

Introduction and Mathematical Preliminaries

In these lecture notes we introduce some preliminary material for the course. We discuss the famous St. Petersburg paradox and it's resolution through the use of utility functions. We also use this example to motivate the difficulty of valuing random cash-flows and then introduce European call and put options in the context of the binomial model. We spend considerable time on the mechanics of futures contracts and markets as they are of fundamental importance. We then review some important ideas from Monte-Carlo simulation - a key computational tool in finance - and end with some key results and models from probability including Brownian motion (BM) and geometric Brownian motion (GBM). GBM is a classic model in finance that is often used to model security prices and in fact underlies the famous Black-Scholes model for pricing European call and put options.

1 The St. Petersburg Paradox and Valuing Random Cash-Flows

Consider a game where a fair coin is tossed repeatedly until the first head appears. If the first head appears on the n^{th} toss, then you will receive $\$2^n$. How much would you be willing to pay in order to play this game? Note that the expected payoff satisfies

$$\mathrm{E}[\text{Payoff}] \;=\; \sum_{n=1}^{\infty} 2^n\, \mathbf{P}(1^{st} \text{ head on } n^{th} \text{ toss}) \;=\; \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \infty. \tag{1}$$

Given the result in (1), would you say the fair value of the game is $X = \infty$? Is this how much you would be prepared to pay? This is the St. Petersburg paradox and it refers to the mistaken belief that we ought to be willing to pay the fair, i.e. expected value, of a game in order to play it.

1.1 Utility Functions

The 17^{th} century Swiss mathematician Daniel Bernoulli resolved this paradox by introducing the idea of a utility function. A utility function $u(x)$ is defined on levels of wealth x with the interpretation that $u(x)$ measures how much **utility** or benefit someone obtains from holding x units of wealth. Different people will have different utility functions. A utility function $u(.)$ should have the following properties:

1. It should be **increasing** in x, i.e. $u'(x) > 0$ where $u'(x)$ is the derivative of u w.r.t. x. This simply reflects the fact people prefer more money to less money.

2. It should be **concave**, i.e. $u''(x) < 0$ where $u''(x)$ is the second derivative of u w.r.t. x. This property reflects the fact that the marginal benefit of an additional dollar decreases in the wealth level x.

A **risk averse** individual / investor will have an increasing and concave utility function. Common examples of such utility functions include:

- **Log** utility with $u(x) = \log(x)$
- **Power** utility with $u(x) = x^{1-\gamma}/(1-\gamma)$ for $\gamma > 0$ and $\gamma \neq 1$
- **Exponential** utility with $u(x) = -e^{-\alpha x}/\alpha$ for $\alpha \geq 0$.

Many other utility functions[1] are of course possible. The particular utility function that Bernoulli introduced was the $\log$ utility function. In particular, if an individual has log utility then her **expected utility** of the payoff is given by

$$\mathrm{E}[u(\text{Payoff})] = \sum_{n=1}^{\infty} \log(2^n)\,\mathbf{P}(1^{st} \text{ head on } n^{th} \text{ toss}) = \log(2)\sum_{n=1}^{\infty} \frac{n}{2^n} < \infty.$$

Exercise 1 *How much do you think the opportunity to play the game is worth to you if you have log utility?*

1.1.1 Certainty Equivalent

In order to answer Exercise 1 we need to compute the certainty equivalent of the game. This is the **fixed amount** of money x_{ce} say whereby having x_{ce} for certain has the same utility as playing the game. That is, x_{ce} satisfies

$$u(x_{\mathsf{ce}}) = \mathrm{E}[u(\text{Payoff})]$$

which in the game example above simplifies to

$$\log(x_{\mathsf{ce}}) = \log(2)\sum_{n=1}^{\infty} \frac{n}{2^n}.$$

We can then solve for x_{ce} and for a player with $\log$ utility, x_{ce} is how much he/she should be willing to pay to play the game.

1.1.2 Valuing Random Cash-Flows

Many problems in finance require us to model the preferences of the decision-maker. If the decision-maker is an individual, we often assume she is *risk-averse* and endow her with an appropriate utility function. But different individuals will have different utility functions reflecting their different levels of risk aversion. This means they will have different values, i.e. certainty equivalents, of random cash-flows. So how then should random cash-flows be valued? This is a central problem in financial economics and in general **equilibrium models** are required. One exception to this, however, is the pricing of derivative securities in so-called **complete-market** models. The binomial model is a complete-market model.

1.2 Option Pricing in the Binomial Model

Consider the following example of a financial market. There are 3 periods, one risky asset and one risk-free asset. At any time t the value of the risky asset S_t will either increase by a factor u (with probability p) or decrease by a factor d (with probability $1-p$) over the next period. The possible evolutions of S_t are given in Figure 1 where $u = 1.06$, $d = 1/u$ and $S_0 = 100$. The risk-free asset is a **cash account** so that $\$1$ invested in it at $t = 0$ will be worth R^t dollars at time t, where R is the gross interest rate per period.

Suppose we are interested in computing the price of a **European call option** in this market. Specifically the option **expires** at time $t = 3$ and has a **strike** of 95. This means the option has a time $t = 3$ payoff of

$$\text{European call option payoff} := (S_3 - 95)^+ := \max(0, S_3 - 95).$$

In contrast the payoff of a **European put option** with the same strike and maturity would be

$$\text{European put option payoff} := (95 - S_3)^+ := \max(0, 95 - S_3).$$

Some interesting questions / observations arise regarding the prices of these options.

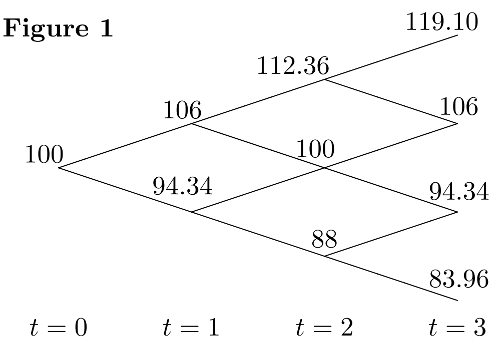
1. Do we have enough information to compute their prices?

[1] The theory of expected utility was formalized in the 1950's by John Von Neumann and Oscar Morgenstern. It has been a cornerstone of economics and finance ever since.

2. As with our earlier coin-tossing game, shouldn't the option prices somehow depend on the utility functions of the buyer and seller?

3. Will the price depend on the probability of an *up-move* in each period?

Figure 1



In fact, we will see later in the course that we can indeed compute unique arbitrage-free prices for all securities in the binomial model, including in particular European call and put option prices. This will be achieved via the construction of a **self-financing trading strategy** that **replicates** the payoff of the security we wish to price. We note for now that a trading strategy is simply **a rule** telling us exactly what positions to hold in the stock and cash account at each time t and at each node. A positive (negative) position in the stock means we hold (have short-sold) the stock whereas a positive (negative) position in the cash-account means that we have lent to (borrowed from) the "bank". The "self-financing" property implies that the value of the strategy **immediately after** trading at any node is identical to its value **immediately before** trading at the node. Finally the rule telling us what to hold at each node can only depend on information that is available to us at that node. In particular, it cannot depend on future information. We will discuss these properties in further definition later in the course but we include here a brief aside on the mechanics of short-selling.

1.2.1 Short-Selling

The ability to short-sell is extremely important[2] for the "price-discovery" role of financial markets. It is also an extremely important component of the "plumbing" of financial markets and is required for the pricing and hedging of derivatives securities. It's therefore important that participants under how short-selling works. The specific steps are:

1. Find a broker who will lend the stock to you. This typically requires you to pay a (usually) small fee to the broker.

2. Sell the stock in the market.

3. At some point you buy the stock back in the market and return it to the broker. (Occasionally the broker may demand the stock back in which case you will have to buy the stock back at that point or find an alternative broker who will lend the stock to you.)

Note that short-selling allows you to gain a negative exposure to the stock. For example suppose the stock is worth $100 when you sell but only $60 when you buy it back. This means you will have made a profit of $40 from your short-sale. On the other hand if the stock goes up between selling and buying it back then you will incur a loss. And as a stock's price is unbounded, your potential losses from a short-sale are also unbounded. This makes short-selling a very risky activity and it's only suitable for professional investors and market participants.

[2] In times of crisis there are usually some participants who call for short-selling to be banned as it can add to the price pressure on distressed securities. The vast majority of objective observers support the ability to short-sell, however, even in times of crisis because of the important roles that it plays. Typically when it is banned the decision is due to political pressure rather than objective or economic reasons.

2 Futures Markets and their Mechanics

Perhaps the best way to understand the mechanics of a futures market is by example and we will do so via an imaginary futures market based on the outcome of a test match in cricket which typically lasts five days.

Example 1 (Cricket Futures)
We consider an example of a futures market where the futures contracts are *not* written on an underlying financial asset or commodity. Instead, they are written on the total number of runs that are scored by the two teams in a cricket test match. The market opens before the cricket match takes place and expires at the conclusion of the match. Similar futures markets do exist in practice and this example simply demonstrates that in principle, futures markets can be created where just about any underlying and ultimately observable variable can serve as the underlying asset.

In the table below we present one possible evolution of the futures market between June 3 and June 19. The initial position is long 100 contracts and it is assumed that this position is held until the test match ends on June 19. An initial balance of $\$10,000$ is assumed and this balance earns interest at a rate of $.005\%$ per day. It is also important to note that when the futures position is initially adopted the cost is zero, i.e. initially there is no exchange of cash.

CRICKET FUTURES CONTRACTS

Date	Price	Position	Profit	Interest	Balance	
June 3	720.00	100	0	0	10,000	
June 4	721.84	100	184	1	10,184	
June 5	721.52	100	-31	1	10,153	
June 6	711.88	100	-964	1	9,190	
June 7	716.67	100	479	0	9,669	
June 8	720.04	100	337	0	10,006	
June 9	672.45	100	-4,759	1	5,248	Any explanation?
June 10	673.25	100	80	0	5,328	
June 11	687.04	100	1,379	0	6,708	
June 12	670.56	100	-1,648	0	5,060	
June 13	656.25	100	-1,431	0	3,630	
June 14	647.14	100	-912	0	2,718	
June 15	665.57	100	1,843	0	4,561	Test Match Begins
June 16	673.48	100	791	0	5,353	
June 17	672.88	100	-60	0	5,293	
June 18	646.63	100	-2,625	0	2,669	
June 19	659.00	100	1,294	0	3,963	Test Match Ends
Total			-6,042		3,963	

The particular details of the cricket futures market are as follows:

- The futures market opens on June 3rd and the test match itself begins on June 15th. The market closes when the match is completed on June 19th. The final price of the futures contract is by definition then set equal to the total number of runs that were scored in the test match.
- The closing "price" on the first day of the market was 720. This can be interpreted as the market forecast for the total number of runs that will be scored by both teams in the test match. This value varies

through time as new events occur and new information becomes available. Examples of such events include information regarding player selection and fitness, current form of players, weather forecast updates, umpire selection, condition of the field etc.

- The **contract size** is $\$1$. This means if you go **long** one contract and the price increases by two say, then you will have $2 \times 1 \times \$1 = \$2$ added to your cash balance on the futures exchange. On the other hand, if the price had decreased by 8, say, and you were **short** 5 contracts then your balance would increase by $(-8) \times (-5) \times \$1 = \$40$. This process of **marking-to-market** is usually done on a daily basis. Moreover, **the value of your futures position immediately after marking-to-market is always zero**, as any accrued profits or losses have already been added to or subtracted from your cash balance. ∎

Remark 1 *A convenient way of thinking about a futures contract with price process F_t is as follows. A futures contract is a security that is always worth zero but it pays a "dividend" every day. In particular the dividend per contract at the end of period t is $\pm(F_t - F_{t-1})$ with someone who is long the futures contract receiving $+(F_t - F_{t-1})$ and someone who is short the contract receiving $-(F_t - F_{t-1})$. Note that these dividends are random and can be positive or negative.*

In Example 1 we did not discuss the details of *margin requirements* which are intended to protect against the risk of default. A typical margin requirement would be that the futures trader maintain a minimum *balance* in her trading account. This minimum balance will often be a function of the contract value (perhaps 5% to 10%) multiplied by the position, i.e., the number of contracts that the trader is long or short. When the balance drops below this minimum level a **margin call** is made after which the trader must deposit enough funds so as to meet the balance requirement. Failure to satisfy this margin call will result in the futures position being **closed out**.

2.1 Strengths and Weaknesses of Futures Markets

Futures markets are useful for a number of reasons:

- It is easy to take a position using futures markets without having to purchase the underlying asset. Indeed, it is not even possible to buy the underlying asset in some cases, e.g., equity indices, interest rates, cricket matches and presidential elections. In fact, when people invest in indices such as the S&P 500, they typically[3] invest in futures contracts written on the S&P 500.
- Futures markets allow you to *leverage* your position. That is, you can dramatically increase your exposure to the underlying security by using the futures market instead of the spot market.
- They are well organized and designed to eliminate counter-party risk as well as the so-called double-coincidence-of-wants problem.
- The mechanics of a futures market are generally independent of the underlying 'security' so they are easy to "operate" and easily understood by investors.

Futures markets also have some weaknesses:

- The fact that they are so useful for leveraging a position also makes them dangerous for unsophisticated and / or rogue investors.
- Futures prices are (more or less) *linear* in the price of the underlying security. This limits the types of risks that can be perfectly hedged using futures markets. Nonetheless, non-linear risks can still be partially hedged using futures.

[3]Nowadays, it is also quite common to invest in indices via *exchange-traded funds*, i.e. ETFs.

2.2 Hedging with Futures: the Perfect and Minimum-Variance Hedges

Futures markets are of great importance for hedging against risk. They are particularly suited to hedging risk that is *linear* in the underlying asset. This is because the final payoff at time T from holding a futures contract is linear[4] in the terminal price of the underlying security, S_T. In this case we can achieve a perfect hedge by taking an equal and opposite position in the futures contract.

Example 2 (Perfect Hedge)
Suppose a wheat producer knows that she will have $100,000$ bushels of wheat available to sell in three months time. She is concerned that the spot price of wheat will move against her, i.e. fall, in the intervening three months and so she decides to lock in the sale price now by hedging in the futures markets. Since each wheat futures contract is for $5,000$ bushels, she therefore decides to sell 20 three-month futures contracts. Note that as a result, the wheat producer has a perfectly hedged position. ∎

In general, perfect hedges are not available for a number of reasons:

1. None of the expiration dates of available futures contracts may exactly match the expiration date of the payoff Z_T that we want to hedge.
2. Z_T may not correspond exactly to an *integer* number of futures contracts.
3. The security underlying the futures contract may be different to the security underlying Z_T. For example, airlines often wish to hedge against the price of jet fuel but futures contracts are only available on oil.
4. Z_T may be a non-linear function of the security price underlying the futures contract.
5. Combinations of all the above are also possible.

When perfect hedges are not available, we often use the *minimum-variance* hedge to identify a good hedging position in the futures markets. To derive the minimum-variance hedge, we let Z_T be the cash flow that occurs at date T that we wish to hedge, and we let F_t be the time t price of the futures contract. At date $t = 0$ we adopt a position[5] of h in the futures contract and hold this position until time T. Since the initial cost of a futures position is zero, we can (if we ignore issues related to interest on the margin account) write the terminal cash-flow, Y_T, as

$$Y_T \;=\; Z_T \;+\; h(F_T - F_0).$$

Our objective then is to minimize

$$\mathrm{Var}(Y_T) \;=\; \mathrm{Var}(Z_T) \;+\; h^2\mathrm{Var}(F_T) \;+\; 2h\mathrm{Cov}(Z_T,\ F_T)$$

and we find that the minimizing h and minimum variance are given by

$$\begin{aligned} h^* &= -\frac{\mathrm{Cov}(Z_T, F_T)}{\mathrm{Var}(F_T)} \\ \mathrm{Var}(Y_T^*) &= \mathrm{Var}(Z_T) \;-\; \frac{\mathrm{Cov}(Z_T,\ F_T)^2}{\mathrm{Var}(F_T)}. \end{aligned}$$

[4]The final payoff is $\pm x(F_T - F_0) = \pm x(S_T - F_0)$ depending on whether or not we are long or short x futures contracts and this position is held for the entire period, $[0, T]$. This assumes that we are ignoring the costs and interest payments associated with the margin account. As they are of secondary importance, we usually do this when determining what hedging positions to take.

[5]A positive value of h implies that we are long the futures contract while a negative value implies that we are short. More generally, we could allow h to vary stochastically as a function of time. We might want to do this, for example, if Z_T is *path dependent* or if it is a non-linear function of the security price underlying the futures contract. When we allow h to vary stochastically, we say that we are using a *dynamic hedging strategy.* Such strategies are often used for hedging options and other derivative securities with non-linear payoffs.

Such *static hedging strategies* are often used in practice, even when dynamic hedging strategies are capable of achieving a smaller variance. Note also, that unless $\mathrm{E}[F_T] = F_0$, it will not be the case that $\mathrm{E}[Z_T] = \mathrm{E}[Y_T^*]$. It is also worth noting that the mean-variance hedge is not in general the same as the *equal-and-opposite* hedge.

Example 3 (From Luenberger's *Investment Science*)
Assume that the cash flow is given by $y = S_T W + (F_T - F_0)h$. Let $\sigma_S^2 = \mathrm{Var}(S_T)$, $\sigma_F^2 = \mathrm{Var}(F_T)$ and $\sigma_{ST} = \mathrm{Cov}(S_T, F_T)$. In an equal and opposite hedge, h is taken to be an opposite equivalent dollar value of the hedging instrument. Therefore $h = -kW$, where k is the price ratio between the asset and the hedging instrument. Express the standard deviation of y with the equal and opposite hedge in the form

$$\sigma_y \;=\; W\sigma_S \;\times\; B.$$

That is, find B.

Solution: We have $y = S_T W - (F_T - F_0)Wk$ where $k = S_0/F_0$. Note that h is determined at date 0 and is therefore a function of date 0 information only. It is easy to obtain

$$\begin{aligned}
\sigma_y^2 \;&=\; W^2\sigma_S^2 \;+\; \frac{W^2 S_0^2}{F_0^2}\,\sigma_F^2 \;-\; 2\frac{W^2 S_0}{F_0}\,\sigma_{S,F} \\
\Rightarrow\; \sigma_y \;&=\; W\sigma_S\,\sqrt{1 \;+\; \left(\frac{S_0\sigma_F}{F_0\sigma_S}\right)^2 \;-\; 2\frac{S_0\sigma_{S,F}}{F_0\sigma_S^2}}
\end{aligned}$$

which implicitly defines B.

As a check, suppose that S_T and F_T are perfectly correlated. We then obtain (check) that

$$\sigma_y \;=\; W\sigma_S\left(1 \;-\; \frac{S_0\sigma_F}{F_0\sigma_S}\right)$$

which is not in general equal to 0! However, if F_t and S_t are scaled appropriately (alternatively we could scale h), then we can obtain a perfect hedge.

∎

---

3 Monte-Carlo Simulation

Monte-Carlo is a key computational tool in finance and so we review / introduce the key ideas here. Suppose we want to estimate $\theta := \mathbb{E}[h(\mathbf{X})]$ where $\mathbf{X} \in \mathbb{R}^n$ via Monte-Carlo simulation. We can do this by first simulating $\mathbf{X}_1, \ldots, \mathbf{X}_n$ IID from the distribution of $\mathbf{X}$ and then setting

$$\widehat{\theta}_n = \frac{h(\mathbf{X}_1) + \ldots + h(\mathbf{X}_n)}{n}.$$

The Strong Law of Large Numbers (SLLN) then implies

$$\widehat{\theta}_n \to \theta \;\text{ as } n \to \infty \text{ with probability (w.p.) } 1.$$

This is a very important result since it guarantees that if we take n sufficiently large then our estimator $\widehat{\theta}_n$ will get very close to θ.

3.1 How to Simulate Random Variables?

Many courses and textbooks on Monte-Carlo variables spend considerable time explaining how to simulate IID random variables from various distributions. Rather than dive into this sub-topic we will simply note that modern programming languages and software (including `R`, `Python` etc.) have functions that will do this for you. As a result, we will bypass this important aspect of Monte-Carlo and instead focus here on the construction of (approximate) confidence intervals for the quantity of interest θ.

3.2 Constructing a Confidence Interval for θ

But how large should n be so that we can have confidence in $\widehat{\theta}_n$ as an estimator of θ. Put another way, for a fixed value of n, what can we say about the quality of $\widehat{\theta}_n$? We can answer this question through the use of **confidence intervals** and to simplify our notation we will take $Y_i := h(\mathbf{X}_i)$.
Suppose then that we want to estimate θ and we have a random vector $\mathbf{Y} = (Y_1, \ldots, Y_n)$ whose distribution depends on θ. Then we seek $L(\mathbf{Y})$ and $U(\mathbf{Y})$ such that

$$\mathbf{P}\left(L(\mathbf{Y}) \leq \theta \leq U(\mathbf{Y})\right) = 1 - \alpha$$

where $0 \leq \alpha \leq 1$ is a pre-specified number. We then say that $[L(\mathbf{Y}), U(\mathbf{Y})]$ is a $100(1-\alpha)\%$ confidence interval for θ. Note that $[L(\mathbf{Y}), U(\mathbf{Y})]$ is a random interval. However, once we replace $\mathbf{Y}$ with a sample vector, $\mathbf{y}$, then $[L(\mathbf{y}), U(\mathbf{y})]$ becomes a real interval. We can use the Central Limit Theorem to obtain better estimates of $\mathbf{P}\left(|\widehat{\theta}_n - \theta| \geq k\right)$ and as a result, narrower confidence intervals for θ.

The Central Limit Theorem

The Central Limit Theorem is among the most important theorems in probability theory and we state it here for convenience with the symbol "$\stackrel{d}{\rightarrow}$" denoting convergence in distribution.

Theorem 1 (Central Limit Theorem)
Suppose $Y_1, \ldots, Y_n$ are IID and $\mathbb{E}[Y_i^2] < \infty$. Then

$$\frac{\widehat{\theta}_n - \theta}{\sigma/\sqrt{n}} \stackrel{d}{\rightarrow} \mathsf{N}(0,1) \quad \text{as} \quad n \rightarrow \infty$$

where $\widehat{\theta}_n = \sum_{i=1}^n Y_i/n$, $\theta := \mathbb{E}[Y_i]$ and $\sigma^2 := \mathrm{Var}(Y_i)$. $\square$

Note that we assume nothing about the distribution of the Y_i's in Theorem 1 other than that $\mathbb{E}[Y_i^2] < \infty$. If n is sufficiently large in our simulations, then we can use the CLT to construct approximate confidence intervals for $\theta := \mathbb{E}[Y]$. Specifically. let $z_{1-\alpha/2}$ be the the $(1-\alpha/2)$ percentile point of the $\mathsf{N}(0,1)$ distribution so that

$$\mathbf{P}(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha$$

when $Z \sim \mathsf{N}(0,1)$. Suppose now that we have simulated IID samples, Y_i, for $i = 1, \ldots, n$, and that we want to construct a $100(1-\alpha)\%$ CI for $\theta = \mathbb{E}[Y]$. That is, we want $L(\mathbf{Y})$ and $U(\mathbf{Y})$ such that

$$\mathbf{P}\left(L(\mathbf{Y}) \leq \theta \leq U(\mathbf{Y})\right) = 1 - \alpha.$$

The CLT implies $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)/\sigma$ is approximately $\mathsf{N}(0,1)$ for large n so we have

$$\begin{aligned}
\mathbf{P}\left(-z_{1-\alpha/2} \;\leq\; \frac{\sqrt{n}(\widehat{\theta}_n - \theta)}{\sigma} \;\leq\; z_{1-\alpha/2}\right) \;&\approx\; 1-\alpha \\
\Rightarrow\; \mathbf{P}\left(-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \;\leq\; \widehat{\theta}_n - \theta \;\leq\; z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \;&\approx\; 1-\alpha \\
\Rightarrow\; \mathbf{P}\left(\widehat{\theta}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \;\leq\; \theta \;\leq\; \widehat{\theta}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \;&\approx\; 1-\alpha.
\end{aligned}$$

Our approximate $100(1-\alpha)\%$ CI for θ is therefore given by

$$[L(\mathbf{Y}),\ U(\mathbf{Y})] \;=\; \left[\widehat{\theta}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \ \widehat{\theta}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]. \tag{2}$$

Recall that $\widehat{\theta}_n = (Y_1 + \ldots + Y_n)/n$, so L and U are indeed functions of $\mathbf{Y}$. There is still a problem, however, as we do not usually know σ^2. We resolve this issue by estimating σ^2 with

$$\widehat{\sigma}_n^2 = \frac{\sum_{i=1}^n (Y_i - \widehat{\theta}_n)^2}{n-1}.$$

It is easy to show that $\widehat{\sigma}_n^2$ is an unbiased estimator of σ^2 and that $\widehat{\sigma}_n^2 \rightarrow \sigma^2$ w.p. 1 as $n \rightarrow \infty$. So now we replace σ with $\widehat{\sigma}_n$ in (2) to obtain

$$[L(\mathbf{Y}),\ U(\mathbf{Y})] = \left[\widehat{\theta}_n - z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}},\ \ \widehat{\theta}_n + z_{1-\alpha/2}\frac{\widehat{\sigma}_n}{\sqrt{n}}\right] \tag{3}$$

as our *approximate* $100(1-\alpha)\%$ CI for θ when n is large. **This is the interval you can and should compute every time you run a Monte-Carlo simulation!**

Remark 2 *Note that when we obtain sample values of* $\mathbf{y} = (y_1, \ldots, y_n)$*, then* $[L(\mathbf{y}),\ U(\mathbf{y})]$ *becomes a real interval. Then we can no longer say (why not?) that*

$$\mathbf{P}\left(\theta \in [L(\mathbf{y}),\ U(\mathbf{y})]\right) = 1 - \alpha.$$

Instead, we say that we are $100(1-\alpha)\%$ confident *that* $[L(\mathbf{y}),\ U(\mathbf{y})]$ *contains* θ*. Furthermore, the smaller the value of* $U(\mathbf{y}) - L(\mathbf{y})$*, the more confidence we will have in our estimate of* θ*.*

Properties of the Confidence Interval

The *width* of the confidence interval is given by

$$U - L = \frac{2\widehat{\sigma}_n z_{1-\alpha/2}}{\sqrt{n}}$$

and so the *half-width* then is $(U-L)/2$. The width clearly depends on α, $\widehat{\sigma}_n$ and n. However, $\widehat{\sigma}_n \rightarrow \sigma$ w.p. 1 as $n \rightarrow \infty$, and σ is a constant. Therefore, for a fixed α, we need to increase n if we are to decrease the width of the confidence interval. Indeed, since $U - L \propto 1/\sqrt{n}$, we can see for example that we would need to increase n by a factor of four in order to decrease the width of the confidence interval by only a factor of two.

---

4 Brownian Motion and Geometric Brownian Motion

A key aspect of finance is modeling the dynamics of security prices, e.g. stock prices, currency prices, commodity prices etc. as well as interest rates, credit spreads etc. In order to do this we must work with stochastic processes. Loosely speaking a stochastic process is a set of random variables $\{X_t\}_{t\geq 0}$ with X_t representing the value of the process at time t and where time can be continuous or discrete. The security price S_t in the binomial model of Section 1.2, for example, provides an example of a discrete-time stochastic process. In this section we will introduce two very important continuous-time stochastic processes, namely Brownian and geometric Brownian motion.

4.1 Brownian Motion

Definition 1 *A stochastic process* X_t *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 random variables.

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 with probability 1.*

We say that X is a $B(\mu, \sigma)$ Brownian motion with drift μ and volatility σ. When $X_0 = 0$, $\mu = 0$ and $\sigma = 1$ we have a **standard** Brownian motion (SBM). We will use B_t to denote a SBM. Note that if $X \sim B(\mu, \sigma)$ and $X_0 = x$ then we can write

$$X_t = x + \mu t + \sigma B_t$$

where B is a SBM. We will usually write a $B(\mu, \sigma)$ Brownian motion in this way.

4.1.1 Simulating a Standard Brownian Motion

It is not possible to simulate an entire sample path of Brownian motion between 0 and T as this would require an infinite number of random variables. This is usually not a problem, however, since we often only wish to simulate the value of Brownian motion at certain fixed points in time. For example, we may wish to simulate B_{t_i} for $t_1 < t_2 < \ldots < t_n$, as opposed to simulating B_t for every $t \in [0, T]$.

Suppose then that we need to simulate B_{t_i} for $t_1 < t_2 < \ldots < t_n$ and for a fixed n. We will now see how to do this. The first observation we make is that

$$(B_{t_2} - B_{t_1}),\ (B_{t_3} - B_{t_2}), \ldots,\ \ (B_{t_n} - B_{t_{n-1}})$$

are mutually independent, and for $s > 0$, $B_{t+s} - B_t \sim \mathsf{N}(0, s)$. The idea then is as follows: we begin with $t_0 = 0$ and $B_{t_0} = 0$. We then generate B_{t_1} which we can do since $B_{t_1} \sim \mathsf{N}(0, t_1)$. We now generate B_{t_2} by first observing that $B_{t_2} = B_{t_1} + (B_{t_2} - B_{t_1})$. Then since $(B_{t_2} - B_{t_1}) \sim \mathsf{N}(0, t_2 - t_1)$ is independent of B_{t_1}, we can generate B_{t_2} by generating an $\mathsf{N}(0, t_2 - t_1)$ random variable and simply adding it to B_{t_1}. More generally, if we have already generated B_{t_i} then we can generate $B_{t_{i+1}}$ by generating an $\mathsf{N}(0, t_{i+1} - t_i)$ random variable and adding it to B_{t_i}. We have the following algorithm.

Simulating a Standard Brownian Motion

> **set** $t_0 = 0$, $B_{t_0} = 0$
> **for** $i = 1$ to n
>
>   **generate** $Z \sim \mathsf{N}(0, t_i - t_{i-1})$)
>   **set** $B_{t_i} = B_{t_{i-1}} + Z$

Remark 3 *It is very important that when you generate $B_{t_{i+1}}$, you do so* conditional *on the value of B_{t_i}. If you generate B_{t_i} and $B_{t_{i+1}}$ independently of one another then you are effectively simulating from* different *sample paths of the Brownian motion. This is not correct! In fact when we generate $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$ we are actually generating a* random vector *that does* not *consist of IID random variables.*

Simulating a $\mathbf{B}(\mu, \sigma)$ Brownian Motion

Suppose now that we want to simulate a $B(\mu, \sigma)$ BM X_t at the times $t_1, t_2, \ldots, t_{n-1}, t_n$. Then all we have to do is simulate an SBM, $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$, and use our earlier observation that $X_t = x + \mu t + \sigma B_t$.

4.1.2 Brownian Motion as a Model for Stock Prices?

There are a number of reasons why Brownian motion is **not** a good model for stock prices. They include

1. Since a $B(\mu, \sigma)$ BM can go negative this immediately suggests it is not a good model for stock prices which can never go negative.

2. The fact that people care about *returns*, not absolute prices so the IID increments property of BM should *not* hold for stock prices.

As a result, geometric Brownian Motion (GBM) is a much better model for stock prices.

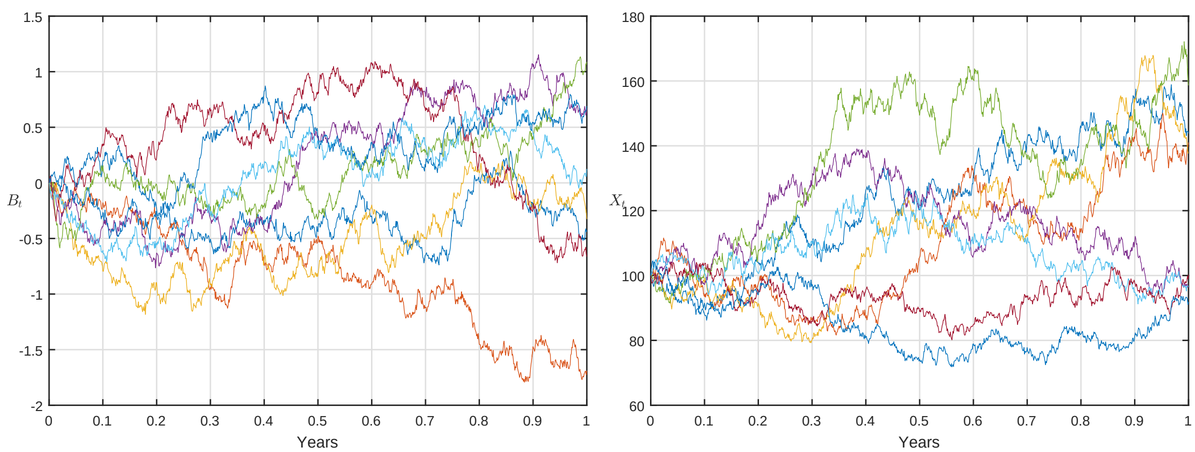
4.2 Geometric Brownian Motion

Definition 2 *A stochastic process X_t is a (μ, σ) geometric Brownian motion (GBM) if* $\log(X) \sim B(\mu - \sigma^2/2,\ \sigma)$. *We write $X \sim GBM(\mu, \sigma)$.*

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

1. Fix $t_1, t_2, \ldots, t_n$. Then $\frac{X_{t_2}}{X_{t_1}}, \frac{X_{t_3}}{X_{t_3}}, \frac{X_{t_n}}{X_{t_{n-1}}}$ are mutually independent.
2. For $s > 0$, $\log\left(\frac{X_{t+s}}{X_t}\right) \sim \mathsf{N}\left((\mu - \sigma^2/2)s,\ \sigma^2 s\right)$.
3. X_t is continuous with probability 1.

Again, we call μ the drift and σ the volatility. If $X \sim GBM(\mu, \sigma)$, then $\log(X_t)$ has a normal distribution. In Figure 1 we have plotted some sample paths of Brownian and geometric Brownian motions.



(a) Brownian motion (b) Geometric Brownian motion

Figure 1: Sample paths of Brownian motion, B_t, and geometric Brownian motion (GBM), $X_t = X_0\, e^{(\mu-\sigma^2/2)t+\sigma B_t}$. Parameters for the GBM were $X_0 = 100$, $\mu = 10\%$ and $\sigma = 30\%$.

Question: How would you simulate a sample path of $GBM(\mu, \sigma^2)$ at the fixed times $0 < t_1 < t_2 < \ldots < t_n$?

Answer: Simulate $\log(X_{t_i})$ first and then take exponentials!

4.2.1 Modelling Stock Prices as Geometric Brownian Motion

Suppose $X \sim \mathsf{GBM}(\mu, \sigma)$. Note the following:

1. If $X_t > 0$, then X_{t+s} is always positive for any $s > 0$ so limited liability is not violated.
2. The distribution of $\frac{X_{t+s}}{X_t}$ only depends on s so the distribution of *returns* from one period to the next only depends on the length of the period.

This suggests that GBM might be a reasonable model for stock prices. In fact, we will often model stock prices as GBM's and we will use the following notation:

- S_0 is the known stock price at $t = 0$
- S_t is the random stock price at time t and

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

where B is a standard BM. The drift is μ, σ is the volatility and S is a therefore a $GBM(\mu, \sigma)$ process that begins at S_0.

Exercise 2 *Show* $S_{t+\Delta t} = S_t e^{(\mu - \sigma^2/2)\Delta t + \sigma(B_{t+\Delta t} - B_t)}$.

Suppose now that we wish to simulate $S \sim GBM(\mu, \sigma)$. Then we know

$$S_{t+\Delta t} = S_t e^{(\mu - \sigma^2/2)\Delta t + \sigma(B_{t+\Delta t} - B_t)}$$

so that we can simulate $S_{t+\Delta t}$ conditional on S_t for any $\Delta t > 0$ by simply simulating an $\mathsf{N}(0, \Delta t)$ random variable.

4.3 Pricing Options via Monte-Carlo

As mentioned earlier, Monte-Carlo is a key computational tool in finance. One of its many applications is in the pricing of derivative securities. While the topic of derivative security pricing will be discussed later in the course, we can see here how Monte-Carlo can be used to do this.

Example 4 (Pricing a European Call Option)
Suppose we want to estimate the price C_0 of a call option on a non-dividend-paying stock whose price process S_t is a $GBM(\mu, \sigma)$. The theory of derivatives pricing tells us that

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

where Q denotes the so-called risk-neutral[6] probability distribution and where $S_t \sim GBM(r, \sigma)$ under Q. That is, we assume $S_T = S_0 \exp\big((r - \sigma^2/2)T + \sigma Z\big)$ where $Z \sim \mathsf{N}(0, T)$. The relevant parameters are $r = .05$, $T = 0.5$ years, $S_0 = \$100$, $\sigma = 0.2$ and strike $K = \$110$.

Though we can compute C_0 analytically, we can also estimate C_0 using Monte Carlo with (3) yielding an approximate $100(1-\alpha)\%$ CI for C_0 with $Y_i := e^{-rT} \max(S_T^{(i)} - K, 0)$ denoting the i^{th} discounted sample payoff of the option. Based on $n = 100$k samples, we obtain $[15.16,\ 15.32]$ as our approximate 95% CI for C_0. ∎

4.4 An Aside on the Tower Property of Conditional Expectations

Let I_t denote all the information that is available in a given probabilistic model up to and including time t. Then for any random variable X, $\mathrm{E}_t[X] := \mathrm{E}[X \mid I_t]$ denotes the expected value of X conditional on I_t. Note that when viewed from any time $u \leq t$, $\mathrm{E}_t[X]$ is also a random variable. Consider for example $X = S_v$ where $S \sim GBM(\mu, \sigma)$ and $v \geq t$. Then

$$\begin{aligned}
\mathrm{E}_t[S_v] &= \mathrm{E}_t\left[S_t e^{(\mu-\sigma^2/2)(v-t) + \sigma(B_v - B_t)}\right] && \text{by Exercise 2 with } \Delta t = v - t \\
&= S_t e^{(\mu-\sigma^2/2)(v-t)}\, \mathrm{E}_t\left[e^{\sigma(B_v - B_t)}\right] \\
&= S_t e^{\mu(v-t)} && (4)
\end{aligned}$$

where the second equality follows because S_t is known at time t and the third equality follows from the moment generating function[7] (MGF) of a normal random variable. Now note that $\mathrm{E}_t[S_v] = S_t e^{\mu(v-t)}$ is itself a random

[6] We will discuss derivatives pricing, i.e. , i.e. risk-neutral pricing, later in the course.

[7] If $Z \sim \mathrm{N}(a, b^2)$ then $\mathrm{E}[e^{cZ}] = e^{ca + \frac{1}{2}c^2 b^2}$ for any c. This is the MGF of the normal distribution.

variable when viewed from time $u \leq t$ and so we can take its expectation conditional on I_u. We obtain

$$\begin{aligned}
\mathrm{E}_u[\mathrm{E}_t[S_v]] &= \mathrm{E}_u[S_t e^{\mu(v-t)}] \qquad \text{by (4)} \\
&= e^{\mu(v-t)}\,\mathrm{E}_u[S_u e^{(\mu-\sigma^2/2)(t-u)+\sigma(B_t-B_u)}] \qquad \text{by Exercise 2} \\
&= S_u\, e^{\mu(v-u)} e^{-\sigma^2(t-u)/2}\,\mathrm{E}_u[e^{\sigma(B_t-B_u)}] \\
&= S_u\, e^{\mu(v-u)} \qquad \text{by MGF of a normal r.var} \\
&= \mathrm{E}_u[S_v] \qquad \text{by (4).}
\end{aligned}$$

We have therefore shown $\mathrm{E}_u[\mathrm{E}_t[S_v]] = \mathrm{E}_u[S_v]$ for any $u \leq t$. This is not an accident and indeed it can be shown that

$$\mathrm{E}_u[\mathrm{E}_t[X]] = \mathrm{E}_u[X]$$

for any random variable X whenever $u \leq t$. This is an example of the so-called **tower property of conditional expectations**. It will prove useful later in the course when we price options in the binomial model as well as futures.