Lecture-17: Random Processes

1 Stochastic Processes

Let $(\Omega, \mathcal{F}, P)$ be a probability space. For an arbitrary index set T and state space $\mathcal{X} \subseteq \mathbb{R}$, a **random process** is a measurable map $X : (\Omega, T) \to \mathcal{X}$. For each $t \in T$, we have $X_t \triangleq \{X(t,\omega) : \omega \in \Omega\}$ is a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$, and random process X is a collection of random variables $X = (X_t \in \mathcal{X} : t \in T)$. For each $\omega \in \Omega$, we have a sample path $X_\omega \triangleq (X_t(\omega) : t \in T)$ of the process X.

1.1 Classification

State space $\mathcal{X}$ can be countable or uncountable, corresponding to discrete or continuous valued process. If the index set T is countable, the stochastic process is called **discrete**-time stochastic process or random sequence. When the index set T is uncountable, it is called **continuous**-time stochastic process. The index set T doesn't have to be time, if the index set is space, and then the stochastic process is spatial process. When $T = \mathbb{R}^n \times [0,\infty)$, stochastic process $X(t)$ is a spatio-temporal process.

Example 1.1. We list some examples of each such stochastic process.

i. Discrete random sequence: brand switching, discrete time queues, number of people at bank each day.

ii. Continuous random sequence: stock prices, currency exchange rates, waiting time in queue of nth arrival, workload at arrivals in time sharing computer systems.

iii. Discrete random process: counting processes, population sampled at birth-death instants, number of people in queues.

iv. Continuous random process: water level in a dam, waiting time till service in a queue, location of a mobile node in a network.

1.2 Specification

To define a measure on a random process, we can either put a measure on sample paths, or equip the collection of random variables with a joint measure. We are interested in identifying the joint distribution $F : \mathbb{R}^T \to [0,1]$. To this end, for any $x \in \mathbb{R}^T$ we need to know

$$F(x) = P\left(\bigcap_{t\in T}\{\omega \in \Omega : X_t(\omega) \le x_t\}\right) = P(\bigcap_{t\in T} X_t^{-1}(-\infty, x_t]) = P \circ X^{-1} \bigtimes_{t\in T}(-\infty, x_t].$$

However, even for a simple independent process with countably infinite T, any function of the above form would be zero if x_t is finite for all $t \in T$. Therefore, we only look at the values of $F(x)$ when $x_t \in \mathbb{R}$ for indices t in a finite set S and $x_t = \infty$ for all $t \notin S$. That is, for any finite set $S \subseteq T$ we focus on the product sets of the form

$$\bigtimes_{s\in S}(-\infty, x_s] \bigtimes_{s\notin S}\mathbb{R}.$$

We can define a **finite dimensional distribution** for any finite set $S \subseteq T$ and $x_S = \{x_s \in \mathbb{R} : s \in S\}$,

$$F_S(x_S) = P\left(\bigcap_{s\in S}\{\omega \in \Omega : X_s(\omega) \leq x_s\}\right) = P(\bigcap_{s\in S} X_s^{-1}(-\infty, x_s]).$$

Set of all finite dimensional distributions of the stochastic process $\{X_t : t \in T\}$ characterizes its distribution completely. Simpler characterizations of a stochastic process $X(t)$ are in terms of its moments. That is, the first moment such as mean, and the second moment such as correlations and covariance functions.

$$m_X(t) \triangleq \mathbb{E}X_t, \qquad R_X(t,s) \triangleq \mathbb{E}X_tX_s, \qquad C_X(t,s) \triangleq \mathbb{E}(X_t - m_X(t))(X_s - m_X(s)).$$

Example 1.2. Some examples of simple stochastic processes.

i_ $X_t = A\cos 2\pi t$, where A is random.

ii_ $X_t = \cos(2\pi t + \Theta)$, where Θ is random and uniformly distributed between $(-\pi, \pi]$.

iii_ $X_n = U^n$ for $n \in \mathbb{N}$, where U is uniformly distributed in the open interval $(0,1)$.

iv_ $Z_t = At + B$ where A and B are independent random variables.

1.3 Independence

Recall, given the probability space $(\Omega, \mathcal{F}, P)$, two events $A, B \in \mathcal{F}$ are **independent events** if

$$P(A \cap B) = P(A)P(B).$$

Random variables X, Y defined on the above probability space, are **independent random variables** if for all $x, y \in \mathbb{R}$

$$P\{X(\omega) \leq x, Y(\omega) \leq y\} = P\{X(\omega) \leq x\}P\{Y(\omega) \leq y\}.$$

A stochastic process X is said to be **independent** if for all finite subsets $S \subseteq T$, we have

$$P(\{X_s \leq x_s, s \in S\}) = \prod_{s\in S} P\{X_s \leq x_s\}.$$

Two stochastic process X, Y for the common index set T are **independent random processes** if for all finite subsets $I, J \subseteq T$

$$P\left(\{X_i \leq x_i, i \in I\} \cap \{Y_j \leq y_j, j \in J\}\right) = P\left(\{X_i \leq x_i, i \in I\}\right) P\left(\{Y_j \leq y_j, j \in J\}\right).$$

1.4 Conditional Expectation

Let $(\Omega, \mathcal{F}, P)$ be the probability space. Let X be a measurable random variable on this probability space denoted as $X \in \mathcal{F}$, if the event $X^{-1}(-\infty, x] = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for each $x \in \mathbb{R}$. Let $\mathcal{E} \subseteq \mathcal{F}$ be a σ-algebra, then the **conditional expectation** of X given $\mathcal{E}$ is denoted $\mathbb{E}[X|\mathcal{E}]$ and is a random variable $Y = \mathbb{E}[X|\mathcal{E}]$ where

i_ $Y \in \mathcal{E}$,

ii_ for each event $A \in \mathcal{E}$, we have $\mathbb{E}[X1_A] = \mathbb{E}[Y1_A]$.

Intuitively, we think of the σ-algebra $\mathcal{E}$ as describing the information we have. For each $A \in \mathcal{E}$, we know whether or not A has occurred. The conditional expectation $\mathbb{E}[X|\mathcal{E}]$ is then the "best guess" of the value of X given the information $\mathcal{E}$. Let X, Y be two random variables defined on this probability space. Then, the conditional expectation of X given Y is defined as

$$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)].$$

A random variable X is **independent** of the σ-algebra $\mathcal{E}$, if for all $x \in \mathbb{R}$ and $A \in \mathcal{E}$,

$$\mathbb{E}[1_{\{X\leq x\}}1_A] = P\{X \leq x\} \cap A = P\{X \leq x\}P(A) = \mathbb{E}1_{\{X\leq x\}}\mathbb{E}1_A.$$

Lemma 1.3. *Let $(\Omega, \mathcal{F}, P)$ be a probability space with $\mathcal{E} \subseteq \mathcal{F}$ a σ-algebra. If $X \in \mathcal{E}$ is a random variable, then $\mathbb{E}[X|\mathcal{E}] = X$.*

Proof. First condition is true by hypothesis, and the second condition holds for any $A \in \mathcal{E}$. □

Lemma 1.4. *Let $(\Omega, \mathcal{F}, P)$ be a probability space with $\mathcal{E} \subseteq \mathcal{F}$ a σ-algebra. If $X \in \mathcal{F}$ be a random variable independent of $\mathcal{E}$. Then, $\mathbb{E}[X|\mathcal{E}] = \mathbb{E}[X]$.*

Proof. This follows since $\mathbb{E}X \in \mathcal{E}$ and the random variables X and 1_A are independent for any $A \in \mathcal{E}$, which implies

$$\mathbb{E}[X1_A] = \mathbb{E}X\mathbb{E}1_A = \mathbb{E}[(\mathbb{E}X)1_A].$$

□

One can partition the state space $\mathbb{R}$ into measurable sets $E_1, E_2, \dots$ for the random variable X defined on the given probability space. Then $\Omega_i \triangleq X^{-1}(E_i)$ is a partition of the sample space Ω. Let Y be a random variable defined as the partition index for the random variable X. That is,

$$Y = \sum_{i \in \mathbb{N}} i \cdot 1_{\{X \in E_i\}}.$$

Let $\mathcal{E} \triangleq \sigma(\Omega_1, \Omega_2, \dots)$, then one can check that $Y \in \mathcal{E}$ or $\sigma(Y) = \mathcal{E}$. Hence, $\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X|\mathcal{E}]$. Clearly, $\mathbb{E}[X|Y]$ would be a function of Y and since Y takes countably many values, we have $Z = \mathbb{E}[X|Y]$ taking countably many values, with $Z_i = Z1_{\{Y=i\}}$ being a constant on the corresponding partition Ω_i of the sample space. One can compute this conditional expectation using joint distribution directly as

$$\mathbb{E}[X|Y = i] = \int_{\mathbb{R}} x dF_{X|Y=i}(x) = \frac{1}{P(\Omega_i)} \int_{E_i} x dF(x) = \frac{\mathbb{E}X1_{\Omega_i}}{P(\Omega_i)}$$

Lemma 1.5. *Suppose $\{\Omega_i : i \in \mathbb{N}\}$ be a countable partition of the sample space Ω, and $\mathcal{E} = \sigma(\Omega_1, \Omega_2, \dots)$ is the σ-field generated by this partition. Then,*

$$\mathbb{E}[X|\mathcal{E}] = \frac{\mathbb{E}[X1_{\Omega_i}]}{P(\Omega_i)} \text{ on } \Omega_i.$$

Proof. It is easy to see that the RHS is constant on each partition Ω_i and hence is measurable with respect to $\mathcal{E}$. Further, for each $\Omega_i \in \mathcal{E}$, we have

$$\int_{\Omega_i} \frac{\mathbb{E}[X1_{\Omega_i}]}{P(\Omega_i)} dP = \mathbb{E}[X1_{\Omega_i}] = \int_{\Omega_i} X dP.$$

□

Corollary 1.6. $P(A|B)P(B) = P(A \cap B)$.

Proof. Taking $X = 1_A$ and $\mathcal{E} = \{\emptyset, \Omega, B, B^c\}$, from the previous Lemma we get

$$P(A|B) = \mathbb{E}[1_A|1_B] = \mathbb{E}[1_A|\mathcal{E}] = \frac{\mathbb{E}[1_A 1_B]}{P(B)} = \frac{P(A \cap B)}{P(B)}.$$

□

Theorem 1.7 (Bayes' Formula). *For a σ-algebra $\mathcal{E} \subseteq \mathcal{F}$, and for any events $G \in \mathcal{E}$ and $A \in \mathcal{F}$, we have*

$$P(G|A) = \frac{\mathbb{E}[1_G P(A|\mathcal{E})]}{\mathbb{E}P(A|\mathcal{E})}.$$

Proof. It is easy to check that numerator is $\mathbb{E}1_G \mathbb{E}[1_A|\mathcal{E}] = \mathbb{E}[1_{A \cap G}|\mathcal{E}]$. It suffices to show that $\mathbb{E}\mathbb{E}[1_A|\mathcal{E}] = \mathbb{E}1_A$, which follows from definition. □

Corollary 1.8. *For the countable partition $(\Omega_1, \Omega_2, \dots)$, if the σ-algebra $\mathcal{E} = \sigma(\Omega_1, \Omega_2, \dots)$, then for any events $G \in \mathcal{E}$ and $A \in \mathcal{F}$, we have*

$$P(\Omega_i|A) = \frac{P(A|\Omega_i)P(\Omega_i)}{\sum_{j \in \mathbb{N}} P(A|\Omega_j)P(\Omega_j)}.$$

Proof. Result follows from the fact that $P(A|\mathcal{E}) \in \mathcal{E}$ and hence is a constant on each partition Ω_j. □

1.5 Filtration

A net of σ-algebras $\mathscr{F} = \{\mathcal{F}_t \subseteq \mathcal{F} : t \in T\}$ is called a **filtration** when the index set T is totally ordered and the net is non-decreasing, that is for all $s \leqslant t \in T$ implies $\mathcal{F}_s \subseteq \mathcal{F}_t$. Consider a random process X indexed by the ordered set T on the probability space $(\Omega, \mathcal{F}, P)$. The process X is called **adapted** to the filtration $\mathscr{F}$, if for each $t \in T$, we have the random variable $X_t \in \mathcal{F}_t$. For a random process X with an ordered index set T, we can define a natural filtration $\mathscr{F} = \{\mathcal{F}_t \subseteq \mathcal{F} : t \in T\}$ indexed by T, where $\mathcal{F}_t \triangleq \sigma(X_s, s \leqslant t)$ is the information about the process till index t and the process X is adapted to its natural filtration by definition.

If $X = (X_t : t \in T)$ is an independent process with the associated natural filtration $\mathscr{F}$, then for any $t > s$ and events $A \in \mathcal{F}_s$, X_t is independent of the event A. This is just a fancy way of saying X_t is independent of $(X_u, u \leqslant s)$. Hence, for any random variable $Y \in \mathcal{F}_s$, we have

$$\mathbb{E}[\mathbb{E}[X_t Y | \mathcal{F}_s]] = \mathbb{E}[\mathbb{E}[X_t]Y] = \mathbb{E}X_t \mathbb{E}Y.$$