Lecture 11: 
Introduction to Random Processes 
-cont’d

ELE 525: Random Processes in Information Systems

Hisashi Kobayashi
Department of Electrical Engineering
Princeton University
October 16, 2013

12.2.5.3 Continuous-time, continuous-space (CTCS) Markov process

A CTCS Markov process is also known as a **diffusion process**

Its state transition probability (or conditional probability) distribution function

\[ F \big|_{X(t_1)X(t_0)} (x_0, x_1; t_0, t_1) \]

satisfies a partial differential equation known as a diffusion equation. The diffusion equations (forward and backward) for a general class of diffusion processes were derived by **Andrey N. Kolmogorov** (pp. 500-501).

The diffusion equation for Brownian motion (a special case of diffusion processes) was derived by **Albert Einstein** (p. 499).
12.2.5.4 Discrete-Time, Continuous-Space (DTCS) Markov Process

Examples:

- **Autoregressive** (AR) time series (Section 13.4.1)
- State sequence associated an **autoregressive moving average** (ARMA) forms a multidimensional a DTCS Markov process. (Section 13.4.3)
- Observation of a diffusion process at discrete-time moments
- A sequence of random variates generated in a **Markov chain Monte Carlo** (MCMC) simulation technique.
12.2.5.5 Semi-Markov process and embedded Markov chain

If the interval \( \tau_k = t_{k+1} - t_k \) in (12.5), i.e.,

\[
X(t) = i, \text{ for } t_k \leq t < t_{k+1}, \text{ where } i = X(t_k), \text{ and } X(t_{k+1}) = j(\neq i).
\]  

(12.5)

Is not exponentially distributed, \( X(t) \) is called a semi-Markov process (Section 16.1)

For a given CTMC or semi-Markov process \( X(t) \), a DTMC \( \{ X_k \} \) defined by observing at the epochs of state transition in \( X(t) \) is said to be embedded in the process \( X(t) \).

An embedded Markov chain (EMC) \( \{ X_k \} \) may often be amenable to analysis, whereas the original process \( X(t) \) may not (Section 23.3.5).
12.2.6 Point processes and renewal processes

A **point process** is a random process that consists of a sequence of epochs
\[ t_1 \quad t_2 \quad t_3 \quad \ldots \quad t_n \quad \ldots \]
where “point events” occur.

A one dimensional point process can be represented by a **counting process** \( N(t) \),
\[ N(t) = \max\{ n : t_n \leq t \} \quad (12.6) \]

The difference of the event points
\[ X_n = t_n - t_{n-1} \quad (12.7) \]
represents the interval between the \((n-1)st\) and the \(n\)th point events.

If the \( X_n \) are i.i.d., \( N(t) \) is called a **renewal process**, and the event points \( t_n \) are called **renewal points**, and the intervals \( X_n \) are the **lifetimes**.

The **Poisson process** is a point process, which is a **renewal process** (because the \( X_n \) are independent) and is a **CTMC** (because the \( X_n \) are exponentially distributed).
12.2.7 Real-valued versus complex-valued processes

In communication systems, the class of carrier-modulated data transmission systems that adopt linear modulation

- amplitude shift keying (ASK)
- amplitude-phase shift keying (APSK)
- quadrature amplitude modulation (QAM)
- amplitude modulation (AM)
- phase modulation (PM)
- single sideband modulation (SSB)
- vestigial sideband modulation (VSB)

can be concisely represented in terms of complex-valued processes, called analytic signals.

Gaussian noise which goes through a bandpass filter can be compactly represented by a complex-valued Gaussian process.

Fading due to multipath propagation in a radio channel introduces a multiplicative factor, which can be compactly represented by a complex-valued Gaussian RV.
12.2.8 One-dimensional versus vector processes

Vector representation of a random process

Type 1: A scalar or one-dimensional process is observed at multiple instants $t = (t_1, t_2, ..., t_n)$. $X = (X(t_1), X(t_2), ..., X(t_n))$

Type 2: A multidimensional process, such as in diversity reception. Multiple-input, multiple-output (MIMO) in communication systems can be compactly represented as complex-valued vector processes.

In statistics and econometrics, vector time series or multivariate time series are often used.

An observed sample path of such time series is represented as panel data, which is a two-dimensional array.
12.3 Stationary Random Process

A **stationary random process** is one for which all the distribution functions are **invariant** under a shift of the time origin.

12.3.1 Strict Stationarity and Wide-Sense Stationarity

**Definition 12.1 (Strictly stationary random process).** A real-valued random process $X(t)$ is said to be **strictly stationary**, **strict-sense stationary (SSS)**, or **strongly stationary**, if for every finite set of time instants $\{t_i; i = 1, 2, \ldots, n\}$ and for every constant $h$, the joint distribution functions of $X_i = X(t_i)$ and those of $X'_i = X(t_i + h); i = 1, 2, \ldots, n$ are the same:

$$F_{X_1X_2\cdots X_n}(x_1, x_2, \ldots, x_n) = F_{X'_1X'_2\cdots X'_n}(x_1, x_2, \ldots, x_n).$$  \hspace{1cm} (12.8)

Let the **autocorrelation function** be defined as

$$R_X(t_1, t_2) \triangleq E[X(t_1)X(t_2)].$$  \hspace{1cm} (12.10)
Definition 12.2 (Wide-sense stationary (WSS) random process). A random process $X(t)$ is said to be wide-sense stationary (WSS), weakly stationary, covariance stationary or second-order stationary, if it has

1. a constant mean: $E[X(t)] = \mu_X$ for all $-\infty < t < \infty$, 
2. finite second moments $E[X(t)^2] < \infty$, for all $-\infty < t < \infty$, and 
3. the covariance

$$E[(X(s) - \mu)(X(t) - \mu)]$$

that depends only on the time difference $|s - t|$.

**Note:** SSS implies WSS, but the converse is not true.

**Example 12.1: An i.i.d. sequence.**

Let $\{X_n\}$ be a sequence of i.i.d. real-valued RVs with a common mean $\mu_X$ and variance $\sigma_X^2$.

$$R_X(k) = E[X_{n+k}X_n] = \begin{cases} \sigma_X^2 + \mu_X^2, & k = 0, \\ \mu_X^2, & k \neq 0. \end{cases} \quad (12.12)$$

Thus $\{X_n\}$ is wide-sense stationary. It is also strict stationary since it satisfies (12.8).

From the SLLN we can assert

$$\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{a.s.} \mu_X.$$
Example 12.3 Identical sequence

Let \( Z \) be a RV with known distribution.
Set
\[
Z_n = Z, \text{ for all } n.
\]
Then clearly \( \{Z_n\} \) is SSS. It is also WSS with the autocorrelation function
\[
R_Z[k] \triangleq E[Z_n Z_{n+k}] = \sigma^2, \text{ for all } k.
\]
\[\text{(12.14)}\]
We also readily see
\[
\frac{\sum_{i=1}^{n} Z_i}{n} \xrightarrow{a.s.} Z,
\]
\[\text{(12.15)}\]
Since each term in the sum is the same as the limit \( Z \)

\[\checkmark \text{ Comparison of Examples 12.1 and 12.3}\]

Observing \( X_1, X_2, \ldots, X_n \) provides no information to predict \( X_{n+1} \).

Observing \( Z_1 \) allows us to predict \( Z_2, Z_3, \ldots \) exactly.
\[
\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n}
\]
converges a.s. to \( \mu_X \).
\[
\overline{Z_n}
\]
contains as much randomness as in the first observation \( Z_1 \).
Example 12.4: Sinusoidal function with random amplitudes

\[ X(t) = A \cos \omega_0 t + B \sin \omega_0 t, \quad (12.16) \]

\[ E[X(t)] = E[A] \cos \omega_0 t + E[B] \sin \omega_0 t, \]

For the mean \( E[X(t)] \) to be constant, it is necessary that

\[ E[A] = E[B] = 0. \quad (12.17) \]

In order for \( X(t) \) to be WSS, the autocorrelation function

\[
E[X(t)X(u)]
= E[A^2 \cos \omega_0 t \cos \omega_0 u + B^2 \sin \omega_0 t \sin \omega_0 u + AB (\cos \omega_0 t \sin \omega_0 U + \sin \omega_0 t \cos \omega_0 u)]
= \frac{E[A^2]}{2} [\cos \omega_0 (t + u) + \cos \omega_0 (t - u)] - \frac{E[B^2]}{2} [\cos \omega_0 (t + u) - \cos \omega_0 (t - u)]
+ E[AB] \sin \omega_0 (t + u) \quad (12.18)
\]

must be a function of \( t-u \) only. Suppose that

\[ E[A^2] = E[B^2] = \sigma^2, \quad \text{and} \quad E[AB] = 0, \quad (12.19) \]

hold. Then it readily follows that
\[ E[X(t)X(u)] = \sigma^2 \cos \omega_0 (t - u) \triangleq R_X(t - u). \]  

(12.21)

Hence, \( X(t) \) is WSS. Conversely, suppose that \( X(t) \) is WSS. Then

\[ E[X(t_1)X(t_1)] = E[X(t_2)X(t_2)] = R_X(0) \]  

(12.22)

for any \( t_1 \) and \( t_2 \). Letting \( t_1 = 0 \) and \( t_2 = \frac{\pi}{2\omega_0} \),

\[ X(t_1) = X(0) = A \quad \text{and} \quad X(t_2) = X \left( \frac{\pi}{2\omega_0} \right) = B. \]

Then

\[ E[A^2] = E[B^2] = R_X(0) = \sigma^2 \]

Furthermore,

\[ E[X(t)X(u)] = \sigma^2 \cos \omega_0 (t - u) + E[AB] \sin \omega_0 (t + u). \]

The last expression is a function of \( t - u \) only if

\[ E[AB] = 0 \]
12.3.2 Gaussian Process

Definition 12.3 (Gaussian process). A real-valued continuous-time process $X(t); -\infty < t < \infty$ is called a Gaussian process, if for every finite set of time instants

$$\mathcal{T} = \{t_1, t_2, \ldots, t_n\},$$

the vector

$$X = (X_1, X_2, \ldots, X_n)^\top,$$

where $X_i = X(t_i), \ i = 1, 2, \ldots, n$

has the multivariate normal distribution with some mean vector $\mu$ and covariance matrix $C$, both of which may depend on $\mathcal{T}$:

$$f_X(x) = \frac{1}{(2\pi)^{n/2}|\det C|^{1/2}} \exp \left\{ -\frac{1}{2}(x - \mu)^\top C^{-1}(x - \mu) \right\}. \tag{12.23} $$

$$
C = [C_{ij}] .
$$

$$
C_{ij} = E \left[ (X_i - \mu_i)(X_j - \mu_j) \right] \\
= E \left[ (X(t_i) - \mu(t_i))(X(t_j) - \mu(t_j)) \right] \\
= R_X(t_i, t_j) - \mu(t_i)\mu(t_j). \tag{12.24}
$$
12.3.2.1 Stationary Gaussian process

Since the mean vector and covariance matrix completely specify the PDF of multivariate normal variables, a WSS Gaussian process is also SSS. Thus,

**Theorem 12.1 (Stationary Gaussian process).** A real-valued continuous-time process \( X(t); -\infty < t < \infty \) is a stationary Gaussian process, if

\[
E[X(t)] = \mu, \quad \text{and} \quad E[X(t)X(s)] = R_X(t - s), \quad \text{for} \quad -\infty < t, s < \infty, \tag{12.27}
\]

with some constant \( \mu \) and autocorrelation function \( R_X(\cdot) \), and if for every finite set of time instants

\[
T = \{t_1, t_2, \ldots, t_n\},
\]

the vector

\[
X = (X_1, X_2, \ldots, X_n)^\top, \quad \text{where} \quad X_i = X(t_i), \quad i = 1, 2, \ldots, n
\]

has the multivariate normal distribution with mean \( \mu = (\mu, \mu, \ldots, \mu) \) and covariance matrix \( C \), whose \((i, j)\) entry \( C_{ij} \) is given by (12.26). \( \square \)
Recall \{ X_n \} of Example 12.1 and \{ Z_n \} of Example 12.3. Most sequences of interest lie somewhere between the above two extreme cases in that the sequences are not uncorrelated but the covariance \( C[k] \) approaches zero as the lag \( k \) increases.

A stationary process \( X_t \) is said to be \textit{ergodic}, if sample averages formed from a single sample process converges to some underlying parameter of the process. In other words, an ergodic process is a stationary process such that its \textit{time average} is equivalent to the \textit{ensemble average} (or \textit{population mean})

An \textit{ergodic theorem} specifies conditions for a stationary process to be ergodic. Law of large numbers for sequences that are not uncorrelated are often called \textit{ergodic theorems}. 