3 Review of probability and random signals

3.1 Probability theory

Axioms of probability theory: A probability system consists of the triplet:

1. A sample space $\mathcal{S}$ of elementary events (experiment outcomes)
2. A class $\mathcal{E}$ of events that are subsets of $\mathcal{S}$
3. A probability measure $P(\cdot)$ assigned to each event $\mathcal{A}$ in the class $\mathcal{E}$, which has the following properties
   (a) $P(\mathcal{S}) = 1$
   (b) $0 \leq P(\mathcal{A}) \leq 1$
   (c) If $\mathcal{A}_1, \mathcal{A}_2, \ldots$ is any countable sequence of mutually disjoint events (i.e. $\mathcal{A}_m \bigcap \mathcal{A}_n = \emptyset$) in the class $\mathcal{E}$, then

$$P\left(\bigcup_{i=1}^{\infty} \mathcal{A}_i\right) = \sum_{i=1}^{\infty} P\left(\mathcal{A}_i\right)$$

3.2 Random variables (RV), random vectors (Rv)

a) random variables

Random variable:

function X: $\mathcal{S} \rightarrow I\!R$ such that for any $x \in I\!R \quad \{s \in \mathcal{S} : \; X(s) \leq x\} \in \mathcal{E}$

Discrete random variable: X takes only a finite number of values in any finite interval.
Continuous random variable: X can take continuous values.

Cumulative or (probability) distribution function (cdf):

$$F_X(x) = P\Big(\{s \in \mathcal{S} : \; X(s) \leq x\}\Big) = P\left(X \leq x\right) \qquad \text{for short}$$

$F_x(x)$ is monotonic, non-decreasing and satisfies $0 \leq F_X(x) \leq 1$.
Probability density function (pdf):

$$f_X(x) = \frac{dF_X(x)}{dx} \quad \text{(Note that } f_X(x) \text{ satisfies } \int_{-\infty}^{\infty} f_X(x) dx = 1\text{)}$$

cdf in terms of pdf for continuous random variables:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(y)dy$$

cdf in terms of pdf for discrete random variables: Let X be a discrete random variable taking on values x_i with probability mass function (pmf) $\{p_i = P(X = x_i)\}$.

$$\begin{aligned}
F_X(x) = P(X \leq x) &= \sum_{y \leq x} P(X = y) && \text{where } P(X = y) = 0 \text{ for } y \neq x_i \\
&= \sum_i p_i u(x - x_i) && \text{where } u(x)\text{: unit step function} \\
&= \int_{-\infty}^{x} f_X(y)dy && \text{with } f_X(x) \triangleq \sum_i p_i \delta(x - x_i)
\end{aligned}$$

b) random vectors

A random vector (Rv) $\boldsymbol{X}$ is a vector of random variables:

$$\boldsymbol{X} = [X_1, \ldots X_N]^T$$

where for $i = 1, \ldots, N$ X_i are random variables.

Joint cumulative distribution (cdf) of a random vector:

$$\begin{aligned}
F_{\boldsymbol{X}}(\boldsymbol{x}) &= P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_N \leq x_N) \\
&= \int_{-\infty}^{x_N} \cdots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{\boldsymbol{X}}(y_1, y_2, \ldots, y_N) dy_1 dy_2 \ldots dy_N = \int_{-\infty}^{\boldsymbol{x}} f_{\boldsymbol{X}}(\boldsymbol{y}) d\boldsymbol{y} \quad \text{(continuous Rv)} \\
&= \sum_{y_N \leq x_N} \cdots \sum_{y_2 \leq x_2} \sum_{y_1 \leq x_1} P(X_1 = y_1, X_2 = y_2, \ldots, X_N = y_N) = \sum_{\boldsymbol{y} \leq \boldsymbol{x}} P(\boldsymbol{X} = \boldsymbol{y}) \quad \text{(discrete Rv)}
\end{aligned}$$

where $\boldsymbol{x} = [x_1, \ldots, x_N]^T$ and $\boldsymbol{y} = [y_1, \ldots, y_N]^T$.

The *marginal probabilities density functions* for a subset of the random vector are obtained by integrating (or summing for discrete random vectors) the other variables out.

Example: N=2

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx \qquad P(X = x) = \sum_y P(X = x, Y = y)$$

Necessary and sufficient conditions for independence of the random variables $\{X_i\}_{i=1,\ldots,N}$*:*

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{N} F_{X_i}(x_i) \qquad \text{(continuous and discrete RV's)}$$

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{N} f_{X_i}(x_i) \qquad \text{(continuous RV's)}$$

$$P(\boldsymbol{X} = \boldsymbol{x}) = \prod_{i=1}^{N} P(X_i = x_i) \qquad \text{(discrete RV's)}$$

Probability that $\boldsymbol{X}$ *is in a set* $\mathcal{Z}$*:* Let $I\!R^N$ be the N-dimensional space of real vectors. Let $\mathcal{Z} \subseteq I\!R^N$ be a subset of $I\!R^N$ for which a probability measure can be defined (i.e. a Borel set or equivalently $\mathcal{Z}$ is a countable union or countable intersection of N-dimensional cells defined as $(a_1, b_1] \times (a_2, b_2] \times \ldots (a_N, b_N]$), then

$$\begin{aligned} P(\boldsymbol{X} \subseteq \mathcal{Z}) &= \int_{\mathcal{Z}} f_{\boldsymbol{X}}(\boldsymbol{y}) d\boldsymbol{y} && \text{(continuous random vector } \boldsymbol{X}\text{)} \\ &= \sum_{y \in \mathcal{Z}} P(\boldsymbol{X} = \boldsymbol{y}) && \text{(discrete random vector } \boldsymbol{X}\text{)} \end{aligned}$$

Conditional cdf, pdf and pmf: Let $\boldsymbol{X}$ be N-dimensional and $\boldsymbol{Y}$ be M-dimensional random vectors. Let $\mathcal{A} \subseteq I\!R^N$, $\mathcal{B} \subseteq I\!R^M$, then

$$P(\boldsymbol{X} \in \mathcal{A} | \boldsymbol{Y} \in \mathcal{B}) = \frac{P(\boldsymbol{X} \in \mathcal{A}, \boldsymbol{Y} \in \mathcal{B})}{(\boldsymbol{Y} \in \mathcal{B})}$$

For continuous random vectors, the conditional cdf and pdf are defined as

$$F_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = P(\boldsymbol{X} \leq \boldsymbol{x} | \boldsymbol{Y} = \boldsymbol{y}) = \int_{-\infty}^{x_N} \ldots \int_{-\infty}^{x_1} f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{u}|\boldsymbol{y}) d\boldsymbol{u} \quad \text{conditional cdf (continuous Rv)}$$

$$f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = \frac{f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x}, \boldsymbol{y})}{f_{\boldsymbol{Y}}(\boldsymbol{y})} \quad \text{if } f_{\boldsymbol{Y}}(\boldsymbol{y}) \neq 0 \qquad \text{conditional pdf (continuous Rv)}$$

Remark: $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent iff $f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = f_{\boldsymbol{X}}(\boldsymbol{x})$.

For discrete random vectors, the conditional cdf and pmf are defined as

$$\begin{aligned} F_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = P(\boldsymbol{X} \leq \boldsymbol{x} | \boldsymbol{Y} = \boldsymbol{y}) &= \frac{P(\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Y} = \boldsymbol{y})}{P(\boldsymbol{Y} = \boldsymbol{y})} = \frac{\sum_{\boldsymbol{z} \leq \boldsymbol{x}} P(\boldsymbol{X} = \boldsymbol{z}, \boldsymbol{Y} = \boldsymbol{y})}{P(\boldsymbol{Y} = \boldsymbol{y})} \\ &= \sum_{\boldsymbol{z} \leq \boldsymbol{x}} P(\boldsymbol{X} = \boldsymbol{z} | \boldsymbol{Y} = \boldsymbol{y}) \quad \text{conditional pdf (discrete Rv)} \end{aligned}$$

$$P\left(\boldsymbol{X}=\boldsymbol{z}|\boldsymbol{Y}=\boldsymbol{y}\right)=\frac{P\left(\boldsymbol{X}=\boldsymbol{z},\boldsymbol{Y}=\boldsymbol{y}\right)}{P\left(\boldsymbol{Y}=\boldsymbol{y}\right)} \quad \text{conditional pmf (given } \boldsymbol{Y}=\boldsymbol{y}\text{) (discrete Rv)}$$

Average and moments:

Let $\boldsymbol{X}$ be an N-dimensional random vector and let $g(\boldsymbol{X})$ and $h(\boldsymbol{X})$ be two functions of $\boldsymbol{X}$ such that $g(\boldsymbol{X}) \in \mathbb{R}$ and $h(\boldsymbol{X}) \in \mathbb{R}^M$. Then $g(\boldsymbol{X})$ is a random variable and $h(\boldsymbol{X})$ is a random vector. Their expectations are given by

$$E\left[g(\boldsymbol{X})\right]=\begin{cases}\displaystyle\int_{\mathbb{R}^N} g(\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x} & \text{continuous } \boldsymbol{X}\\ \displaystyle\sum_{\boldsymbol{y}\in\mathbb{R}^N} g(\boldsymbol{y})P\left(\boldsymbol{X}=\boldsymbol{y}\right) & \text{discrete } \boldsymbol{X}\end{cases} \qquad (E\left[g(\boldsymbol{X})\right] \text{ is a scalar})$$

$$E\left[h(\boldsymbol{X})\right]=\begin{cases}\displaystyle\int_{\mathbb{R}^N} h(\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x} & \text{continuous } \boldsymbol{X}\\ \displaystyle\sum_{\boldsymbol{y}\in\mathbb{R}^N} h(\boldsymbol{y})P\left(\boldsymbol{X}=\boldsymbol{y}\right) & \text{discrete } \boldsymbol{X}\end{cases} \qquad (E\left[h(\boldsymbol{X})\right] \in \mathbb{R}^M \text{ (vector)})$$

Examples:

Mean (vector) ($h(\boldsymbol{X})=\boldsymbol{X}$):

$$E\left[\boldsymbol{X}\right]=\Big[E[X_1]E[X_2]\ldots E[X_N]\Big]^T$$

Correlation matrix ($h(\boldsymbol{X})=\boldsymbol{X}\boldsymbol{X}^T$):

$$E\left[\boldsymbol{X}\boldsymbol{X}^T\right]=\begin{pmatrix} E[X_1^2] & E[X_1X_2] & \ldots & E[X_1X_N]\\ E[X_2X_1] & E[X_2^2] & \ldots & E[X_2X_N]\\ \vdots & \vdots & \ddots & \\ E[X_NX_1] & E[X_NX_2] & \ldots & E[X_N^2]\end{pmatrix} \quad (\text{matrix with } ij^{th} \text{ equal to } E[X_iX_j])$$

Characteristic function ($g(\boldsymbol{X})=e^{j\boldsymbol{\omega}^T\boldsymbol{X}}$):

$$\phi_{\boldsymbol{X}}(\omega_1,\ldots,\omega_N)=E\left[e^{j\boldsymbol{\omega}^T\boldsymbol{X}}\right]=E\left[e^{j\sum_{i=1}^N \omega_i X_i}\right]$$

where $\boldsymbol{\omega}=\left[\omega_1,\ldots,\omega_N\right]^T$.

Special case of one dimensional Rv X:

$$E\left[X\right]=\mu_X \qquad \text{(mean)}$$

$$E\left[(X-\mu_X)^2\right]=\sigma_X^2=\text{var}(X) \qquad \text{(variance)}$$

$$\sigma_X=\sqrt{\text{var}(X)} \qquad \text{(standard deviation)}$$

Special case of 2 random variables with means μ_X, μ_Y:

$$E[XY] \qquad \text{((cross)-correlation of } X \text{ and } Y\text{)}$$

$$E[(X-\mu_X)(Y-\mu_Y)] = \text{cov}(XY) = E[XY] - \mu_X\mu_Y \qquad \text{((cross)-covariance of } X \text{ and } Y\text{)}$$

$$\rho = \frac{\text{cov}(XY)}{\sigma_X\sigma_Y} \qquad \text{(correlation coefficient)}$$

X and Y are uncorrelated iff $\text{cov}(XY) = 0$ (i.e. $\rho = 0$). X and Y are orthogonal iff $E(XY) = 0$.

3.3 Random processes

a) Definition

Consider a random experiment with outcomes $s \in \mathcal{S}$ ($\mathcal{S}$: sample space). Suppose that we assign to each sample point s a function of time $X(t,s)$ where t belongs to an index set $\mathcal{I}$. For any s_0 fixed, $X(t,s_0)$ is called a **realization or sample function** of the random process $X(t,s)$. For every fixed $t_k \in \mathcal{I}$, $X(t_k,s)$ is a random variable (function of S to $I\!R$). Thus a **random process** is an indexed ensemble (family) of random variables $\{X(t,s)\}_{t\in\mathcal{I}}$. Usually $X(t,s)$ is denoted by $X(t)$ where the parameter s has been omitted.

Example: Let θ be a number selected at random in $[-\pi,\pi]$, then $X(t,\theta) = \cos(2\pi t + \theta)$ is a random process. For fixed t_0 $X(t_0,\theta)$ is a random variable. For a particular outcome $\theta = \theta_0$, $\cos(2\pi t + \theta_0)$ is a realization of $X(t,\theta)$.

b) Characterization of a random process

Let $X_i = X(t_i,s)$ be the samples of the random process $X(t,s)$ at the instants t_i. A random process is specified by the collection of the k^{th} order joint cumulative distribution functions

$$F_{X_1,\ldots,X_k}(x_1,x_2,\ldots,x_k) = P(X_1 \le x_1, X_2 \le x_2, \ldots, X_k \le x_k)$$

for any k and any choice of sampling instants $t_1, \ldots, t_k$.

A discrete-valued random process can be specified by the collection of probability mass functions $P(X_1 = x_1, \ldots, X_k = x_k)$.

A continuous-valued random process can be specified by the collection of probability density functions $f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)$.

c) Moments of a random process

Mean:

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

where $f_{X(t)}(x)$ is the pdf of $X(t)$ when t is fixed. Generally since $f_{X(t)}(x)$ depends on t as indicated by the presence of t in $X(t)$, $m_X(t)$ is a function of time.

Let t_0 be fixed, then $X_0 = X(t_0)$ is a random variable with pdf $f_{X(t_0)}(x) = f_{X_0}(x)$ and mean

$$m_x(t_0) = E\left[X(t_0)\right] = \int_{-\infty}^{\infty} x f_{X_0}(x) dx$$

$f_{X_0}(x)$ depends generally on t_0. Note that the notation usually omits the dependence of $f_{X(t)}(x)$ on t.

Autocorrelation:

$$R_X(t_1, t_2) = E\left[X(t_1)X(t_2)\right] = \int_{-\infty}^{\infty} xy f_{X(t_1)X(t_2)}(x, y) dx dy$$

where $f_{X(t_1)X(t_2)}(x, y)$ is the joint pdf of $X(t_1)$ and $X(t_2)$ when t_1 and t_2 are fixed (second order pdf of $X(t)$). Generally $R_X(t_1, t_2)$ is a function of t_1 and t_2.

Auto-covariance:

$$C_X(t_1, t_2) = E\Big[\left(X(t_1) - m_X(t_1)\right)\left(X(t_2) - m_X(t_2)\right)\Big] = R_X(t_1, t_2) - m_X(t_1)m_X(t_2)$$

Cross-correlation of $X(t)$ and $Y(t)$:

$$R_{XY}(t_1, t_2) = E\left[X(t_1)Y(t_2)\right] = \int_{-\infty}^{\infty} xy f_{X(t_1)Y(t_2)}(x, y) dx dy$$

Example: Consider a sinusoidal signal with a random phase defined by

$$X(t) = A\cos\left(2\pi f_c t + \theta\right)$$

where A and f_c are constants and θ is a random variable that is uniformly distributed between $-\pi$ and π.

The mean :

$$m_X(t) = E[X(t)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(2\pi f_c t + y)\, dy = 0$$

The Auto-covariance and autocorrelation :

$$\begin{aligned} C_X(t_1, t_2) = R_X(t_1, t_2) &= E[\cos(2\pi f_c t_1 + \theta)\cos(2\pi f_c t_2 + \theta)] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(2\pi f_c t_1 + y)\cos(2\pi f_c t_2 + y)\, dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}[\cos(2\pi f_c(t_1 - t_2)) + \cos(2\pi f_c(t_1 + t_2) + 2y)]\, dy \\ &= \frac{1}{2}\cos(2\pi f_c(t_1 - t_2)) \end{aligned}$$

d) Stationarity of a random process

A random process with time invariant statistical properties is called a stationary random process. Formally let $X(t_1), X(t_2), \ldots, X(t_k)$ denote the random variables obtained by sampling the random process $X(t)$ at times $t_1, t_2, \ldots, t_k$ and let $X(t_1+\tau), X(t_2+\tau), \ldots, X(t_k+\tau)$ be a new set of random variables obtained by shifting the sampling instants by a fixed amount τ. The random process is said to be **stationary in the strict sense** if

$$F_{X(t_1+\tau),\ldots X(t_k+\tau)}(x_1, \ldots, x_k) = F_{X(t_1),\ldots X(t_k)}(x_1, \ldots, x_k)$$

for all time shifts τ, all k and all possible choices of sampling instants $t_1, \ldots, t_k$, where $F_{X(t_1),\ldots X(t_k)}(x_1, \ldots, x_k)$ is the joint cumulative distribution of $X(t_1), \ldots, X(t_k)$ and $F_{X(t_1+\tau),\ldots X(t_k+\tau)}(x_1, \ldots, x_k)$ is the joint cdf of $X(t_1+\tau), \ldots, X(t_k+\tau)$.

In particular, the first order distribution of a stationary random process must be independent of t and its second order distribution can depend only on the time difference between the samples.

Prove these results by considering special values of τ.

Therefore to prove that a process is **not** stationary in the strict sense, it is enough to show that one of the following condition does **not hold**:

$$
\begin{aligned}
m_X(t) &\triangleq E\left[X(t)\right] = m && \text{for all } t && (1)\\
var(X(t)) &\triangleq E\left[(X(t) - m_X(t))^2\right] = \sigma^2 && \text{for all } t && (2)\\
R_X(t_1, t_2) &\triangleq E\left[X(t_1)X(t_2)\right] = R_X(t_1 - t_2) && \text{for all } t_1, t_2 && (3)\\
C_X(t_1, t_2) &\triangleq E\left[\big(X(t_1) - m_X(t_1)\big)\big(X(t_2) - m_X(t_2)\big)\right] = C_X(t_1 - t_2) && \text{for all } t_1, t_2 && (4)
\end{aligned}
$$

e) Wide-Sense stationarity of a random process

A random process with time invariant mean and autocorrelation that depends only on the time difference $t_1 - t_2$ is called a **Wide-Sense Stationary (WSS) random process**. Its autocorrelation function is denoted by:

$$R_X(\tau) = E\left[X(t+\tau)X(t)\right] = E\left[X(t)X(t-\tau)\right]$$

$X(t)$ and $Y(t)$ are called jointly WSS if each of them is WSS and their cross-correlation function $R_{XY}(t_1, t_2) = E\left[X(t_1)Y(t_2)\right]$ depends only on $t_1 - t_2$.

Give an example of a WSS random process.

f) (Wide-Sense) cyclo stationarity (WS cyclo S) of a random process

A random process is called (wide-sense) cyclo stationary if its mean $m_X(t)$ and autocorrelation function $R_X(t_1, t_2)$ are invariant with respect to shifts in the time origin by integer multiples of some period T, that is, for every integer m,

$$m_X(t + mT) = m_X(t) \tag{5}$$

$$R_X(t_1 + mT, t_2 + mT) = R_X(t_1, t_2) \tag{6}$$

Equivalently, $m_X(t)$ and $R_X(t + \tau, t)$ are periodic in t with the same "period" T, defined as the highest of the two periods of $m_X(t)$ and $R_X(t + \tau, t)$ in case they do not have the same period.

Example: Let $m(t)$ be a WSS random process, show that $X(t) = m(t)\cos(2\pi f_c t)$ is (wide-sense) cyclo stationary with common "period" $\frac{1}{f_c}$.

h) Power spectral density

The Einstein-Wiener-Khinchin theorem:

Let $X(t)$ be a continuous-valued wide-sense stationary random process with mean m_X and autocorrelation function $R_X(\tau)$. The **power spectral density** of $X(t)$ is given by the Fourier transform of $R_X(\tau)$:

$$S_X(f) = \mathcal{F}\left\{R_X(\tau)\right\} = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f \tau} d\tau$$

"Generalization" of the Einstein-Wiener-Khinchin theorem for wide-sense cyclo stationary random processes: Let $X(t)$ be a WS cyclo S random process with "period" T_0. Then $X_S(t) = X(t + \theta)$ where θ is uniformly distributed between $-\frac{T_0}{2}$ and $\frac{T_0}{2}$ is WSS with mean

$$E\left[X_S(t)\right] = < m_X(t) >_t = \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} m_X(t) dt$$

and autocorrelation

$$R_{X_S}(\tau) = E\left[X_S(t+\tau)X_S(t)\right] = \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} R_X(t+\tau, t)dt$$
$$=< R_X(t+\tau, t) >_t = R_X^a(\tau)$$

where $R_X^a(\tau)$ is the time average of the autocorrelation function of $X(t)$. Then the power spectral density of $X(t)$ is defined as

$$S_X(f) = \mathcal{F}\left\{R_X^a(\tau)\right\} = \int_{-\infty}^{\infty} R_X^a(\tau)e^{-j2\pi f\tau}d\tau$$

h) Average power of $X(t)$

The average power of a real random process $X(t)$ is defined as

$$P_X = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T2} E\left[X^2(t)\right]dt$$

If $X(t)$ is WSS then $E\left[X^2(t)\right] = R_X(0)$, thus

$$P_X = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} R_X(0)dt = R_X(0) = E\left[X^2(t)\right] = \int_{-\infty}^{\infty} S_X(f)df$$

For any random process $X(t)$, $E\left[X^2(t)\right] = R_X(t,t)$, thus if $X(t)$ is WS cyclo S

$$P_X = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} R_X(t,t)dt = \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} R_X(t,t)dt = R_X^a(0) = \int_{-\infty}^{\infty} S_X(f)df$$

since $R_X(t,t)$ is periodic in t with period T_0.

i) Ergodicity of a random process

Measurement of averages $E\left[g(X(t))\right]$:

- Repeat the random experiment that yields the random process a large number of times and take arithmetic average of realizations.

  Example: estimate of the mean of $X(t)$

$$\widehat{m_X(t)} = \frac{1}{N}\sum_{i=1}^{N} X(t, s_i)$$

where N is the number of repetitions of the experiment, and $X(t, s_i)$ is the realization observed in the i^{th} repetition.

- Use time average of a single realization $X(t, s_0)$

$$< g(X(t, s_0)) >_t = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} g\left(X(t, s_0)\right) dt$$

If $< g(X(t, s_0) >_t = E\left[g(X(t))\right]$ for all functions $g(\cdot)$, then the random process is called **ergodic**.

Ergodic in the mean:

$$E\left[X(t)\right] = < X(t, s_0) >_t = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t, s_0) dt$$

Give an example of a process ergodic in the mean.

Ergodic in the autocorrelation function:

$$E\left[X(t+\tau)X(t)\right] = < X(t+\tau, s_0)X(t, s_0) >_t = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t+\tau, s_0)X(t, s_0) dt$$

Hence the average power of a process ergodic in the autocorrelation function is

$$P_X = E\left[X^2(t)\right] = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} X^2(t, s_0) dt \quad \text{(independent of } t\text{)}$$

Ergodicity implies stationarity in the strict sense but the converse is not true.

j) Gaussian random processes

A random process $X(t)$ is called a **Gaussian random process** if the samples $X_i = X(t_i), i = 1, \ldots k$ are jointly Gaussian random variables for all k, and all choices of $t_1, \ldots, t_k$, i.e.

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{k/2}\left[\det(C)\right]^{1/2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{m}\right)^T C^{-1}\left(\boldsymbol{x}-\boldsymbol{m}\right)\right\}$$

where $\boldsymbol{x} = \left[x_1, \ldots, x_k\right]^T$ and

$$\begin{aligned} \boldsymbol{m} &= \left[m_X(t_1), \ldots, m_X(t_k)\right]^T \\ C &= E\left[\left(\boldsymbol{X}-\boldsymbol{m}\right)\left(\boldsymbol{X}-\boldsymbol{m}\right)^T\right] \end{aligned}$$

$$
\begin{aligned}
&= \begin{pmatrix} C_X(t_1,t_1) & C_X(t_1,t_2) & \dots & C_X(t_1,t_k) \\ C_X(t_2,t_1) & C_X(t_2,t_2) & \dots & C_X(t_2,t_k) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_k,t_1) & C_X(t_k,t_2) & \dots & C_X(t_k,t_k) \end{pmatrix} \\
\boldsymbol{X} &= [X(t_1), \dots, X(t_k)]^T
\end{aligned}
$$

Properties of Gaussian random processes:

- A Gaussian random process is completely specified by its mean $m_X(t)$ and its autocovariance function $C_X(t_1, t_2)$.
- A wide-sense stationary Gaussian random process is stationary in the strict sense.
- If a Gaussian random process is applied to a stable linear system, the response is also a Gaussian random process.
- A weighted sum of jointly Gaussian random processes is a Gaussian random process, a linear combination of Gaussian random variables is also a Gaussian random variable.

k) White process

A process is said to be white if

$$
\begin{aligned}
S_X(f) &= K \quad \text{constant} \\
R_X(t_1, t_2) &= K\delta(t_1 - t_2)
\end{aligned}
$$

l) Response of a linear time invariant system to a WSS or WS cyclo S random process

Let $X(t)$ be a wide-sense stationary or cyclostationary random process with power sprectral density $S_X(f)$. Let $h(t)$ be the impulse response of a stable linear time-invariant filter. Assume that $E[X(t)] < \infty$, $E[X^2(t)] < \infty$, then the ouput $Y(t)$ of the linear filter $h(t)$ when $X(t)$ is applied at its input is also a WSS or WS cyclo S with power spectral density

$$S_Y(f) = |H(f)|^2 S_X(f)$$

Furthermore

$$E[X(t+\tau)Y(t)] = R_X(\tau) * h(-\tau)$$

Exercise: Prove these results for WSS random processes by using autocorrelation functions.

Exercise: Find the power spectral density and autocorrelation of the Hilbert transform of a random process.

m) Bandpass random processes

A random process $X(t)$ whose power spectrum exists is a bandpass process if

$$S_X(f) \neq 0 \quad \text{only if } |f \pm f_c| \leq \frac{B}{2} \text{ where } B < 2f_c$$

or equivalently if $S_X(f) = 0$ if $|f - f_c| \geq \frac{B}{2}$.

Assume that $X(t)$ is a **WSS bandpass process with zero mean** then $X(t)$ admits Rice's representation:

$$X(t) = X_I(t)\cos(2\pi f_c t) - X_Q(t)\sin(2\pi f_c t) \tag{7}$$

where $X_I(t)$ is the in-phase component of $X(t)$ and $X_Q(t)$ is the quadrature component of $X(t)$. $X_I(t)$ and $X_Q(t)$ are jointly WSS and low pass. Taking Hilbert's transform of (7),

$$\hat{X}(t) = X_I(t)\sin(2\pi f_c t) + X_Q(t)\cos(2\pi f_c t) \tag{8}$$

Combining (7) with (8) yields

$$\begin{aligned} X_I(t) &= X(t)\cos(2\pi f_c t) + \hat{X}(t)\sin(2\pi f_c t) \\ X_Q(t) &= \hat{X}(t)\cos(2\pi f_c t) - X(t)\sin(2\pi f_c t) \end{aligned}$$

Hence

$$\begin{aligned} &R_{X_I}(\tau) \stackrel{\triangle}{=} E\left[X_I(t+\tau)X_I(t)\right] \\ &= E\left[\big(X(t+\tau)\cos(2\pi f_c(t+\tau)) + \hat{X}(t+\tau)\sin(2\pi f_c(t+\tau))\big)\big(X(t)\cos(2\pi f_c t) + \hat{X}(t)\sin(2\pi f_c t)\big)\right] \\ &= R_X(\tau)\cos 2\pi f_c(t+\tau)\ \cos 2\pi f_c t + R_{X\hat{X}}(\tau)\cos 2\pi f_c(t+\tau)\ \sin 2\pi f_c t \\ &\quad + R_{\hat{X}X}(\tau)\sin 2\pi f_c(t+\tau)\ \cos 2\pi f_c t + R_{\hat{X}\hat{X}}(\tau)\sin 2\pi f_c(t+\tau)\ \sin 2\pi f_c t \end{aligned}$$

where $R_X(\tau) \stackrel{\triangle}{=} E\left[X(t+\tau)X(t)\right]$,

$$R_{X\hat{X}}(\tau) \stackrel{\triangle}{=} E\left[X(t+\tau)\hat{X}(t)\right] = R_X(\tau) * \frac{-1}{\pi\tau} = -\hat{R}_X(\tau)$$

$$R_{\hat{X}X}(\tau) \triangleq E\left[\hat{X}(t+\tau)X(t)\right] = R_X(\tau) * \frac{1}{\pi\tau} = \hat{R}_X(\tau)$$
$$R_{\hat{X}\hat{X}}(\tau) \triangleq E\left[\hat{X}(t+\tau)\hat{X}(t)\right] = R_X(\tau)$$

$$\begin{aligned} R_{X_I}(\tau) &= R_X(\tau)\Big[\cos 2\pi f_c(t+\tau)\ \cos 2\pi f_c t + \sin 2\pi f_c(t+\tau)\ \sin 2\pi f_c t\Big] \\ &\quad + \hat{R}_X(\tau)\Big[\sin 2\pi f_c(t+\tau)\ \cos 2\pi f_c t - \cos 2\pi f_c(t+\tau)\ \sin 2\pi f_c t\Big] \\ &= R_X(\tau)\cos 2\pi f_c\tau + \hat{R}_X(\tau)\sin 2\pi f_c\tau \end{aligned}$$

Similarly prove that

$$\begin{aligned} R_{X_Q}(\tau) &= R_X(\tau)\cos 2\pi f_c\tau + \hat{R}_X(\tau)\sin 2\pi f_c\tau = R_{X_I}(\tau) \\ R_{X_I X_Q}(\tau) &= R_X(\tau)\sin 2\pi f_c\tau - \hat{R}_X(\tau)\cos 2\pi f_c\tau \end{aligned}$$

Taking Fourier transforms of the autocorrelation functions

$$\begin{aligned} S_{X_I}(f) &\triangleq \mathcal{F}\left\{R_{X_I}(\tau)\right\} = S_{X_Q}(f) \triangleq \mathcal{F}\left\{R_{X_Q}(\tau)\right\} \\ &= \frac{1}{2}\Big\{S_X(f-f_c) + S_X(f+f_c)\Big\} + \frac{1}{2j}\Big\{\hat{S}_X(f-f_c) - \hat{S}_X(f+f_c)\Big\} \\ &= \frac{1}{2}\left[\big(1-\ \mathrm{sgn}(f-f_c)\big)S_X(f-f_c) + \big(1+\ \mathrm{sgn}(f+f_c)\big)S_X(f+f_c)\right] \\ &= \begin{cases} S_X(f-f_c) + S_X(f+f_c) & -\frac{B}{2} \le f \le \frac{B}{2} \\ 0 & \text{else} \end{cases} \\ S_{X_I X_Q}(f) &= \frac{1}{2j}\left\{\big(1-\ \mathrm{sgn}(f-f_c)\big)S_X(f-f_c) - \big(1+\ \mathrm{sgn}(f+f_c)\big)S_X(f+f_c)\right\} \\ &= \begin{cases} j\left\{S_X(f+f_c) - S_X(f-f_c)\right\} & -\frac{B}{2} \le f \le \frac{B}{2} \\ 0 & \text{else} \end{cases} \end{aligned}$$

If $S_X(f)$ is locally symmetric around $\pm f_c$ (i.e. $S_X(f+f_c) = S_X(f-f_c), |f| \le \frac{B}{2}$), then

$R_{X_I X_Q}(\tau) = 0$. Furthermore if $X(t)$ is Gaussian and locally symmetric around $\pm f_c$ then $X_I(t)$ and $X_Q(t)$ are statistically independent.

Extraction of $X_I(t)$ and $X_Q(t)$ from $X(t)$:

Draw the corresponding block diagram.

Special case of bandpass white noise: