Lecture 9 :

The Convergence Theorems

Assumptions:

I : Irreducible
A: Aperiodic
R: Recurrent
S: existence of a stationary distribution $\vec{\pi}$

The following theorem answers Q3 from Lecture 8.

Theorem 9.1. (Convergence Theorem) If I, A & S hold, then

$$\lim_{n\to\infty} [P^n]_{xy} = \pi_y \quad , \forall\, x,y \in \mathcal{X} .$$

Recall (Convergence of Random Variables)

①. A sequence of random variables $\{X_n\}_{n\geq 0}$ converges in probability towards the random variable X if

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0 .$$

This is denoted as $X_n \xrightarrow{P} X$.

②. A sequence of random variables $\{X_n\}_{n\geq 0}$ converges almost surely towards the random variable X if

$$\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1 .$$

This is denoted as $X_n \xrightarrow{a.s.} X$.

③. Almost sure convergence implies convergence in probability.

Let $N_n(y) := \#$ of visits to y before time n

Theorem 9.2. (Asymptotic Frequency). If I & R hold, then

$$\frac{N_n(y)}{n} \xrightarrow{a.s.} \frac{1}{\mathbb{E}_y \tau_y}, \quad \forall y \in \mathcal{X}.$$

Theorem 9.3. If I & S hold, then $\vec{\pi}_y = \frac{1}{\mathbb{E}_y \tau_y}, \quad \forall y \in \mathcal{X}$.

Corollary 9.1 If I holds, then existence of stationary distributions implies uniqueness.

Corollary 9.1 answers Q2 from Lecture 8.

Corollary 9.2. If I, R & S hold, then

$$\frac{N_n(y)}{n} \xrightarrow{a.s.} \vec{\pi}_y, \quad \forall y \in \mathcal{X}.$$

Theorem 9.4. If I & S hold, and $f : \mathcal{X} \to \mathbb{R}$ has $\sum_{x \in \mathcal{X}} |f(x)| \, \vec{\pi}_x < \infty$, then

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \xrightarrow{a.s.} \sum_{x \in \mathcal{X}} f(x) \cdot \vec{\pi}_x = \mathbb{E}_{X \sim \pi}[f(X)].$$

Ex1 (Inventory chain) We have a TV shop with the potential for sales of 0, 1, 2, 3 of these units each day with probability 0.3, 0.4, 0.2, 0.1. Each night at the close of business, if the number of units $\le s$, new units can be ordered such that the number of units will be S when the store opens in the next morning. Let X_n be the number of units on hand at the end of the day n. Suppose we make \$120 profit on each unit sold but it costs \$20 a day to store each item.

Q: What is the long-run profit per day if $s=1$, $S=5$?

A: The transition matrix

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 0 & 0.1 & 0.2 & 0.4 & 0.3 \\ 0 & 0 & 0.1 & 0.2 & 0.4 & 0.3 \\ 0.3 & 0.4 & 0.3 & 0 & 0 & 0 \\ 0.1 & 0.2 & 0.4 & 0.3 & 0 & 0 \\ 0 & 0.1 & 0.2 & 0.4 & 0.3 & 0 \\ 0 & 0 & 0.1 & 0.2 & 0.4 & 0.3 \end{bmatrix} \end{array}$$

why?

Notice that this chain is irreducible, aperiodic, recurrent and has a unique stationary distribution

$$\vec{\pi} = \frac{1}{9740}(885, 1516, 2250, 2100, 1960, 1029).$$

We first compute the average number of units sold per day. Note that average demand is

$$\mathbb{E}(\text{Demand}) = 0.4 \times 1 + 0.2 \times 2 + 0.1 \times 3 = 1.1.$$

And average number of lost sales per day is

$$\pi(2) \times \mathbb{P}(\text{Demand} = 3) \times (3-2) = \frac{225}{9740} = \frac{45}{1948}.$$

Therefore, the average number of sales per day is

$1.1 - \frac{45}{1948} = \frac{20978}{19480} = \frac{10489}{9740}$, with a profit

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

$$\frac{10489}{9740} \times \$120 = \frac{62934}{487}.$$

Taking $f(x) = 20x$ in Theorem 9.4 (notice that $\sum_{x \in \mathcal{X}} |f(x)| \vec{\pi}_x < \infty$), we get the average hold costs per day is

$$\frac{1}{9740} \times (1516 \times 20 + 2250 \times 40 + 2100 \times 60 + 1960 \times 80 + 1029 \times 100)$$

$$= \frac{25301}{487}.$$

Thus, the long-run net profit for $s = 1, S = 5$ is

$$\frac{62934}{487} - \frac{25301}{487} = \frac{37633}{487} = 77.275.$$

Recall: Weak Law of Large Numbers (WLLN)

Let $\{X_k\}_{k\in\mathbb{N}}$ be i.i.d. Lebesgue integrable random variables with $\mathbb{E}(X_k)=\mu$, then the sample average converges in probability to the expected value:

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow{P} \mu, \quad \text{as } n \to \infty.$$

That is, $\lim\limits_{n\to\infty} \mathbb{P}\left(\left| \frac{1}{n}\sum_{k=0}^{n-1} X_k - \mu \right| > \varepsilon \right) = 0, \ \forall \varepsilon > 0.$

Recall: Strong Law of Large Numbers (SLLN)

Let $\{X_k\}_{k\in\mathbb{N}}$ be i.i.d. Lebesgue integrable random variables with $\mathbb{E}(X_k)=\mu$, then the sample average converges almost surely to the expected value:

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow{a.s.} \mu, \quad \text{as } n \to \infty.$$

That is, $\mathbb{P}\left(\lim\limits_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} X_k = \mu \right) = 1.$

Proof of Theorem 9.2.

Suppose we start at $y \in \mathcal{X}$, define $R(k) := \tau_y^k = \min\{n \geq 1 : N_n(y) = k\}$ as the time of k-th return.

Since the times between returns are i.i.d., the Strong Law of Large Numbers tells

$$\frac{1}{k} \cdot R(k) \xrightarrow{a.s.} \mathbb{E}_y \tau_y \leq \infty, \text{ as } k \longrightarrow \infty.$$

Thus, $\mathbb{P}\left(\omega \in \Omega : \lim_{k\to\infty} \frac{1}{k} \cdot R(k)(\omega) = \mathbb{E}_y \tau_y \right) = 1$.

From the definition of $R(k)$, it follows that

$$R(N_n(y)) \leq n < R(N_n(y)+1).$$

Thus, $\dfrac{R(N_n(y))}{N_n(y)} \leq \dfrac{n}{N_n(y)} < \dfrac{R(N_n(y)+1)}{N_n(y)+1} \cdot \dfrac{N_n(y)+1}{N_n(y)}$.

Taking $n \longrightarrow \infty$, then $N_n(y) \longrightarrow \infty$ (because y is recurrent), and for any $\omega \in \Omega$ such that

$$\lim_{k\to\infty} \frac{1}{k} \cdot R(k)(\omega) = \mathbb{E}_y \tau_y,$$

the squeeze theorem implies

$$\lim_{n\to\infty} \frac{n}{N_n(y)}(\omega) = \lim_{n\to\infty} \frac{R(N_n(y))}{N_n(y)}(\omega) = \lim_{k\to\infty} \frac{R(k)}{k}(\omega) = \mathbb{E}_y \tau_y ,$$

and thus $\lim_{n\to\infty} \frac{N_n(y)}{n}(\omega) = \frac{1}{\mathbb{E}_y \tau_y}$.

Therefore,

$$\mathbb{P}\left(\omega\in\Omega : \lim_{n\to\infty} \frac{N_n(y)}{n}(\omega) = \frac{1}{\mathbb{E}_y \tau_y}\right) = 1 ,$$

i.e.,

$$\frac{N_n(y)}{n} \xrightarrow{a.s.} \frac{1}{\mathbb{E}_y \tau_y} .$$

$\square$

Remark 9.1. If the chain starts from $x \neq y$, define

$\alpha_0 :=$ the time of first visit to y, and for $i \geq 1$,

$\alpha_i :=$ the time of first return to y starting from α_{i-1}.

Then $\alpha_0, \alpha_1, \alpha_2, \cdots$ are independent and $\alpha_1, \alpha_2, \cdots$ are identical distributed. And the Strong Law of Large Numbers would imply the same result.

This is the end of this lecture!