \Rightarrow \quad F = \frac{S_0}{d(0,T)}$$

Why is F strictly greater than the spot price S_0?

- Cost of carry

Examples of forward contracts

Example. Forward contract on a non-dividend paying stock that matures in 6 months. The current stock price is \$50 and the 6-month interest rate is 4% per annum.

Solution. Assuming semi-annual compounding, the discount factor

$$d(0,.5) = \frac{1}{1+\frac{0.04}{2}} = 0.9804.$$

Therefore,

$$F = 50/0.9804 = 51.0$$

Forward value f_t for $t > 0$

Recall the value of a long forward position

- at time 0: $f_0 = 0$
- at time T: $f_T = S_T - F$

- F_0: Forward price at time 0 for delivery at time T
- F_t: Forward price at time t for delivery at time T

Pricing via the no-arbitrage arguments

Cash flow	$t=t$	$t=T$
Short F_t contract	0	$F_t - S_T$
Long F_0 contract	$-f_t$	$S_T - F_0$
Net cash flow	$-f_t$	$F_t - F_0$

The portfolio has a deterministic cash flow. Therefore,

$$f_t = (F_t - F_0)\,d(t,T)$$

Financial Engineering and Risk Management

Futures

Martin Haugh   Garud Iyengar

Columbia University

Industrial Engineering and Operations Research

Problems with forward contracts

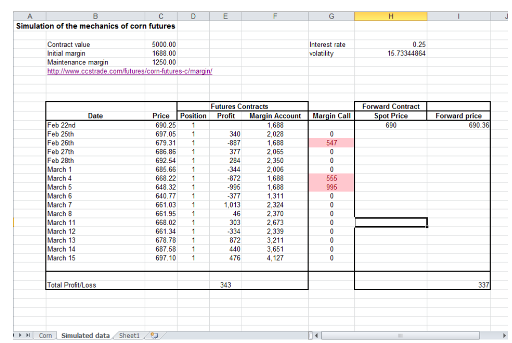
- Not organized through an exchange.
- Consequently, no price transparency!
- Double-coincidence-of-wants: need someone to take the opposite side!
- Default risk of the counterparty.

Futures contract

- Solves the problem of a multitude of prices for the same maturity by marking-to-market
  - disbursing profits/losses at the end of each day
- Now contracts can be organized through an exchange.
- Can be written on any underlying security with a settlement price
  - Commodities
  - Broad based indicies, e.g. S & P 500, Russel 2000, etc.
  - Volatility of the market, e.g. VIX futures
- http://www.cmegroup.com/market-data/delayed-quotes/commodities.html

Mechanics of a futures contract

- Individuals open a margin account with a broker
- Enter into N futures contracts with price F_0
- Deposit initial margin into the account $\approx 5 - 10\%$ of contract value
- All profit/loss settled using margin account
- Margin call if balance is low



Simulation of the mechanics of corn futures

Contract value	5000.00	Interest rate	0.25
Initial margin	1688.00	volatility	15.73344864
Maintenance margin	1250.00		

http://www.ccstrade.com/futures/corn-futures-c/margin/

		Futures Contracts				Forward Contract	
Date	Price	Position	Profit	Margin Account	Margin Call	Spot Price	Forward price
Feb 22nd	690.25	1		1,688		690	690.36
Feb 25th	697.05	1	340	2,028	0		
Feb 26th	679.31	1	-887	1,688	547		
Feb 27th	686.86	1	377	2,065	0		
Feb 28th	692.54	1	284	2,350	0		
March 1	685.66	1	-344	2,006	0		
March 4	668.22	1	-872	1,688	555		
March 5	648.32	1	-995	1,688	995		
March 6	640.77	1	-377	1,311	0		
March 7	661.03	1	1,013	2,324	0		
March 8	661.95	1	46	2,370	0		
March 11	668.02	1	303	2,673	0		
March 12	661.34	1	-334	2,339	0		
March 13	678.78	1	872	3,211	0		
March 14	687.58	1	440	3,651	0		
March 15	697.10	1	476	4,127	0		
Total Profit/Loss			343				337

Corn | Simulated data | Sheet1

Pros/cons of futures

Pros:

- High leverage: high profit
- Very liquid
- Can be written on a wide variety of underlying assets

Cons:

- High leverage: high risk
- Futures prices are approximately linear function of the underlying – only linear payoffs can be hedged
- May not be flexible enough; back to Forwards!

Pricing futures

- Need martingale pricing formalism
- Deterministic interest rates: forward price = futures price
- At maturity futures price F_T = price of underlying S_T

Hedging using Futures: Long hedge

Today is Sept. 1st. A baker needs $500,000$ bushels of wheat on December 1st. So, the baker faces the risk of an uncertain price on Dec. 1st.

Hedging strategy: buy 100 futures contracts maturing on Dec. 1st – each for 5000 bushels

Cash flow on Dec. 1st

- Futures position at maturity: $F_T - F_0 = S_T - F_0$
- Buy in the spot market: S_T
- Effective cash flow: $S_T - F_0 - S_T = -F_0$

Price fixed at F_0!

Did this cost anything? Cash flows associated with margin calls.

Perfect hedges are not always possible

Why?

- The date T may not be a futures expiration date.
- P_T may not correspond to an integer number of futures contracts
- A futures contract on the underlying may not be available
- The futures contract might not be liquid
- The payoff P_T may be nonlinear in the underlying

Basis = Spot price of underlying - futures price

- Perfect hedge: basis = 0 at time T
- Basis risk: basis $\neq$ 0 at time T
- Basis risk arises because the futures contract is on a related but different asset, or expires at a different time.

Hedging problem with basis risk

Today is Sept. 1st. A taco company needs $500,000$ bushels of kidney beans on December 1st. So, the taco company faces the risk of an uncertain price.

Problem: No kidney bean futures available. Basis risk inevitable.

Hedge: Go long y soybean futures each for 5000 bushels of soybeans

Cash Flow in 90 days

- Futures position at maturity: $(F_T - F_0)y$
- Buy kidney beans in the spot market: P_T
- Effective cash flow: $C_T = y(F_T - F_0) - P_T$

$P_T \neq yF_T$ for any y: Perfect hedge impossible!

Minimum variance hedging

Variance of the cash flow

$$\begin{aligned}\mathbf{var}(C_T) &= \mathbf{var}(P_T) + \mathbf{var}\big(y(F_T - F_0)\big) \\ &\quad - 2\mathbf{cov}\big(y(F_T - F_0), P_T\big) \\ &= \mathbf{var}(P_T) + y^2\mathbf{var}(F_T) - 2y\mathbf{cov}(F_T, P_T)\end{aligned}$$

Set the derivative with respect to y to zero:

$$\frac{d\mathbf{var}(C_T(y))}{dy} = 2y\mathbf{var}(F_T) - 2\mathbf{cov}(F_T, P_T) = 0$$

Optimal number of Futures contracts:

$$y^* = + \frac{\mathbf{cov}(F_T, P_T)}{\mathbf{var}(F_T)}$$

Financial Engineering and Risk Management

Swaps

Martin Haugh Garud Iyengar
Columbia University
Industrial Engineering and Operations Research

Swaps

Definition. Swaps are contracts that transform one kind of cash flow into another.

Example.

- Plain vanilla swap: fixed interest rate vs floating interest rates
- Commodity swaps: exchange floating price for a fixed price. e.g. gold swaps, oil swaps.
- Currency swaps

Why swaps?

- Change the nature of cash flows
- Leverage strengths in different markets

Example of leveraging strengths

Two companies

Company	Fixed	Floating
A	4.0%	LIBOR + 0.3%
B	5.2%	LIBOR + 1.0%

Company A is "better" in both but relatively weaker in the floating rate market

Company A

- Borrows in fixed market at 4.0%
- Swap with B: pays LIBOR and receives 3.95%

Company B

- B borrow in the floating market at LIBOR + 1.0%
- Swap with A: pays 3.95% and receives LIBOR

Effective rates:

- A: $-4\% + 3.95\% - \text{LIBOR} = -(\text{LIBOR} + 0.05\%)$
- B: $-\text{LIBOR} - 1\% + \text{LIBOR} - 3.95\% = -4.95\%$

Both gain!

Role of financial intermediaries

Same two companies

Company	Fixed	Floating
A	4.0%	LIBOR + 0.3%
B	5.2%	LIBOR + 1.0%

Financial intermediary that constructs the swap.

Company A

- Borrows in fixed market at 4.0%
- Swap with Intermediary: pays LIBOR and receives 3.93%

Company B

- B borrow in the floating market at LIBOR + 1.0%
- Swap with Intermediary: pays 3.97% and receives LIBOR

Financial intermediary makes 0.04% or 4 basis points. Why?

- Compensation for taking on counterparty risk and providing a service

Pricing interest rate swaps

r_t = floating (unknown) interest rate at time t

Cash flows at time $t = 1, \ldots, T$

- Company A (long): receives Nr_{t-1} and pays NX
- Company B (short): receives NX and pays Nr_{t-1}

Value of swap to company A

- $N(r_0, \ldots, r_{T-1})$ = Cash flow of floating rate bond - Face value. Therefore, value of swap to company A

$$V_A = N\big(1 - d(0,T)\big) - NX \sum_{t=1}^{T} d(0,t)$$

- Set X so that $V_A = 0$, i.e.

$$X = \frac{1 - d(0,T)}{\sum_{t=1}^{T} d(0,t)}$$

Financial Engineering and Risk Management

Options

Martin Haugh   Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Options

Definition. A European Call Option gives the buyer the right but not the obligation to purchase 1 unit of the underlying at specified price K (strike price) at a specified time T (expiration).

Definition. An American Call Option gives the buyer the right but not the obligation to purchase 1 unit of the underlying at specified price K (strike price) at any time until a specified time T (expiration).

Definition. A European Put Option gives the buyer the right but not the obligation to sell 1 unit of the underlying at specified price K (strike price) at a speficied time T (expiration).

Definition. An American Put Option gives the buyer the right but not the obligation to sell 1 unit of the underlying at specified price K (strike price) at any time until a specified time T (expiration).

Payoff and intrinsic value of a call option

Payoff of a European call option at expiration T

- price $S_T < K$: do not exercise the option, payoff $= 0$
- price $S_T > K$: exercise option and sell in the spot market, payoff $= S_T - K$

Payoff $= \max\{S_T - K, 0\}$ ... nonlinear in the price S_T

Intrinsic value of a call option at time $t \leq T = \max\{S_t - K, 0\}$

- In the money: $S_t > K$
- At the money: $S_t = K$
- Out of the money: $S_t < K$

Everything works the other way for put options.

Payoff and intrinsic value of a put option

Payoff of a European put option at expiration T

- price $S_T < K$: exercise the option and see in spot market, payoff $= K - S_T$
- price $S_T > K$: do not exercise option, payoff $= 0$

Payoff $= \max\{K - S_T, 0\}$ ... nonlinear in the price S_T

Intrinsic value of a put option at time $t \leq T = \max\{K - S_t, 0\}$

- In the money: $S_t < K$
- At the money: $S_t = K$
- Out of the money: $S_t > K$

Pricing options

Nonlinear payoff ... cannot price without a model for the underlying

Prices of options

- European put/call with strike K and expiration T: $p_E(t;K,T)$, $c_E(t;K,T)$
- American put/call with strike K and expiration T: $p_A(t;K,T)$, $c_A(t;K,T)$

European put-call parity at time t for non-dividend paying stock:

$$p_E(t;K,T) + S_t = c_E(t;K,T) + Kd(t,T)$$

Trading strategy

- At time t buy European call with strike K and expiration T
- At time t sell European put with strike K and expiration T
- At time t (short) sell 1 unit of underlying and buy at time T
- Lend $K \cdot d(t,T)$ dollars up to time T

No-arbitrage argument

- Cash flow at time T: $\max\{S_T - K, 0\} - \max\{K - S_T, 0\} - S_T + K = 0$
- Cash flow at time t: $-c_E(t;K,T) + p_E(t;K,T) + S_t - Kd(t,T) = 0$

Bounds on prices of European options

Price of American option $\geq$ Price of European option

- $c_A(t;K,T) \geq c_E(t;K,T)$, and $p_A(t;K,T) \geq p_E(t;K,T)$

Lower bound on European options as a function of stock price S_t

- $c_E(t;K,T) = \max\{S_t + p_E(t;K,T) - Kd(t,T), 0\} \geq \max\{S_t - Kd(t,T), 0\}$
- $p_E(t;K,T) = \max\{Kd(t,T) + c_E(t;K,T) - S_t, 0\} \geq \max\{Kd(t,T) - S_t, 0\}$

Upper bound on European options as a function of stock price S_t

- $\max\{S_T - K, 0\} \leq S_T$ implies $c_E(t;K,T) \leq S_t$
- $\max\{K - S_T, 0\} \leq K$ implies $p_E(t;K,T) \leq Kd(t,T)$

Effect of dividends $p_E(t;K,T) + S_t - D = c_E(t;K,T) + Kd(t,T)$

- $D =$ present value of all dividends until maturity

Bounds on prices of American options

Price of American call as function of stock price S_t:

- $c_A(t,K,T) \geq c_E(t;K,T) \geq \max\left\{S_t - Kd(t,T), 0\right\} > \max\left\{S_t - K, 0\right\}$

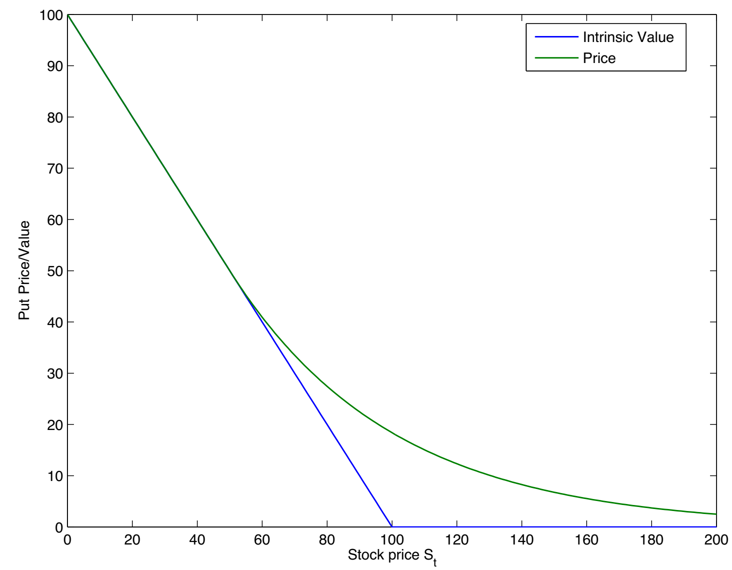
Thus, the price of an American call is always strictly greater than the exercise value of the call option.

Never optimal to exercise an American call on a non-dividend paying stock early!
$c_A(t;K,T) = c_E(t,K,T)$

Price of American put as function of stock price S_t:

- Bound $p_A(t,K,T) \geq p_E(t;K,T) \geq \max\left\{Kd(t,T) - S_t, 0\right\}$
- But the exercise value of a American put option is $\max\left\{K - S_t, 0\right\}$

Bounds do not tell us much!



Financial Engineering & Risk Management

Option Pricing and the Binomial Model

M. Haugh **G. Iyengar**

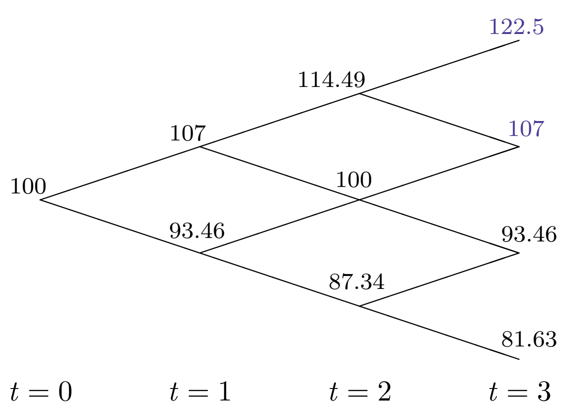
Department of Industrial Engineering and Operations Research
Columbia University

A Brief Overview of Option Pricing

In the next series of modules we'll study:

1. The 1-period binomial model
2. The multi-period binomial model
3. Replicating strategies
4. Pricing European and American options in the binomial lattice
5. The Black-Scholes formula

Stock Price Dynamics in the Binomial Model



- A risk-free asset or cash account also available
  - \$1 invested in cash account at $t = 0$ worth R^t dollars at time t

Some Questions

1. How much is an option that pays $\max(0,\ S_3 - 100)$ at $t = 3$ worth?
   (i) do we have enough information to answer this question?
   (ii) should the price depend on the utility functions of the buyer and seller?
   (iii) will the price depend on the true probability, p, of an up-move in each period? Perhaps the price should be

$$\mathsf{E}_0^{\mathbb{P}}[R^{-3} \max(0,\ S_3 - 100)]? \tag{1}$$

2. Suppose now that:
   (i) you stand to lose a lot at date $t = 3$ if the stock is worth 81.63
   (ii) you also stand to earn a lot at date $t = 3$ if the stock is worth 122.49.

   If you don't want this risk exposure could you do anything to eliminate it?

The St. Petersberg Paradox

- Consider the following game
  - a fair coin is tossed repeatedly until first head appears
  - if first head appears on the n^{th} toss, then you receive $\$2^n$
- How much would you be willing to pay in order to play this game?
- The expected payoff is given by

$$
\begin{aligned}
\mathsf{E}_0^{\mathbb{P}}[\text{Payoff}] &= \sum_{n=1}^{\infty} 2^n \, \mathsf{P}(1^{st} \text{ head on } n^{th} \text{ toss}) \\
&= \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} \\
&= \infty
\end{aligned}
$$

- But would you pay an infinite amount of money to play this game?
  - clear then that (1) does not give correct option price.

The St. Petersberg Paradox

- Daniel Bernouilli resolved this paradox by introducing a utility function, $u(\cdot)$
  - $u(x)$ measures how much utility or benefit you obtains from x units of wealth
  - different people have different utility functions
  - $u(.)$ should be increasing and concave
- Bernouilli introduced the $\log(\cdot)$ utility function so that

$$
\mathsf{E}_0^{\mathbb{P}}[u(\text{Payoff})] = \sum_{n=1}^{\infty} \log(2^n) \frac{1}{2^n} = \log(2) \sum_{n=1}^{\infty} \frac{n}{2^n} < \infty
$$

- So maybe just need to figure out appropriate utility function and use it to compute option price
  – maybe, but who's utility function?
  – in fact we'll see there's a much simpler way.

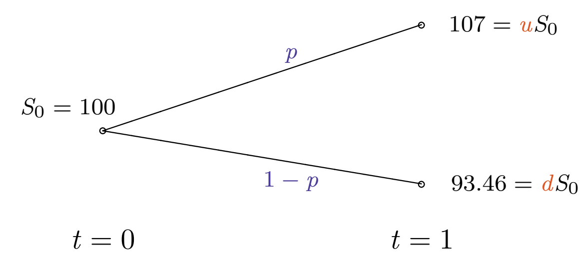
Financial Engineering & Risk Management

The 1-Period Binomial Model

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

The 1-Period Binomial Model



- Can borrow or lend at gross risk-free rate, R
  - so $\$1$ in cash account at $t = 0$ is worth $\$R$ at $t = 1$
- Also assume that short-sales are allowed.

The 1-Period Binomial Model

Questions:

1. How much is a call option that pays $\max(S_1 - 107,\ 0)$ at $t = 1$ worth?

2. How much is a call option that pays $\max(S_1 - 92,\ 0)$ at $t = 1$ worth?

Type A and Type B Arbitrage

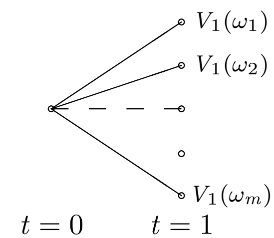
Earlier definitions of weak and strong arbitrage applied in a deterministic world.
Need more general definitions when we introduce randomness.

Definition. A type A arbitrage is a security or portfolio that produces immediate positive reward at $t = 0$ and has non-negative value at $t = 1$.

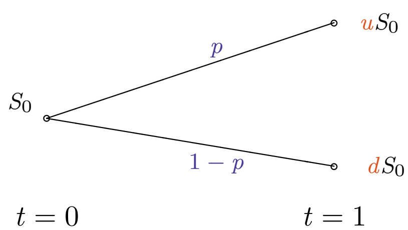
i.e. a security with initial cost, $V_0 < 0$, and time $t = 1$ value $V_1 \geq 0$.

Definition. A type B arbitrage is a security or portfolio that has a non-positive initial cost, has positive probability of yielding a positive payoff at $t = 1$ and zero probability of producing a negative payoff then.

i.e. a security with initial cost, $V_0 \leq 0$, and $V_1 \geq 0$ but $V_1 \neq 0$.



Arbitrage in the 1-Period Binomial Model



- Recall we can borrow or lend at gross risk-free rate, R, per period.
- And short-sales are allowed.

Theorem. There is no arbitrage if and only if $d < R < u$.

Proof: (i) Suppose $R < d < u$. Then borrow S_0 and invest in stock.

(ii) Suppose $d < u < R$. Then short-sell one share of stock and invest proceeds in cash-account.

Both case give a type B arbitrage.

Will soon see other direction, i.e. if $d < R < u$, then there can be no-arbitrage.

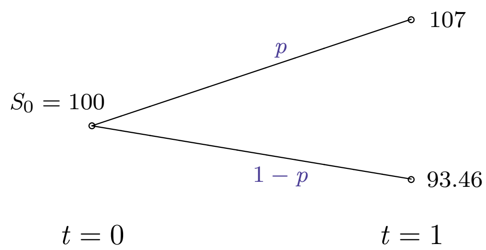
Financial Engineering & Risk Management

Option Pricing in the 1-Period Binomial Model

M. Haugh **G. Iyengar**

Department of Industrial Engineering and Operations Research
Columbia University

Option Pricing in the 1-Period Binomial Model



Assume now that $R = 1.01$.

1. How much is a call option that pays $\max(S_1 - 102,\ 0)$ at $t = 1$ worth?
2. How will the price vary as p varies?

To answer these questions, we will construct a replicating portfolio.

The Replicating Portfolio

- Consider buying x shares and investing \$$y$ in cash at $t = 0$
- At $t = 1$ this portfolio is worth:

$$\begin{array}{rl} 107x + 1.01y & \text{when } S = 107 \\ 93.46x + 1.01y & \text{when } S = 93.46 \end{array}$$

- Can we choose x and y so that portfolio equals option payoff at $t = 1$?
- If so, then we must solve

$$\begin{aligned} 107x + 1.01y &= 5 \\ 93.46x + 1.01y &= 0 \end{aligned}$$

The solution is

$$\begin{aligned} x &= 0.3693 \\ y &= -34.1708 \end{aligned}$$

So yes, we can construct a replicating portfolio!

The Replicating Portfolio

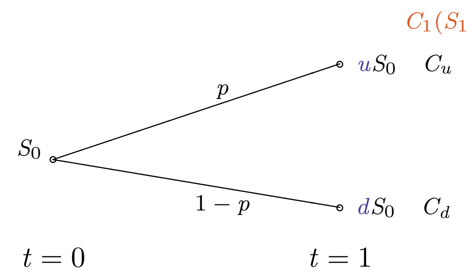
Question: What does a negative value of y mean?
Question: What would a negative value of x mean?

- The cost of this portfolio at $t = 0$ is

$$0.3693 \times 100 - 34.1708 \times 1 \approx 2.76$$

- So the fair value of the option is 2.76
  - indeed 2.76 is the arbitrage-free value of the option.
- So option price does not directly depend on buyer's (or seller's) utility function.

Derivative Security Pricing



- Can use same replicating portfolio argument to find price, C_0, of any derivative security with payoff function, $C_1(S_1)$, at time $t = 1$.
- Set up replicating portfolio as before:

$$\begin{aligned} uS_0x + Ry &= C_u \\ dS_0x + Ry &= C_d \end{aligned}$$

- Solve for x and y as before and then must have $C_0 = xS_0 + y$.

Derivative Security Pricing

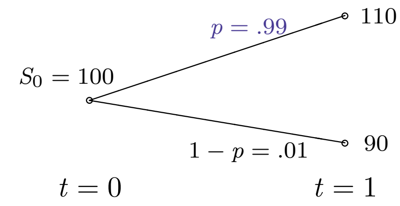
- Obtain

$$
\begin{aligned}
C_0 &= \frac{1}{R}\left[\frac{R-d}{u-d}C_u + \frac{u-R}{u-d}C_d\right] \\
&= \frac{1}{R}\left[q\,C_u + (1-q)\,C_d\right] \\
&= \frac{1}{R}\mathsf{E}_0^{\mathbb{Q}}[C_1]. \qquad (2)
\end{aligned}
$$

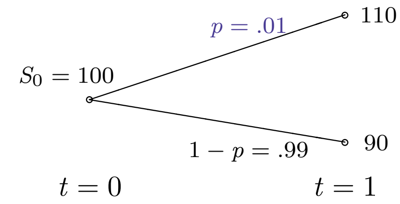
- Note that if there is no-arbitrage then $q > 0$ and $1 - q > 0$
  - we call (2) risk-neutral pricing
  - and $(q, 1-q)$ are the risk-neutral probabilities.
- So we now know how to price any derivative security in this 1-period model.
- Can also answer earlier question: "How does the option price depend on p?"
  - but is the answer crazy?!

What's Going On?

- Stock ABC



- Stock XYZ



Question: What is the price of a call option on ABC with strike $K = \$100$?
Question: What is the price of a call option on XYZ with strike $K = \$100$?

Financial Engineering & Risk Management

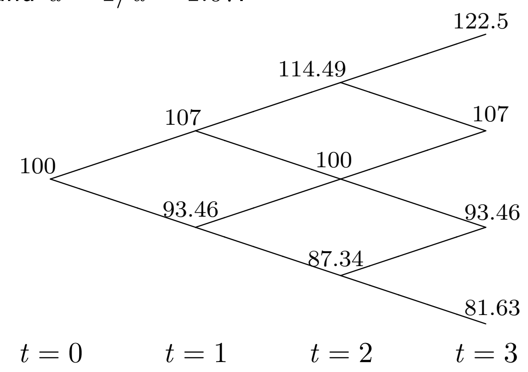
The Multi-Period Binomial Model

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

A 3-period Binomial Model

Recall $R = 1.01$ and $u = 1/d = 1.07$.

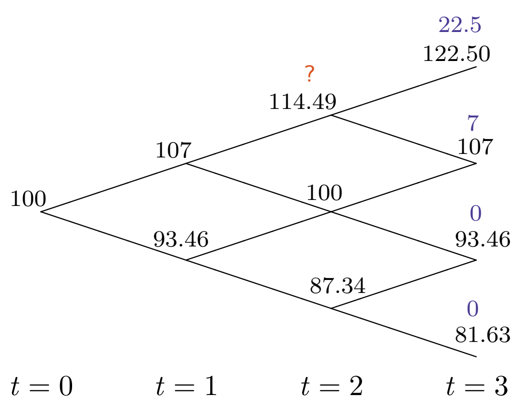


Just a series of 1-period models spliced together!

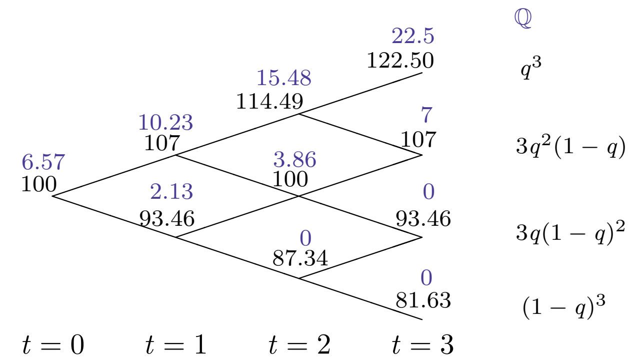
- all the results from the 1-period model apply
- just need to multiply 1-period probabilities along branches to get probabilities in multi-period model.

Pricing a European Call Option

Assumptions: expiration at $t = 3$, strike $= \$100$ and $R = 1.01$.



Pricing a European Call Option



- We can also calculate the price as

$$C_0 = \frac{1}{R^3} \mathsf{E}_0^{\mathbb{Q}} \left[\max(S_T - 100,\ 0) \right] \tag{1}$$

  - this is risk-neutral pricing in the binomial model
  - avoids having to calculate the price at every node.
- How would you find a replicating strategy?
  - to be defined and discussed in another module.

Financial Engineering & Risk Management

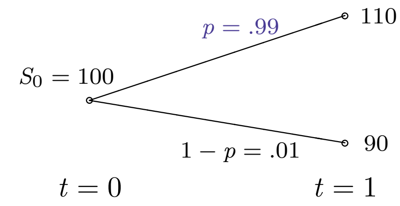
What's Going On?

M. Haugh G. Iyengar

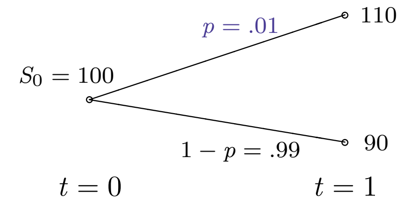
Department of Industrial Engineering and Operations Research
Columbia University

What's Going On?

- Stock ABC



- Stock XYZ



Question: What is the price of a call option on ABC with strike $K = \$100$?

Question: What is the price of a call option on XYZ with strike $K = \$100$?

What's Going On?

- Saw earlier

$$
\begin{aligned}
C_0 &= \frac{1}{R}\left[\frac{R-d}{u-d}C_u + \frac{u-R}{u-d}C_d\right] \\
&= \frac{1}{R}\left[q C_u + (1-q) C_d\right] \\
&= \frac{1}{R}\mathsf{E}_0^{\mathbb{Q}}[C_1]
\end{aligned}
$$

- So it appears that p doesn't matter!
- This is true ...
- ... but it only appears surprising because we are asking the wrong question!

Another Surprising Result?

$R = 1.02$

Stock Price

			119.10
		112.36	106.00
	106.00	100.00	94.34
100.00	94.34	89.00	83.96
t=0	t=1	t=2	t=3

European Option Price: K = 95

			24.10
		19.22	11.00
	14.76	7.08	0.00
11.04	4.56	0.00	0.00
t=0	t=1	t=2	t=3

$R = 1.04$

Stock Price

			119.10
		112.36	106.00
	106.00	100.00	94.34
100.00	94.34	89.00	83.96
t=0	t=1	t=2	t=3

European Option Price: K = 95

			24.10
		21.01	11.00
	18.19	8.76	0.00
15.64	6.98	0.00	0.00
t=0	t=1	t=2	t=3

Question: So the option price increases when we increase R. Is this surprising?

(See *"Investment Science"* (OUP) by D. G. Luenberger for additional examples on the binomial model.)

Existence of Risk-Neutral Probabilities $\Leftrightarrow$ No-Arbitrage

Recall our analysis of the binomial model:

- no arbitrage $\Leftrightarrow d < R < u$
- any derivative security with time T payoff, C_T, can be priced using

$$C_0 = \frac{1}{R^n} \mathsf{E}_0^{\mathbb{Q}}[C_T] \tag{2}$$

where $q > 0$, $1 - q > 0$ and $n = \#$ of periods.
(If Δt is the length of a period, then $T = n \times \Delta t$.)

In fact for any model if there exists a risk-neutral distribution, $\mathbb{Q}$, such that (2) holds, then arbitrage cannot exist. Why?

Reverse is also true: if there is no arbitrage then a risk-neutral distribution exists.

Together, these last two statements are often called the first fundamental theorem of asset pricing.

Financial Engineering & Risk Management

Pricing American Options

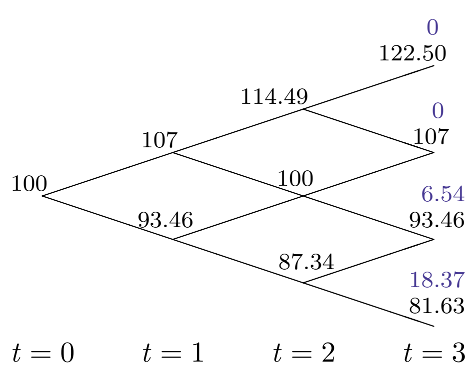
M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

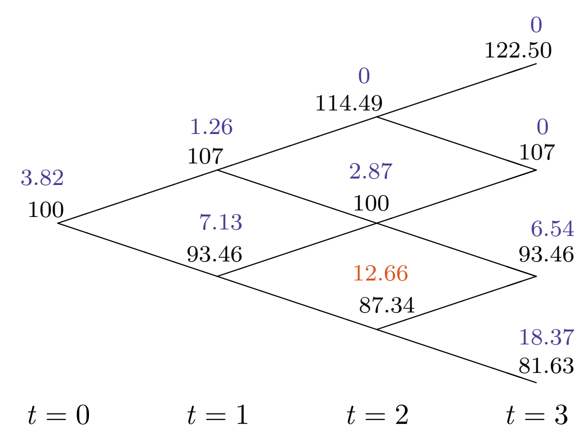
Pricing American Options

- Can also price American options in same way as European options
  - but now must also check if it's optimal to early exercise at each node.
- But recall never optimal to early exercise an American call option on non-dividend paying stock.

e.g. Price American put option: expiration at $t = 3$, $K = \$100$ and $R = 1.01$.



Pricing American Options



- Price option by working backwards in binomial the lattice.

e.g. $\quad 12.66 = \max\left[12.66,\ \frac{1}{R}\left(q \times 6.54 + (1-q) \times 18.37\right)\right]$

A Simple Die-Throwing Game

Consider the following game:

1. You can throw a fair 6-sided die up to a maximum of three times.
2. After any throw, you can choose to 'stop' and obtain an amount of money equal to the value you threw.

   e.g. if 4 thrown on second throw and choose to 'stop', then obtain $\$4$.

Question: If you are risk-neutral, how much would you pay to play this game?

Solution: Work backwards, starting with last possible throw:

1. You have just 1 throw left so fair value is 3.5.
2. You have 2 throws left so must figure out a strategy determining what to do after 1^{st} throw. We find

$$\text{fair value} = \frac{1}{6} \times (4+5+6) + \frac{1}{2} \times 3.5 = 4.25.$$

3. Suppose you are allowed 3 throws. Then ...

Question: What if you could throw the die 1000 times?

Financial Engineering & Risk Management

Replicating Strategies in the Binomial Model

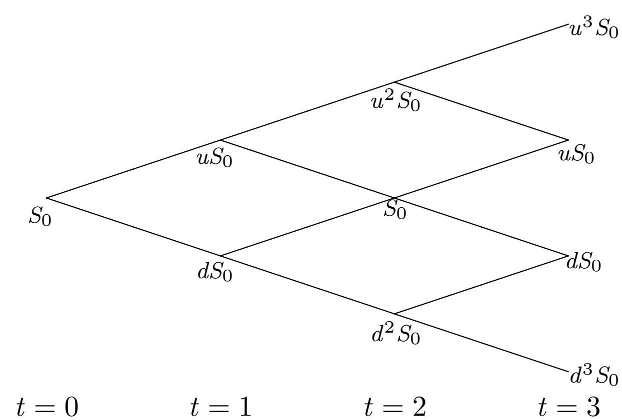
M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Trading Strategies in the Binomial Model

- Let S_t denote the stock price at time t.
- Let B_t denote the value of the cash-account at time t
  - assume without any loss of generality that $B_0 = 1$ so that $B_t = R^t$
  - so now explicitly viewing the cash account as a security.
- Let x_t denote # of shares held between times $t-1$ and t for $t = 1, \ldots, n$.
- Let y_t denote # of units of cash account held between times $t-1$ and t for $t = 1, \ldots, n$.
- Then $\boldsymbol{\theta}_t := (x_t, y_t)$ is the portfolio held:
  - (i) immediately **after** trading at time $t-1$ so it is known at time $t-1$
  - (ii) and immediately **before** trading at time t.
- $\boldsymbol{\theta}_t$ is also a random process and in particular, a **trading strategy**.

Trading Strategies in the Binomial Model



Self-Financing Trading Strategies

Definition. The value process, $V_t(\boldsymbol{\theta})$, associated with a trading strategy, $\boldsymbol{\theta}_t = (x_t, y_t)$, is defined by

$$V_t = \begin{cases} x_1 S_0 + y_1 B_0 & \text{for } t = 0 \\ x_t S_t + y_t B_t & \text{for } t \geq 1. \end{cases} \tag{3}$$

Definition. A self-financing trading strategy is a trading strategy, $\boldsymbol{\theta}_t = (x_t, y_t)$, where changes in V_t are due entirely to trading gains or losses, rather than the addition or withdrawal of cash funds. In particular, a self-financing strategy satisfies

$$V_t \;=\; x_{t+1} S_t + y_{t+1} B_t\,, \qquad t = 1, \ldots, n-1. \tag{4}$$

The definition states that the value of a self-financing portfolio just before trading is equal to the value of the portfolio just after trading

– so no funds have been deposited or withdrawn.

Self-Financing Trading Strategies

Proposition. If a trading strategy, $\boldsymbol{\theta}_t$, is self-financing then the corresponding value process, V_t, satisfies

$$V_{t+1} - V_t \;=\; x_{t+1}\,(S_{t+1} - S_t) \;+\; y_{t+1}\,(B_{t+1} - B_t)$$

so that changes in portfolio value can only be due to capital gains or losses and not the injection or withdrawal of funds.

Proof. For $t \geq 1$ we have

$$\begin{aligned} V_{t+1} - V_t &= (x_{t+1} S_{t+1} + y_{t+1} B_{t+1}) \;-\; (x_{t+1} S_t + y_{t+1} B_t) \\ &= x_{t+1}(S_{t+1} - S_t) \;+\; y_{t+1}(B_{t+1} - B_t) \end{aligned}$$

and for $t = 0$ we have

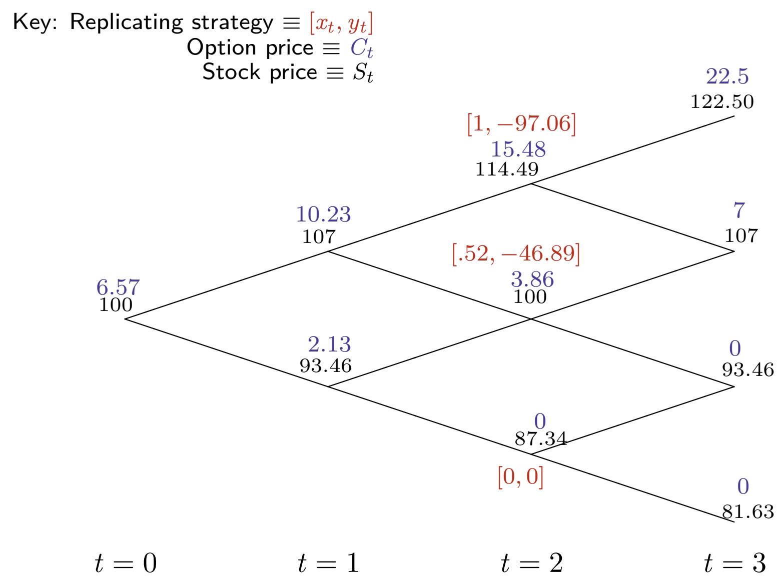
$$\begin{aligned} V_1 - V_0 &= (x_1 S_1 + y_1 B_1) - (x_1 S_0 + y_1 B_0) \\ &= x_1(S_1 - S_0) + y_1(B_1 - B_0). \end{aligned}$$

□

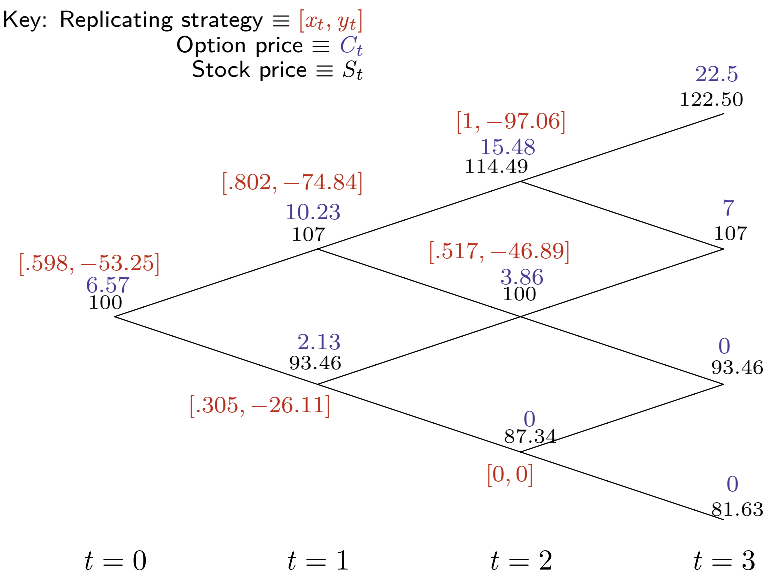
Risk-Neutral Price ≡ Price of Replicating Strategy

- We have seen how to price derivative securities in the binomial model.
- The key to this was the use of the 1-period risk neutral probabilities.
- But we first priced options in 1-period models using a replicating portfolio
  - and we did this without needing to define risk-neutral probabilities.
- In the multi-period model we can do the same, i.e., can construct a self-financing trading strategy that replicates the payoff of the option
  - this is called dynamic replication.
- The initial cost of this replicating strategy must equal the value of the option
  - otherwise there's an arbitrage opportunity.
- The dynamic replication price is of course equal to the price obtained from using the risk-neutral probabilities and working backwards in the lattice.
- And at any node, the value of the option is equal to the value of the replicating portfolio at that node.

The Replicating Strategy For Our European Option



The Replicating Strategy For Our European Option



e.g. $.802 \times 107 + (-74.84) \times 1.01 = 10.23$ at upper node at time $t = 1$

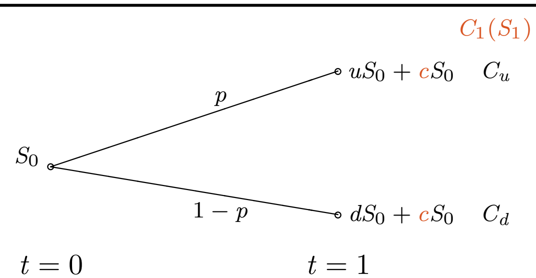
Financial Engineering & Risk Management

Including Dividends

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Including Dividends



- Consider again 1-period model and assume stock pays a proportional dividend of cS_0 at $t = 1$.
- No-arbitrage conditions are now $d + c < R < u + c$.
- Can use same replicating portfolio argument to find price, C_0, of any derivative security with payoff function, $C_1(S_1)$, at time $t = 1$.
- Set up replicating portfolio as before:

$$
\begin{aligned}
uS_0x + cS_0x + Ry &= C_u \\
dS_0x + cS_0x + Ry &= C_d
\end{aligned}
$$

Derivative Security Pricing with Dividends

- Solve for x and y as before and then must have $C_0 = xS_0 + y$.
- Obtain

$$
\begin{aligned}
C_0 &= \frac{1}{R}\left[\frac{R-d-c}{u-d}C_u + \frac{u+c-R}{u-d}C_d\right] \\
&= \frac{1}{R}\left[qC_u + (1-q)C_d\right] \\
&= \frac{1}{R}\mathsf{E}_0^{\mathbb{Q}}[C_1].
\end{aligned} \tag{5}
$$

- Again, can price any derivative security in this 1-period model.
- Multi-period binomial model assumes a proportional dividend in each period
  - so dividend of cS_i is paid at $t = i + 1$ for each i.
- Then each embedded 1-period model has identical risk-neutral probabilities
  - and derivative securities priced as before.
- In practice dividends are not paid in every period
  - and are therefore just a little more awkward to handle.

The Binomial Model with Dividends

- Suppose the underlying security does not pay dividends. Then

$$S_0 = \mathsf{E}_0^{\mathbb{Q}}\left[\frac{S_n}{R^n}\right] \qquad (6)$$

  – this is just risk-neutral pricing of European call option with $K = 0$.
- Suppose now underlying security pays dividends in each time period.
- Then can check (6) no longer holds.
- Instead have

$$S_0 = \mathsf{E}_0^{\mathbb{Q}}\left[\frac{S_n}{R^n} + \sum_{i=1}^{n}\frac{D_i}{R^i}\right] \qquad (7)$$

  - D_i is the dividend at time i
  - and S_n is the ex-dividend security price at time n.
- Don't need any new theory to prove (7)
  - it follows from risk-neutral pricing and observing that dividends and S_n may be viewed as a portfolio of securities.

Viewing a Dividend-Paying Security as a Portfolio

- To see this, we can view the i^{th} dividend as a separate security with value

$$P_i = \mathsf{E}_0^{\mathbb{Q}}\left[\frac{D_i}{R^i}\right].$$

- Then owner of underlying security owns a "portfolio" of securities at time 0
  - value of this "portfolio" is $\sum_{i=1}^{n} P_i + \mathsf{E}_0^{\mathbb{Q}}\left[\frac{S_n}{R^n}\right]$.
- But value of underlying security is S_0.
- Therefore must have

$$S_0 = \sum_{i=1}^{n} P_i + \mathsf{E}_0^{\mathbb{Q}}\left[\frac{S_n}{R^n}\right]$$

which is (7).

Financial Engineering & Risk Management

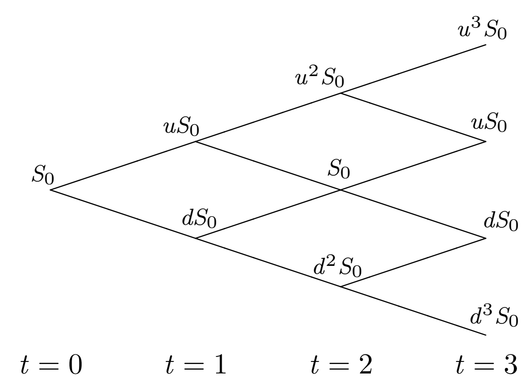
Pricing Forwards and Futures

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Pricing Forwards in the Binomial Model

- Have an n-period binomial model with $u = 1/d$.



- Consider now a forward contract on the stock that expires after n periods.
- Let G_0 denote date $t = 0$ "price" of the contract.
- Recall G_0 is chosen so that contract is initially worth zero.

Pricing Forwards in the Binomial Model

- Therefore obtain

$$0 \;=\; \mathsf{E}_0^{\mathbb{Q}}\left[\frac{S_n - G_0}{R^n}\right]$$

so that

$$G_0 \;=\; \mathsf{E}_0^{\mathbb{Q}}\left[S_n\right]. \tag{8}$$

- Again, (8) holds whether the underlying security pays dividends or not.

What is a Futures "Price"?

- Consider now a futures contract on the stock that expires after n periods.
- Let F_t be the date t "price" of the futures contract for $0 \leq t \leq n$.
- Then $F_n = S_n$. Why?
- A common misconception is that:
  (i) F_t is how much you must pay at time t to buy one contract
  (ii) or how much you receive if you sell one contract

  This is **false!**
- A futures contract always costs nothing.
- The "price", F_t is only used to determine the cash-flow associated with holding the contract
  - so that $\pm(F_t - F_{t-1})$ is the payoff received at time t from a long or short position of one contract held between $t-1$ and t.
- In fact a futures contract can be characterized as a security that:
  (i) is always worth zero
  (ii) and that pays a dividend of $(F_t - F_{t-1})$ at each time t.

Pricing Futures in the Binomial Model

- Can compute time $t = n-1$ futures price, F_{n-1}, by solving

$$0 \;=\; \mathsf{E}_{n-1}^{\mathbb{Q}}\left[\frac{F_n - F_{n-1}}{R}\right]$$

to obtain $F_{n-1} = \mathsf{E}_{n-1}^{\mathbb{Q}}[F_n]$.

- In general we have $F_t = \mathsf{E}_t^{\mathbb{Q}}[F_{t+1}]$ for $0 \leq t < n$ so that

$$\begin{aligned} F_t &= \mathsf{E}_k^{\mathbb{Q}}[F_{t+1}] \\ &= \mathsf{E}_t^{\mathbb{Q}}[\mathsf{E}_{t+1}^{\mathbb{Q}}[F_{t+2}]] \\ &\vdots \qquad \vdots \\ &= \mathsf{E}_t^{\mathbb{Q}}[\mathsf{E}_{t+1}^{\mathbb{Q}}[\cdots\ \mathsf{E}_{n-1}^{\mathbb{Q}}[F_n]]]. \end{aligned}$$

Pricing Futures in the Binomial Model

- Law of iterated expectations then implies $F_t \;=\; \mathsf{E}_t^{\mathbb{Q}}[F_n]$
  - so the futures price process is a $\mathbb{Q}$-martingale.
- Taking $t = 0$ and using $F_n = S_n$ we also have

$$F_0 \;=\; \mathsf{E}_0^{\mathbb{Q}}\left[S_n\right]. \tag{9}$$

- Note that (9) holds whether the security pays dividends or not
  - dividends only enter through $\mathbb{Q}$.
- Comparing (8) and (9) and we see that $F_0 = G_0$ in the binomial model
  - not true in general.

Financial Engineering & Risk Management

The Black-Scholes Model

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

The Black-Scholes Model

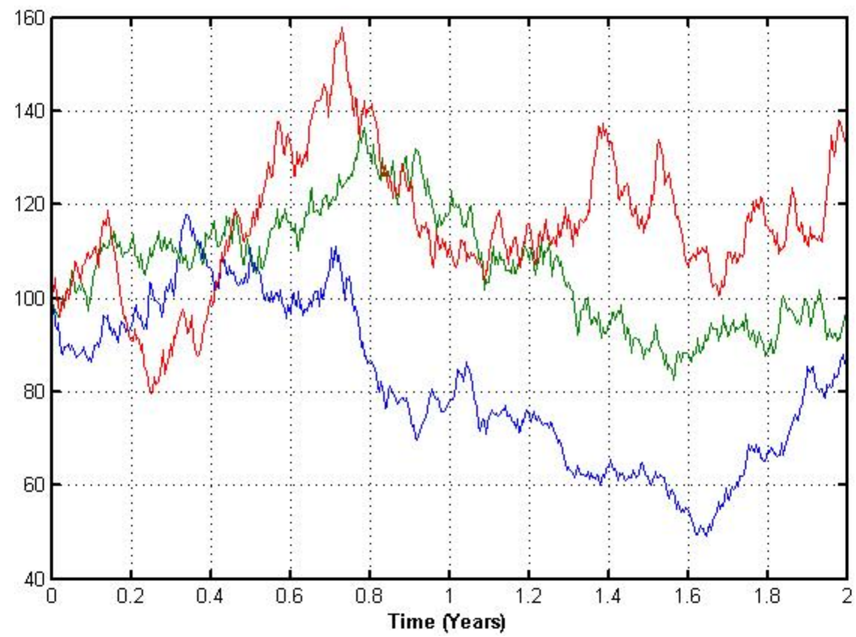
Black and Scholes assumed:

1. A continuously-compounded interest rate of r.
2. Geometric Brownian motion dynamics for the stock price, S_t, so that

$$S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t}$$

   where W_t is a standard Brownian motion.
3. The stock pays a dividend yield of c.
4. Continuous trading with no transactions costs and short-selling allowed.

Sample Paths of Geometric Brownian Motion



The Black-Scholes Formula

- The Black-Scholes formula for the price of a European call option with strike K and maturity T is given by

$$C_0 = S_0 e^{-cT} N(d_1) - K e^{-rT} N(d_2)$$

  where

$$d_1 = \frac{\log(S_0/K) + (r - c + \sigma^2/2)T}{\sigma\sqrt{T}},$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

  and $N(d) = \mathrm{P}(N(0,1) \leq d)$.
- Note that μ does not appear in the Black-Scholes formula
  - just as p is not used in option pricing calculations for the binomial model.
- European put option price, P_0, can be calculated from put-call parity

$$P_0 + S_0 e^{-cT} = C_0 + K e^{-rT}.$$

The Black-Scholes Formula

- Black-Scholes obtained their formula using a similar replicating strategy argument to the one we used for the binomial model.
- In fact, can show that under the Black-Scholes GBM model

$$C_0 = \mathsf{E}_0^{\mathbb{Q}}\left[e^{-rT}\max(S_T - K,\, 0)\right]$$

where under $\mathbb{Q}$

$$S_t = S_0 e^{(r-c-\sigma^2/2)t+\sigma W_t}.$$

Calibrating a Binomial Model

- Often specify a binomial model in terms of Black-Scholes parameters:
  1. r, the continuously compounded interest rate.
  2. σ, the annualized volatility.
- Can convert them into equivalent binomial model parameters:
  1. $R_n = \exp\left(r\frac{T}{n}\right)$, where $n =$ number of periods in binomial model
  2. $R_n - c_n = \exp\left((r-c)\frac{T}{n}\right) \approx 1 + r\frac{T}{n} - c\frac{T}{n}$
  3. $u_n = \exp\left(\sigma\sqrt{\frac{T}{n}}\right)$
  4. $d_n = 1/u_n$

  and now price European and American options, futures etc. as before.
- Then risk-neutral probabilities calculated as

$$q_n = \frac{e^{(r-c)\frac{T}{n}} - d_n}{u_n - d_n}.$$

- Spreadsheet calculates binomial parameters this way
  - binomial model prices converge to Black-Scholes prices as $n \to \infty$.

The Binomial Model as $\Delta t \to 0$

- Consider a binomial model with n periods
  - each period corresponds to time interval of $\Delta t := T/n$.
- Recall that we can calculate European option price with strike K as

$$C_0 = \frac{1}{R^n}\mathsf{E}_0^{\mathbb{Q}}\left[\max(S_T - K,\ 0)\right] \qquad (10)$$

- In the binomial model can write (10) as

$$\begin{aligned} C_0 &= \frac{1}{R_n^n}\sum_{j=0}^{n}\binom{n}{j} q_n^j(1-q_n)^{n-j}\ \max(S_0 u_n^j d_n^{n-j} - K,\ 0) \\ &= \frac{S_0}{R_n^n}\sum_{j=\eta}^{n}\binom{n}{j} q_n^j(1-q_n)^{n-j}\ u_n^j d_n^{n-j} \ - \ \frac{K}{R_n^n}\sum_{j=\eta}^{n}\binom{n}{j} q_n^j(1-q_n)^{n-j} \end{aligned}$$

  where $\eta := \min\{j : S_0 u_n^j d_n^{n-j} \geq K\}$.
- Can show that if $n \to \infty$ then C_0 converges to the Black-Scholes formula.

Some History

- Bachelier (1900) perhaps first to model Brownian motion
  - modeled stock prices on the Paris Bourse
  - predated Einstein by 5 years.
- Samuelson (1965) rediscovered the work of Bachelier
  - proposed geometric Brownian motion as a model for security prices
  - succeeded in pricing some kinds of warrants
  - was Merton's doctoral adviser
- Itô (1950's) developed the Itô or stochastic calculus
  - the main mathematical tool in finance
  - Itô's Lemma used later by Black-Scholes-Merton
  - Doeblin (1940) recently credited with independently developing stochastic calculus
- Black-Scholes-Merton (early 1970's) published their papers
- Many other influential figures
  - Thorpe (card-counting and perhaps first to discover Black-Scholes formula?)
  - Cox and Ross
  - Harrison and Kreps
  - . . .

Financial Engineering & Risk Management

An Example: Pricing a European Put on a Futures Contract

M. Haugh G. Iyengar

Department of Industrial Engineering and Operations Research
Columbia University

Pricing a European Put on a Futures Contract

- We can also price an option on a futures contract.
- In fact many of the most liquid options are options on futures contracts
  - **e.g.** S&P 500, Eurostoxx 50, FTSE 100 and Nikei 225.
  - in these cases the underlying security is not actually traded.
- Consider the following parameters:
  - $S_0 = 100$, $n = 10$ periods, $r = 2\%$, $c = 1\%$ and $\sigma = 20\%$
  - futures expiration = option expiration = $T = .5$ years.
- Futures price lattice obtained using $S_n = F_n$ and then

$$F_t = \mathsf{E}_t[F_{t+1}] \quad \text{for } 0 \leq t < n.$$

- Obtain a put option value of 5.21.

Pricing a European Put on a Futures Contract

- In practice we don't need a model to price liquid options
  - market forces, i.e. supply and demand, determines the price
  - which in this case amounts to determining σ or the implied volatility.
- Models are required to hedge these options however
  - and price exotic or illiquid derivative securities.
- Will return to this near end of course.