Additional Materials: Convergence of Random Variables

Sample Paths and Random Variables — Revisited

Sample paths

When we observe a stochastic process, we are actually running a random experiment, where the sample space Ω of the experiment is the set of all possible outcomes. For example:

- In a discrete time Markov chain with state space $\mathcal{S}$, the sample space Ω is the set of all sequences $\{x_i\}_{i=1}^{\infty}$ with $x_i \in \mathcal{S}$.
- In a continuous time Markov chain with state space $\mathcal{S}$, the sample space Ω is the set of all right-continuous functions $x(t)$ which map $t \in [0, \infty) \to \mathcal{S}$.

A single element ω of Ω is known as a sample path of the process. There are uncountably many sample paths (even if the state space is finite). Therefore, in general the probability of any given sample path occurring is 0, just like the probability of picking a real number randomly between 0 and 1. There are also sets of sample paths that have probability zero.

Random variables

A random variable defined on a stochastic process is a mapping from Ω to $\mathbb{R}$ (or sometimes a subset of $\mathbb{R}$). For example:

- In a discrete-time Markov chain with state space $\mathcal{S} = \{0, 1, \ldots, n\}$, the function

$$\begin{aligned} X_j &: \Omega \to \mathbb{R} \\ (x_i)_{i=1}^{\infty} &\to x_j \end{aligned}$$

  is the random variable that gives the state at time point j.

- In a continuous-time Markov chain with state space $\mathcal{S} = \{0, 1, \ldots, n\}$, the function

$$\begin{aligned} X_s &: \Omega \to \mathbb{R} \\ x(t) : t \in [0, \infty) &\to x_s \end{aligned}$$

  is the random variable that gives the state at time point s.

Types of Convergences

Let $Y_1, Y_2, \ldots$ be a sequence of random variables defined on the sample space of a stochastic process and Y be another random variable defined on the same sample space. When we consider the convergence of the sequence of random variables Y_n to Y (written $\lim_{n\to\infty} Y_n = Y$), we could mean one of at least three things.

1. **Convergence with probability 1 (almost sure convergence):** If there is a set $\Omega^* \subseteq \Omega$ that has probability 1 (which may not contain all the sample paths as we saw before) such that

$$\lim_{n\to\infty} Y_n(\omega) \to Y(\omega) \quad \text{for all } \omega \in \Omega^*$$

then we say

$$\lim_{n\to\infty} Y_n = Y \text{ with probability } 1.$$

2. **Convergence in probability:** If for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr\left[\omega : \left|Y_n(\omega) - Y(\omega)\right| > \epsilon\right] = 0,$$

then we say $\lim_{n\to\infty} Y_n = Y$ in probability.

3. **$\mathcal{L}_p$ – convergence, or convergence in the pth mean:** The set of random variables Y_n converges to the random variable Y in the pth mean or $\mathcal{L}_p$ if

$$\lim_{n\to\infty} \mathbb{E}\left[\left|Y_n - Y\right|^p\right] = 0,$$

where the expectation is taken with respect to the probability measure associated with the sample paths.

The most important case is when $p = 2$, when we say $Y_n \to Y$ in $\mathcal{L}_2$ (or mean square).

The three types of convergence are related as follows:

- convergence with probability 1 $\longrightarrow$ convergence in probability
- convergence in $\mathcal{L}_p$ $\longrightarrow$ convergence in probability.

Convergence in probability is weaker than the other two, which do not imply each other.

Laws of Large Numbers

Theorem (Weak law of large numbers.). *Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables and suppose that $\mathbb{E}[Y_i] = \mu < \infty$. Then, the sequence of random variables*

$$\overline{Y}(n) = \frac{Y_1 + Y_2 + \cdots + Y_n}{n} \qquad \textit{converges in probability to } \mu.$$

That is, for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr\left[|\overline{Y}(n) - \mu| > \varepsilon\right] = 0.$$

Theorem (Strong law of large numbers.). *Let $Y_1, Y_2, \ldots$ be a sequence of independent, identically distributed random variables and suppose that $\mathbb{E}[Y_i] = \mu < \infty$. Then, the sequence of random variables*

$$\overline{Y}(n) = \frac{Y_1 + Y_2 + \cdots + Y_n}{n} \qquad \textit{converges to } \mu \textit{ with probability } 1.$$

That is,

$$\Pr\left[\lim_{n\to\infty} \overline{Y}(n) = \mu\right] = 1.$$