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 \(X \subseteq \mathbb{R}\), a random process is a measurable map \(X : (\Omega, T) \rightarrow X\). 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 \(X = (X_t \in X : t \in T)\). 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 \(X\) 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 \(T = \mathbb{R}^d \times [0, \infty)\), 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 \(n\)th 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 \(F : \mathbb{R}^T \rightarrow [0, 1]\). To this end, for any \(x \in \mathbb{R}^T\) we need to know

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

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

\[
\times_{s \in S} \times_{s \notin S} \mathbb{R}.
\]

1
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.

\[
m_X(t) \triangleq \mathbb{E}X_t, \quad R_X(t, s) \triangleq \mathbb{E}X_tX_s, \quad C_X(t, s) \triangleq \mathbb{E}(X_t - m_X(t))(X_s - m_X(s)).
\]

**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 \( \Theta \) 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 \( I, J \subseteq T \)

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

### 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 \( \sigma \)-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 \( \sigma \)-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 \( \sigma \)-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 \(\sigma\)-algebra. If \(X \in \mathcal{E}\) is a random variable, then \(E[X|\mathcal{E}] = X\).

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 \(\sigma\)-algebra. If \(X \in \mathcal{F}\) be a random variable independent of \(\mathcal{E}\). Then, \(E[X|\mathcal{E}] = E[X]\).

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

\[
E[X1_A] = EXE1_A = E(E[X]1_A).
\]

One can partition the state space \(\mathbb{R}\) into measurable sets \(E_1, E_2, \ldots\) 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 \(\Omega\). 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 \(\mathcal{E} \triangleq \sigma(\Omega_1, \Omega_2, \ldots)\), then one can check that \(Y \in \mathcal{E}\) or \(\sigma(Y) = \mathcal{E}\). Hence, \(E[X|Y] = E[X|\sigma(Y)] = E[X|\mathcal{E}]\). Clearly, \(E[X|Y]\) would be a function of \(Y\) and since \(Y\) takes countably many values, we have \(Z = E[X|Y]\) taking countably many values, with \(Z_i = Z1_{\{Y = i\}}\) being a constant on the corresponding partition \(\Omega_i\) of the sample space. One can compute this conditional expectation using joint distribution directly as

\[
E[X|Y = i] = \int_{\mathbb{R}} xdF_{X|Y=i}(x) = \frac{1}{P(\Omega_i)} \int_{E_i} xdF(x) = \frac{EX1_{\Omega_i}}{P(\Omega_i)}.
\]

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

\[
E[X|\mathcal{E}] = \frac{EX1_{\Omega_i}}{P(\Omega_i)}\] on \(\Omega_i\).

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

\[
\int_{\Omega_i} \frac{EX1_{\Omega_i}}{P(\Omega_i)} dP = E[X1_{\Omega_i}] = \int_{\Omega_i} X dP.
\]

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

\[
P(A|B) = E[1_A|B] = E[1_A|\mathcal{E}] = \frac{E[1_A1_B]}{P(B)} = \frac{P(A \cap B)}{P(B)}.
\]

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

\[
P(G|A) = \frac{E[1_GP(A|\mathcal{E})]}{EP(A|\mathcal{E})}.
\]

Proof. It is easy to check that numerator is \(E1_GP[1_A|\mathcal{E}] = E[1_{A \cap G}|\mathcal{E}]\). It suffices to show that \(EP[1_A|\mathcal{E}] = EP[1_A]\), which follows from definition.

Corollary 1.8. For the countable partition \((\Omega_1, \Omega_2, \ldots)\), if the \(\sigma\)-algebra \(\mathcal{E} = \sigma(\Omega_1, \Omega_2, \ldots)\), 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 \(\Omega_j\).
1.5 Filtration

A net of \(\sigma\)-algebras \(\mathcal{F} = \{\mathcal{F}_t \subseteq \mathcal{F} : t \in T\}\) is called a filtration when the index set \(T\) is totally ordered and the net is non-decreasing, that is for all \(s \leq t \in T\) implies \(\mathcal{F}_s \subseteq \mathcal{F}_t\). Consider a random process \(X\) indexed by the ordered set \(T\) on the probability space \((\Omega, \mathcal{F}, \mathcal{P})\). The process \(X\) is called adapted to the filtration \(\mathcal{F}\), if for each \(t \in T\), we have

\[\mathcal{F}_t = \sigma(X_s, s \leq t)\]

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 \(\mathcal{F}\), 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 \((X_u, u \leq s)\). Hence, for any random variable \(Y \in \mathcal{F}_s\), we have

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