Random (Stochastic) Processes

- Random (or Stochastic) Process (or Signal)
  - A random process is random 'sequence', not only a single random variable.
  - A sequence $x[n],\ -\infty < n < \infty$. Each individual sample $x[n]$ is assumed to be an outcome of some underlying *random variable* $\boldsymbol{X}_n$.
  - Difference from a single random variable:
    - for a random variable the outcome of a random sample is a number, whereas for a random process the outcome is a sequence.

Random (Stochastic) Processes

- Consider a random process $x[n], -\infty < n < \infty$, where each $x[n]$ is drawn from the *random variable* $\boldsymbol{X}_n$.
- Hence, there are infinite random variables, $\boldsymbol{X}_n, -\infty < n < \infty$.
- Strictly speaking, its joint distribution,

$$\boldsymbol{p}(\ldots, \boldsymbol{X}_{-2}=x[-2], \boldsymbol{X}_{-1}=x[-1],, \boldsymbol{X}_0=x[0],, \boldsymbol{X}_1=x[1], \boldsymbol{X}_2=x[2], \ldots)$$

is a probability distribution in an infinite-dimensional space.

Random (Stochastic) Processes

- However, it is infeasible to represent the random (or stochastic) process as a distribution in an infinite-dimensional space.
- The most common way to describe a random process is to characterize the distribution for its finite samples, say, $\{n_1, n_2, \ldots, n_k\}$, and specify their probability distribution in a finite-dimensional space:

$$\boldsymbol{p}(\boldsymbol{X}_{n1}=x[n_1], \boldsymbol{X}_{n2}=x[n_2], \ldots, \boldsymbol{X}_{n\mathrm{k}}=x[n_\mathrm{k}])$$

Example: Gaussian Process

- For example, the Gaussian process is defined as follows:
- If for any set of samples $n_1, n_2, \ldots, n_k$ $(n_i \in Z, k \in N^+)$, the random process satisfies that the joint distribution of these samples

$$\boldsymbol{p}(\boldsymbol{X}_{n1}=x[n_1], \boldsymbol{X}_{n2}=x[n_2], \ldots, \boldsymbol{X}_{n\mathrm{k}}=x[n_\mathrm{k}])$$

is a multivariate Gaussian distribution, then this process is called a Gaussian process.

- Gaussian process is often used in machine learning for nonlinear regression.

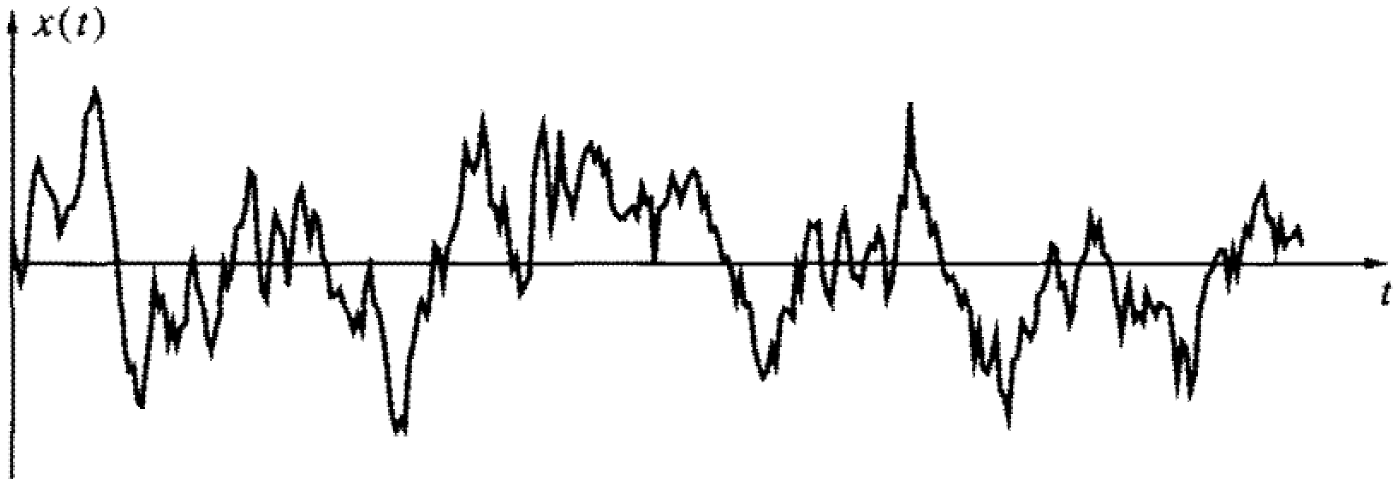
Random process in signal processing

- In signal processing, the random process considered focuses more on the joint distributions of two samples, $\boldsymbol{p}(\boldsymbol{X}_{n1}=x[n_1], \boldsymbol{X}_{n2}=x[n_2])$.
- In addition, the **shift-invariant** property is further imposed (called **stationary** process).
- Details will be given below.

Continuous-time Random Signals

- Random signals appear often in real life.
- Examples include:
  - 1. The noise heard from a radio receiver that is not tuned to an operating channel
  - 2. Electrical signals recorded from a human brain through electrodes put in contact with the skull (these are called electroencephalograms, or EEGs).
  - 3. Mechanical vibrations sensed in a vehicle moving on a rough terrain.

- 4. Angular motion of a boat in the sea caused by waves and wind.
- 5. Television signal
- 6. Radar signal



Definitions
(Oppenheim, Appendix)

- **Probability density function** (pdf) of $x[n]$: $p(x_n, n)$
  - The pdf is at time n.

- **Joint distribution** of two samples $x[n]$ and $x[m]$: $p\left(x_n, n, x_m, m\right)$
  - The joint pdf of times n and m

- Eg., $x_1[n] = A_n \cos(wn + \phi_n)$, where A_n and ϕ_n are random variables, $-\infty < n < \infty$, then $x_1[n]$ is a random process.

Independence and Stationary

- $x[n]$ and $x[m]$ are *independent* iff

$$p(x_n, x_m) = p(x_n)p(x_m)$$

- x is a **stationary process** iff

$$p(x_n, x_m) = p(x_{n+k}, x_{m+k})$$

for all k.

- That is, the joint distribution of $x[n]$ and $x[m]$ **depends only on the time difference** $m - n$.

Stationary (continue)

- In particular, the above definition of stationary process is also applicable to the situation of $m = n$.
- Hence, a stationary random process should also satisfy

$$p(x_n) = p(x_{n+k})$$

- That is, the probability density function (pdf) of a stationary process is fixed for all time n.

Stochastic Processes vs. Deterministic Signal

- In many applications, random processes serve as signal-source models.
  - A particular signal can be considered a sample sequence of a random process.
  - The signal sampled from a random process is not random any more, and is referred to as a deterministic signal.

Average Ensembles: Expectation

- Let us denote the *expectation* operator, $\varepsilon\{\cdot\}$, as

$$\varepsilon\{g(x_n)\} = \int_{-\infty}^{\infty} g(x_n)p(x_n)dx_n$$

- Mean (or average)

$$m_{x_n} = \varepsilon\{x_n\} = \int_{-\infty}^{\infty} x_n p(x_n)dx_n$$

Independent Variables

- For independent random variables, we have

$$\varepsilon\left\{x_n y_m\right\} = \varepsilon\left\{x_n\right\}\varepsilon\left\{y_m\right\}$$

Statistics: Mean Square Value and Variance

- Mean squared value (also called power of the random signal)

$$\varepsilon\{|x_n|^2\} = \int_{-\infty}^{\infty} |x_n|^2 p(x_n) dx_n$$

- Variance:

$$\boldsymbol{var}\{x_n\} = \varepsilon\left\{\left|x_n - m_{x_n}\right|^2\right\}$$

Autocorrelation and Autocovariance

Besides the above single-variable statistics, we can also define the statistics of joint variables between times n and m.

- Autocorrelation
  - correlation between time indices m and n of a random process

$$\phi_{xx}\{n,m\} = \varepsilon\{x_n x_m^*\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_n x_m^* p(x_n, x_m)\, dx_n\, dx_m$$

- Autocovariance
  - covariance between time indices m and n of a random process

$$\gamma_{xx}\{n,m\} = \varepsilon\left\{(x_n - m_{x_n})(x_m - m_{x_m})^*\right\}$$
$$= \phi_{xx}\{n,m\} - m_{x_n} m_{x_m}^*$$

Stationary Process

- According to the definition of stationary process, the autocorrelation of a stationary process is dependent only on the time difference $m - n$.
- Hence, for stationary process, we have the property that the mean and variance are independent to the time n:

$$m = m_{x_n} = \varepsilon\{x_n\}$$

$$\sigma_x^2 = \varepsilon\{(x_n - m_x)^2\}$$

- Aotocorrelation is dependent only to the **time difference** k,

$$\phi_{xx}(k) = \phi_{xx}\{n+k, n\} = \varepsilon\{x_{n+k} x_n^*\}$$

Wide-sense Stationary

- In the above, the stationary property is defined in the **strict sense** that the **pdf should remain the same for all time**.
- However, the definition is too strict sometimes. In most cases we require only the statistics up to the second order (i.e., mean, variance, covariance) invariant with time.
- To **relax the definition**, we call the process ***wide-sense stationary*** (W. S. S.) if the following equations hold (even though the joint distributions between time n and $n+k$ are not invariant with n.)

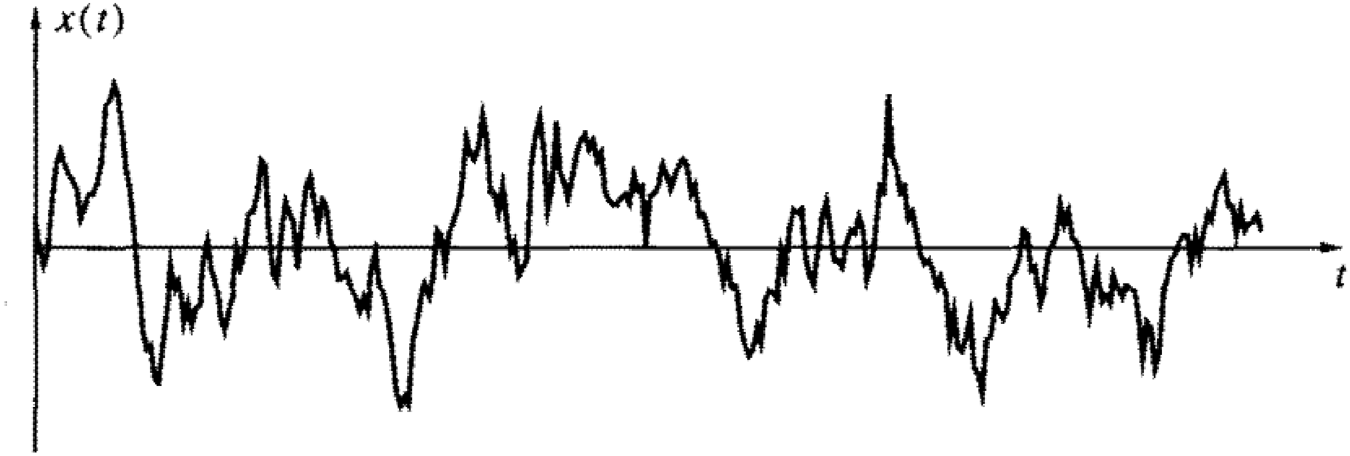
$$m_x = m_{x_n} = \varepsilon\{x_n\}$$

$$\sigma_x^2 = \varepsilon\{(x_n - m_x)^2\}$$

$$\phi_{xx}(n+k,n) = \phi_{xx}(k) = \varepsilon\{x_{n+k}x_n^*\}$$

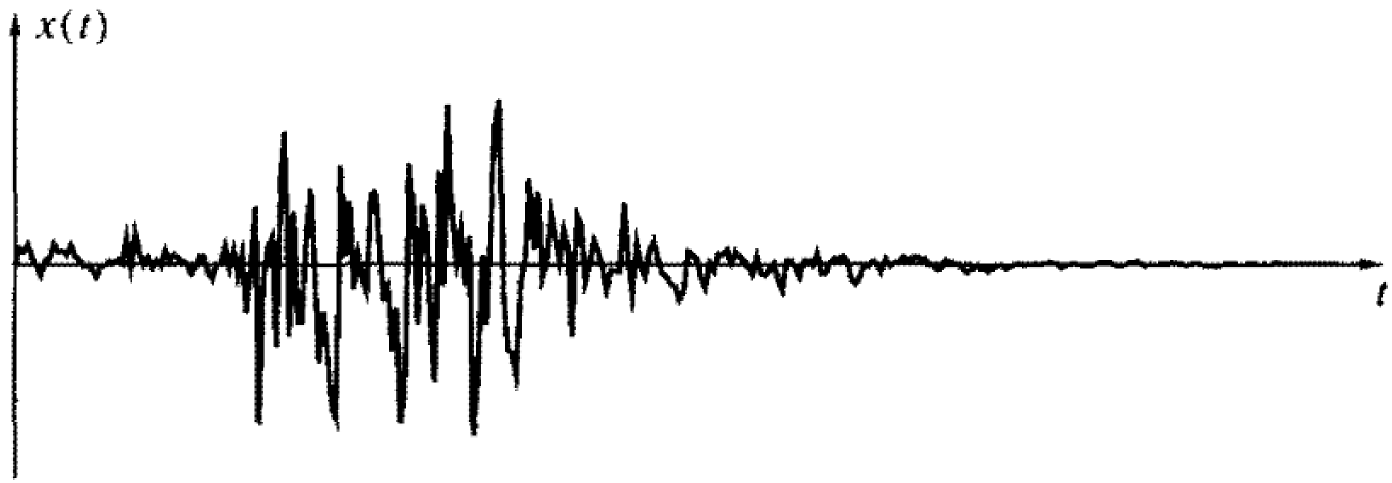
Example of Stationary Process

- As can be seen in the illustration, intuitively, a WSS signal looks more or less the same at different time intervals.
- Although its detailed form varies, its overall shape does not.



Example of nonstationary

- An example of a random signal that is not WSS is a seismic wave during an earthquake.
- As we see, the amplitude of the wave shortly before the beginning of the earthquake is small. At the start of the earthquake the amplitude grows suddenly, sustains its amplitude for a certain time, then decays.



W. S. S. correlation sequences

- **Definition**: Autocorrelation and cross-correlation sequences (or functions) of W. S. S. random process

$$\varphi_{xx}[m] = \varepsilon\left\{x_{n+m}x_n^*\right\}$$ autocorrelation sequence

$$\varphi_{xy}[m] = \varepsilon\left\{x_{n+m}y_n^*\right\}$$ cross-correlation sequence

- Due to the stationary property, the above sequences exist.

Autocorrelation sequence

- Autocorrelation sequence (or function) is a **deterministic signal** (not a random signal), which cannot be well defined for a random process that is not W. S. S.

$$\varphi_{xx}[m] = \varepsilon\{x_{n+m}x_n^*\}$$

- A W. S. S. random process can be well described via its autocorrelation sequence. The autocorrelation sequence exactly specifies a W. S. S. random process.

Properties of autocorrelation and cross-correlation sequences of stationary process

- For simplicity, in the following, when we use the term stationary process, we mean that the process is WSS.
- **Property of stationary process** (similar properties have already been shown in the deterministic case)

$$\left|\varphi_{xy}[m]\right|^2 \le \varphi_{xx}[0]\,\varphi_{yy}[0]$$

- The above implies

$$\boxed{\left|\varphi_{xx}[m]\right| \le \left|\varphi_{xx}[0]\right|}$$

Eg. Autocorrelation sequence of a First-order Markov (Autoregressive(1); AR(1)) Process

- $\frac{\varphi_{xx}(m)}{\varphi_{xx}(0)}$ is called the normalized autocorrelation sequence, its absolute value is in [0,1].

- If a zero-mean W. S. S. process has the following normalized autocorrelation sequence,

$$\boxed{\frac{\varphi_{xx}(m)}{\varphi_{xx}(0)} = \begin{cases} \rho^m, & m \geq 0 \\ 0, & \text{otherwise} \end{cases}}$$

we call the process an AR(1) process.

- $\rho \in [0,1]$ is called the correlation coefficient in the AR(1) process.

Eg. Autocorrelation sequence of a First-order Markov (Autoregressive(1); AR(1)) Process

- Since $\varphi_{xx}(0) = \varepsilon(xx^*) = \varepsilon(\|x\|^2) = \sigma^2$
  - i.e., $\varphi_{xx}(0)$ is the variance of the zero-mean W.S.S. signal.
- In two-sided AR(1), the autocorrelation sequence $\varphi_{xx}(m)$ is

$$\sigma^2[\cdots, \rho^3, \rho^2, \rho, 1, \rho, \rho^2, \rho^3, \cdots]$$

  - Eg., if $\rho = 0.95$ for a row in an image, we say that the image has a correlation coefficient of 0.95 between the adjacent pixels.

Eg. Autocorrelation sequence of a First-order Markov (Autoregressive(1); AR(1)) Process

- Why called AR(1)?
  - Because this process has the first-order Markov property, where the signal at locations i and j depends only on the location difference $m = |i - j|$ for $m = 1$ (i.e., adjacent locations).
  - The normalized autocorrelation between adjacent locations (i.e., $m = 1$) is ρ. For the other locations ($m = 2, 3, \cdots$) the autocorrelation is obtained by assuming the independence of conditional probability $\rightarrow \rho^m$. (Markov property)
  - AR(1) is quite often used to model real-world signals in time-series analysis.

Properties of autocorrelation and cross-correlation sequences (continue)

- **Property:** autocorrelation is shift invariant
  - If $y_n = x_{n-n_0}$

$$\varphi_{yy}\left[m\right] = \varphi_{xx}\left[m\right]$$

Properties of autocorrelation and cross-correlation sequences (continue)

- **Property** (definition of the **power** of a stationary signal)

$$\varphi_{xx}[0] = \varepsilon\left[\left|x_n\right|^2\right] =$$

The **mean squared value** (i.e., **power**) of the random signal;
Or the **variance** of a zero-mean random signal (i.e., the stationary process with $m_{x_n} = 0$).

$\boxed{\varphi_{xx}[0] \text{ is the power (i.e., } \varepsilon([x_n]^2)\text{) of the random signal}}$

Fourier Transform Representation of Random Signals

- Because the autocorrelation sequence of a random process is a deterministic signal, its DTFT exists.

- Let the DTFT of the autocorrelation sequence be

$$\varphi_{xx}[m] \leftrightarrow \Phi_{xx}\left(e^{j\omega}\right)$$

- By doing so, we can view a **W. S. S. random process** in the **spectral domain**.

Fourier Transform Representation of Random Signals (continue)

- Applying the inverse DTFT:

$$\varphi_{xx}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{xx}\left(e^{jw}\right)e^{jwm}dw$$

- Recall that the autocorrelation sequence at time 0:

$$\varphi_{xx}[0] = \varepsilon\left\{x[n]^2\right\}$$

- Consequently, when $m = 0$,

$$\underbrace{\varepsilon\left\{x[n]^2\right\}}_{\text{the power of the random signal.}} = \varphi_{xx}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{xx}\left(e^{jw}\right)dw$$

the **power** of the random signal.

Fourier Transform Representation of Random Signals (continue)

- Denote $P_{xx}(\omega) = \Phi\left(e^{j\omega}\right)$,

  i.e., the DTFT of the autocorrelation sequence is referred to as the **power spectral density (p.s.d.)**; power density spectrum (or **power spectrum**) of the WSS random process x.

Fourier Transform Representation of Random Signals (continue)

$$P_{xx}(\omega) = \Phi\left(e^{j\omega}\right)$$

- Hence, we have

$$\varepsilon\{x[n]^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\omega)\,d\omega$$

That is, the total area under power density in $[-\pi, \pi]$ is the total power of the random stationary signal.

Power Spectral Density

- In sum, from

$$\varepsilon\{x[n]^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\omega)d\omega$$

- $P_{xx}(\omega)$ can be treated as the "density" at the frequency ω of the total power.
- Integrating all the densities from $-\pi$ to π then constitutes the total power of a stationary random signal.
  - This is why $P_{xx}(\omega)$ is called power spectral density (or power density, power density spectrum).

White Noise

- A white noise is a stationary random signal for which

$$\varphi_{xx}[m] = \sigma_x^2 \delta[m]$$ delta function

i.e.,
$$\varphi_{xx}[m] = \begin{cases} \sigma_x^2 & m = 0 \\ 0 & m \neq 0 \end{cases}$$

- White noise is also a special case of the AR(1) process with the correlation coefficient $\rho = 0$.

- Because **its autocorrelation is a delta function**, its samples at different times (i.e., $m \neq 0$) are **uncorrelated** (i.e., equal to zero).

White Noise

- The power spectral density (i.e., DTFT of the autocorrelation) of a white noise is therefore a **constant**

$$\Phi_{xx}(e^{j\omega}) = \sigma_x^2, \quad \text{for all } \omega$$

- This is analogous to white light. Thus, it is called **white noise**.

White Gaussian Noise

- The average power of a white-noise is

$$\varphi_{xx}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{xx}(e^{j\omega})d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sigma_x^2 d\omega = \sigma_x^2$$

- **White Gaussian noise**: if a random process is both white noise and Gaussian process it is called a white Gaussian noise.