2 The binomial tree methods (2 lectures)

The analysis made on option pricing in the previous lecture is solely based in the assumption that the market is arbitrage-free, without referring to any price model of the underlying asset. Without a price model of the underlying asset, only qualitative discussions on option pricing are possible. Quantitative pricing of the derivatives requires specific model on price movement of the underlying asset.

In 1900, Louis Bachelier, a young French mathematician, completed a thesis called "the theory of speculation". Bachelier developed the now universally used concept of stochastic process and proposed a model on how the stock price change with respect to time. This enable him to make the first theoretical attempt to value options [7].

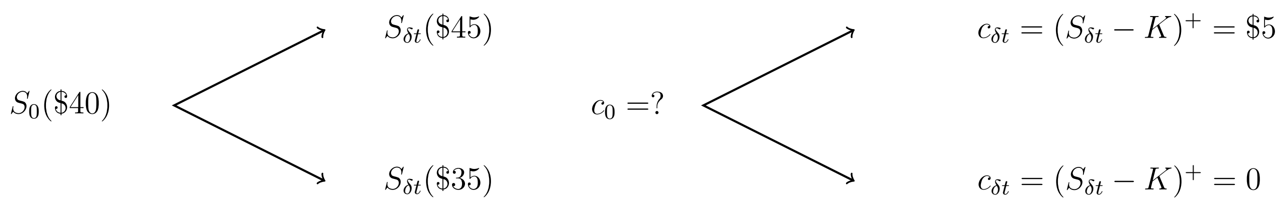
Bachelier's thesis was not well received. The pioneering nature of his work was recognized only after several decades, first by Kolmogorov[8] who pointed out his work to Lévy, then by L. J. Savage who brought the work of Bachelier to the attention of Paul Samuelson in 1950s (no economist at the time had ever heard of Bachelier).

An intensive period of development in financial economics followed, first at MIT, where Samuelson worked, leading to the Nobel Prize-winning solution of the option pricing problem by Fischer Black, Myron Scholes and Robert Merton in 1973. In the same year, the world's first listed options exchange opened its doors in Chicago.

In 1979, Cox, Ross and Rubinstein introduced the N-period binomial tree model for stock price and N-period binomial method for option pricing [9]. They are easy to understand and implement in practice. We will also show that resulting option price converges to the Black-Scholes-Merton formula for option pricing as $N \to \infty$.

We will start with the 1-period binomial tree model for the stock price. Then we move to N-period model, and finally the geometric Brownian motion model proposed by Samuelson and adapted by Black-Scholes-Merton.

2.1 An example



[7] A 1688 treatise on the workings of the Amsterdam stock exchange (established in 1602 by the Dutch East India Company) reveals that options were already dominating trading activities at the time. Amsterdam was the most sophisticated and important financial center of the seventeenth century.

[8] Born in 1903. In 1933, Kolmogorov published his book, Foundations of the Theory of Probability, laying the modern axiomatic foundations of probability theory.

[9] J. C. Cox, S. A. Ross, and M. Rubinstein, Option pricing: A simplified approach, Journal of Financial Economics 7 (1979) 229–263.

Let the price of a stock be \$40 at $t = 0$, and suppose a month later ($t = \delta t$) the stock price will either goes up to \$45 or down to \$35. Now consider buying a European call option of the stock at $t = 0$ with strike price \$40 and 1 month maturity. If the risk-free annual interest rate is 12%, how much should the price for the call option be?

Now, construct two portfolios

$$\Phi_1 = S - 2c, \qquad \Phi_2 = B \text{ with } B_0 = \frac{35}{e^{0.01}} = 34.652.$$

At $t = \delta t$,

$$V_{\delta t}(\Phi_1) = S_{\delta t} - 2(S_{\delta t} - K)^+ = \begin{cases} 45 - 2 \times (45 - 40)^+ = 35 & \text{if } S \text{ goes up} \\ 35 - 2 \times (35 - 40)^+ = 35 & \text{if } S \text{ goes up} \end{cases}.$$

$$V_{\delta t}(\Phi_2) = \frac{35}{e^{0.01}} e^{0.12 \times \frac{1}{12}} = 35$$

$V_{\delta t}(\Phi_1) = V_{\delta t}(\Phi_2)$. Hence

$$V_0(\Phi_1) = V_0(\Phi_2).$$

So, $S_0 - 2c_0 = B_0$ which leads to $c_0 = \frac{40-34.652}{2} \approx \$2.67$. We conclude that the investor should pay \$2.67 for the stock option.

This example reveals the idea of hedging: it is possible to construct a risk-free investment portfolio Φ with c and its underlying asset S.

2.2 The one period binomial tree method

Now we give a more systematic analysis.

- Let T be the length of the time interval we are considering.
- Let r denote the risk-free interest rate and let $\rho = e^{rT}$ so that \$1 deposit in the bank becomes \$$\rho$ after T units of time. Here we assume that the interest is compounded continuously. Otherwise, we can change the formula of ρ accordingly.
- Assume that the underlying Stock S, whose value is S_0 at $t = 0$, can have two values $S^u = uS_0$ and $S^d = dS_0$ at time T with $d < u$. Under the no arbitrage principle, we must have

  $$d < \rho < u. \tag{2.1}$$

  This is because if $\rho \leq d$ then no investor would deposit his/her money in the bank (or buy treasury bills), and if $u \leq \rho$ then no investor would invest in the stock market.
- Option price $ⓟ_t$ has two values $ⓟ^u$ and $ⓟ^d$ at time T depending on whether S goes up or goes down [10].

[10] If T is the length between now and the expiration date, and if it is a call option, then $ⓟ^u = \max(uS_0 - K, 0)$ and $ⓟ^d = \max(dS_0 - K, 0)$. If it is a put option, then $ⓟ^u = \max(K - uS_0, 0)$ and $ⓟ^d = \max(K - dS_0, 0)$.

Now, suppose we know S_0, u, d, $\textcircled{p}^u$, $\textcircled{p}^d$, ρ, and we want to determine $\textcircled{p}_0$, which is the option price at $t = 0$.

To do that, let us build a portfolio Φ as follows: buy an option $\textcircled{p}$, short sell Δ shares of the underlying stock S[11]:

$$\Phi = \textcircled{p} - \Delta S. \tag{2.2}$$

Δ is determined so that Φ takes the same value at T no matter whether S is going up or going down. Hence we find Δ by solving

$$\textcircled{p}^u - \Delta S_0 u = \textcircled{p}^d - \Delta S_0 d \qquad \Rightarrow \qquad \Delta = \frac{\textcircled{p}^u - \textcircled{p}^d}{S_0(u-d)}. \tag{2.3}$$

Hence $V_T(\Phi) = \textcircled{p}^u - \frac{\textcircled{p}^u - \textcircled{p}^d}{u-d} u = \frac{-d}{u-d}\textcircled{p}^u + \frac{u}{u-d}\textcircled{p}^d$.

Suppose there is a bank deposit B_0 made at $t = 0$. We choose B_0 so that

$$V_T(\Phi) = V_T(B) = \rho B_0.$$

Thus

$$B_0 = \frac{1}{\rho}\left(\frac{-d}{u-d}\textcircled{p}^u + \frac{u}{u-d}\textcircled{p}^d\right). \tag{2.4}$$

By Corollary 1.30, $V_0(\Phi) = V_0(B)$, which means $\textcircled{p}_0 - \Delta S_0 = B_0$, i.e.,

$$\textcircled{p}_0 = B_0 + \Delta S_0 = \frac{1}{\rho}\left(\frac{\rho - d}{u-d}\textcircled{p}^u + \frac{u-\rho}{u-d}\textcircled{p}^d\right) \overset{\text{def}}{=} \frac{1}{\rho}\left(q_u \textcircled{p}^u + q_d \textcircled{p}^d\right). \tag{2.5}$$

Please note that $q_u + q_d = 1$. Hence (2.5) says that the option price is a weighted average of the value of the call option in the up state and the down state. Please also note that the definitions of q_u and q_d imply that

$$\frac{1}{\rho}\left(q_u S^u + q_d S^d\right) = \frac{1}{\rho}\left(\frac{\rho - d}{u-d}u + \frac{u-\rho}{u-d}d\right) S_0 = S_0. \tag{2.6}$$

Example 2.1 *For the example at the beginning of this section, we get* $q_u = \frac{e^{0.01} - \frac{35}{40}}{\frac{10}{40}}$ *and* $q_d = \frac{\frac{45}{40} - e^{0.01}}{\frac{10}{40}}$. $\textcircled{p}_0 = \frac{1 - \frac{35}{40}e^{-0.01}}{\frac{10}{40}} 5 = \$2.67$.

The formula (2.5) is astonishing. It is astonishing because what it says is that as long as we agree on the value of u and d, even if we disagree on the probability of how the stock is going up or down tomorrow, yet we still are going to agree on what the value of a call option on that stock is.

For example, you may think that the price is more likely to go up, therefore, you want to have that kind of a call option bet, but I may think that the price is more likely to go down, so I'm happy to sell it to you. As long as we agree on the value of u and d, we will agree on the price of the option. That's what drives the market.

[11] A negative Δ from (2.3) implies that we should buy.

2.3 The multiperiod binomial tree method

In the previous one period model, we assume that after say, 5 minutes, or 5 days, or 5 months, the stock price can either be uS_0 or vS_0. If we can't agree on that there are only two prices, let's agree that between now and 5 minutes from now, there are 100 possible outcomes for the stock price. Do you agree on that? If you agree, then all we need to do is to have enough steps between now and 5 minutes from now to have 100 possible prices.

The multiperiod method is simply that we now have a bunch of possibilities, and you are figuring out what the price of the option is at date 0 when it pays off at date N.

Given an option with expiration date T, we consider a multiperiod binomial tree of length N, obtained by stringing together single period binomial trees, where the length of the time interval for each single period binomial tree is $\delta t = T/N$.

See Figure 2.1 for an illustration.

Image we toss a coin repeatedly (we do not need to worry whether the coin is fair or biased). Whenever we get a head (H), the stock price moves up by a factor u, and whenever we get a tail (T), the stock price moves down by a factor d. In addition to this stock, there is a money market asset with an interest rate r. We assume $d < \rho = e^{r\delta t} < u$ as in (2.1).

Let S_0 be the initial stock price. We denote the stock price at time δt by $S_1(H) = uS_0$ if the first toss result in head, and by $S_1(T) = dS_0$ if the first toss result in tail. After the second toss, the stock price will be one of

$$S_2(HH) = uS_1(H) = u^2 S_0, \quad S_2(HT) = dS_1(H) = duS_0,$$
$$S_2(TH) = uS_1(T) = udS_0, \quad S_2(TT) = dS_1(T) = d^2 S_0.$$

After 3 tosses, there are 8 possible coin sequences, and 4 different stock price at $t = 3\delta t$. This process will repeat until we reach $t = N\delta t = T$. See Figure 2.1 for an illustration.

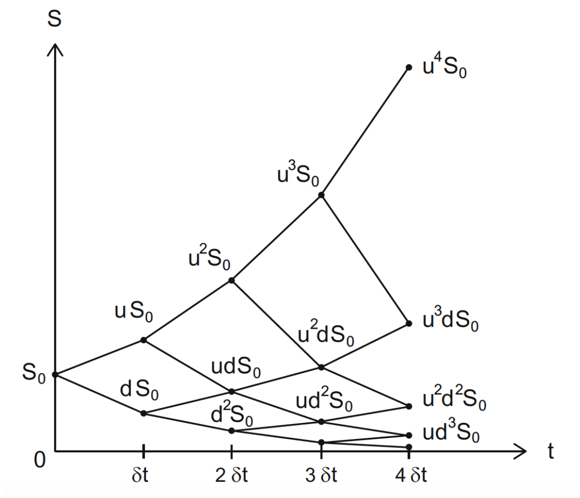


Figure 2.1: A multiperiod binomial tree of length $N = 4$ with $u = 3/2$ and $d = 1/2$. At time $t = N\delta t$, the stock price can take one of $N + 1$ possible values.

To price an option, we start from the nodes at $t = N\delta t$ and travel backward, repeatedly using the formula (2.5)

$$\textcircled{p}_0 = \frac{1}{\rho}\left(q_u \textcircled{p}^u + q_d \textcircled{p}^d\right)$$

or more precisely, using our current notation,

$$\textcircled{p}_n(\omega_1\omega_2\cdots\omega_n) = \frac{1}{\rho}\left(q_u \textcircled{p}_{n+1}(\omega_1\omega_2\cdots\omega_n H) + q_d \textcircled{p}_{n+1}(\omega_1\omega_2\cdots\omega_n T)\right). \tag{2.7}$$

Here $\rho \overset{\text{def}}{=} e^{r\delta t}$, ω_i is either H or T, $\textcircled{p}_n(\omega_1\omega_2\cdots\omega_n)$ denotes the option price at $t = n\delta t$ after observing the head-tail sequence $(\omega_1\omega_2\cdots\omega_n)$ (i.e., knowing the underlying stock price up to $t = n\delta t$).

Example 2.2 *We now compute the price of a European call option with the strike price $K = 30$ and $S_0 = 32$ using a three-step binomial tree. $u = \frac{3}{2}$ and $d = \frac{1}{2}$. For the sake of computational convenience, we assume $r = 0$ (i.e., $\rho = 1$). Note that $q_u = \frac{\rho-d}{u-d} = \frac{1}{2}$ and $q_d = \frac{1}{2}$. The solution is shown by Figures 2.2 to 2.4.*

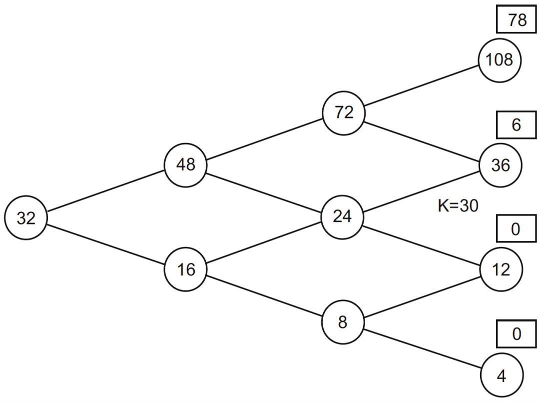


Figure 2.2: Payoff of a European call option in a binomial tree of length $N = 3$. $K = 30$. $\boxed{\text{option price}}$. stock price

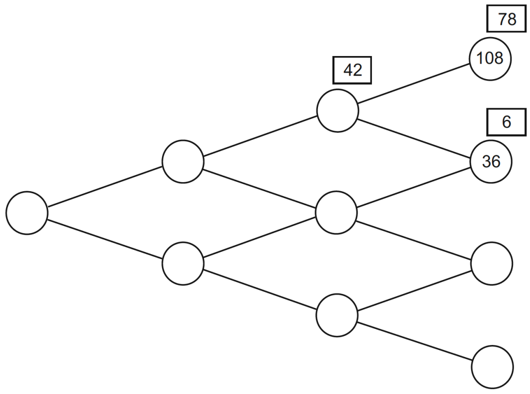


Figure 2.3: The first step in pricing of a European call option. $K = 30$. $q_u = \frac{1}{2} = q_d$. $\rho = 1$. $\boxed{\text{option price}}$. stock price . The number 42 in $\boxed{42}$ is derived using $\textcircled{p}_0 = \frac{1}{\rho}\left(q_u \textcircled{p}^u + q_d \textcircled{p}^d\right)$ in (2.5) (or (2.7)) with $\textcircled{p}^u = 78$, $\textcircled{p}^d = 6$.

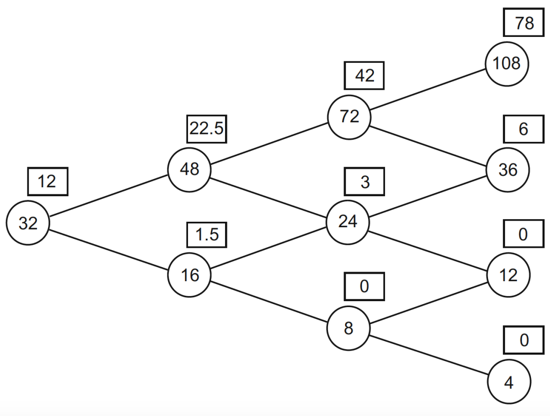


Figure 2.4: Pricing of a European call option in a binomial tree of length $N = 3$. $K = 30$. The result is 12. $\boxed{\text{option price}}$. stock price

Now, we want to derive the general formula for the option price. It means that we want to derive a formula of $ⓟ_0$ in terms of $\{ⓟ_N(\omega_1\omega_2\cdots\omega_N), \omega_i \in \{H,T\}\}$ using the backward recursive relation

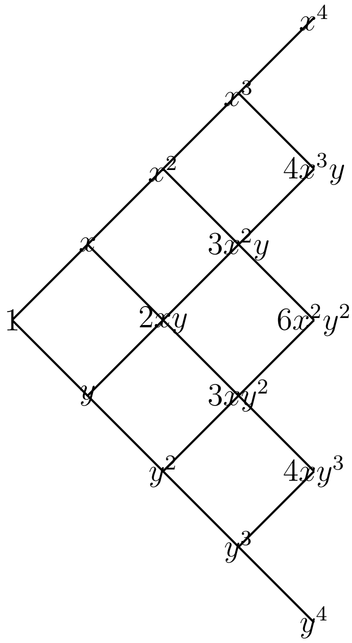
$$ⓟ_n(\omega_1\omega_2\cdots\omega_n) = \frac{1}{\rho}\left(q_u ⓟ_{n+1}(\omega_1\omega_2\cdots\omega_n H) + q_d ⓟ_{n+1}(\omega_1\omega_2\cdots\omega_n T)\right). \tag{2.8}$$

Obviously, each $ⓟ_N(\omega_1\omega_2\cdots\omega_N)$ contributes to $ⓟ_0$. By (2.8), if $(\omega_1\omega_2\cdots\omega_N)$ contains a H, it contributes $\frac{q_u}{\rho}$. If it contains a T, it contributes $\frac{q_d}{\rho}$. We should keeping in mind that, say, $ⓟ_4(HTTT) = ⓟ_4(THTT) = \cdots = ⓟ_4(TTTH) = \max(ud^3S_0 - K, 0)$ if we are considering a 4 period European call option.

Recall $\binom{N}{j} = \frac{N!}{(N-j)!j!}$ gives the number of ways, disregarding order, that j objects can be chosen from among N objects. Using it, we can have the binomial expansion formula

$$(x+y)^N = \underbrace{(x+y)\,(x+y)\,(x+y)\,\cdots(x+y)}_{N\text{terms}} = \sum_{j=0}^{N}\binom{N}{j}x^j y^{N-j} \tag{2.9}$$

and the corresponding Pascal's triangle



n \ j	0	1	2	3	4	5	6	…
0	1							
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
5	1	5	10	10	5	1		
6	1	6	15	20	15	6	1	…
⋮			⋮					

Figure 2.5: Pascal's triangle

The stock price paths, moving from left to right, containing $N-j$ down's and j up's, always go to the same terminal point. The number of such paths is $\binom{N}{j}$. Hence we find that the contribution of $ⓟ_N$(ω contains j H's and $N-j$ T's) to $ⓟ_0$ is precisely $\frac{1}{\rho^N}\binom{N}{j}q_u^j q_d^{N-j}$. So we have proved the following theorem (which can also be proved by induction)

Theorem 2.1 *The price of an European call option with strike price K, expiration date T, interests rate r, and stock price S_0 at $t=0$, is*

$$\begin{aligned}\textcircled{p}_0 =& \frac{1}{\rho^N}\sum_{i=0}^{N}\binom{N}{j}q_u^j q_d^{N-j}\textcircled{p}_N(\omega \textit{ contains } j \ H\textit{'s and } N-j \ T\textit{'s}) \\ =& e^{-rT}\sum_{j=0}^{N}\binom{N}{j}q_u^j q_d^{N-j}\left(S_0u^jd^{N-j}-K\right)^+. \end{aligned} \tag{2.10}$$

2.4 Convergence to the Black-Scholes-Merton formula

In this section we show that the European call option price (2.10) converges to the solution from the Black-Scholes-Merton partial differential equation (which will be discussed later) as $N\to\infty$.

Note that as $N\to\infty$, $\delta t = T/N \to 0$. In (2.10), we need to decide when $\left(S_0u^id^{N-i}-K\right)^+ \neq 0$.

So, let m be the smallest integer such that

$$S_0u^md^{N-m} > K.$$

Then

$$\textcircled{p}_0 = e^{-rT}\sum_{i=m}^{N}\binom{N}{i}q_u^i q_d^{N-i}\left(S_0u^id^{N-i}-K\right).$$

Let $A = e^{-rT}\sum_{i=m}^{N}\binom{N}{i}q_u^iq_d^{N-i}u^id^{N-i}$ and $B=\sum_{i=m}^{N}\binom{N}{i}q_u^iq_d^{N-i}$. Then

$$\textcircled{p}_0 = AS_0 - Ke^{-rT}B.$$

One can then manage to prove that (see Page 204 of "Option markets" by of Cox and Rubinstein or Section 14.4 the book of G. H. Choe)

$$\lim_{N\to\infty}\textcircled{p}_0 = N(d_1)S_0 - Ke^{-rT}N(d_2), \tag{2.11}$$

where $N(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy = \int_{-x}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$,

$$d_1 = \lim_{N\to\infty}\frac{\frac{\log\frac{S_0}{K}+N\log d}{\log\frac{u}{d}}+Nq^*}{\sqrt{Nq^*(1-q^*)}} \qquad d_2 = \lim_{N\to\infty}\frac{\frac{\log\frac{S_0}{K}+N\log d}{\log\frac{u}{d}}+Nq_u}{\sqrt{Nq_u(1-q_u)}} \tag{2.12}$$

with $q^* = e^{-r\delta t}q_uu$.

If we set

$$u = e^{\sigma\sqrt{\delta t}}, \qquad d = e^{-\sigma\sqrt{\delta t}},$$

Then

$$d_1 = \frac{\log\frac{S_0}{K} + (r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \qquad d_2 = \frac{\log\frac{S_0}{K} + (r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \tag{2.13}$$

The right hand side of (2.11) is then exactly the Black-Scholes-Merton (1973) option pricing formula.

Remark: For your information, $u, d = e^{\pm\sigma\sqrt{\delta t}}$ and $q_u = \frac{e^{r\delta t}-d}{u-1}$ is proposed by Cox, Ross and Rubinstein in their 1979 paper. Another option is to choose $u = e^{(r-\frac{1}{2}\sigma^2)\delta t+\sigma\sqrt{\delta t}}$, $d = e^{(r-\frac{1}{2}\sigma^2)\delta t-\sigma\sqrt{\delta t}}$, $q_u = 1/2$ is proposed by Jarrow and Rudd in 1980 [12]. In both case, the σ is the volatility of the stock price which will be discussed later.

2.4.1 Limit distribution of the stock price in the risk-neutral world as $\delta t \to 0$ [13]

In the multiperiod binomial tree model, the stock price satisfies $S_n = S_0 u^i d^{n-i}$ if the first n coin toss contains i heads and $n-i$ tails. Here we take $\delta t = 1/N$ (i.e., N steps per unit time), $t_n = n\delta t$, $u = e^{\sigma\sqrt{\delta t}}$, $d = e^{-\sigma\sqrt{dt}}$. I like to explain why as $N \to \infty$, S_{Nt} converges to the distribution of

$$S(t) = S(0)e^{\sigma W(t)+(r-\frac{1}{2}\sigma^2)t} \tag{2.14}$$

with $W(t) \sim N(0,t)$ [14] , $u = e^{\sigma\sqrt{\delta t}}$, $d = e^{-\sigma\sqrt{\delta t}}$. By Taylor expansion, $q_u = \frac{e^{r\delta t}-e^{-\sigma\sqrt{\delta t}}}{e^{\sigma\sqrt{\delta t}}-e^{-\sigma\sqrt{\delta t}}} \approx \frac{1+r\delta t-(1-\sigma\sqrt{\delta t}+\frac{1}{2}\sigma^2\delta t)}{2\sigma\sqrt{\delta t}} = \frac{(r-\frac{1}{2}\sigma^2)\delta t}{2\sigma\sqrt{\delta t}} + \frac{1}{2}$. From $S_n = S_0u^id^{n-i} = S_0e^{\xi_1}e^{\xi_2}\cdots e^{\xi_n}$, we see that

$$\log(S_{Nt}/S_0) = i\log u + (Nt-i)\log d = \sum_{k=1}^{Nt}\xi_k.$$

Here i is a binomial distributed variable $\sim B(Nt, q_u)$ (i.e. among the Nt trials, one obtains i heads with q_u probability to get a head in each trial). $e^{\xi_k} = u$ if the kth coin toss is a head. $e^{\xi_k} = d$ if the kth coin toss is a tail. In other words, $\mathbb{P}(\xi_k = \log u) = q_u$, $\mathbb{P}(\xi_k = \log d) = q_d$.

$$\begin{aligned}\mathbb{E}\xi_k =& q_u\log u + q_d\log d \approx \left(\frac{(r-\frac{1}{2}\sigma^2)\delta t}{2\sigma\sqrt{\delta t}} + \frac{1}{2}\right)\sigma\sqrt{\delta t} + \left(-\frac{(r-\frac{1}{2}\sigma^2)\delta t}{2\sigma\sqrt{\delta t}} + \frac{1}{2}\right)(-\sigma\sqrt{\delta t}) \\ =& (r-\frac{1}{2}\sigma^2)\delta t.\end{aligned}$$

[12] R. Jarrow and A. Rudd, Approximate option valuation for arbitrary stochastic process, Journal of Financial Economics 10 (1982) 347–369.

[13] For your information only. It won't be tested. Please skip in the first reading if you have not learned probability yet.

[14] It is not a fully rigorous proof as I use $\approx$ instead of $=$ in the calculation. But it is enough for you to get the idea. Please pay attention to where $r - \frac{1}{2}\sigma^2$ comes from. More rigorous proofs can be found in Shreve II, Section 3.2.7 for $r = 0$ and Exercise 3.8 for general r.

Denote $\mathbb{E}\xi_k \stackrel{\text{def}}{=} a$. Then

$$\begin{aligned}
\mathrm{Var}(\xi_k) =& \mathbb{E}(\xi_k^2) - (\mathbb{E}\xi_k)^2 \\
\approx& \left(\frac{(r-\frac{1}{2}\sigma^2)\delta t}{2\sigma\sqrt{\delta t}} + \frac{1}{2}\right)\sigma^2\delta t + \left(-\frac{(r-\frac{1}{2}\sigma^2)\delta t}{2\sigma\sqrt{\delta t}} + \frac{1}{2}\right)\sigma^2\delta t - a^2 \\
=& \sigma^2\delta t + \text{ higher order terms w.r.t. } \delta t.
\end{aligned}$$

By the central limit theorem[15], we know $\frac{\sum_{k=1}^{Nt}(\xi_k - \mathbb{E}\xi_k)}{\sqrt{Nt}\sqrt{\mathrm{Var}(\xi_k)}} \to N(0,1)$, or

$$\frac{\sum_{k=1}^{Nt}\left(\xi_k - (r-\frac{1}{2}\sigma^2)\delta t\right)}{\sqrt{Nt}\sigma\sqrt{\delta t}} \to W(1)$$

where $W(1) \sim N(0,1)$. Note that $N\delta t = 1$. $Nt\delta t = t$. Hence

$$\frac{\left(\sum_{k=1}^{Nt}\xi_k\right) - (r-\frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}} \to W(1).$$

Since $\sqrt{t}W(1) \sim N(0,t)$ and can be denoted as $W(t)$,

$$\log(S_{Nt}/S_0) = \sum_{k=1}^{Nt}\xi_k \to \sigma W(t) + (r - \frac{1}{2}\sigma^2)t.$$

This proves (2.14). The resulting $S(t)$ is said to satisfy log-normal distribution (which means that its log has normal distribution).

2.5 Computer experiments

The Matlab code in this section is taken from Choe's book but with slight modification.

Example 2.3 *Now consider buying a European call option of the stock at $t = 0$ with strike price \$110 and 1 year maturity. If the risk-free annual interest rate is 5% and $u = e^{\sigma\sqrt{dt}}$ and $d = e^{-\sigma\sqrt{dt}}$ with $\sigma = 0.3$, how much should the price for the call option be?*

```
S0 = 100;
K = 110;
T = 1;
r = 0.05;
sigma = 0.3;
M = 10; % number of time steps
```

[15]Suppose $\{X_1, \cdots\}$ is a sequence of independent identically distributed random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$. Then as $n \to \infty$, the random variable $\frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{n}\sigma}$ converge in the distribution to a standard normal random variable $N(0,1)$.

```
dt = T/M;
u = exp(sigma*sqrt(dt));
d = exp(-sigma*sqrt(dt));
q = (exp(r*dt)-d)/(u-d);
% We compute asset prices.
for i = 0:1:M
    fprintf('time = %i\n', i)
    S = S0*u.^([i:-1:0]).*d.^([0:1:i])
end
fprintf('Payoff at expiry\n')
Call = max(S0*u.^([M:-1:0]).*d.^([0:1:M])- K,0)
% We proceed backward to compute option value at time 0.
for i = M:-1:1
    fprintf('time = %i\n', i-1)
    Call = exp(-r*dt)*(q*Call(1:i) + (1-q)*Call(2:i+1))
end
```

If we increase M, the option priced obtained by the above code converges to the result given by the classic Black-Scholes-Merton formula.

```
S0 = 100;
K = 110;
T = 1;
r = 0.05;
sigma = 0.3;

M_values = [50:1:1000];
Call_prices = zeros(length(M_values),1);

for j = 1:length(M_values)
    M = M_values(j);
    dt = T/M;
    u = exp(sigma*sqrt(dt));
    d = exp(-sigma*sqrt(dt));
    q = (exp(r*dt)-d)/(u-d);
    Call = max(S0 *u.^([M:-1:0]) .* d.^([0:1:M]) - K,0);

    for i = M:-1:1
        Call = exp(-r*dt)*(q*Call(1:i)+(1-q)*Call(2:i+1));
    end

    Call_prices(j) = Call;
end
```

```
plot(M_values,Call_prices,'.');
hold on;
% the Black-Scholes-Merton formula.
d1 = (log(S0/K)+(r+0.5*sigma^2)*T)/(sigma*sqrt(T));
d2 = d1-sigma*sqrt(T);
Call_BSM = S0*normcdf(d1) - K*exp(-r*T)*normcdf(d2);
x=0:1000;
plot(x,Call_BSM*ones(size(x)),'r');
```

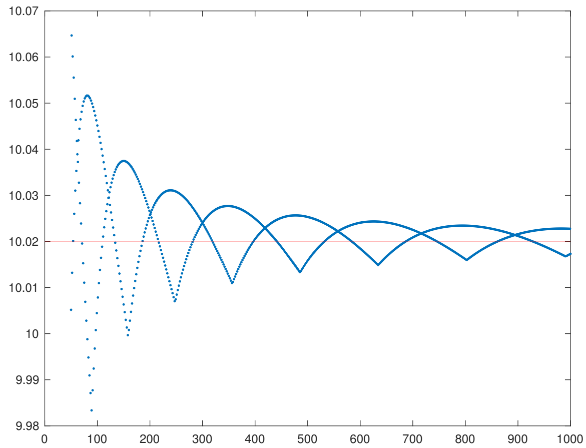


Figure 2.6: Convergence to the Black-Scholes-Merton price.

2.6 Finite probability space

So far, we have used little probability. Let us now put what we have discussed so far into the standard probability framework. Besides Shreve I and II which we are going to follow, another standard textbook for probability is "A First Course in Probability" by Sheldon Ross.

Recall that in our multiperiod binomial model, at each $t_n = n\delta t$, a coin is tossed, and the outcome of the coin toss determine how S_{n-1} should change to S_n. We are interested in the stock price S_n. But there is a correspondence between the stock price and the sequence of the coin tosses. Let Ω be the set of all possible outcomes of the coin tosses. For example, if we toss the coin three times, the set of all possible outcome is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}. \tag{2.15}$$

Suppose that on each toss, the probability of a head is p_u and the probability of a tail is p_d. We assume the tosses are independent. So the probabilities of the individual elements ω in Ω are

$$\mathbb{P}(HHH) = p_u^3, \mathbb{P}(HHT) = p_u^2 p_d, \mathbb{P}(HTH) = p_u^2 p_d, \mathbb{P}(HTT) = p_u p_d^2, \ldots\ldots$$

In probability, the subset of Ω are called events. For example, the event

$$\text{“the first toss is a head”} = \{\omega \in \Omega; \omega_1 = H\} = \{HHH, HHT, HTH, HTT\}$$

and we can compute

$$\begin{aligned}\mathbb{P}(\text{“the first toss is a head”}) =& \mathbb{P}(HHH) + \mathbb{P}(HHT) + \mathbb{P}(HTH) + \mathbb{P}(HTT) \\ =& (p_u^3 + p_u^2 p_d) + (p_u^2 p_d + p_u p_d^2) = p_u^2 + p_u p_d = p_u.\end{aligned}$$

Definition 2.1 *(Shreve I, Defintion 2.1.1) A finite probability space consists of a sample space Ω and a probability measure $\mathbb{P}$. The sample space Ω is a nonempty finite set and the probability measure $\mathbb{P}$ is a function that assigns to each element ω of Ω a number in $[0,1]$ so that*

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1. \tag{2.16}$$

An event is a subset of Ω, and we define the probability of an event A to be

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega). \tag{2.17}$$

Remark: By definition, $\mathbb{P}(\Omega) = \sum_{\omega\in\Omega} \mathbb{P}(\omega) = 1$. If A and B are disjoint subsets of Ω, then

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \tag{2.18}$$

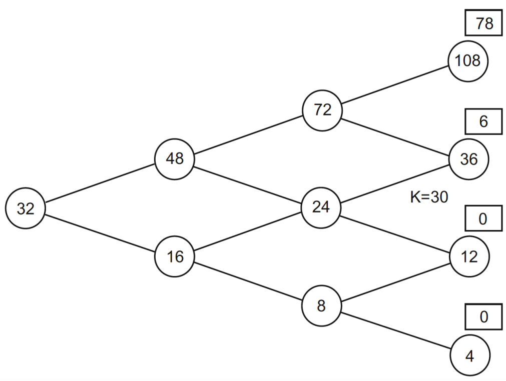
Definition 2.2 *Let $(\Omega, \mathbb{P})$ be a finite probability space. A random variable X is a real-valued function defined on Ω. The expectation of X is defined to be*

$$\mathbb{E}X = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega). \tag{2.19}$$

The variance of X is

$$Var(X) = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[X^2] - (\mathbb{E}X)^2. \tag{2.20}$$

Example 2.4 *(Stock prices) Recall the space Ω in (2.15) which consists of sequence of 3 independent coin tosses. As in Figure 2.2, let us define the Stock prices by the formula*



$$
\begin{aligned}
&S_0(w_1w_2w_3) = 32 \quad \text{for all } \omega = \omega_1\omega_2\omega_3 \in \Omega,\\
&S_1(w_1w_2w_3) = \begin{cases} 48 & \text{if } \omega_1 = H \\ 16 & \text{if } \omega_1 = T \end{cases}\\
&S_2(w_1w_2w_3) = \begin{cases} 72 & \text{if } \omega_1\omega_2 = HH \\ 24 & \text{if } \omega_1\omega_2 = HT \text{ or } TH \\ 8 & \text{if } \omega_1\omega_2 = TT \end{cases}\\
&S_3(w_1w_2w_3) = \begin{cases} 108 & \text{if } \omega_1\omega_2\omega_3 = HHH \\ 36 & \text{if there are two } H\text{'s and one } T \\ 12 & \text{if there are one } H \text{ and two } T\text{'s} \\ 4 & \text{if } \omega_1\omega_2\omega_3 = TTT \end{cases}
\end{aligned}
$$

Here we have written the arguments of S_0, S_1, S_2, and S_3 as $\omega_1\omega_2\omega_3$, even though some of these random variables do not depend on all the coin tosses.

Example 2.5 *Recall that in the one period binomial model. Define $\Omega = \{H, T\}$ and a probability measure $\mathbb{P}$ with $\mathbb{P}(H) = p_u$, $\mathbb{P}(T) = p_d$. They satisfy $p_u + p_d = 1$, $q_u, q_d > 0$. Then the option price at time T, denoted by $Ⓟ_T$, can be $Ⓟ^u$ or $Ⓟ^d$ after one period, depending on whether event H or event T happen.*

$$\mathbb{E}[Ⓟ_T] = p_u Ⓟ^u + p_d Ⓟ^d$$

This $\mathbb{P}$ is the real-world probability. $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$.

Example 2.6 *Recall that in the one period binomial model, we have (see (2.5)):*

$$Ⓟ_0 = \frac{1}{e^{rT}} \left(q_u Ⓟ^u + q_d Ⓟ^d \right).$$

Now define $\Omega = \{H, T\}$ and a probability measure $\tilde{\mathbb{P}}$ with $\tilde{\mathbb{P}}(H) = q_u$, $\tilde{\mathbb{P}}(T) = q_d$. They satisfy $q_u + q_d = 1$, $q_u, q_d > 0$. Then

$$Ⓟ_0 = e^{-rT} \tilde{\mathbb{E}}[Ⓟ_T] \tag{2.21}$$

This $\tilde{\mathbb{P}}$ is called a risk neutral probability. $\tilde{\mathbb{E}}$ is the expectation with respect to $\tilde{\mathbb{P}}$.

Example 2.7 *Recall that in the N-step binomial model, we have (see (2.10) with $\rho = e^{r\delta t}$):*

$$
\begin{aligned}
Ⓟ_0 &= \frac{1}{\rho^N} \sum_{i=0}^{N} \binom{N}{i} q_u^i q_d^{N-i} Ⓟ_N(\omega \text{ contains } i \ H\text{'s and } N-i \ T\text{'s})\\
&= e^{-rT} \sum_{i=0}^{N} \binom{N}{i} q_u^i q_d^{N-i} \left(S_0 u^i d^{N-i} - K \right)^+.
\end{aligned}
$$

Now define $\Omega = \{\omega : \omega = \omega_1 \cdots \omega_N, \omega_i = H \text{ or } T\}$ and a probability measure $\tilde{\mathbb{P}}$ with $\tilde{\mathbb{P}}(\omega) = q_u^{H(\omega)} q_d^{N-H(\omega)}$ where $H(\omega) =$ the number of H's in ω. Define the set $A_i = \{\omega :$

ω contains i H's and $N-i$ T's$\}$. Then A_i contains $\binom{N}{i}$ elements and on those elements, $Ⓓ_N$ takes the same value $\big(S_0 u^i d^{N-i} - K\big)^+$. Hence

$$\tilde{\mathbb{P}}(A_i) = \binom{N}{i} q_u^i q_d^{N-i}.$$

and

$$Ⓓ_0 = e^{-rT} \sum_{i=0}^{N} \tilde{\mathbb{P}}(A_i) Ⓓ_N(A_i) = e^{-rT} \tilde{\mathbb{E}}[Ⓓ_N].$$

This $\tilde{\mathbb{P}}$ is called a risk neutral probability. $\tilde{\mathbb{E}}$ is the expectation with respect to $\tilde{\mathbb{P}}$.

Remark: Recall that the capital asset pricing model (CAPM) proposed by Sharpe in 1964 (see for example, the books of John Hull or Brealey-Myers-Allen) says that

$$\text{Expected return on an asset } = r_f + \beta(r_m - r_f), \tag{2.22}$$

where r_m is the return on the market, and r_f is the return on a risk-free investment. r_m is usually approximated as the return on a well-diversified stock index such as S&P 500. β is a parameter measuring systematic risk and depends on the asset.

Once we have introduced probability and expectation, let's look at what we have found very early in Section 2.2. If we live in a "risk neutral world" which is an imaginary world with q_u, q_d being the probabilities of the price of the stock to go up and down respectively, and if we want to compute the expected values of stock and option in this imaginary world after one time interval δt, we would get

$$\tilde{\mathbb{E}}[Ⓓ_{\delta t}] = q_u Ⓓ^u + q_d Ⓓ^d \overset{(2.5)}{=} \rho Ⓓ_0$$

and

$$\tilde{\mathbb{E}}[S_{\delta t}] = q_u u S_0 + q_d d S_0 \overset{(2.6)}{=} \rho S_0$$

with $\rho = e^{r\delta t}$ where r is the risk-free interest rate. So, the expected return for the option and the stock are the same as that of the risk-free bank account. In other words, in this "risk neutral world", people do not get extra pay or reward for taking the risk to invest in option and stock. The introduction of such an imaginary world makes the option pricing easier. In the real world, the probability is different from the risk-neutral probability, investors do get a risk premium which is the $\beta(r_m - r_f)$ part in (2.22).

Example 2.8 *Recall (2.3)*

$$\Delta = \frac{Ⓓ^u - Ⓓ^d}{S_0(u-d)}$$

is the hedge ratio in a one period binomial tree model so that $\Phi = \textcircled{p} - \Delta S$ is risk-free. It is the change in the value of the option relative to the change of the value of the stock. If we wish to make this comparison in terms of percentage change, we get the option's elasticity

$$\Omega = \frac{(\textcircled{p}^u - \textcircled{p}^d)/\textcircled{p}_0}{(uS_0 - dS_0)/S_0} = \Delta \frac{S_0}{\textcircled{p}_0}. \tag{2.23}$$

The rate of the return of an option is $\frac{\textcircled{p}_{\delta t} - \textcircled{p}_0}{\textcircled{p}_0}$, which could be $\frac{\textcircled{p}^u - \textcircled{p}_0}{\textcircled{p}_0}$ or $\frac{\textcircled{p}^d - \textcircled{p}_0}{\textcircled{p}_0}$, depending on S_t going up or down from $t = 0$ to $t = \delta t$. In the risk neutral world, the mean (denoted by $m_{\textcircled{p}}$) and the standard deviation or volatility (which is the square root of the variance, and is denoted by $\sigma_{\textcircled{p}}$) of the rate of return of an option are therefore

$$m_{\textcircled{p}} = q_u \frac{\textcircled{p}^u - \textcircled{p}_0}{\textcircled{p}_0} + (1 - q_u) \frac{\textcircled{p}^d - \textcircled{p}_0}{\textcircled{p}_0}$$

and

$$\sigma_{\textcircled{p}} = \left[q_u \left(\frac{\textcircled{p}^u - \textcircled{p}_0}{\textcircled{p}_0} - m_{\textcircled{p}} \right) + (1 - q_u) \left(\frac{\textcircled{p}^d - \textcircled{p}_0}{\textcircled{p}_0} - m_{\textcircled{p}} \right) \right]^{1/2}.$$

Similarly, we can define the mean m_S and standard diviation σ_S of the rate of return of the stock. **Show** *that*

$$\sigma_{\textcircled{p}} = \sqrt{q_u(1 - q_u)} \left| \frac{\textcircled{p}^u - \textcircled{p}^d}{\textcircled{p}_0} \right| = |\Omega| \sigma_S. \tag{2.24}$$

Here, $|\Omega|$ is the absolute value of Ω. This equation relates the risk of a call to the risk of the underlying stock: The risk of a call (the standard deviation of its rate of return) equals its elasticity times its underlying stock volatility.

Solution: By Question 8 of Homework II, if $X = \begin{cases} X_a & \text{with probability } q \\ X_b & \text{with probability } 1 - q. \end{cases}$ Then $E[X] = qX_a + (1 - q)X_b$ and

$$\sigma(X) = \sqrt{q(1 - q)} |X_a - X_b|.$$

Hence $\sigma_{\textcircled{p}} = \sqrt{q_u(1 - q_u)} \left| \frac{\textcircled{p}^u - \textcircled{p}^d}{\textcircled{p}_0} \right|$ and

$$\sigma_S = \sqrt{q_u(1 - q_u)} \, |u - d| \,, \qquad \sigma_{\textcircled{p}} / \sigma_S = |\Omega|.$$

2.7 Conditional expectations

Recall that in the binomial pricing model, we have $S_0 = \frac{1}{\rho} \left(q_u S^u + q_d S^d \right)$ ((2.6)) which then implies

$$S_n(\omega_1 \cdots \omega_n) = e^{-r\delta t} \left(q_u S_{n+1}(\omega_1 \cdots \omega_n H) + q_d S_{n+1}(\omega_1 \cdots \omega_n T) \right). \tag{2.25}$$

Define

$$\tilde{\mathbb{E}}_n[S_{n+1}](\omega_1 \cdots \omega_n) = q_u S_{n+1}(\omega_1 \cdots \omega_n H) + q_d S_{n+1}(\omega_1 \cdots \omega_n T). \tag{2.26}$$

Then

$$S_n(\omega_1\cdots\omega_n) = e^{-r\delta t}\tilde{\mathbb{E}}_n[S_{n+1}](\omega_1\cdots\omega_n). \tag{2.27}$$

We call $\tilde{\mathbb{E}}_n[S_{n+1}]$ the conditional expectation of S_{n+1} based on the information at time n. It is still a random variable depending on $(\omega_1\cdots\omega_n)$. It is an estimate of the value of S_{n+1} based on the information of the first n coin tosses.

Definition 2.3 *Consider an N-period binomial model. Let n satisfy $1 \le n \le N$, and let $\omega_1, ..., \omega_n$ be given and, for the moment, fixed. There are 2^{N-n} possible continuations $\omega_{n+1}\cdots\omega_N$ of the sequence $\omega_1\cdots\omega_n$. Denote by $\#H(\omega_{n+1}\cdots\omega_N)$ the number of H's in the continuation $\omega_{n+1}\cdots\omega_N$ and by $\#T(\omega_{n+1}\cdots\omega_N)$ the number of T's. Define*

$$\tilde{\mathbb{E}}_n[X](\omega_1\cdots\omega_n) = \sum_{\omega_{n+1}\cdots\omega_N} q_u^{\#H(\omega_{n+1}\cdots\omega_N)} q_d^{\#T(\omega_{n+1}\cdots\omega_N)} X(\omega_1\cdots\omega_n\omega_{n+1}\cdots\omega_N) \tag{2.28}$$

and call $\tilde{\mathbb{E}}_n[X]$ the conditional expectation of X based on the information at time $n\delta t$.

The two extreme cases of conditioning are $\tilde{\mathbb{E}}_0[X]$, the conditional expectation of X based on no information, which we define by

$$\tilde{\mathbb{E}}_0[X] = \tilde{\mathbb{E}}[X],$$

and $\tilde{\mathbb{E}}_N[X]$, the conditional expectation of X based on information of the full trajectory, which we define by

$$\tilde{\mathbb{E}}_N[X] = X.$$

Example 2.9 *Consider Example 2.4 with $q_u = 1/2$,*

$$\begin{aligned}\tilde{\mathbb{E}}_1[S_3](H) =& \frac{1}{4}\times 108 + \frac{1}{4}\times 36 + \frac{1}{4}\times 36 + \frac{1}{4}\times 12 = 48,\\ \tilde{\mathbb{E}}_1[S_3](T) =& \frac{1}{4}\times 36 + \frac{1}{4}\times 12 + \frac{1}{4}\times 12 + \frac{1}{4}\times 4 = 16.\end{aligned}$$

But if $q_u = 1/3$,

$$\begin{aligned}\tilde{\mathbb{E}}_1[S_3](H) =& \frac{1}{9}\times 108 + \frac{2}{9}\times 36 + \frac{2}{9}\times 36 + \frac{4}{9}\times 12 = \frac{100}{3},\\ \tilde{\mathbb{E}}_1[S_3](T) =& \frac{1}{9}\times 36 + \frac{2}{9}\times 12 + \frac{2}{9}\times 12 + \frac{4}{9}\times 4 = \frac{100}{9}.\end{aligned}$$

Theorem 2.2 *(See Theorem 2.3.2 of Shreve, volume I.) Let N be a positive integer, and let X and Y be random variables depending on the first N steps of the trajectory. Let $0 \le n \le N$ be given. The following properties hold*

i) (Linearity of conditional expectations) For all constants c_1 and c_2, we have

$$\tilde{\mathbb{E}}_n[c_1X + c_2Y] = c_1\tilde{\mathbb{E}}_n[X] + c_2\tilde{\mathbb{E}}_n[Y]. \tag{2.29}$$

ii) (Taking out what is known) If X actually depends only on the first n coin toss, then

$$\tilde{\mathbb{E}}_n[XY] = X\tilde{\mathbb{E}}_n[Y]. \tag{2.30}$$

iii) (Iterated conditioning) If $0 \leq n \leq m \leq N$, then

$$\tilde{\mathbb{E}}_n[\tilde{\mathbb{E}}_m[X]] = \tilde{\mathbb{E}}_n[X]. \tag{2.31}$$

In particular, $\tilde{\mathbb{E}}[\tilde{\mathbb{E}}_m[X]] = \tilde{\mathbb{E}}[X]$.

iv) (Independence) If X depends only on ω_{n+1} through ω_N, then

$$\tilde{\mathbb{E}}_n X = \tilde{\mathbb{E}} X. \tag{2.32}$$

v) (Conditional Jensen's inequality) If $\varphi(x)$ is a convex function of the dummy variable x (e.g. $\varphi(x) = x^2$), then

$$\tilde{\mathbb{E}}_n[\varphi(X)] \geq \varphi(\tilde{\mathbb{E}}_n[X]). \tag{2.33}$$

Example 2.10 *Consider Example 2.4 with $q_u = 1/3$, calculate $\tilde{\mathbb{E}}_2[S_3]$ and $\tilde{\mathbb{E}}_1[\tilde{\mathbb{E}}_2[S_3]]$ by definition.*

Solution:

$$\begin{aligned}
\tilde{\mathbb{E}}_2[S_3](HH) &= \frac{1}{3} \times 108 + \frac{2}{3} \times 36 = 60 \\
\tilde{\mathbb{E}}_2[S_3](HT) &= \frac{1}{3} \times 36 + \frac{2}{3} \times 12 = 20 \\
\tilde{\mathbb{E}}_2[S_3](TH) &= \frac{1}{3} \times 36 + \frac{2}{3} \times 12 = 20 \\
\tilde{\mathbb{E}}_2[S_3](TT) &= \frac{1}{3} \times 12 + \frac{2}{3} \times 4 = \frac{20}{3}.
\end{aligned}
\qquad
\begin{aligned}
\tilde{\mathbb{E}}_1[\tilde{\mathbb{E}}_2[S_3]](H) &= \frac{1}{3} \times 60 + \frac{2}{3} \times 20 = \frac{100}{3} \\
\tilde{\mathbb{E}}_1[\tilde{\mathbb{E}}_2[S_3]](T) &= \frac{1}{3} \times 20 + \frac{2}{3} \times \frac{20}{3} = \frac{100}{9}.
\end{aligned}$$

Comparing with the second part of Example 2.9, we have verified that $\tilde{\mathbb{E}}_1[\tilde{\mathbb{E}}_2[S_3]] = \tilde{\mathbb{E}}_1[S_3]$.

Proof of Theorem 2.2: We only show the proof of iii). The proof won't be tested at all. It is presented so that one can get a feeling of why iii) makes sense.

Keep in mind that $n \leq m \leq N$. Denote $Z = \tilde{\mathbb{E}}_m[X]$. Then Z actually depends on

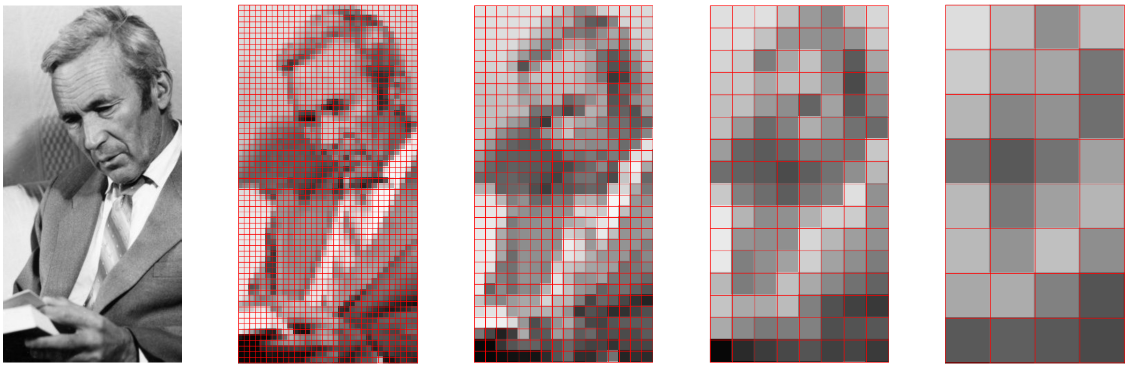


Figure 2.7: $X \leftarrow \tilde{\mathbb{E}}_4[X] \leftarrow \tilde{\mathbb{E}}_3[X] \leftarrow \tilde{\mathbb{E}}_2[X] \leftarrow \tilde{\mathbb{E}}_1[X]$.

$\omega_1 \cdots \omega_m$ only.

$$
\begin{aligned}
&\tilde{\mathbb{E}}_n[\tilde{\mathbb{E}}_m[X]](\omega_1 \cdots \omega_n) = \tilde{\mathbb{E}}_n[Z](\omega_1 \cdots \omega_n) \\
=& \sum_{\omega_{n+1},\cdots,\omega_N} q_u^{\#H(\omega_{n+1}\cdots\omega_N)} q_d^{\#T(\omega_{n+1}\cdots\omega_N)} Z(\omega_1 \cdots \omega_n \omega_{n+1} \cdots \omega_N) \\
=& \sum_{\omega_{n+1},\cdots,\omega_N} q_u^{\#H(\omega_{n+1}\cdots\omega_N)} q_d^{\#T(\omega_{n+1}\cdots\omega_N)} Z(\omega_1 \cdots \omega_n \omega_{n+1} \cdots \omega_m) \\
=& \sum_{\omega_{n+1},\cdots,\omega_N} \left(q_u^{\#H(\omega_{n+1}\cdots\omega_m)} q_d^{\#T(\omega_{n+1}\cdots\omega_m)} Z(\omega_1 \cdots \omega_m) \times q_u^{\#H(\omega_{m+1}\cdots\omega_N)} q_d^{\#T(\omega_{m+1}\cdots\omega_N)} \right) \\
=& \left(\sum_{\omega_{n+1},\cdots,\omega_m} q_u^{\#H(\omega_{n+1}\cdots\omega_m)} q_d^{\#T(\omega_{n+1}\cdots\omega_m)} Z(\omega_1 \cdots \omega_m) \right) \times \left(\sum_{\omega_{m+1},\cdots,\omega_N} q_u^{\#H(\omega_{m+1}\cdots\omega_N)} q_d^{\#T(\omega_{m+1}\cdots\omega_N)} \right) \\
=& \sum_{\omega_{n+1},\cdots,\omega_m} q_u^{\#H(\omega_{n+1}\cdots\omega_m)} q_d^{\#T(\omega_{n+1}\cdots\omega_m)} Z(\omega_1 \cdots \omega_m).
\end{aligned}
$$

In the last step, we have used $\sum_{\omega_{m+1},\cdots,\omega_N} q_u^{\#H(\omega_{m+1}\cdots\omega_N)} q_d^{\#T(\omega_{m+1}\cdots\omega_N)} = (q_u + q_d)^{N-m} = 1$ which is true because of (2.9) or the Pascal's triangle in Figure 2.5.

Then we continue to get

$$
\begin{aligned}
&\tilde{\mathbb{E}}_n[\tilde{\mathbb{E}}_m[X]](\omega_1 \cdots \omega_n) \\
=& \sum_{\omega_{n+1},\cdots,\omega_m} q_u^{\#H(\omega_{n+1}\cdots\omega_m)} q_d^{\#T(\omega_{n+1}\cdots\omega_m)} \times \sum_{\omega_{m+1}\cdots\omega_N} q_u^{\#H(\omega_{m+1}\cdots\omega_N)} q_d^{\#T(\omega_{m+1}\cdots\omega_N)} X(\omega_1 \cdots \omega_N) \\
=& \sum_{\omega_{n+1},\cdots,\omega_N} q_u^{\#H(\omega_{n+1}\cdots\omega_N)} q_d^{\#T(\omega_{n+1}\cdots\omega_N)} X(\omega_1 \cdots \omega_N) \\
=& \tilde{\mathbb{E}}_n[X](\omega_1 \cdots \omega_n). \qquad \square
\end{aligned}
$$

Remark:

Conditional expectation of X is the average of X over certain subset of Ω. One can consider X as the 1st image of Figure 2.7 which is the graph of a function g defined on a rectangle Ω. It contains all the information about X. But $\tilde{\mathbb{E}}_k[X]$ contains less information about X. For example, $\tilde{\mathbb{E}}_4[X]$ could be an image of resolution 64×32, $\tilde{\mathbb{E}}_2[X]$ could be an image of resolution 16×8, and so on.

For example, (2.31) with $n = 2$, $m = 3$, or $\tilde{\mathbb{E}}_2[\tilde{\mathbb{E}}_3[X]] = \tilde{\mathbb{E}}_2[X]$, says that image-4 in Figure 2.7 can be obtained in two different ways. (1) $\tilde{\mathbb{E}}_2[X]$: we can directly average image-1 to obtain image-4; (2) $\tilde{\mathbb{E}}_2[\tilde{\mathbb{E}}_3[X]]$: we can first average image-1 to image-3, and then we average image-3 to obtain image-4.

2.8 Margingales

Definition 2.4 *Consider the binomial asset-pricing model. Let M_0, M_1, ... , M_N be a sequence of random variables, with each M_n depending only on the first n coin tosses (and M_0 constant).* [16]

i) If

$$M_n = \tilde{\mathbb{E}}_n[M_{n+1}], \qquad n = 0, 1, ..., N-1,$$

we say that this process is a *<u>martingale</u>*.

ii) If

$$M_n \le \tilde{\mathbb{E}}_n[M_{n+1}], \qquad n = 0, 1, ..., N-1,$$

we say that this process is a *<u>submartingale</u>*.

iii) If

$$M_n \ge \tilde{\mathbb{E}}_n[M_{n+1}], \qquad n = 0, 1, ..., N-1,$$

we say that this process is a *<u>supermartingale</u>*.

Remark: The martingale property is a "one-step-ahead" condition. It implies a similar condition for any number of steps:

$$M_n = \tilde{\mathbb{E}}_n[M_{n+1}] = \tilde{\mathbb{E}}_n[\tilde{\mathbb{E}}_{n+1}[M_{n+2}]] \overset{(2.31)}{=} \tilde{\mathbb{E}}_n[M_{n+2}].$$

Iterating this argument, we can show that whenever $0 \le n \le m \le N$,

$$M_n = \tilde{\mathbb{E}}_n[M_m]. \tag{2.34}$$

[16] Such a sequence of random variables is called an <u>adapted stochastic process</u>.

Theorem 2.3 *Consider the general binomial model. Let the risk-neutral probabilities be q_u, q_d. Then, in the risk-neutral world, the discounted stock price and the discounted option price are martingale, i.e.,*

$$\tilde{\mathbb{E}}_n[\frac{S_{n+1}}{e^{r(n+1)\delta t}}] = \frac{S_n}{e^{rn\delta t}}, \tag{2.35}$$

$$\tilde{\mathbb{E}}_n[\frac{ⓟ_{n+1}}{e^{r(n+1)\delta t}}] = \frac{ⓟ_n}{e^{rn\delta t}}, \tag{2.36}$$

Proof: (2.35) follows from (2.6) or (2.27) . (2.36) follows from (2.7) . □

Example 2.11 *Toss a coin repeatedly. Assume the probability of head on each toss is $\frac{1}{2}$, so is the probability of tail. Let $X_j = 1$ if the jth toss results in a head and $X_j = -1$ if the jth toss results in a tail. Consider $M_1, M_1, M_2, \cdots$ (which is an example of stochastic process) defined by $M_0 = 0$ and*

$$M_n = \sum_{j=1}^{n} X_j, \qquad n \geq 1.$$

This is called a symmetric random walk; with each head, it steps up one, and with each tails, it steps down one. Define $I_0 = 0$ and

$$I_n = \sum_{j=0}^{n-1} M_j(M_{j+1} - M_j), \quad n = 1, 2, \cdots . \tag{2.37}$$

Show that

$$I_n = \frac{1}{2}M_n^2 - \frac{n}{2}. \tag{2.38}$$

Let n be an arbitrary nonnegative integer, and let $f(y)$ be an arbitrary function of a variable y. In terms of n and f, find the function $g(x)$ satisfying

$$\tilde{\mathbb{E}}_n[f(I_{n+1})] = g(I_n). \tag{2.39}$$

Solution: $M_j(M_{j+1}-M_j) = \frac{1}{2}\left(M_{j+1}^2 - M_j^2 - (M_{j+1} - M_j)^2\right) = \frac{1}{2}\left(M_{j+1}^2 - M_j^2 - (X_{n+1})^2\right) = \frac{1}{2}\left(M_{j+1}^2 - M_j^2 - 1\right)$. Hence

$$I_n = \sum_{j=0}^{n-1} M_j(M_{j+1} - M_j) = \frac{1}{2}M_n^2 - \frac{n}{2}.$$

$$\begin{aligned}
&\tilde{\mathbb{E}}_n[f(I_{n+1})](\omega_1 \cdots \omega_n) \\
=&q_u f(I_{n+1})(\omega_1 \cdots \omega_n H) + q_d f(I_{n+1})(\omega_1 \cdots \omega_n T) \\
=&\frac{1}{2}f(I_n + M_n X_{n+1}(H))(\omega_1 \cdots \omega_n) + \frac{1}{2}f(I_n + M_n X_{n+1}(T))(\omega_1 \cdots \omega_n) \\
=&\frac{1}{2}f(I_n + M_n)(\omega_1 \cdots \omega_n) + \frac{1}{2}f(I_n - M_n)(\omega_1 \cdots \omega_n) \\
=&\frac{1}{2}f(I_n + \sqrt{2I_n + n})(\omega_1 \cdots \omega_n) + \frac{1}{2}f(I_n - \sqrt{2I_n + n})(\omega_1 \cdots \omega_n).
\end{aligned}$$

Hence $g(I_n) = \frac{1}{2}\left(f(I_n + \sqrt{2I_n + n}) + f(I_n - \sqrt{2I_n + n})\right)$
or $g(x) = \frac{1}{2}\left(f(x + \sqrt{2x + n}) + f(x - \sqrt{2x + n})\right)$.

2.9 Information, σ-algebra, and filtration

(From Chapter 2 of Shreve II.) Image that some random experiment is performed and the outcome is a particular ω in the set of all possible outcomes Ω. We might be given some information – not enough to know the precise value of ω, but enough to narrow down the possibilities. For example, the true ω might be the result of three coin tosses, and we are told only the first one. Or we are only told the price of the stock S at time $t = 2$ without being told any of the coin toss. In such a situation, we can make a list of sets that are sure to contain the true ω and other sets that are sure not to contain the true ω. These are the sets that are **resolved** by the information. Please note that the word "resolve" does not mean solve or make a decision here, but means you can separate the whole thing (the sample space Ω) into two or more parts and you are able to distinguish between them. More precisely, a set of outcomes is resolved by certain information if and only if we can tell whether ω belongs to the set or not once we know that information.

Suppose Ω is the possible outcomes of three coin tosses. If we are told the outcome of the first toss only, the following sets of outcomes are resolved

$$A_H = \{HHH, HHT, HTH, HTT\}, \quad A_T = \{THH, THT, TTH, TTT\}.$$

The empty set $\emptyset$ and the whole sample space Ω (see (2.15)) are always resolved, even without any information; the true ω does not belong to $\emptyset$ and does belong to Ω. The four sets that are resolved by the first coin toss from a set of sets (whose elements are themselves all sets, it will be called a σ-algebra later in Definition 2.5)[17]

$$\mathcal{F}_1 = \{\emptyset, \Omega, A_H, A_T\}.$$

We shall think of $\mathcal{F}_1$ as containing the information learned by observing the first coin toss.

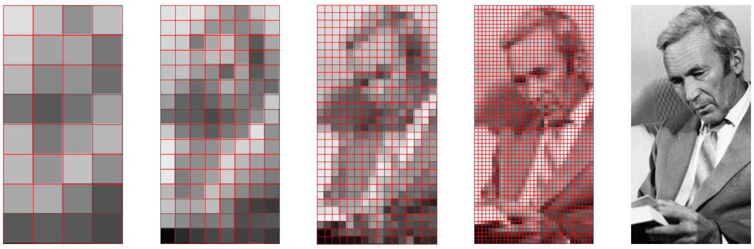


Figure 2.8: More is resolved as you move to the right.

If we are told the first two coin tosses, four additional sets

$$\begin{aligned} A_{HH} &= \{HHH, HHT\}, \ A_{HT} = \{HTH, HTT\}, \\ A_{TH} &= \{THH, THT\}, \ A_{TT} = \{TTH, TTT\} \end{aligned} \tag{2.40}$$

[17] If set A is resolved, so is its complement A^c. If sets A and B are resolved, so are $A \cup B$.

are resolved. Of course, the sets in $\mathcal{F}_1$ are still resolved. But we now have a higher resolution. Whenever a set is resolved, so is its complement, which means A^c_{HH}, A^c_{HT}, A^c_{TH}, and A^c_{TT} are resolved. Whenever two sets are resolved, so is their union. Finally, we have 16 resolved sets that together form a set of sets called $\mathcal{F}_2$; i.e.,

$$\mathcal{F}_2 = \left\{ \begin{array}{l} \emptyset, \Omega, A_H, A_T, A_{HH}, A_{HT}, A_{TH}, A_{TT}, A^c_{HH}, A^c_{HT}, A^c_{TH}, A^c_{TT} \\ A_{HH} \cup A_{TH}, A_{HH} \cup A_{TT}, A_{HT} \cup A_{TH}, A_{HT} \cup A_{TT} \end{array} \right\} \tag{2.41}$$

We shall think of $\mathcal{F}_2$ as containing the information learned by observing the first two coin tosses. In other words, (I) if we know the first two coin tosses, we can tell for each set in $\mathcal{F}_2$, whether the true ω belongs to it; (II) $\mathcal{F}_2$ contains all such sets.

Example 2.12 *Question: does the set $\{HHH\}$ belongs to $\mathcal{F}_2$? The answer is no because if the first two coin tosses is HH, we cannot tell whether the true ω belongs to $\{HHH\}$ or not. We are only able to tell the true $\omega \in \{HHH, HHT\}$ which is A_{HH}. To be able to tell whether the true $\omega \in \{HHH\}$ or not (to get higher resolution), we need to observe all three coin tosses.*

The way we construct $\mathcal{F}_2$ motivates the following definition:

Definition 2.5 *Let Ω be a nonempty set, and let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is a* **σ-algebra** *(or* **tribe***) if the following three conditions are satisfied*

i) the empty set $\emptyset$ belongs to $\mathcal{F}$.

ii) whenever a set A belongs to $\mathcal{F}$, its complement A^c also belongs to $\mathcal{F}$.

iii) whenever a sequence of sets $A_1, A_2, \cdots$ belongs to $\mathcal{F}$, their union $\cup_{n=1}^{\infty} A_n$ also belongs to $\mathcal{F}$.

Remark: From *ii*) and *iii*), we also have that

iv) whenever a sequence of sets $A_1, A_2, \cdots$ belongs to $\mathcal{F}$, their intersection $\cap_{n=1}^{\infty} A_n = (\cup_{n=1}^{\infty} A^c_n)^c$ also belongs to $\mathcal{F}$.

Hence the resolved sets form a σ-algebra.

If we are told all three coin tosses, we know the true ω and every subset of Ω is resolved. There are $2^{2^3} = 256$ subsets of Ω, and taken all together, they constitute the σ-algebra $\mathcal{F}_3$:

$$\mathcal{F}_3 = \text{ The set of all subsets of } \Omega.$$

If we are told nothing about the coin tosses, the only resolved sets are $\emptyset$ and Ω. We form the so called trivial σ-field $\mathcal{F}_0$ with these two sets:

$$\mathcal{F}_0 = \{\emptyset, \Omega\}.$$

We have then four σ-algebra, $\mathcal{F}_0$, $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{F}_3$, indexed by time. As time moves forward, we obtain finer resolution. In other words, of $n < m$, $\mathcal{F}_n \subset \mathcal{F}_m$. This means $\mathcal{F}_m$ contains more information than $\mathcal{F}_n$. The collection of σ-algebras $\mathcal{F}_0$, $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ is an example of a filtration.

Even though we have only discussed the discrete-time version, we now directly give a continuous-time formulation of filtration.

Definition 2.6 *(Definition 2.1.1 of Shreve II) Let Ω be a nonempty set. Let T be a fixed positive number, and assume that for each $t \in [0, T]$ there is a σ-algebra $\mathcal{F}_t$ of subset of Ω. Assume further that if $s \leq t$, then every set in $\mathcal{F}_s$ is also in $\mathcal{F}_t$. Then we call the collection of σ-algebras $\mathcal{F}_t$, $0 \leq t \leq T$, a filtration.*

A filtration tells us the information we will have at future times. More precisely, when we get to time t, we will know for each set in $\mathcal{F}_t$ whether the true ω lies in that set.

Example 2.13 *Suppose our sample space is $\Omega = C_0[0, T]$, the set of continuous functions defined on $[0, T]$ taking the value zero at time zero. The sets that are resolved by time t are just those sets that can be described in terms of the path of ω up to time t. For example,*

- *the set $\{\omega \in \Omega : \max_{0 \leq s \leq t} \omega(s) \leq 1\}$ is resolved and belongs to $\mathcal{F}_t$.*
- *if $t < T$, the set $\{\omega \in \Omega : \omega(T) > 0\}$ is not resolved by time t and does not belong to $\mathcal{F}_t$.*

Besides observing the evolution of an economy over time, there is a second way we might acquire information about the value of ω. Let X be a random variable that depends on ω. Suppose that rather than being told the value of ω, we are told only the value of $X(\omega)$, say, 5, then we know whether ω is in a set, say, $\{\omega : X(\omega) \leq 1\}$. For example, when we know $X(\omega) = 5$, we know ω is not in $\{\omega : X(\omega) \leq 1\}$.

Definition 2.7 *Let X be a random variable defined on a nonempty sample space Ω. The σ-algebra generated by X, denoted by $\sigma(X)$, is the collection of all subsets of Ω of the form $\{\omega : X(\omega) \in B\}$ where B ranges over the Borel subset of $\mathbb{R}^n$.*

For technical reasons, in the above definition, we require B to be a Borel set[18]

Example 2.14 *We return to Example 2.4 for the three-period model. Ω is the set of eight possible outcomes of three coin tosses and S_2 is defined in Example 2.4. Let $B = \{72\}$, then $\{\omega : S_2 \in B\} = A_{HH} = \{HHH, HHT\}$. Let $B = \{72, 24\}$, then $\{\omega : S_2 \in B\} = A_{HH} \cup A_{HT} \cup A_{TH}$. If we let B range over the Borel sets of $\mathbb{R}$, we will obtain the list of the set*

$$\emptyset, \Omega, A_{HH}, A_{TT}, A_{HT} \cup A_{TH}$$

[18]The σ-algebra obtained by beginning with cross products of open intervals $(a_1, b_1) \times (a_2, b_2) \cdots \times (a_n, b_n)$ and then adding everything else necessary in order to form a σ-algebra is called the Borel σ-algebra of $\mathbb{R}^n$. (In most case, we just need $n = 1$.) If B is a set in the Borel σ-algebra, it is called a Borel set.

and all unions and complements of them. This is the σ-algebra $\sigma(S_2)$. One can think of $\sigma(S_2)$ as the information on $\omega_1\omega_2\omega_3$ contained in knowing the value of S_2. Note that it is not precisely the value of $\omega_1\omega_2$ since we can not distinguish HT and TH by knowing the value of S_2.

Recall the $\mathcal{F}_2$ defined in (2.41). It represent the information contained in the first two coin tosses. Note that $\sigma(S_2) \subset \mathcal{F}_2$. This means that there is enough information in $\mathcal{F}_2$ to determine the value of S_2 [19] and even more. We say that S_2 is $\mathcal{F}_2$ measurable.

By the way, in this example, $\mathcal{F}_2$ is strictly larger than $\sigma(S_2)$. For example, $A_{HT} \in \mathcal{F}_2$ but $\sigma(S_2)$. This means that knowing the value of S_2 does not tell us everything about the first two coin tosses. In fact, we cannot tell HT from TH if all we know is the value of S_2.

Example 2.15 *In Question 5 of Homework II, $\Phi_n - \Delta_n S_n$ is the amount of money invested in the money market at $t_n = n\delta t$, while Δ_n is the number of stocks held in the time interval $[t_n, t_{n+1}]$. Note that the Δ_n defined by (2.45) is a function of $\omega_1 \cdots \omega_n$ and is $\mathcal{F}_n$ measurable.*

Remark: Note that S_2 is not $\mathcal{F}_1$ measurable as $\mathcal{F}_1 = \{\emptyset, \Omega, A_H, A_T\}$ is too coarse to distinguish the different values S_2 can take.

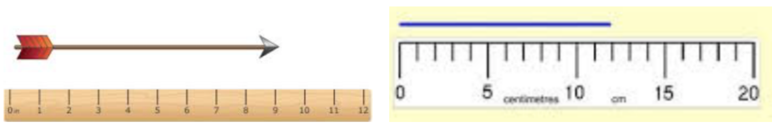


Figure 2.9: Pictures from the internet. Left: the arrow is not measurable by the ruler. To be able to measure the arrow (X), the ruler ($\mathcal{F}$ is the collection of the ticks on the ruler) needs to have finer scale ticks. Right: measurable.

Definition 2.8 *Let X be a random variable (i.e. a function) defined on a nonempty sample space Ω. Let $\mathcal{G}$ be a σ-algebra of subsets of Ω. If every set in $\sigma(X)$ is also in $\mathcal{G}$, we say that X is $\mathcal{G}$-measurable.*

Remark: Combining Definitions 2.7, 2.8 and the definition of Boreal set, we can say that random variable X, which is a function that maps a $\omega \in \Omega$ to a number in $\mathbb{R}$, is $\mathcal{G}$-measurable if

$$X^{-1}(U) \stackrel{\text{def}}{=} \{\omega \in \Omega : X(\omega) \in U\} \in \mathcal{G} \tag{2.42}$$

for all closed interval $U \in \mathbb{R}$. Equivalently, we can also require U be to all closed interval of $\mathbb{R}$.

We shall give another interpretation of measurability at the beginning of Chapter 4.

A random variable X is $\mathcal{G}$-measurable if and only if the information in $\mathcal{G}$ is sufficient to determine the value of X. If X is $\mathcal{G}$-measurable, then $f(X)$ is also $\mathcal{G}$-measurable for any

[19] If for every subset (also called event) A in $\mathcal{F}_2$ (or $\sigma(S_2)$), I know whether ω is in it or not (meaning I know whether A happens or not), then I can determine the value of S_2.

Borel-measurable function f; if the information in $\mathcal{G}$ is sufficient to determine the value of X, it will also determine the value of $f(X)$. If X and Y are $\mathcal{G}$-measurable, then $f(X,Y)$ is $\mathcal{G}$-measurable for any Borel-measurable function $f(x,y)$ of two variables.

Why should investors know about measurability? A portfolio Φ at time t must be $\mathcal{F}_t$-measurable, as investors must depend solely on information available to the investor at time t to adjust the investment strategy. Think about the last term in (1.6) (with $n=1$): $\phi_{t_m}[S_{t_{m+1}} - S_{t_m}]$. At time t_m, the invested decided to hold ϕ_{t_m} shares of stock. Then by t_{m+1}, his wealth increases by $\phi_{t_m}[S_{t_{m+1}} - S_{t_m}]$. We cannot use $\phi_{t_{m+1}}[S_{t_{m+1}} - S_{t_m}]$ as the investment strategy cannot depend on information not yet available. Later we will introduce Itô integral $\int_0^T \phi_t dS_t$ as a limiting process of $\sum_{m=0}^{N/\Delta t-1} \phi_{it_m}[S_{it_{m+1}} - S_{it_m}]$ where we require ϕ_t to be $\mathcal{F}_t$-measurable.

See (2.46) of Question 5 Homework II for an example of how to adjust the investment strategy.

Definition 2.9 *(Definition 2.1.6 of Shreve II) Let Ω be a nonempty sample space equipped with a filtration $\mathcal{F}_t$, $0 \le t \le T$. Let $\{X_t, 0 \le t \le T\}$ be a collection of random variables indexed by t. We say this collection of random variables is an $\mathcal{F}_t$-adapted stochastic process if, for each t, the random variable X_t is $\mathcal{F}_t$-measurable.*

Remark: See the first half of Definition 2.4 for the definition of discrete adapted stochastic process.

In continuous-time finance models, asset prices, portfolio processes will all be adapted to a filtration that we regard as a model of the flow of public information.

2.10 Homework II

(Only submit solutions to Questions 1,2,3,6,10,15.)

1. Using the following three-step binomial tree to compute the price of an European call option with strike price \$90. The initial price for the underlying stock is \$80. $r = 0.05$. $u = \frac{5}{4}$. $d = \frac{4}{5}$.

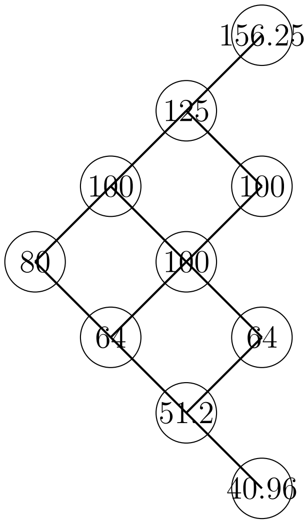


Figure 2.10: A binomial tree

2. (stochastic volatility, random interest rate) Consider a binomial pricing model, but at each time $n \geq 1$, the "up factor" $u_n(\omega_1 \cdots \omega_n)$, the "down factor" $d_n(\omega_1 \cdots \omega_n)$, and the interest rate $r_n(\omega_1 \cdots \omega_n)$ are allowed to depend on n and on the first n coin tosses $\omega_1 \cdots \omega_n$. The initial up factor u_0, the initial down factor d_0, and the initial interest rate r_0 are not random. More specifically, the stock price at time one is given by

$$S_1(\omega_1) = \begin{cases} u_0 S_0 & \text{if } \omega_1 = H, \\ d_0 S_0 & \text{if } \omega_1 = T, \end{cases}$$

and, for $n \geq 1$, the stock price at time $n+1$ is given by

$$S_{n+1}(\omega_1 \cdots \omega_{n+1}) = \begin{cases} u_n(\omega_1 \cdots \omega_n) S_n(\omega_1 \cdots \omega_n) & \text{if } \omega_{n+1} = H, \\ d_n(\omega_1 \cdots \omega_n) S_n(\omega_1 \cdots \omega_n) & \text{if } \omega_{n+1} = T. \end{cases}$$

One dollar invested in or borrowed from the money market at time $t = 0$ grows to an invest or debt of $\rho = e^{r_0 \delta t}$ at time $t_1 = \delta t$, and, for $n \geq 1$, one dollar invested in or borrowed from the money market at time n grows to an investment or debt of $e^{r_n(\omega_1 \cdots \omega_n)\delta t}$ at time t_{n+1}. We assume that for each n and for all $\omega_1 \cdots \omega_n$, the no-arbitrage condition

$$0 < d_n(\omega_1 \cdots \omega_n) < \rho_n = e^{r_n(\omega_1 \cdots \omega_n)\delta t} < u_n(\omega_1 \cdots \omega_n)$$

holds. We also assume that $0 < d_0 < \rho_0 = e^{r_0 \delta t} < u_0$.

i) Let N be a positive integer and let U_n be the price at time $t_n = n\delta t$ of a derivative security. Derive the formula that relates $U_n(\omega_1 \cdots \omega_n)$ to random variables $U_{n+1}(\omega_1 \cdots \omega_n H)$, $U_{n+1}(\omega_1 \cdots \omega_n T)$, $q_{u,n}(\omega_1 \cdots \omega_n) = \frac{\rho_n(\omega_1 \cdots \omega_n) - d_n(\omega_1 \cdots \omega_n)}{u_n(\omega_1 \cdots \omega_n) - d_n(\omega_1 \cdots \omega_n)}$, and $q_{d,n}(\omega_1 \cdots \omega_n) = \frac{u_n(\omega_1 \cdots \omega_n) - \rho_n(\omega_1 \cdots \omega_n)}{u_n(\omega_1 \cdots \omega_n) - d_n(\omega_1 \cdots \omega_n)}$.

ii) Suppose the initial stock price is $S_0 = 80$, with each head the stock price increase by 10, and with each tail the stock price decrease by 10. In another words, $S_1(H) = 90$, $S_1(T) = 70$, $S_2(HH) = 100$. Assume that the interest rate is always zero. Consider a European call with strike price 80, expiring at $t_3 = 3\delta t$. What is the price of this call at time $t = 0$?

3. Consider a European call option on an underlying stock with its present price S_0 =\$50 per share. Suppose that at the expiry date T the stock has only two possible values $S^u = \$80$ and $S^d = \$40$. Assume the strike price $K = \$60$ and the risk-free interest rate is $r = 0$.

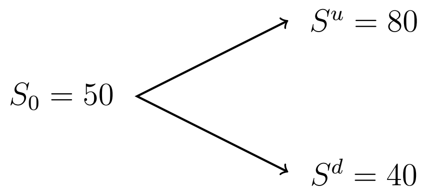


Figure 2.11: A single period binomial tree

Here is how to determine the option price by replication: If we construct a portfolio at $t = 0$ which consists of a debt of \$20 and $\frac{1}{2}$ shares of the stock

$$\Phi = -20 + \frac{1}{2}S.$$

Show that $V_T(\Phi) = V_T(\textcircled{p})$ no matter whether the stock price will go up or go down at $t = T$. We say that such a portfolio replicates the given option. Then use Corollary 1.30 to determine $\textcircled{p}_0 = V_0(\textcircled{p})$. Is this value the same as the one determined by (2.5)?

4. (A generalization of the last problem: Replication in the one period binomial model) Given an option, consider a portfolio Φ consisting of a risk-free asset with interest rate r and an underlying stock:

$$\Phi = B + \Delta S \tag{2.43}$$

where Δ is the number of shares of the underlying stock. We would like to choose $B_0 = V_0(B)$ and Δ so that $V_T(\Phi) = V_T(\textcircled{p}) = \begin{cases} \textcircled{p}^u & \text{if } V_T(S) = S^u, \\ \textcircled{p}^d & \text{if } V_T(S) = S^d, \end{cases}$. Find the formula for B_0 and Δ using r, T, $\textcircled{p}^u$, $\textcircled{p}^d$, S_0, S^u, and S^d. Can you determine $\textcircled{p}_0$ using Corollary 1.30?

Solution:

$$\begin{cases} e^{rT}B_0 + \Delta S^u = \textcircled{p}^u \\ e^{rT}B_0 + \Delta S^d = \textcircled{p}^d \end{cases}$$

or equivalently

$$\begin{bmatrix} e^{rT} & S^u \\ e^{rT} & S^d \end{bmatrix} \begin{bmatrix} B_0 \\ \Delta \end{bmatrix} = \begin{bmatrix} \textcircled{p}^u \\ \textcircled{p}^d \end{bmatrix}.$$

Hence

$$\begin{bmatrix} B_0 \\ \Delta \end{bmatrix} = \frac{1}{e^{rT}(S^d - S^u)} \begin{bmatrix} S^d & -S^u \\ -e^{rT} & e^{rT} \end{bmatrix} \begin{bmatrix} \textcircled{p}^u \\ \textcircled{p}^d \end{bmatrix}.$$

Thus

$$B_0 = \frac{1}{e^{rT}} \left(\frac{-S^d}{S^u - S^d} \textcircled{p}^u + \frac{S^u}{S^u - S^d} \textcircled{p}^d \right),$$
$$\Delta = \frac{\textcircled{p}^u - \textcircled{p}^d}{S^u - S^d}.$$

We mention in passing that the Δ used to build the portfolio (2.43) that replicate a option equals to the Δ used to build a risk-free portfolio in (2.2).

$$\textcircled{p}_0 = V_0(\Phi) = B_0 + \Delta S_0 = \frac{1}{\rho} \left(\frac{\rho S_0 - S^d}{S^u - S^d} \textcircled{p}^u + \frac{S^u - \rho S_0}{S^u - S^d} \textcircled{p}^d \right)$$

with $\rho = e^{rT}$. The above formula for $\textcircled{p}_0$ is the same as the formula we derived in (2.5).

5. (Replication in the multiperiod binomial model) Consider the multiperiod binomial model introduced in Section 2.3. Suppose we define

$$\textcircled{p}_n(\omega_1 \cdots \omega_n) = e^{-r\delta t} \left(q_u \textcircled{p}_{n+1}(\omega_1 \cdots \omega_n H) + q_d \textcircled{p}_{n+1}(\omega_1 \cdots \omega_n T) \right) \tag{2.44}$$

and

$$\Delta_n(\omega_1 \cdots \omega_n) = \frac{\textcircled{p}_{n+1}(\omega_1 \cdots \omega_n H) - \textcircled{p}_{n+1}(\omega_1 \cdots \omega_n T)}{S_{n+1}(\omega_1 \cdots \omega_n H) - S_{n+1}(\omega_1 \cdots \omega_n T)}, \tag{2.45}$$

with $n = N-1, \cdots, 0$. Prove that if we set $\Phi_0 = V_0$ and define recursively forward in time the portfolio values $\Phi_1, \Phi_2, \cdots \Phi_N$ by

$$\Phi_{n+1} = e^{r\delta t} (\Phi_n - \Delta_n S_n) + \Delta_n S_{n+1}, \tag{2.46}$$

then

$$\Phi_N(\omega_1 \cdots \omega_N) = \textcircled{p}_N(\omega_1 \cdots \omega_N) \qquad \text{for all } \omega_1 \cdots \omega_N. \tag{2.47}$$

Proof: We prove by induction in n that

$$\Phi_n(\omega_1 \cdots \omega_n) = \textcircled{p}_n(\omega_1 \cdots \omega_n) \qquad \text{for all } \omega_1 \cdots \omega_n \tag{2.48}$$

for $n = 0, ..., N$.

We know (2.48) is true when $n = 0$. We assume (2.48) is true for n and show that it is true for $n+1$.

By (2.46),

$$\Phi_{n+1}(\omega_1\cdots\omega_n H) = e^{r\delta t}\left(\Phi_n(\omega_1\cdots\omega_n) - \Delta_n(\omega_1\cdots\omega_n)S_n(\omega_1\cdots\omega_n)\right) + \Delta_n(\omega_1\cdots\omega_n)uS_n(\omega_1\cdots\omega_n).$$

To simplify the notation, we suppress $\omega_1\cdots\omega_n$ and write the equation simply as

$$\Phi_{n+1}(H) = e^{r\delta t}\left(\Phi_n - \Delta_n S_n\right) + \Delta_n u S_n. \tag{2.49}$$

With $\omega_1\cdots\omega_n$ suppressed, (2.45) can be written as

$$\Delta_n = \frac{\textcircled{D}_{n+1}(H) - \textcircled{D}_{n+1}(T)}{S_{n+1}(H) - S_{n+1}(T)} = \frac{\textcircled{D}_{n+1}(H) - \textcircled{D}_{n+1}(T)}{(u-d)S_n}.$$

Plugging the above equation into (2.49), we get

$$\begin{aligned}
\Phi_{n+1}(H) &= e^{r\delta t}\Phi_n + \left(u - e^{r\delta t}\right)\Delta_n S_n = e^{r\delta t}\Phi_n + \left(u - e^{r\delta t}\right)\frac{\textcircled{D}_{n+1}(H) - \textcircled{D}_{n+1}(T)}{u-d} \\
&\overset{\text{induction assumption and def of } q_d}{=} e^{r\delta t}\textcircled{D}_n + q_d\left(\textcircled{D}_{n+1}(H) - \textcircled{D}_{n+1}(T)\right) \\
&\overset{(2.44)}{=} q_u\textcircled{D}_{n+1}(H) + q_d\textcircled{D}_{n+1}(T) + q_d\left(\textcircled{D}_{n+1}(H) - \textcircled{D}_{n+1}(T)\right) \\
&= \textcircled{D}_{n+1}(H).
\end{aligned}$$

A similar argument shows that $\Phi_{n+1}(T) = \textcircled{D}_{n+1}(T)$.

6. (The call can never be less risky than the underlying stock) With the same setup as in the last question, consider a one period binomial tree model. Show that for put option, $\Omega \le 0$, and for call option $\Omega \ge 1$.

   [Hint: For call option, you need to prove $\textcircled{D}^u - \textcircled{D}^d \ge (u-d)\textcircled{D}_0 > 0$. Use (2.5) to show that you only need to prove $d\textcircled{D}^u - u\textcircled{D}^d \ge 0$. Now, use $\textcircled{D}^u = \max(uS_0 - K, 0)$ and $\textcircled{D}^d = \max(dS_0 - K, 0)$ for the one period model.] As a matter of fact, this result is also true for the multi-period model. See Page 187 of "Options Markets" by Cox and Rubinstein.

7. (Bernoulli Random Variable) An experiment, whose outcome can be classified as either a success or a failure is performed. Let $X = 1$ when the outcome is a success, and $X = 0$ if the outcome is a failure. Then the probability mass function of X is given by

$$\mathbb{P}(X=0) = 1-p$$
$$\mathbb{P}(X=1) = p$$

   where $p \in [0,1]$ is the probability that the trial is a success. A random variable X is said to be a **Bernoulli random variable** if its probability mass function is given as above. Prove that $E[X] = p$ and $\mathrm{Var}[X] = p(1-p)$.

   Solution:

$$\begin{aligned}
E[X] &= 0\times\mathbb{P}(X=0) + 1\times\mathbb{P}(X=1) = p. \\
E[X^2] &= 0^2\times\mathbb{P}(X=0) + 1^2\times\mathbb{P}(X=1) = p. \\
\mathrm{Var}[X] &= E[X^2] - (E[X])^2 = p - p^2 = p(1-p).
\end{aligned}$$

8. If X is a discrete random variable with

$$\mathbb{P}(X = a) = p$$
$$\mathbb{P}(X = b) = 1 - p$$

where $p \in [0, 1]$. Prove that $E[X] = ap + b(1 - p)$ and $\mathrm{Var}[X] = (a - b)^2 p(1 - p)$.

Solution:

$$E[X] = a \times \mathbb{P}(X = a) + b \times \mathbb{P}(X = b) = ap + b(1 - p).$$
$$E[X^2] = a^2 p + b^2(1 - p).$$

$$\begin{aligned}\mathrm{Var}(X) = E[X^2] - (E[X])^2 &= a^2 p + b^2(1 - p) - (ap + b(1 - p))^2 \\ &= (a - b)^2 p(1 - p).\end{aligned}$$

9. (Binomial Random Variable) Suppose n independent trials, each results in a success with probability p or in a failure with probability $1 - p$, are to be performed. If Y represents the number of successes occur in the n trials, then Y is said to be **binomial random variable** with parameters (n, p), and denoted as $Y \sim B(n, p)$. Prove that $E[Y] = np$ and $\mathrm{Var}[Y] = np(1 - p)$.

Solution: Note that

$$P(Y = i) = \binom{n}{k} p^i (1 - p)^{n-i} \quad \text{with } i = 0, 1, 2, ..., n.$$

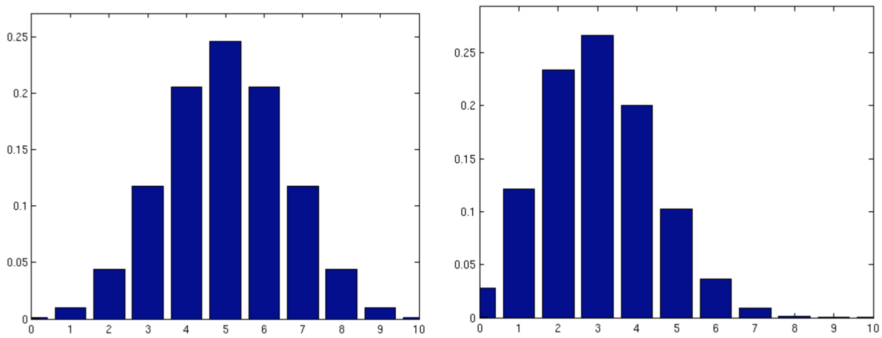


Figure 2.12: $\binom{n}{i} p^i (1 - p)^{n-i}$ with $n = 10$, $i = 0, 1, 2, \ldots, n$. Left: $p = 0.5$. Right: $p = 0.3$

Then, we can prove the statements using definition. We start with computing the kth order moment

$$
\begin{aligned}
E[Y^k] &= \sum_{i=0}^{n} i^k \mathbb{P}(Y=i) = \sum_{i=0}^{n} i^k \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=0}^{n} i^{k-1} np \binom{n-1}{i-1} p^{i-1} (1-p)^{(n-1)-(i-1)} \\
&= \sum_{i=\color{red}{1}}^{n} i^{k-1} np \binom{n-1}{i-1} p^{i-1} (1-p)^{(n-1)-(i-1)} \\
&\overset{j=i-1}{=} np \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j} \\
&= npE[(Z+1)^{k-1}] \qquad (2.50)
\end{aligned}
$$

with $Z \sim B(n-1, p)$. Hence

$$E[Y] = npE[(Z+1)^0] = npE[1] = np. \qquad (2.51)$$

$$
\begin{aligned}
E[Y^2] &= npE[(Z+1)] = np\,(E[Z]+1) \overset{(2.51)}{=} np\,((n-1)p+1)\,. \qquad (2.52) \\
\mathrm{Var}[Y] &= E[Y^2] - (E[Y])^2 = np\,((n-1)p+1) - (np)^2 = np(1-p).
\end{aligned}
$$

For later reference, I would like to present another (easier) proof using a decomposition:

$$Y = X_1 + X_2 + \cdots + X_n$$

where $X_i = \begin{cases} 1 & \text{if the } i\text{th trial is a success} \\ 0 & \text{if the } i\text{th trial is a failure} \end{cases}$ and X_i's are independent. The definition of independence will be introduced in the next Chapter where one can show that if U and V are independent, then $\mathrm{Var}[U+V] = \mathrm{Var}[U] + \mathrm{Var}[V]$. Note that we already have $E[U+V] = E[U] + E[V]$ no matter U, V are independent or not. Since each X_i is a Bernoulli random variable, $E[X] = p$ and $\mathrm{Var}[X] = p(1-p)$. Hence

$$E[Y] = E[X_1] + E[X_2] + \cdots + E[X_n] = np,$$

$$\mathrm{Var}[Y] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \cdots + \mathrm{Var}[X_n] = np(1-p).$$

10. (Skewness) The skewness of a random variable X is defined to be

$$Sk = E\left[\left(\frac{X - E[X]}{\sigma}\right)^3\right]$$

where $\sigma = \sqrt{\mathrm{Var}[X]}$. Prove that the skewness of $B(n,p)$ distribution is

$$Sk(n,p) = \frac{1-2p}{\sqrt{np(1-p)}}.$$

[Hint: Use (2.50) and $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$.] [20]

11. Suppose X is a random variable. Define $M_k = \tilde{\mathbb{E}}_k[X]$. Prove that $\{M_k\}$ is a martingale.

    Proof: $\tilde{\mathbb{E}}_n[M_{n+1}] = \tilde{\mathbb{E}}_n[\tilde{\mathbb{E}}_{n+1}[X]] \overset{(2.31)}{=} \tilde{\mathbb{E}}_n[X] = M_n$.

12. (random walk) Toss a coin repeatedly. Assume the probability of head on each toss is $\frac{1}{2}$, so is the probability of tail. Let $X_j = 1$ if the jth toss results in a head and $X_j = -1$ if the jth toss results in a tail. Consider $M_1, M_1, M_2, \cdots$ (which is an example of stochastic process) defined by $M_0 = 0$ and

$$M_n = \sum_{j=1}^{n} X_j, \qquad n \geq 1.$$

    This is called a symmetric random walk; with each head, it steps up one, and with each tails, it steps down one. Using Theorem 2.2 to show that $M_1, M_1, M_2, \cdots, M_n, \cdots$ is a martingale.

    Proof: M_n only depends on $\omega_1 \cdots \omega_n$.

$$\tilde{\mathbb{E}}_n[M_{n+1}] = \tilde{\mathbb{E}}_n[M_n + X_{n+1}] = \tilde{\mathbb{E}}_n[M_n] + \tilde{\mathbb{E}}_n[X_{n+1}] = M_n + \mathbb{E}[X_{n+1}] = M_n.$$

13. (discrete-time stochastic integral) Suppose $M_0, M_1, \cdots, M_N$ is a martingale, and let $\Delta_0, \Delta_1, \cdots, \Delta_{N-1}$ be an adapted process (see definition 2.4). Define the discrete-time stochastic integral $I_0, I_1, \cdots, I_N$ by setting $I_0 = 0$ and

$$I_n = \sum_{j=0}^{n-1} \Delta_j (M_{j+1} - M_j), \qquad n = 1, ..., N. \tag{2.53}$$

    Show that $I_0, I_1, \cdots, I_N$ is a martingale.

    Proof: I_n only depends on $\omega_1 \cdots \omega_n$ and $I_0, I_1, \cdots, I_N$ is therefore an adapted stochastic process.

$$\begin{aligned}\tilde{\mathbb{E}}_n[I_{n+1}] =& \tilde{\mathbb{E}}_n[I_n + \Delta_n(M_{n+1} - M_n)] \overset{(2.29)}{=} \tilde{\mathbb{E}}_n[I_n] + \tilde{\mathbb{E}}_n[\Delta_n(M_{n+1} - M_n)] \\ \overset{(2.30)}{=}& I_n + \Delta_n \tilde{\mathbb{E}}_n[M_{n+1} - M_n] \\ =& I_n.\end{aligned}$$

[20] For your information, the kurtosis of a random variable X is defined to be

$$K = E\left[\left(\frac{X - E[X]}{\sigma}\right)^4\right]$$

where $\sigma = \sqrt{\mathrm{Var}[X]}$. Using the same idea as in the computation of $Sk(n,p)$, one can prove that the kurtosis of $B(n,p)$ distribution is

$$K(n,p) = 3 + \frac{1 - 6p(1-p)}{np(1-p)}.$$

You are not asked to prove the above formula of $K(n,p)$ in the homework.

In the last step, we have used $\tilde{\mathbb{E}}_n[M_{n+1} - M_n] = \tilde{\mathbb{E}}_n[M_{n+1}] - M_n \overset{M_n \text{ is martingale}}{=} 0$.

14. (i) Consider the dice-toss space similar to the coin-toss space. A typical element ω in this space is an infinite sequence $\omega = \omega_1\omega_2\omega_3\cdots$ with $\omega_i \in \{1, 2, \cdots, 6\}$. Define a random variable

$$X(\omega) = \omega_1$$

and a function

$$f(x) = \begin{cases} 1 & \text{if } x \geq 4 \\ 0 & \text{if } x < 4 \end{cases}.$$

Recall the definition of $\sigma(X)$ (Defintion 2.7). Let $\Omega = \{\omega : \omega = \omega_1\omega_2\omega_3\cdots\}$ and $A_i = \{\omega : \omega_1 = i\}$. It is not hard to see that $\sigma(X) = \{\emptyset, \Omega, A_i,\ A_i \cup A_j,\ A_i \cup A_j \cup A_k,\ A_i \cup A_j \cup A_k \cup A_l,\ A_i \cup A_j \cup A_k \cup A_l \cup A_m,\ i, j, k, l, m$ are not equal pairwisely, $i, j, k, l, m = 1, \cdots, 6\}$.

Since $f(X)$ is also a random variable defined on Ω, we can also define $\sigma(f(X))$. What is $\sigma(f(X))$?

(ii) In general, if X is a random variable, can the σ-algebra generated by $f(X)$ ever be strictly larger than the σ-algebra generated by X?

Solution: (i) $\sigma(f(X)) = \{\emptyset, \Omega, \{\omega : \omega_1 = 1, \text{or } 2, \text{or } 3\}, \{\omega : \omega_1 = 4, \text{or } 5, \text{or } 6\}\}$.

(ii) No. $\sigma(f(X))$ is always a subset of $\sigma(X)$.

15. Consider the symmetric random walk $M_1, M_1, M_2, \cdots$ defined in Example 2.11. Let $\sigma > 0$ be a constant.

a) Define $J_0 = 0$ and

$$J_n = \sum_{j=0}^{n-1} e^{\sigma M_j}(M_{j+1} - M_j), \quad n = 1, 2, \cdots. \tag{2.54}$$

Show that $J_0, J_1, \cdots, J_N$ is a martingale, which means, you need to prove that $\mathbb{E}_n[J_{n+1}] = J_n$.

b) Define

$$K_n = \sum_{j=0}^{n-1} M_{j+1}(M_{j+1} - M_j), \quad n = 1, 2, \cdots. \tag{2.55}$$

Show that

$$K_n = \frac{1}{2}M_n^2 + \frac{n}{2}.$$