Topic 7: Random Processes

- Definition, discrete and continuous processes
- Specifying random processes
  - Joint cdf's or pdf's
  - Mean, auto-covariance, auto-correlation
  - Cross-covariance, cross-correlation
- Stationary processes and ergodicity

Random processes

- A *random process*, also called a *stochastic process*, is a family of random variables, indexed by a parameter t from an indexing set $\mathcal{T}$. For each experiment outcome $\omega \in \Omega$, we assign a function X that depends on t

$$X(t, \omega) \quad t \in \mathcal{T}, \omega \in \Omega$$

  - t is typically time, but can also be a spatial dimension
  - t can be discrete or continuous
  - The range of t can be finite, but more often is infinite, which means the process contains an infinite number of random variables.
- Examples:
  - The wireless signal received by a cell phone over time
  - The daily stock price
  - The number of packets arriving at a router in 1-second intervals
  - The image intensity over 1cm^2 regions

- We are interested in specifying the joint behavior of the random variables within a family, or the behavior of a process. This joint behavior helps in studying
  - The dependencies among the random variables of the process (e.g. for prediction)
  - Long-term averages
  - Extreme or boundary events (e.g. outage)
  - Estimation/detection of a signal corrupted by noise

Two ways of viewing a random process

Consider a process $X(t, \omega)$

- At a fixed t, $X(t, \omega)$ is a random variable and is called a *time sample.*
- For a fixed ω, $X(t, \omega)$ is a *deterministic* function of t and is called a *realization* (or a sample path or sample function)

$\Rightarrow$ ω induces the randomness in $X(t, \omega)$. In the subsequent notation, ω is implicitly implied and therefore is usually suppressed.

- When t comes from a countable set, the process is *discrete-time.* We then usually use n to denote the time index instead and write the process as $X(n, \omega)$, or just X_n, $n \in \mathbb{Z}$.
  - For each n, X_n is a r.v., which can be continuous, discrete, or mixed.
  - Examples: $X_n = Z^n, \quad n \geq 1,\ Z \sim U[0, 1]$.
    Others: sending bits over a noisy channel, sampling of thermal noise.
- When t comes from an uncountably infinite set, the process is *continuous-time.* We then often denote the random process as $X(t)$. At each t, $X(t)$ is a random variable.
  - Examples: $X(t) = \cos(2\pi f t + \theta), \quad \theta \sim U[-\pi, \pi]$.

Specifying a random process

- A random process can be *completely specified* by the collection of joint cdf among the random variables

$$\{X(t_1), X(t_2), \ldots, X(t_n)\}$$

  for any set of sample times $\{t_1, t_2, \ldots, t_n\}$ and any order n.
  Denote $X_k = X(t_k)$,

  - If the process is continuous-valued, then it can also be specified by the collection of joint pdf

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

  - If the process is discrete-valued, then a collection of joint pmf can be used

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P[X_1 = x_1, \ldots, X_n = x_n]$$

- This method requires specifying a vast collection of joint cdf's or pdf's, but works well for some important and useful models of random processes.

Mean, auto-covariance, and auto-correlation functions

The moments of time samples of a random process can be used to *partly specify* the process.

- Mean function:

$$m_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x \, f_{X(t)}(x) \, dx$$

  $m_X(t)$ is a function of time. It specifies the average behavior (or the trend in the behavior) of $X(t)$ over time.

- Auto-correlation function: $R_X(t_1, t_2)$ is defined as the correlation between the two time samples $X_{t_1} = X(t_1)$ and $X_{t_2} = X(t_2)$

$$R_X(t_1, t_2) = E[X_{t_1} X_{t_2}]$$

  Properties:

  - In general, $R_X(t_1, t_2)$ depends on both t_1 and t_2.
  - For real processes, $R_X(t_1, t_2)$ is *symmetric*

$$R_X(t_1, t_2) = R_X(t_2, t_1)$$

– For any t, t_1 and t_2

$$\begin{aligned} R_X(t,t) &= E[X_t^2] \geq 0 \\ |R_X(t_1,t_2)| &\leq \sqrt{E[X_{t_1}^2]E[X_{t_2}^2]} \end{aligned}$$

Processes with $E[X_t^2] < \infty$ for all t is called *second-order*.

- Auto-covariance function: $C_x(t_1,t_2)$ is defined as the covariance between the two time samples $X(t_1)$ and $X(t_2)$

$$\begin{aligned} C_X(t_1,t_2) &= E\left[\{X_{t_1} - m_X(t_1)\}\{X_{t_2} - m_X(t_2)\}\right] \\ &= R_X(t_1,t_2) - m_X(t_1)m_X(t_2) \end{aligned}$$

– The *variance* of $X(t)$ can be obtained as

$$\text{var}(X_t) = E[\{X(t) - m_X(t)\}^2] = C_X(t,t)$$

$\text{var}(X_t)$ is a function of time and is always non-negative.

– The *correlation coefficient function*:

$$\rho_X(t_1,t_2) = \frac{C_X(t_1,t_2)}{\sqrt{C_X(t_1,t_1)}\sqrt{C_X(t_2,t_2)}}$$

$\rho_X(t_1,t_2)$ is a function of times t_1 and t_2. It is also symmetric.

- Examples: Find the mean and autocorrelation functions of the following processes:

a) $X(t) = \cos(2\pi f t + \theta)$, $\quad \theta \sim U[-\pi,\pi]$

b) $X_n = Z_1 + \ldots + Z_n$, $\quad n = 1,2,\ldots$
where Z_i are i.i.d. with zero mean and variance σ^2.

Multiple random processes: Cross-covariance and cross-correlation functions

For multiple random processes:

- Their joint behavior is completely specified by the joint distributions for all combinations of their time samples.

Some simpler functions can be used to partially specify the joint behavior. Consider two random processes $X(t)$ and $Y(t)$.

- Cross-correlation function:

$$R_{X,Y}(t_1, t_2) = E[X_{t_1} Y_{t_2}]$$

  - If $R_{X,Y}(t_1, t_2) = 0$ for all t_1 and t_2, processes $X(t)$ and $Y(t)$ are *orthogonal.*
  - Unlike the auto-correlation function, the cross-correlation function is not necessarily symmetric.

$$R_{X,Y}(t_1, t_2) \neq R_{X,Y}(t_2, t_1)$$

- Cross-covariance function:

$$\begin{aligned} C_{X,Y}(t_1, t_2) &= E[\{X_{t_1} - m_X(t_1)\}\{Y_{t_2} - m_Y(t_2)\}] \\ &= R_{X,Y}(t_1, t_2) - m_X(t_1) m_Y(t_2) \end{aligned}$$

  - If $C_{X,Y}(t_1, t_2) = 0$ for all t_1 and t_2, processes $X(t)$ and $Y(t)$ are *uncorrelated.*

- Two processes $X(t)$ and $Y(t)$ are *independent* if any two vectors of time samples, one from each process, are independent.
  - If $X(t)$ and $Y(t)$ are independent then they are uncorrelated: $C_{X,Y}(t_1, t_2) = 0 \quad \forall\, t_1, t_2$ (the reverse is not always true).

- Example: Signal plus noise

$$Y(t) = X(t) + N(t)$$

where $X(t)$ and $N(t)$ are independent processes.

Stationary random processes

In many random processes, the statistics do not change with time. The behavior is *time-invariant*, even though the process is random. These are called *stationary* processes.

- Strict-sense stationarity:
  - A process is *nth order stationary* if the joint distribution of any set of n time samples is independent of the placement of the time origin.

    $$[X(t_1), \ldots, X(t_n)] \sim [X(t_1+\tau), \ldots, X(t_n+\tau)] \quad \forall \tau$$

    For a discrete process:

    $$[X_1, \ldots, X_n] \sim [X_{1+m}, \ldots, X_{n+m}] \quad \forall m$$

  - A process that is nth order stationary for every integer $n > 0$ is said to be *strictly stationary*, or just stationary for short.
  - Example: The i.i.d. random process is stationary.
- Strict stationarity is a strong requirement.

  - First-order stationary processes: $f_{X(t)}(x) = f_X(x)$ for all t. Thus

    $$\begin{aligned} m_X(t) &= m \quad \forall t \\ \text{var}(X_t) &= \sigma^2 \quad \forall t \end{aligned}$$

  - Second-order stationary processes:

    $$f_{X(t_1),X(t_2)}(x_1, x_2) = f_{X(t_1+\tau),X(t_2+\tau)}(x_1, x_2) \quad \forall \tau$$

    The second-order joint pdf (pmf) depends only on the time difference $t_2 - t_1$. This implies

    $$\begin{aligned} R_X(t_1, t_2) &= R_X(t_2 - t_1) \\ C_X(t_1, t_2) &= C_X(t_2 - t_1) \end{aligned}$$

Wide-sense stationary random processes

- $X(t)$ is *wide-sense stationary* (WSS) if the following two properties both hold:

$$\begin{aligned} m_X(t) &= m \quad \forall t \\ R_X(t_1, t_2) &= R_X(t_2 - t_1) \quad \forall t_1, t_2 \end{aligned}$$

  - WSS is a much more relaxed condition than strict-sense stationarity.
  - All stationary random processes are WSS. A WSS process is not always strictly stationary.
  - Example: Sequence of independent r.v.'s
    $X_n = \pm 1$ with probability $\frac{1}{2}$ for n even
    $X_n = -1/3$ and 3 with probabilities $\frac{9}{10}$ and $\frac{1}{10}$ for n odd
- Properties of a WSS process:
  - $R_X(0)$ is the *average power* of the process

$$R_X(0) = E[X(t)^2] \geq 0$$

    $R_X(0)$ thus is always positive.

  - $R_X(\tau)$ is an even function

$$R_X(\tau) = R_X(-\tau)$$

  - $R_X(\tau)$ is maximum at $\tau = 0$

$$|R_X(\tau)| \leq R_X(0)$$

  - If $R_X(0) = R_X(T)$ then $R_X(\tau)$ is periodic with period T

$$\text{if} \quad R_X(0) = R_X(T) \quad \text{then} \quad R_X(\tau) = R_X(\tau + T) \quad \forall \tau$$

  - $R_X(\tau)$ measures the rate of change of the process

$$P\left[|X(t+\tau) - X(t)| > \epsilon\right] \leq \frac{2\left(R_X(0) - R_X(\tau)\right)}{\epsilon^2}$$

- If a Gaussian process is WSS, then it is also strictly stationary.
  - A WWS Gaussian process is completely specified by the constant mean m and covariance $C_X(\tau)$.
- WSS processes play a crucial role in linear time-invariant systems.

Cyclostationary random processes

- Many processes involves the repetition of a procedure with period T.
- A random process is *cyclostationary* if the joint distribution of any set of samples is invariant over a time shift of mT (m is an integer)

$$[X(t_1), \ldots, X(t_n)] \sim [X(t_1 + mT), \ldots, X(t_n + mT)] \quad \forall m, n, t_1, \ldots, t_n$$

- A process is *wide-sense cyclostationary* if for all integer m

$$\begin{aligned} m_X(\tau + mT) &= m_X(\tau) \\ R_X(t_1 + mT, t_2 + mT) &= R_X(t_1, t_2) \end{aligned}$$

  - If $X(t)$ is WSS, then it is also wide-sense cyclostationary.
- We can obtain a stationary process $X_s(t)$ from a cyclostationary process $X(t)$ as

$$X_s(t) = X(t + \theta)\ , \quad \theta \sim U[0, T]$$

  - If $X(t)$ is wide-sense cyclostationary then $X_s(t)$ is WSS.

Time averages and ergodicity

- Sometimes we need to estimate the parameters of a random process through measurement.
- A quantity obtainable from measurements is the *ensemble average*. For example, an estimate of the mean is

$$\hat{m}_X(t) = \frac{1}{N} \sum_{i=1}^{N} X(t, \omega_i)$$

  where ω_i is the ith outcome of the underlying random experiment.
  - In general, since $m_X(t)$ is a function of time, we need to perform N repetitions of the experiment at each time t to estimate $m_X(t)$.
- If the process is stationary, however, then $m_X(t) = m$ for all t. Then we may ask if m can be estimated based on the realization (over time) of a single outcome ω alone.
- We define the *time average* over an interval $2T$ of a single realization as

$$\langle X(t) \rangle_T = \frac{1}{2T} \int_{-T}^{T} X(t, \omega)\, dt$$

Question: When does the time average converge to the ensemble average?

- Example 1: If $X_n = X(n, \omega)$ is a stationary, i.i.d. discrete-time random process with mean $E[X_n] = m$, then by the strong LLN

$$\frac{1}{N}\sum_{i=1}^{N} X_i \xrightarrow{\text{a.s.}} m \qquad \text{as } N \to \infty$$

$\Longrightarrow$ Convergence.

- Example 2: $X(t) = A$ for all t, where A is a zero-mean random variable. Then $X(t)$ is stationary and $m_X = E[A] = 0$ for all t, but

$$\langle X(t) \rangle_T = \frac{1}{2T}\int_{-T}^{T} A\, dt = A$$

$\Longrightarrow$ No convergence.

- Ergodicity let us characterize this convergence for a larger class of random processes.

Ergodicity of a WSS process

- Consider a WSS random process $X(t)$ with mean m. $X(t)$ is *mean-ergodic* if

$$\langle X(t) \rangle_T \longrightarrow m \qquad \text{as} \qquad T \to \infty$$

Notes:

  – Because of stationarity, the expected value of $\langle X(t) \rangle_T$ is

$$E\left[\langle X(t) \rangle_T\right] = E\left[\frac{1}{2T}\int_{-T}^{T} X(t)\, dt\right] = \frac{1}{2T}\int_{-T}^{T} E[X(t)]\, dt = m$$

  – Mean-ergodic definition therefore implies that $\langle X(t) \rangle_T$ approaches its mean as $T \to \infty$.

- Mean-ergodic theorem: The WSS process $X(t)$ is mean-ergodic in the mean-square sense, that is

$$\lim_{T\to\infty} E\left[\left(\langle X(t) \rangle_T - m\right)^2\right] = 0$$

if and only if its covariance satisfies

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \left(1 - \frac{|u|}{2T}\right) C_X(u)\, du = 0$$