

EE6617: Rudiments of Random Signals

In engineering and science, signals are often treated and modeled as random process (or random signals). In this note, we briefly overview how to characterize random signals using cumulative distribution function (CDF), probability density function (PDF), mean, variance, etc.

Conditional probability and Bayes' rule

Consider a sample space $\mathcal{S}$ that consists of two events A and B .

- The **conditional probability** $P(A|B)$ is the probability that the event A occurs given that the event B has occurred.

$$P(A|B) = \frac{P(A, B)}{P(B)}.$$

- $P(A, B) = P(A \cap B)$ is the joint probability of the events A and B .
- If the events A and B are independent, $P(A, B) = P(A)P(B)$ and

$$P(A|B) = P(A),$$

i.e., the occurrence of B does not change the probability of A .

Let events A_i , $i = 1, \dots, n$ form a sample space $\cup_{i=1}^n A_i = \mathcal{S}$. Let $B \in \mathcal{S}$ be some event with non-zero probability.

- Marginal probability $P(B)$ has the equality

$$P(B) = \sum_{i=1}^n P(B, A_i) = \sum_{i=1}^n P(B|A_i)P(A_i).$$

Bayes' theorem is

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i, B)}{P(B)} \\ &= \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^n P(B|A_i)P(A_i)}. \end{aligned}$$

Note that many of classical estimators and detectors are obtained by properly manipulating conditional probabilities using Bayes's rule. We will see this throughout the whole lectures.

CDF and PDF

When the measurement $x[n]$ with n being the time index is a realization of a random variable, we denote the random variable by $\bar{x}[n]$. The density function of $\bar{x}[n]$ is written as $p_{\bar{x}[n]}(x[n])$. Note that

$$\int_{-\infty}^{\infty} p(x[n])dx[n] = 1 \quad \text{and} \quad p(x[n]) \geq 0$$

- **Cumulative distribution function (CDF)**

$$\Pr(\bar{x}[n] \leq x[n]) = F_{\bar{x}[n]}(x[n])$$

Note that $x[n]$ is a realization of $\bar{x}[n]$ and $0 \leq \Pr(\bar{x}[n] \leq x[n]) \leq 1$.

Consider two random process $\bar{x}[n]$ and $\bar{y}[m]$. The joint CDF of $\bar{x}[n]$ and $\bar{y}[m]$ is

$$F_{\bar{x}[n], \bar{y}[m]}(x[n], y[m]) = \Pr(\bar{x}[n] \leq x[n], \bar{y}[m] \leq y[m]).$$

Note that CDF is always monotonically increasing.

- **Probability density function (PDF)**

$$p_{\bar{x}[n]}(x[n]) = \frac{dF_{\bar{x}[n]}(x[n])}{dx[n]} \Rightarrow F_{\bar{x}[n]}(x[n]) = \int_{-\infty}^{x[n]} p_{\bar{x}[n]}(t) dt$$

The joint PDF of $\bar{x}[n]$ and $\bar{y}[m]$ is

$$p_{\bar{x}[n], \bar{y}[m]}(x[n], y[m]) = \frac{\partial^2 F_{\bar{x}[n], \bar{y}[m]}(x[n], y[m])}{\partial x[n] \partial y[m]} \Rightarrow F_{\bar{x}[n], \bar{y}[m]}(x[n], y[m]) = \int_{-\infty}^{x[n]} \int_{-\infty}^{y[m]} p_{\bar{x}[n], \bar{y}[m]}(t, s) dt ds.$$

- **Marginal PDFs**

$$p_{\bar{x}[n]}(x[n]) = \int_{-\infty}^{\infty} p_{\bar{x}[n], \bar{y}[m]}(x[n], s) ds, \quad p_{\bar{y}[m]}(y[m]) = \int_{-\infty}^{\infty} p_{\bar{x}[n], \bar{y}[m]}(t, y[m]) dt.$$

Note that PDF is always positive.

In what follows, we will omit subscripts $\bar{x}[n]$ and/or $\bar{y}[m]$ attached to CDF and PDF for simplicity.

Mean, Moments, and Variance

- The **mean** value of $\bar{x}[n]$ at time n is defined as

$$\mu[n] = E[\bar{x}[n]] = \int x[n] p(x[n]) dx[n]$$

Note that, in general, $\mu[n] \neq \mu[m]$, $m \neq n$. The mean value is also called **expected value** and **ensemble mean**.

- **m-th order moment** is the generalization of the mean value

$$E[\bar{x}[n]^m] = \int x[n]^m p(x[n]) dx[n]$$

- The **variance** of $\bar{x}[n]$ at time n is defined as

$$\sigma^2[n] = E[(\bar{x}[n] - \mu[n])^2] = \int (\bar{x}[n] - \mu[n])^2 p(x[n]) dx[n] = E[\bar{x}[n]^2] - \mu[n]^2$$

Someone also calls this as **second central moment**.

Autocorrelation and Covariance

- The **autocorrelation** of $\bar{x}[n]$ is defined as

$$R_{\bar{x}\bar{x}}(m, n) = E[\bar{x}[n]\bar{x}[m]] = \int x[n]x[m]p(x[n], x[m]) dx[n]dx[m].$$

It measures the dependence between $\bar{x}[n]$ and $\bar{x}[m]$. In particular,

- $R_{\bar{x}\bar{x}}(n, n) = E[\bar{x}[n]^2]$
- When $E[\bar{x}[n]] = 0$, $\sigma^2[n] = R_{\bar{x}\bar{x}}(n, n) = E[\bar{x}[n]^2]$.

- The **covariance** of $\bar{x}[n]$ is defined as

$$\begin{aligned} C_{\bar{x}\bar{x}}(m, n) &= E[(\bar{x}[n] - \mu[n])(\bar{x}[m] - \mu[m])] \\ &= E[\bar{x}[n]\bar{x}[m]] - \mu[n]\mu[m] = R_{\bar{x}\bar{x}}(m, n) - \mu[n]\mu[m] \end{aligned}$$

In particular,

- $C_{\bar{x}\bar{x}}(n, n) = E[(\bar{x}[n] - \mu[n])^2] = \sigma^2(n)$, i.e., it is the variance.
- When $E[\bar{x}[n]] = 0$, $C_{\bar{x}\bar{x}}(m, n) = R_{\bar{x}\bar{x}}(m, n)$

Crosscorrelation and Crosscovariance

- The **crosscorrelation** of $\bar{x}[m]$ and $\bar{y}[n]$ is defined as

$$R_{\bar{x}\bar{y}}(m, n) = E[\bar{x}[m]\bar{y}[n]] = \int x[m]y[n]p(x[m], y[n]) dx[m]dy[n]$$

It measures the correlation between $\bar{x}[m]$ and $\bar{y}[n]$.

- Similarly, the **crosscovariance** of $\bar{x}[m]$ and $\bar{y}[n]$ is defined as

$$C_{\bar{x}\bar{y}}(m, n) = E[(\bar{x}[m] - E[\bar{x}[m]])(\bar{y}[n] - E[\bar{y}[n]])] = R_{\bar{x}\bar{y}}(m, n) - E[\bar{x}[m]]E[\bar{y}[n]]$$

If $C_{\bar{x}\bar{y}}(m, n) = 0$, the signals $\bar{x}[m]$ and $\bar{y}[n]$ are said to be **uncorrelated**.

Uncorrelated vs. Independent

- Two random variables $\bar{x}[m]$ and $\bar{y}[n]$ are said to be **uncorrelated** if $C_{\bar{x}\bar{y}}(m, n) = 0$.
- Two random variables $\bar{x}[m]$ and $\bar{y}[n]$ are said to be **independent** if $p(x[m], y[n]) = p(x[m])p(y[n])$.
- The uncorrelated signal satisfies

$$R_{\bar{x}\bar{y}}(m, n) = E[\bar{x}[m]]E[\bar{y}[n]], \tag{1}$$

while the independent signal satisfies

$$E[\bar{x}[m]\bar{y}[n]] = E[\bar{x}[m]]E[\bar{y}[n]]. \tag{2}$$

Between (1) and (2), which one is stronger than the other?

Stationary

- If k -degree joint PDF $p(x[n_1], x[n_2], \dots, x[n_k])$ is **shift-invariant** for any k and any set $\{n_i\}_{i=1}^k$, the $x[n]$ is said to be **strictly stationary**, i.e.,

$$p(x[n_1], x[n_2], \dots, x[n_k]) = p(x[n_1 + n_0], x[n_2 + n_0], \dots, x[n_k + n_0])$$

- A random process is said to be **wide-sense stationary (WSS)**, if the mean and autocorrelation are shift invariant. This is

$$\begin{aligned}\mu &= E[\bar{x}[n]] = E[\bar{x}[m]] \\ R_{\bar{x}\bar{x}}(m, n) &= E[\bar{x}[m]\bar{x}[n]] = R_{\bar{x}\bar{x}}(m - n) = R_{\bar{x}\bar{x}}(i)\end{aligned}$$

where $i = m - n$ is called the **correlation lag**.

- Important properties of WSS random process

- $R_{\bar{x}\bar{x}}(i)$ is symmetric about the origin

$$R_{\bar{x}\bar{x}}(i) = R_{\bar{x}\bar{x}}(-i)$$

- The power, $R_{\bar{x}\bar{x}}(0)$, is greater than or equal to the magnitude of the correlation for any other lag, i.e.,

$$R_{\bar{x}\bar{x}}(0) \geq |R_{\bar{x}\bar{x}}(i)|, \quad i \neq 0.$$

This can be shown through the following steps

$$\begin{aligned}0 &\leq E[(\bar{x}[n+i] - \bar{x}[n])^2] \\ &= E[\bar{x}[n+i]^2] + E[\bar{x}[n]^2] - 2E[\bar{x}[n+i]\bar{x}[n]] \\ &= 2R_{\bar{x}\bar{x}}(0) - 2R_{\bar{x}\bar{x}}(i)\end{aligned}$$

(3)

- When $\bar{x}[n]$ is zero-mean, then

$$\sigma^2 = C_{\bar{x}\bar{x}}(0) = R_{\bar{x}\bar{x}}(0)$$

- Same arguments hold for $C_{\bar{x}\bar{x}}(i)$

Ergodicity

- A **stationary** process is said to be **ergodic**, if its time average with infinite samples equals to its ensemble average (i.e., $E[\cdot]$).
- The statistical property of the process can be determined by time averaging over **a single sample function** (i.e., $x[n]$) of **the process** $\bar{x}[n]$.
- Ergodic in the mean if

$$\mu[n] = E[\bar{x}[n]] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x[n]$$

- Ergodic in the autocorrelation if

$$R_{\bar{x}\bar{x}}(i) = E[\bar{x}[n]\bar{x}[n-i]] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_0^{N-1} x[n]x[n-i]$$

Unless stated otherwise, we assume **ergodic (and thus stationary) random process** in the course. Thus, in what follows, we use x to represent the random process.

Gaussian Distribution - Scalar $x \in \mathbb{R}$

- The PDF of **Gaussian distribution** with mean μ and variance σ^2

$$p(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

We write $x \sim \mathcal{N}(\mu, \sigma^2)$.

- If $x \sim \mathcal{N}(0, 1)$, it is called **normal distributed** and we have

$$p(x) = \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{x^2}{2}\right)$$

- In particular, we will often deal with **right tail probability** function:

$$\Pr(x < a) = \int_{-\infty}^a \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{t^2}{2}\right) dt = \Phi(a)$$

and **Gaussian Q-function**

$$\Pr(x \geq a) = 1 - \Phi(a) = \int_a^{\infty} \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{t^2}{2}\right) dt = Q(a)$$

Multivariate Gaussian Distribution - Vector $\mathbf{x} \in \mathbb{R}^{N \times 1}$

We often deal with multiple random variables

$$\begin{aligned} \mathbf{x} &= [x[1] \ x[2] \ \dots \ x[N-1]]^T \in \mathbb{R}^{N \times 1} \\ E[\mathbf{x}] &= \boldsymbol{\mu} \text{ and } E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] = \mathbf{C} \in \mathbb{R}^{N \times N} \end{aligned}$$

where $\boldsymbol{\mu}_n = E[x[n]]$ and $\mathbf{C}$ is the covariance matrix. Notice $\mathbf{C} = \mathbf{C}^T$ by WSS.

- The multivariate Gaussian PDF is

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}(\det(\mathbf{C}))^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$$

and we write $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$

- If $\{x[n]\}_{n=0}^{N-1}$ are mutually independent and identically distributed (i.i.d.) with $x \sim \mathcal{N}(\mu, \sigma^2)$, then

$$p(\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - \mu)^2 \right)$$

and we write $\mathbf{x} \sim \mathcal{N}(\mu \mathbf{1}, \sigma^2 \mathbf{I})$, where $\mathbf{1} = [1 \ 1 \dots 1]^T$ and $\mathbf{I}$ denotes the identity matrix.

Multivariate Gaussian Statistics

We often deal with functions of i.i.d. random variables

- Sum of i.i.d $\mathbf{x}$ with $x \sim \mathcal{N}(\mu, \sigma^2)$: $T(\mathbf{x}) = \sum_{n=0}^{N-1} x[n]$

$$\begin{aligned} E[T(\mathbf{x})] &= N\mu \\ \text{var}(T(\mathbf{x})) &= N\sigma^2 \end{aligned}$$

- Sample mean of i.i.d $\mathbf{x}$ with $x \sim \mathcal{N}(\mu, \sigma^2)$: $S(\mathbf{x}) = \frac{1}{N}T(\mathbf{x})$

$$\begin{aligned} E[S(\mathbf{x})] &= \mu \\ \text{var}(S(\mathbf{x})) &= \frac{\sigma^2}{N} \end{aligned}$$

- Linear function of $T(\mathbf{x})$, $L(\mathbf{x}) = aT(\mathbf{x}) + b$ for $a, b \in \mathcal{R}$

$$\begin{aligned} E[L(\mathbf{x}) = aT(\mathbf{x}) + b] &= aE[T(\mathbf{x})] + b = aN\mu + b \\ \text{var}(L(\mathbf{x})) &= a^2 \text{var}(T(\mathbf{x})) = a^2 N\sigma^2 \end{aligned}$$