Probability Theory on Coin Toss Space

Peiliang Guo

January 10, 2017

1 Finite Probability Space

A finite probability space is used to model a situation in which a random experiment with finitely many possible outcomes is conducted.
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

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

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

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

2 Random Variables, Distributions, and Expectations

Let $(\Omega, \mathbb{P})$ be a finite probability space. A random variable is a real-valued function defined on Ω.
The distribution of a random variable is a specification of the probabilities that the random variable takes various values. Let X be a random variable defined on a finite probability space $(\Omega, \mathbb{P})$. The expectation of X is defined to be

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

The variance of X is

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

If $l(x) = ax + b$, then $\mathbb{E}[l(X)] = l(\mathbb{E}X)$.
Jensen's inequality. Let X be a random variable on a finite probability space, and let $\varphi(x)$ be a convex function of a dummy variable x. Then

$$\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}X)$$

3 Conditional Expectations

Let n satisfy $1 \leq n \leq N$, and let $\omega_1 ... \omega_n$ be given and fixed. There are 2^{N-n} possible continuations $\omega_{n+1} ... \omega_N$ of sequence fixed $\omega_1 ... \omega_n$. Denote by $\#H(\omega_{n+1} ... \omega_N)$ the number of heads in the continuation $\omega_{n+1} ... \omega_N$ and by $\#T(\omega_{n+1} ... \omega_N)$ the number of tails. We define

$$\tilde{\mathbb{E}}_n[X](\omega_1 ... \omega_n) = \sum_{\omega_{n+1} ... \omega_N} \tilde{p}^{\#H(\omega_{n+1} ... \omega_N)} \tilde{q}^{\#T(\omega_{n+1} ... \omega_N)} X(\omega_1 ... \omega_N)$$

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

$$\tilde{\mathbb{E}}_0[X] = \tilde{E}X \text{ and } \tilde{\mathbb{E}}_N[X] = X.$$

Theorem 1. *Let N be a positive integer, and let X and Y be random variables depending on first N coin tosses. 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*

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

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

$$\mathbb{E}_n[XY] = X \cdot \mathbb{E}_n[Y]$$

(iii) Iterated conditioning. *If $0 \le n \le m \le N$, then*

$$\mathbb{E}_n\left[\mathbb{E}_m[X]\right] = \mathbb{E}_n[X]$$

(iv) Independence. *If X depends only on tosses $n+1$ through N, then*

$$\mathbb{E}_n[X] = \mathbb{E}X$$

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

$$\mathbb{E}_n[\varphi(X)] \ge \varphi(\mathbb{E}_n[X])$$

4 Martingales

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). Such a sequence of random variables is called an adapted stochastic process.
(i) Martingale:

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

(ii) Submartingale:

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

(iii) Supermartingale:

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

Theorem 2. *Consider the general binomial model with $0 < d < 1 + r < u$. Let the risk-neutral probabilities be given by*

$$\tilde{p} = \frac{1+r-d}{u-d}, \tilde{q} = \frac{u-1-r}{u-d}.$$

Then, under risk-neutral measure, the discounted stock price is a martingale.

Theorem 3. *Consider the binomial model with N periods. Let $\Delta_0, \Delta_1, ..., \Delta_{N-1}$ be an adapted portfolio process, let X_0 be a real number, and let the wealth process $X_1, ..., X_N$ be generated recursively by*

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n), n = 0, 1, ..., N-1$$

Then the discounted wealth process $\frac{X_n}{(1+r)^n}, n = 0, 1, ..., N$, is a martingale under the risk-neutral measure; i.e.

$$\frac{X_n}{(1+r)^n} = \tilde{\mathbb{E}}_n\left[\frac{X_{n+1}}{(1+r)^{n+1}}\right], n = 0, 1, ..., N-1$$

Theorem 4. *Consider an N-period binomial asset-pricing model with $0 < d < 1 + r < u$ and with risk-neutral probability measure $\tilde{\mathbb{P}}$. Let V_N be a random variable depending on the coin tosses. Then, for n between 0 and N, the price of the derivative security at time n given by the risk-neutral pricing formula*

$$V_n = \tilde{\mathbb{E}}_n\left[\frac{V_N}{(1+r)^{N-n}}\right]$$

Furthermore, the discounted price of the derivative security is a martingale under $\tilde{\mathbb{P}}$; i.e.

$$\frac{V_n}{(1+r)^n} = \tilde{\mathbb{E}}_n\left[\frac{V_{n+1}}{(1+r)^{n+1}}\right], n = 0, 1, ..., N-1$$

Theorem 5. *Consider an N-period binomial asset pricing model with $0 < d < 1 + r < u$, and risk-neutral probability measure $\tilde{\mathbb{P}}$. Let $C_0, C_1, ..., C_N$ be a sequence of random variables such that C_n depends only on $\omega_1...\omega_n$. The price at time n of the derivative security that makes payments $C_n, ..., C_N$ at times $n, ..., N$, respectively, is*

$$V_n = \tilde{\mathbb{E}}_n \left[\sum_{k=n}^{N} \frac{C_k}{(1+r)^{k-n}} \right], n = 0, 1, ..., N.$$

The price process $V_n, n = 0, 1, ..., N$, satisfies

$$C_n(\omega_1...\omega_n) = V_n(\omega_1...\omega_n) - \frac{1}{1+r} \left[\tilde{p} V_{n+1}(\omega_1...\omega_n H) + \tilde{q} V_{n+1}(\omega_1...\omega_n T) \right]$$

We define

$$\Delta_n(\omega_1...\omega_n) = \frac{V_{n+1}(\omega_1...\omega_n H) - V_{n+1}(\omega_1...\omega_n T)}{S_{n+1}(\omega_1...\omega_n H) - S_{n+1}(\omega_1...\omega_n T)},$$

where n ranges from 0 and $N-1$. If we set $X_0 = V_0$ and define recursively forward in time the portfolio values $X_1, .., X_N$ by

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - C_n - \Delta_n S_n),$$

then we have

$$X_n(\omega_1...\omega_n) = V_n(\omega_1...\omega_n)$$

for all n and all $\omega_1...\omega_n$.

5 Markov Processes

The problem is significantly reduced if we only consider what information is relevant.
Consider the binomial asset pricing model. Let $X_0, X_1, ..., X_N$ be an adapted process. If, for every n between 0 and $N-1$ and for every function $f(x)$, there is another function $g(x)$ (depending on n and f) such that

$$\mathbb{E}_n[f(X_{n+1})] = g(X_n),$$

we say that $X_0, X_1, ..., X_N$ is a Markov process.
Sometimes, when we encounter a non-Markov process, we can sometimes recover the Markov property by adding one or more state variables.
Consider the binomial asset-pricing model. Let $\{(X_n^1, ..., X_n^K); n = 0, 1, ..., N\}$ be a K-dimensional adapted process; i.e., K one-dimensional processes. If, for every n between 0 and $N-1$ and for every function $f(x^1, ..., x^K)$, there is another function $g(x^1, ..., x^K)$ such that

$$\mathbb{E}_n[f(X_{n+1}^1, ..., X_{n+1}^K)] = g(X_n^1, ..., X_n^K),$$

we say that $\{(X_n^1, ..., X_n^K); n = 0, 1, ..., N\}$ is a K-dimensional Markov process.
Note that in the definition, only the "one-step-ahead" Markov property is focused. However, "one-step-ahead" Markov property implies that for every function h, there is a function f such that

$$\mathbb{E}_{n+1}[h(X_{n+2})] = f(X_{n+1})$$

Taking conditional expectation on both sides based on information at time n and using iterated conditioning property, we obtain

$$\mathbb{E}_n[h(X_{n+2})] = \mathbb{E}_n[\mathbb{E}_{n+1}[h(X_{n+2})]] = \mathbb{E}_n[f(X_{n+1})]$$

Because of the "one-step-ahead" Markov property, the right-hand side is $g(X_n)$ for some function g, and we obtain the "two-step-ahead" Markov property $\mathbb{E}_n[h(X_{n+2})] = g(X_n)$.

5.1 Feynman Kac's Theorem

In the binomial pricing model, suppose we have a Markov process, $X_0, X_1, ..., X_N$ under the risk-neutral probability measure $\tilde{\mathbb{P}}$, and we have a derivative security whose payoff V_N at time N is a function v_N of X_N, i.e.

$$V_N(\omega_1...\omega_N) = v_N(X_N(\omega_1...\omega_N)) \text{ for all } \omega_1...\omega_N$$

The risk-neutral pricing formula says that the price of this derivative security at time n is

$$V_n(\omega_1...\omega_n) = \tilde{\mathbb{E}}_n\left[\frac{V_N}{(1+r)^{N-n}}\right](\omega_1...\omega_n) \text{ for all } \omega_1...\omega_n.$$

On the other hand, the Markov property implies that there is a function v_n such that

$$\tilde{E}_n\left[\frac{V_N}{(1+r)^{N-n}}\right](\omega_1...\omega_n) = v_n(X_n(\omega_1...\omega_n)) \text{ for all } \omega_1...\omega_n$$

Therefore, the price of the the derivative at time n is a function of X_n,

$$V_n = v_n(X_n)$$

In the look-back option example, the payoff of the derivative is $V_3 = M_3 - S_3$, the difference between the stock price at time three and its maximum between time zero and three. Because only the stock price and its maximum-to-date appear in the payoff, we can use the two-dimensional Markov process $\{(S_n, M_n); n = 0, 1, 2, 3\}$ to treat this problem.
According to the Markov property, for any n between 0 and N, there is a function $v_n(s, m)$ such that the price of option at time n is

$$V_n = v_n(S_n, M_n) = \tilde{\mathbb{E}}_n\left[\frac{v_N(S_N, M_N)}{(1+r)^{N-n}}\right]$$

Suppose that for some n between zero and $N-1$, we have computed the function v_{n+1} such that $V_{n+1} = v_{n+1}(S_{n+1}, M_{n+1})$. Then

$$\begin{aligned} V_n =& \frac{1}{1+r}\tilde{\mathbb{E}}_n[V_{n+1}] \\ =& \frac{1}{1+r}\tilde{\mathbb{E}}_n\left[v_{n+1}\left(S_n \cdot \frac{S_{n+1}}{S_n}, M_n \vee \left(S_n \cdot \frac{S_{n+1}}{S_n}\right)\right)\right]. \end{aligned}$$

To compute the expression, we replace S_n and M_n by dummy variables s and m (the values depend only on the outcome of the first n tosses), so

$$v_n(s, m) = \frac{1}{1+r}[\tilde{p}v_{n+1}(us, m \vee (us)) + \tilde{q}v_{n+1}(ds, m \vee (ds))]$$

In this example, we have $V_3 = v_3(s, m)$, where $v_3(s, m) = m - s$, and we can use equation above recursively to obtain v_2, v_1, and v_0.
In continuous time, we shall see that the analogue of recursive equations become partial differential equations, which is called the *Feynman-Kac's Theorem.*

Theorem 6. *Let $X_0, .., X_N$ be a Markov process under the risk-neutral probability measure $\mathbb{P}$ in the binomial model. Let $v_N(x)$ be a function of dummy variable x, and consider a derivative security whose payoff at time N is $v_N(X_N)$. Then, for each n between 0 and N, the price V_n of this derivative security is some function v_n of X_n, i.e.*

$$V_n = v_n(X_n), n = 0, 1, ..., N.$$

There is a recursive algorithm for computing v_n whose exact formula depends on the underlying Markov process $X_0, ..., X_N$. Analogous results hold if the underlying Markov process is multidimensional.