AI 5090: Stochastic Processes

Lecture 13

Scribe: Eswar Venkata Sai K, A. S. Aditya

18 February 2025

Contents

The lecture introduces random processes, their interpretation, finite-dimensional distributions, stationary processes, and IID processes.

Topics Covered: Random processes, Interpretation of Random Processes, Finite-Dimensional Distributions (FDDS), Consistency of FDDS, Properties of Random Processes, Stationary Processes, IID Processes.

Note: This scribe considers $(\Omega, \mathscr{F}, \mathbb{P})$ as the underlying probability space for all definitions.

1 Random Processes

1.1 Motivation

Many real-world systems evolve over time with inherent randomness. A **random process** (also called a **stochastic process**) helps us describe how the state of a system changes probabilistically over time.
Generalizing, we can understand random process as:

- A random variable describes a single uncertain outcome.
- A random process extends this idea to describe a sequence or collection of random variables indexed by a set. (This set is often represented as time in most real-time applications).

Stochastic processes are ubiquitous, and are encountered commonly in signal processing, robot motion planning, and other related domains.

1.2 Definition

We now formally define a random process.

Definition 1. *Consider an index set $\mathcal{T}$. The collection of random variables $\{X_t : t \in \mathcal{T}\}$ indexed by elements of $\mathcal{T}$ and all defined w.r.t. $\mathscr{F}$ is called a random process.*

Point to Remember: The cardinality of $\mathcal{T}$ is equal to the number of random variables in the collection.

Note:

- If $\mathcal{T}$ is finite, say $\mathcal{T} = \{1, 2, ..., d\}$, then the random process is a d-dimensional vector of random variables, represented by $X = [X_1, ..., X_d]^\top$.
- If $\mathcal{T}$ is a countably infinite set, we say that $\{X_t : t \in \mathcal{T}\}$ is a discrete-parameter process or a discrete-time process.
- If $\mathcal{T}$ is uncountably infinite, we say that $\{X_t : t \in \mathcal{T}\}$ is a continuous-parameter process or a continuous-time process.

1.3 Interpretation of Random Processes

We can try to interpret the definition of the random process in the following different ways.

- A random process can be viewed as a collection of random variables indexed by a set $\mathcal{T}$, represented as $X_t : t \in \mathcal{T}$, where each X_t is a random variable defined as a function $X_t : \Omega \to \mathbb{R}$.
- Another perspective is to fix $\omega \in \Omega$ and consider $X(\omega)$ as a function of $\mathcal{T}$. This can be expressed as:

$$X_{\cdot}(\omega) : \mathcal{T} \to \mathbb{R}$$

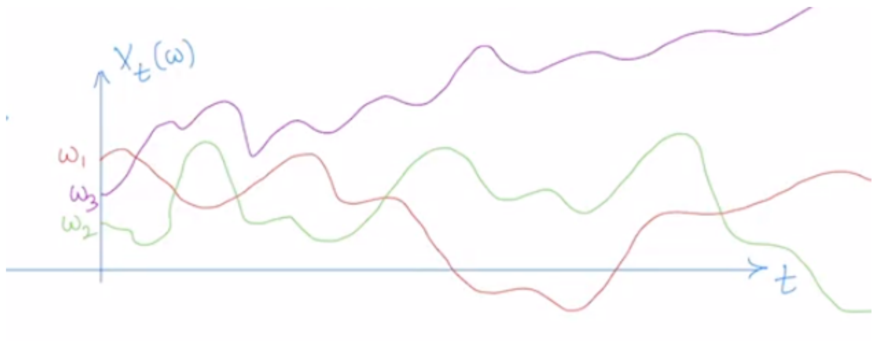

Figure 1: Pictorial representation of sample path of a random process.

- Each line in the above picture represents a sample path $\{X_t(\omega) : t \in \mathcal{T}\}$ for a fixed $\omega \in \Omega$.
- The random process considered in the above picture is a continuous-parameter process.

- A more general way to describe a random process is as a function of the Cartesian product $\mathcal{T} \times \Omega$, represented as:

$$X : \mathcal{T} \times \Omega \to \mathbb{R}$$

Note: For our discussion in this scribe and in the following lectures, we will consider the index set $\mathcal{T}$ to be one of

$$\mathbb{R}_+, \mathbb{R}, \mathbb{Z}, \mathbb{N}.$$

1.4 Finite Dimensional Distributions

Definition 2 (Finite Dimensional Distributions)**.** *Let $X_t : t \in \mathcal{T}$ be a random process w.r.t. $\mathscr{F}$. Given $n \in \mathbb{N}$ and* $\mathbf{t} = (t_1, t_2, ..., t_n) \in \mathcal{T}^n$*, the joint CDF of $X_{t_1}, X_{t_2}, ..., X_{t_n}$ is given by*

$$F_{\mathbf{t}}(\underline{x}) = \mathbb{P}(X_{t_1} \leq x_1, X_{t_2} \leq x_2, ..., X_{t_n} \leq x_n), \qquad \underline{x} = (x_1, x_2, x_3, ..., x_n) \in \mathbb{R}^n$$

The collection of all possible joint CDFs is known as finite dimensional distribution(FDDs) of a random process $\{X_t : t \in \mathcal{T}\}$, and is represented as

$$\{F_t : n \in \mathbb{N}, t \in \mathcal{T}_n\}$$

Point to Remember: FDDs are not simply an arbitrary collection of joint CDFs. They must satisfy a certain "consistency" property outlined in the following example.

Example 1.1. *Suppose $\mathcal{T} = \{1, 2, 3\}$. The FDDs are given by*

$$\left\{F_1, F_2, F_3, F_{1,2}, F_{1,3}, F_{2,3}, F_{1,2,3}\right\}.$$

We can also observe that $F_2(x) = F_{1,2}(\infty, x)$ and $F_{2,3}(x, y) = F_{1,2,3}(\infty, x, y)$. This also means that we could get the marginal distributions from the joint distributions.

The FDDs of any random process satisfy a consistency property as outlined above. We now formalise this in the below definition.

1.4.1 Consistency of Finite Dimensional Distributions

Definition 3 (Consistency of Finite Dimensional Distributions)**.** *FDDs of a process $\{X_t : t \in \mathcal{T}\}$ are said to be consistent if for any $m, n \in \mathbb{N}$ with $m < n$ and $\mathcal{T}_m \subset \mathcal{T}_n \subset \mathcal{T}$ with $|\mathcal{T}_m| = m$ and $|\mathcal{T}_n| = n$, we have*

$$F_t(x_1, x_2, ..., x_m) = F_s\underbrace{(\infty, \infty, ..., \infty, x_1, \infty, \infty, ..., \infty, x_2, ..., x_m, \infty, \infty, ..., \infty)}_{n}$$

where $t \in \mathcal{T}_m$ and $s \in \mathcal{T}_n$. All components other than those of $x_1, \ldots, x_m$ are set to ∞ on the right-hand side above.

Our interest is in the study of random processes whose FDDs are consistent. Some processes with consistent FDDs include:

- IID process, Bernoulli process, Gaussian process, Markov Chains, Poisson process, Levy process, Brownian motion, and diffusion process.

Note: In the context of random processes, the finite-dimensional distributions (FDDs) are consistent because if $|t| > |s|$ (i.e., the index set t contains more elements than s, making s a subset of t), then the distribution function F_s can be obtained from F_t by marginalizing over the indices in t that are not present in s.

1.5 Properties of Random Processes: Mean, Autocorrelation, and Autocovariance

A random process is a collection of random variables indexed by a set $\mathcal{T}$. To characterize or understand a random process, we need a statistical measure that describes its behaviour over the index set. Unlike deterministic functions, random processes exhibit uncertainity, making it essential to analyse their statistical properties rather than specific values at given index (or time). The three properties that are used to describe a random process are:

- Mean
- Autocorrelation
- Autocovariance

These quantities are formally defined below.

Definition 4 (Mean of a random process)**.** *The mean of a random process* $\{X_t : t \in \mathcal{T}\}$ *is a function* $M_X : \mathcal{T} \to [-\infty, \infty]$ *defined as*

$$M_X(t) = \mathbb{E}[X_t], \qquad t \in \mathcal{T}.$$

Definition 5 (Autocorrelation of a random process)**.** *The autocorrelation of a process* $\{X_t : t \in \mathcal{T}\}$ *is a function* $R_X : \mathcal{T} \times \mathcal{T} \to [-\infty, \infty]$ *defined as*

$$R_X(t, s) = \mathbb{E}[X_t X_s], \qquad t, s \in \mathcal{T}.$$

Points to Remember:

- The word auto in autocorrelation refers to the fact that the function measure how the process is correlated with itself at different indices.
- Autocorrelation of a process is symmetric, i.e., $R_X(t, s) = R_X(s, t)$.
- If X_t and X_s are strongly correlated, then $R_X(t, s)$ will be large.
- If X_t and X_s are uncorrelated, then $R_X(t, s) = 0$ will be small.

Definition 6 (Autocovariance of a random process)**.** *The autocovariance of a process* $\{X_t : t \in \mathcal{T}\}$ *is a function* $C_X : \mathcal{T} \times \mathcal{T} \to [-\infty, \infty]$ *defined as*

$$C_X(t, s) = \mathsf{Cov}(X_t, X_s) = \mathbb{E}[(X_t - \mathbb{E}[X_t]\,(X_s - \mathbb{E}[X_s])], \qquad t, s \in \mathcal{T}.$$

- The word auto in autocovariance refers to the fact that the function measure how the deviation from the mean of a random process are related with itself at different indices.
- Autocovariance of a process is symmetric i.e., $C_X(t, s) = C_X(s, t)$.

1.6 Stationary and Wide-Sense Stationary Processes

Different random processes exhibit different statistical behaviour and hence they can be classified based on their properties. We will be discussing about them briefly in what follows. For the ease of exposition, we set $\mathcal{T} = \mathbb{R}_+$.

Definition 7 (Stationary Process)**.** *A process $\{X_t : t \in \mathbb{R}_+\}$ is said to be stationary or strictly stationary if all its FDDs are translation invariant, i.e., for any $n \in \mathbb{N}$, $\mathbf{t} \in \mathbb{R}_+^n$ and $h \in \mathbb{R}_+$,*

$$F_{\mathbf{t}} = F_{\mathbf{t}+h}.$$

Here $\mathbf{t} + h$ is a vector with each coordinate incremented by h with respect to the corresponding coordinate in $\mathbf{t}$.

Note: If a process $\{X_t : t \in \mathbb{R}_+\}$ is stationary, then

$$F_t = F_s, \qquad \forall t, s \in \mathcal{T}.$$

Thus, all random variables in a stationary process possess the same CDF.

Definition 8 (Weakly Stationary process)**.** *A process $\{X_t : t \in \mathbb{R}_+\}$ is said to be weakly stationary or wide-sense stationary if for all $t_1, t_2 \in \mathbb{R}_+$ and $h \in R_+$:*

- $M_X(t_1) = M_X(t_2)$.
- $C_X(t_1, t_2) = C_X(t_1 + h, t_2 + h)$.

That is, for a process to be weakly stationary, the mean of the process should be a constant function, and the autocovariance should be translation invariant.

Remark 1. *A process is weakly stationary if and only if it has constant mean and $C_X(t, t+h) = C_X(0, h)$ for all $t, h \in \mathbb{R}_+$. Indeed, By definition, we note that a weakly stationary process has constant mean. Suppose now that $C_X(t_1, t_2) = C_X(t_1 + h, t_2 + h)$ for all $t_1, t_2, h \in \mathbb{R}_+$. Fix $t, \bar{h} \in \mathbb{R}_+$ arbitrarily. Substituting $t_1 = 0$, $t_2 = \bar{h}$, and $h = t$, we get*

$$C_X(0, \bar{h}) = C_X(t_1, t_2) = C_X(t_1 + h, t_2 + h) = C_X(t, t + \bar{h}).$$

Conversely, suppose that $C_X(0, h) = C_X(t, t + h)$ for all $t, h \in \mathbb{R}_+$. Fix $t_1, t_2, \bar{h} \in \mathbb{R}_+$ arbitrarily. We then have

$$C_X(t_1, t_2) = C_X(t_1, t_1 + t_2 - t_1) = C_X(0, t_2 - t_1) = C_X(t_1 + \bar{h}, t_2 - t_1 + t_1 + \bar{h}) = C_X(t_1 + \bar{h}, t_2 + \bar{h}).$$

Note: For a wide-sense stationary process, to find the autocovariance of two random variables of the process, we need not require the exact value of the two indices, but instead it suffices to know the difference between the two indices. In short, we can say that autocovariance depends only on the time gap between the two random variables.

Remark 2. *Every stationary process with finite variance is wide-sense stationary process.*

Proof. To show that the a strictly stationary process is a weakly stationary process we must verify that strict stationarity ensures the two weak stationarity conditions.

1. **Constant Mean**

   From strict stationarity, the marginal distribution of X_t is same for all $t \in \mathcal{T}$. Hence the expectation doesn't change.

   $$\mathbb{E}[X_t] = \mathbb{E}[X_{t+h}], \quad \forall t, h \in \mathcal{T} \tag{1}$$

   Thus mean is constant.

2. **Autocovariance depends only on lag** The autocovariance of a process is given by

   $$C_X(t_1, t_2) = \mathbb{E}[(X_{t_1} - \mathbb{E}[X_{t_1}])(X_{t_2} - \mathbb{E}[X_{t_2}])].$$

   Using strict stationarity, the joint distribution of (X_{t_1}, X_{t_2}) is same as that of (X_{t_1+h}, X_{t_2+h}) for any h. This implies that

   $$C_X(t_1, t_2) = C_X(t_1 + h, t_2 + h) \tag{2}$$

Hence from equations (1) and (2), we can say that every stationary process with finite variance is wide-sense stationary process. □

1.6.1 IID processes

We now define an independent and identically distributed (IID) process.

Definition 9 (IID process)**.** *A process $\{X_t : t \in \mathbb{R}_+\}$ is said to be an IID process with common CDF F if for any $n \in \mathbb{N}$ and* **t** $\in \mathbb{R}_+$*,*

$$F_{\mathbf{t}}(\underline{x}) = \prod_{i=1}^{n} F(x_i), \quad \underline{x} = (x_1, ..., x_n) \in \mathbb{R}^n.$$

Lemma 1. *Suppose that $\{X_t : t \in \mathbb{R}_+\}$ is an IID process.*

1. *The FDDs of $\{X_t : t \in \mathbb{R}_+\}$ are consistent.*
2. *$\{X_t : t \in \mathbb{R}_+\}$ is strictly stationary. That is, every IID process is stationary.*

Proof. Consider an IID process $\{X_t : t \in \mathbb{R}_+\}$

1. Since every X_t in an IID process has the same distribution (as they are identical), shifting the indices doesn't change their probability distribution.
2. Because the values are independent, the joint distribution of any collection of them remains unchanged under shifts, i.e.,

$$F_{t_1,...,t_n}(x_1, ..., x_n) = \prod_{i=1}^{n} F(x_i) = F_{t_1+h,...,t_n+h}(x_1, ..., x_n)$$

These two points satisfy the definition of strict stationarity. Hence we can say that every **IID process is a stationary process**. □

Remark 3. *Every IID process with finite variance is a Wide-Sense Stationary Process.*

Proof. To be **WSS**, the process must have

1. **Constant Mean:** This is true for IID since each X_t has the same expected value i.e., $\mathbb{E}[X_{t_1}] = \mathbb{E}[X_{t_2}], \quad \forall t_1, t_2$
2. **Autocovariance is dependent on time gap:** Since IID means independence, the covariance is:

$$C_X(t, t+h) = \begin{cases} \sigma^2, & \text{if } h = 0 \\ 0, & \text{if } h \neq 0 \end{cases}$$

   This only depends on h (which is time gap). Hence we can say that IID process with finite variance is WSS procecss

□

1.7 Examples

Example 1.2. *Let $X_1, X_2, ..$ be an $\mathbb{N}$-valued IID process.*

Let $S_0 := 0$, and for each $n \in \mathbb{N}$, let

$$S_n = \sum_{i=1}^{n} X_i.$$

- Determine the mean and autocovariance for the process $\{S_n\}_{n=0}^{\infty}$.

Solution:

1. **Compute Mean Function** $M_S(n)$

$$\begin{aligned} M_S(n) &= \mathbb{E}[S_n] \\ &= \mathbb{E}[\sum_{i=1}^{n} X_i] \\ &= \sum_{i=1}^{n} \mathbb{E}[X_i]. \end{aligned}$$

Since the X_i are IID, they all have the same mean $\mathbb{E} = \mu$, so:

$$M_S(n) = \mathbb{E}[S_n] = n\mathbb{E}[X] = n\mu$$

2. **Computing Autocovariance:** $C_S(n, m)$ The covariance of S_n and S_m is:

$$C_S(n, m) = Cov(S_n, S_m) = \mathbb{E}[S_n S_m] - \mathbb{E}[S_n]\mathbb{E}[S_m].$$

   (a) **Computing** $\mathbb{E}[S_n S_m]$
   Since S_n and S_m are sums of IID variables:

$$S_n = X_1 + X_2 + ... + X_n$$

$$S_m = X_1 + X_2 + ... + X_m$$

   if $n \leq m$, then:

$$S_m = S_n + \sum_{i=n+1}^{m} X_i$$

   Taking expectations:

$$\mathbb{E}[S_n S_m] = \mathbb{E}[S_n(S_n + \sum_{i=n+1}^{m} X_i)].$$

   Expanding:

$$\mathbb{E}[S_n S_m] = \mathbb{E}[S_n^2] + \mathbb{E}[S_n]\mathbb{E}[\sum_{i=n+1}^{m} X_i].$$

   Since $\mathbb{E}[\sum_{i=n+1}^{m} X_i] = (m-n)\mu$. we get:

$$\mathbb{E}[S_n S_m] = \mathbb{E}[S_n^2] + (m-n)\mu\mathbb{E}[S_n].$$

   Using the variance formula:

$$\text{Var}(S_n) = \sum_{i=1}^{n} \text{Var}(X_i) = n\sigma^2,$$

   we obtain:

$$\begin{aligned} \mathbb{E}[S_n^2] &= \text{Var}(S_n) + (\mathbb{E}[S_n])^2 \\ &= n\sigma^2 + n^2\mu^2. \end{aligned}$$

   Thus:

$$\mathbb{E}[S_n S_m] = n\sigma^2 + n^2\mu^2 + (m-n)n\mu^2.$$

(b) **Compute** $\mathbb{E}[S_n]\mathbb{E}[S_m]$

$$\begin{aligned}\mathbb{E}[S_n]\mathbb{E}[S_m] &= (n\mu)(m\mu)\\ &= nm\mu^2.\end{aligned}$$

(c) **Compute** $C_S(n, m)$

$$C_S(n,m) = \mathbb{E}[S_nS_m] - \mathbb{E}[S_n]\mathbb{E}[S_m].$$

Substituting the values:

$$C_S(n,m) = \left(n\sigma^2 + n^2\mu^2 + (m-n)n\mu^2\right) - nm\mu^2.$$

Simplifying:

$$C_S(n,m) = n\sigma^2.$$

Similarly, if $m \leq n$:

$$C_S(n,m) = m\sigma^2$$

Thus, the autocovariance function is:

$$C_S(n,m) = \min(n,m)\sigma^2.$$

3. **Conclusion:**
From steps, 1 and 2 we can conclude that:

$$M_S(n) = n\sigma^2$$

and

$$C_S(n,m) = \min(n,m)\sigma^2.$$

Example 1.3. *Consider a random process $\{X_t : t \in \mathbb{R}\}$ defined as $X_t = \cos(t+U)$, where $U \sim$ Unif$(0, 2\pi)$. Show that X_t is a WSS process.*

Solution:
To show that the given process is Wide-Sense Stationary (WSS), we need to verify:

1. **Constant Mean**: The expected value remains the same $\forall t_1, t_2$, i.e., $\mathbb{E}[t_1] = \mathbb{E}[t_2]$.
2. **Autocovariance depends on gap**: $C_X(t_1, t_2) = C_X(t_1 + h, t_2 + h)$.

If these conditions hold, the process is WSS.

1. **Mean:**

$$\mu_X(t) = \mathbb{E}[X_t] = \mathbb{E}[\cos(t+U)]$$

Since t is fixed here and U is random variable, with $f_U(u) = \frac{1}{2\pi}, \quad u \in [0, 2\pi]$

$$\begin{aligned}\mathbb{E}[\cos(t+U)] &= \int_{-\infty}^{\infty} \cos(t+u) f_u(u)du\\ &= \int_0^{2\pi} \cos(t+u)\frac{1}{2\pi}du\\ &= \frac{1}{2\pi}|\sin(t+u)|_0^{2\pi}\\ &= \frac{1}{2\pi}(\sin(t+2\pi) - \sin(t))\\ &= \frac{1}{2\pi}(\sin(t) - \sin(t)) = 0.\end{aligned}$$

So $\mu_X(t) = 0$ and is constant $\forall t \in \mathbb{R}$.

2. **Autocovariance:** Now

$$
\begin{aligned}
C_X(t,s) &= \mathbb{E}[(X_t - \mu_X(t))(X_s - u_X(s))] \\
&= \mathbb{E}[X_t X_s] \qquad\qquad (\because \mu_X(t) = \mu_X(s) = 0\, \forall t, s \in \mathbb{R} \quad \text{from step 1}) \\
&= \mathbb{E}[\cos(t+U)\cos(s+U)] \\
&= \mathbb{E}[\frac{1}{2}\cos(t+s+2U) + \frac{1}{2}\cos(t-s)] \quad (\because \cos(X)\cos(Y) = \frac{1}{2}[\cos(X+Y) + \cos(X-Y)]) \\
&= \frac{1}{2}\mathbb{E}[\cos(t+s+2U)] + \frac{1}{2}\mathbb{E}[\cos(t-s)] \\
\\
&= \frac{1}{2}\int_0^{2\pi} \cos(t+2u+s)\frac{1}{2\pi}du + \frac{1}{2}\cos(t-s) \\
\\
&= \frac{1}{4\pi} \mid \frac{\sin(t+s+2u)}{2} \mid_0^{2\pi} + \frac{1}{2}\cos(t-s) \\
\\
&= \frac{1}{8\pi}(\sin(4\pi+t+s) - \sin(t+s)) + \frac{1}{2}\cos(t-s) \\
\\
&= 0 + \frac{1}{2}\cos(t-s) \\
&= C_X(0, t-s).
\end{aligned}
$$

So, from steps 1 and 2, we can conclude that the process $\{X_t : t \in \mathbb{R}\}$ where $X_t = \cos(t+U)$ with $U \sim \text{Unif}(0, 2\pi)$ is a WSS process.