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

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

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

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

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

For example, the mean is calculated as

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

Let r = r-1. Doesn't affect infinity, and
lets you change the values of RHS to 1 —>

$$= \;\lambda \sum_{r=1}^{\infty} \frac{\lambda^{r-1} e^{-\lambda}}{(r-1)!}$$

$$= \;\lambda \sum_{r=0}^{\infty} \frac{\lambda^{r} e^{-\lambda}}{r!} \;=\; \lambda.$$

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 $\mathsf{P}(Y_1 \geq 4 \,|\, X \geq 8)$?

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

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

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 $N \sim \text{Poisson}(\lambda)$. 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

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

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

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

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

$$\overset{\text{n x n}}{\mathbf{\Sigma}} := \mathsf{Cov}(\mathbf{X}) \ := \ \mathsf{E}\left[\overset{\text{n x 1}}{(\mathbf{X} - \mathsf{E}[\mathbf{X}])} \ \overset{\text{1 x n}}{(\mathbf{X} - \mathsf{E}[\mathbf{X}])^\top}\right]$$

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.

Remember that corr(xi, xj) = Cov(xi, xj) / sqrt(Var(xi), Var(xj))

Variances and Covariances

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

$$\begin{aligned} \mathsf{E}\left[\mathbf{A}\mathbf{X} + \mathbf{a}\right] &= \mathbf{A}\mathsf{E}\left[\mathbf{X}\right] + \mathbf{a} &\quad (4) \\ \mathsf{Cov}(\mathbf{A}\mathbf{X} + \mathbf{a}) &= \mathbf{A}\ \mathsf{Cov}(\mathbf{X})\ \mathbf{A}^\top. &\quad (5) \end{aligned}$$

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.

Remember that corr(xi, xj) = Cov(xi, xj) / sqrt(Var(xi), Var(xj))

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 $\boldsymbol{\Sigma}$ then we write

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

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

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

where $|\cdot|$ denotes the determinant.

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

- in this case the X_i's are independent.

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

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

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

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

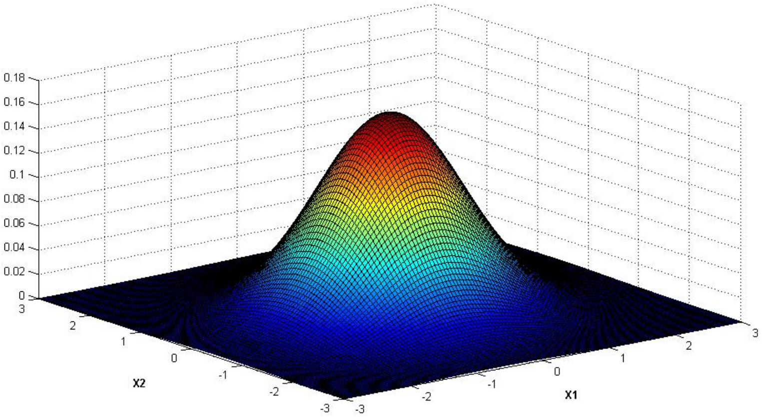
are the mean vector and covariance matrix of $(\mathbf{X_1}, \mathbf{X_2})$.

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, $\mathbf{X_i} \sim \mathrm{MN}(\boldsymbol{\mu}_i, \mathbf{\Sigma}_{ii})$, 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 $\mathbf{\Sigma}_{2.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$.

Due to covariance between x1 and x2, x1 gives you info about x2. Hence, you can use this to adjust your guess of x2, and the range of values x2 can take also decreases hence lower variance

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

fn is just all the information we know at time n

1. $\mathsf{E}^P[|X_n|] < \infty$ for all $n \geq 0$ for integrability
2. $\mathsf{E}^P[X_{n+m}|\mathcal{F}_n] = X_n$ for all $n, m \geq 0$. expected value of x in future is the current value today

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

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

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.

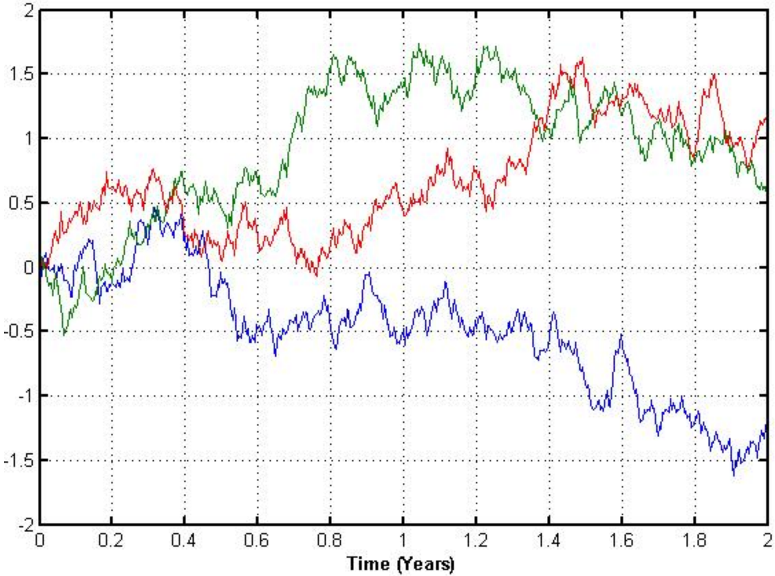
Standard Brownian Motion

- 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

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

Now we know (why?) $\mathsf{E}_0\left[W_s^2\right] = s$.

To calculate first term on r.h.s. of (9) a version of the conditional expectation identity implies

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

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 $X_t \sim \mathsf{GBM}(\mu, \sigma)$.

Note that

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

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

Geometric Brownian Motion

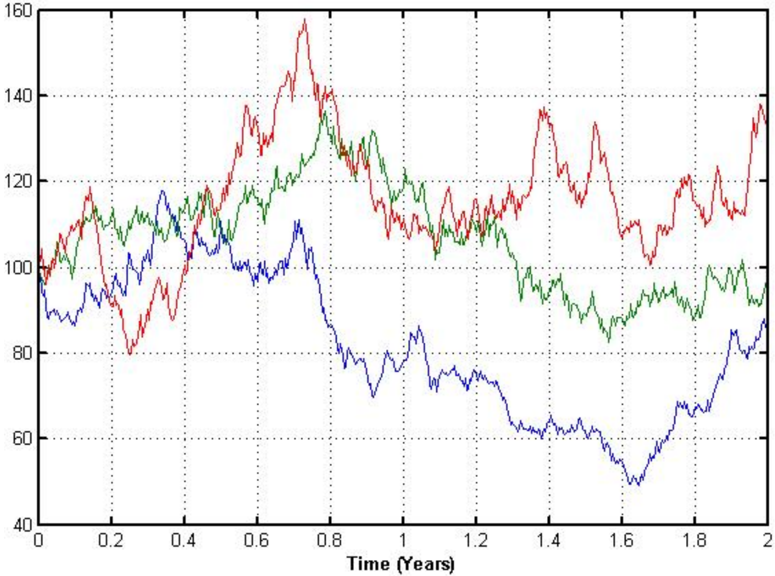
Question: Suppose $X_t \sim \mathsf{GBM}(\mu, \sigma)$. What is $\mathsf{E}_t[X_{t+s}]$?

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 $X_t \sim \text{GBM}(\mu, \sigma)$. 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.