Lecture 5 : Stochastic Processes I

1 Stochastic process

A stochastic process is a collection of random variables indexed by time.

An alternate view is that it is a probability distribution over a space of paths; this path often describes the evolution of some random value, or system, over time. In a deterministic process, there is a fixed trajectory (path) that the process follows, but in a stochastic process, we do not know apriori which path we will be given. One should not regard this as having no information of the path since the information on the path is given by the probability distribution. For example, if the probability distribution is given as one path having probability one, then this is equivalent to having a deterministic process. Also, it is often interpreted that the process evolves over time. However, from the formal mathematical point of view, a better picture to have in mind is that we have some underlying (unknown) path, and are observing only the initial segment of this path.

For example, the function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ given by $f(t) = t$ is a deterministic process, but a 'random function' $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ given by $f(t) = t$ with probability $1/2$ and $f(t) = -t$ with probability $1/2$ is a stochastic process. This is a rather degernerate example and we will later see more examples of stochastic processes.

We are still dealing with a single basic experiment that involves outcomes goverened by a probability law. However, the newly introduced time variable allows us to ask many new interesting questions. We emphasize on the following topics:

(a) We tend to focus on the dependencies in the sequence of values generated by the process. For example, how do future prices of a stock depend on past values?

(b) We are interested in long-term averages involving the entire sequence of generated values. For example, what is the fraction of time that a machine is idle?

(c) We are interested in boundary events. For example, what is the probability that within a given hour all circuits of some telephone system become simultaneously busy?

A stochastic process has discrete-time if the time variable takes positive integer values, and continuous-time if the time variable takes postivie real values. We start by studying discrete time stochastic processes. These processes can be expressed explicitly, and thus are more 'tangible', or 'easy to visualize'. Later we address continuous time processes.

2 Simple random walk

Let $Y_1, Y_2, \cdots$ be i.i.d. random variables such that $Y_i = \pm 1$ with equal probability. Let $X_0 = 0$ and

$$X_k = Y_1 + \cdots + Y_k,$$

for all $k \geq 1$. This gives a probability distribution over the sequences $\{X_0, X_1, \cdots, \}$, and thus defines a discrete time stochastic process. This process is known as the *one-dimensional simple random walk*, which we conveniently refer to as *random walk* from now on.

By the central limit theorem, we know that for large enough n, the distribution of $\frac{1}{\sqrt{n}} X_n$ converges to the normal distribution with mean 0 and variance 1. This already tells us some information about the random walk. We state some further properties of the random walk.

Proposition 2.1. *(i)* $\mathbb{E}[X_k] = 0$ *for all* k.

(ii) (Independent increment) For all $0 = k_0 \leq k_1 \leq \cdots \leq k_r$*, the random variables* $X_{k_{i+1}} - X_{k_i}$ *for* $0 \leq i \leq r-1$ *are mutually independent.*

(iii) (Stationary) For all $h \geq 1$ *and* $k \geq 0$*, the distribution of* $X_{k+h} - X_k$ *is the same as the distribution of* X_h*.*

Proof. The proofs are straightforward and are left as an exercise. Note that these properties hold as long as the increments Y_i are identical and independent and have mean 0. $\square$

Example 2.2. (i) Suppose that a gambler plays the following game. At each turn the dealer throws an un-biased coin, and if the outcome is head the gambler wins \$1, while if it is head she loses \$1. If each coin toss is independent, then the balance of the gambler has the distribution of the simple random walk.

(ii) Random walk can also be used as a (rather inaccurate) model of stock price.

For two positive integers A and B, what is the probability that the random walk reaches A before it reaches $-B$? Let τ be the first time at which the random walk reaches either A or $-B$. Then $X_\tau = A$ or $-B$. Define

$$f(k) = \mathbf{P}(X_\tau = A \mid X_0 = k),$$

and note that our goal is to compute $f(0)$. The recursive formula $f(k) = \frac{1}{2}f(k+1) + \frac{1}{2}f(k-1)$ follows from considering the outcome of the first coin-toss. We also have the boundary conditions $f(A) = 1$, $f(-B) = 0$. If we let $f(-B+1) = \alpha$, then it follows that $f(-B+r) = \alpha r$ for all $r \leq A+B$. Therefore, $\alpha = \frac{1}{A+B}$, and it follows that

$$f(0) = \frac{B}{A+B}.$$

3 Markov Chain

One important property of the simple random walk is that the effect of the past on the future is summarized only by the current state, rather than the whole history. In other words, the distribution of X_{k+1} depended only on the value of X_k, not on the whole set of values of $X_0, X_1, \cdots, X_k$. A stochastic process with such property is called a *Markov chain.*

More formally, let $X_0, X_1, \cdots$ be a discrete-time stochastic process where each X_i takes value in some discrete set S (note that this is not the case in the simple random walk). The set S is called the *state space.* We say that the stochastic process has the *Markov property* if

$$\mathbf{P}(X_{n+1} = i \mid X_n, X_{n-1}, \cdots, X_0) = \mathbf{P}(X_{n+1} = i \mid X_n)$$

for all $n \geq 0$ and $i \in S$. We will discuss the case when S is a finite set. In this case, we let $S = [m]$ for some positive integer m.

A stochastic process with the Markov property is called a Markov chain. Note that a finite Markov chain can be described in terms of the *transition probabilities*

$$p_{ij} = \mathbf{P}(X_{n+1} = j \mid X_n = i) \quad i, j \in S.$$

One can easily see that

$$\sum_{j \in S} p_{ij} = 1 \quad \forall i \in S.$$

All the elements of a Markov chain model can be encoded in a *transition probability matrix*

$$A = \begin{pmatrix} p_{11} & p_{21} & \cdots & p_{m1} \\ p_{12} & p_{22} & \cdots & p_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1m} & p_{2m} & \cdots & p_{mm} \end{pmatrix}.$$

Note that the sum of each column is equal to one.

Example 3.1. (i) A machine can be either working or broken on a given day. If it is working, it will break down in the next day with probability 0.01, and will continue working with probability 0.99. If it breaks down on a given day, it will be repaired and be working in the next day with probability 0.8, and will continue to be broken down with probability 0.2. We can model this machine by a Markov chain with two states: working, and broken down. The transition probability matrix is given by

$$\begin{bmatrix} 0.99 & 0.8 \\ 0.01 & 0.2 \end{bmatrix}.$$

(ii) A simple random walk is an example of a Markov chain. However, there is no transition probability matrix associated with the simple random walk since the sample space is of infinite cardinatlity.

Let $r_{ij}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$ be the *n-th step transition probabilities.* These probabilities satisfy the recurrence relation

$$r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)p_{kj} \quad for\, n > 1,$$

where $r_{ij}(1) = p_{ij}$. Hence the n-step transition probability matrix can easily be shown to be A^n.

A *stationary distribution* of a Markov chain is a probability distribution over the state space S (where $\mathbf{P}(X_0 = j) = \pi_j$) such that

$$\pi_j = \sum_{k=1}^{m} \pi_k \cdot p_{kj} \quad (\forall j \in S).$$

Example 3.2. Let $S = \mathbb{Z}_n$ and $X_0 = 0$. Consider the Markov chain $X_0, X_1, X_2, \cdots$ such that $X_{n+1} = X_n + 1$ with proability $\frac{1}{2}$ and $X_{n+1} = X_n - 1$ with probability $\frac{1}{2}$. Then the stationary distribution of this Markov chain is $\pi_i = \frac{1}{n}$ for all i.

Note that the vector $(\pi_1, \pi_2, \cdots, \pi_m)$ is an eigenvector of A with eigenvalue 1. Hence the following theorem can be deduced from the Perron-Frobenius theorem.

Theorem 3.3. *If $p_{ij} > 0$ for all $i, j \in S$, then there exists a unique stationary distribuion of the system. Moreover,*

$$\lim_{n\to\infty} r_{ij}(n) = \pi_j, \quad \forall i, j \in S.$$

A corresponding theorem is not true if we consider infinite state spaces.

4 Martingale

Definition 4.1. A discrete-time stochastic process $\{X_0, X_1, \cdots\}$ is a *martingale* if

$$X_t = \mathbb{E}[X_{t+1}|\mathcal{F}_t],$$

for all $t \geq 0$, where $\mathcal{F}_t = \{X_0, \cdots, X_t\}$ (hence we are conditioning on the initial segment of the process).

This says that our expectated gain in the process is zero at all times. We can also view this definition as a Mathematical formalization of a game of chance being fair.

Proposition 4.2. *For all $t \geq s$, we have $X_s = \mathbb{E}[X_t|\mathcal{F}_s]$.*

Proof. This easily follows from induction. □

Example 4.3. (i) Random walk is a martingale.

(ii) The balance of a roulette player is not a martingale (we always have $X_k > \mathbb{E}[X_{k+1}|\mathcal{F}_k]$).

(iii) Let $Y_1, Y_2, \cdots$ be i.i.d. random variables such that $Y_i = 2$ with probability $\frac{1}{3}$ and $Y_i = \frac{1}{2}$ with probability $\frac{2}{3}$. Let $X_0 = 0$, $X_k = \prod_{i=1}^k Y_i$. Then $\{X_0, X_1, \cdots\}$ forms a martingale.

5 Optional stopping theorem

Definition 5.1. (Stopping time) Given a stochastic process $\{X_0, X_1, \cdots\}$, a non-negative integer-valued random variable τ is called a *stopping time* if for every integer $k \geq 0$, the event $\tau \leq k$ depends only on the events $X_0, X_1, \cdots, X_k$.

Example 5.2. (i) In the coin toss game, consider a gambler who bets \$1 all the time. Let τ be the first time at which the balance of the gambler becomes \$100. Then τ is a stopping time.

(ii) Consider the same gambler as in (i). Let τ be the time of the first peak (local maximum) of the balance of the gambler. Then τ is not a stopping time.

Theorem 5.3. *(Doob's optional stopping time theorem, weak form) Suppose that $X_0, X_1, X_2, \cdots$ is a martingale sequence and τ is a stopping time such that $\tau \leq T$ for some constant T. Then $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.*

Proof. Note that

$$X_\tau = X_0 + \sum_{i=0}^{T-1} (X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}}.$$

(we used the fact $\tau \leq T$). Since T is a constant, by linear of expectation we have

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0] + \sum_{i=0}^{T-1} \mathbb{E}\Big[(X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big].$$

The main observation is that $\tau \geq i+1$ is determined by $X_0, X_1, \cdots, X_i$. Hence

$$\begin{aligned}
\mathbb{E}\Big[(X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big] &= \mathbb{E}\Big[\mathbb{E}\left[(X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}} \,|\, \mathcal{F}_i\right]\Big] \\
&= \mathbb{E}\Big[\left(\mathbb{E}[X_{i+1} \,|\, \mathcal{F}_i] - X_i\right) \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big] \\
&= \mathbb{E}\Big[0 \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big] = 0.
\end{aligned}$$

Hence $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$. $\square$

The condition can be further weakened (see [4]). The lesson to learn is that 'a mortal being has no winning strategy (when the game is fair)'. On the other hand, if one has some advantage over an opponent in some game, then no matter how small that advantage is, he/she will win in the long run.

Exercise 5.4. In the coin toss game, consider the following strategy. The gambler stops playing the first time at which the balance becomes \$100. Note that by definition, we have $\mathbb{E}[X_\tau] = 100$. Does this contradict the optional stopping theorem?

Optional stopping time theorem can be used to deduce interesting facts.

Exercise 5.5. For two positive integers a and b , consider the following strategy for the coin toss game. A player stops at the first time the balance equals either to a or to $-b$. Let this time be τ. What is the probability distribution of X_τ? (i.e., what are the probabilities that $X_\tau = a$ and $X_\tau = b$?)

References

[1] S. Ross, A first course in probability

[2] D. Bertsekas, J. Tsitsiklis, Introduction to probability

[3] L. Bachelier, Théorie de la spéculation, Annales Scientifiques de l'École Normale Supérieure, **3**, 21-86.

[4] R. Durrett, Probability: Theory and Examples, 3rd edition.

[5] S.R.Srinivasa Varadhan, Lecture Note (http://www.math.nyu.edu/faculty/varadhan/spring06/spring06.1.pdf)

18.S096 Topics in Mathematics with Applications in Finance
Fall 2013