A First Course in Digital Communications

Ha H. Nguyen and E. Shwedyk

CAMBRIDGE UNIVERSITY PRESS

February 2009

Introduction

- The main objective of a communication system is the transfer of information over a channel.
- Message signal is best modeled by a random signal: Any signal that conveys information must have some uncertainty in it, otherwise its transmission is of no interest.
- Two types of imperfections in a communication channel:
  - Deterministic imperfection, such as linear and nonlinear distortions, inter-symbol interference, etc.
  - Nondeterministic imperfection, such as addition of noise, interference, multipath fading, etc.
- We are concerned with the methods used to describe and characterize a random signal, generally referred to as a *random process* (also commonly called *stochastic process*).
- In essence, a random process is a *random variable* evolving in time.

Sample Space and Probability

- *Random experiment*: its outcome, for some reason, cannot be predicted with certainty.
- Examples: throwing a die, flipping a coin and drawing a card from a deck.
- *Sample space*: the set of all possible outcomes, denoted by Ω. Outcomes are denoted by ω's and each ω lies in Ω, i.e., $\omega \in \Omega$.
- A sample space can be *discrete* or *continuous*.
- *Events* are subsets of the sample space for which measures of their occurrences, called probabilities, can be defined or determined.

Three Axioms of Probability

For a discrete sample space Ω, define a probability measure P on Ω as a set function that assigns nonnegative values to all events, denoted by E, in Ω such that the following conditions are satisfied

Axiom 1: $0 \leq P(E) \leq 1$ for all $E \in \Omega$ (on a % scale probability ranges from 0 to 100%. Despite popular sports lore, it is impossible to give more than 100%).

Axiom 2: $P(\Omega) = 1$ (when an experiment is conducted there has to be an outcome).

Axiom 3: For mutually exclusive events[1] E_1, E_2, E_3,... we have $P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$.

[1]The events E_1, E_2, E_3,... are mutually exclusive if $E_i \cap E_j = \oslash$ for all $i \neq j$, where $\oslash$ is the null set.

Important Properties of the Probability Measure

1. $P(E^c) = 1 - P(E)$, where E^c denotes the complement of E. This property implies that $P(E^c) + P(E) = 1$, i.e., something has to happen.
2. $P(\oslash) = 0$ (again, something has to happen).
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$. Note that if two events E_1 and E_2 are mutually exclusive then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$, otherwise the nonzero common probability $P(E_1 \cap E_2)$ needs to be subtracted off.
4. If $E_1 \subseteq E_2$ then $P(E_1) \leq P(E_2)$. This says that if event E_1 is contained in E_2 then occurrence of E_2 means E_1 has occurred but the converse is not true.

Conditional Probability

- We observe or are told that event E_1 has occurred but are actually interested in event E_2: Knowledge that of E_1 has occurred changes the probability of E_2 occurring.
- If it was $P(E_2)$ before, it now becomes $P(E_2|E_1)$, the probability of E_2 occurring given that event E_1 has occurred.
- This *conditional probability* is given by

$$P(E_2|E_1) = \begin{cases} \dfrac{P(E_2 \cap E_1)}{P(E_1)}, & \text{if } P(E_1) \neq 0 \\ 0, & \text{otherwise} \end{cases} .$$

- If $P(E_2|E_1) = P(E_2)$, or $P(E_2 \cap E_1) = P(E_1)P(E_2)$, then E_1 and E_2 are said to be *statistically independent*.
- *Bayes*' rule

$$P(E_2|E_1) = \frac{P(E_1|E_2)P(E_2)}{P(E_1)},$$

Total Probability Theorem

- The events $\{E_i\}_{i=1}^n$ partition the sample space Ω if:

$$\text{(i)} \quad \bigcup_{i=1}^{n} E_i = \Omega \tag{1a}$$

$$\text{(ii)} \quad E_i \cap E_j = \emptyset \quad \text{for all } 1 \leq i, j \leq n \text{ and } i \neq j \tag{1b}$$

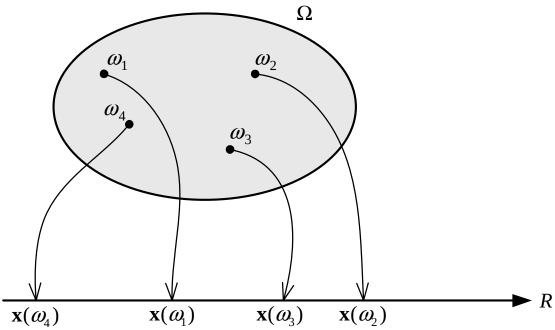
- If for an event A we have the conditional probabilities $\{P(A|E_i)\}_{i=1}^n$, $P(A)$ can be obtained as

$$P(A) = \sum_{i=1}^{n} P(E_i) P(A|E_i).$$

- *Bayes' rule*:

$$P(E_i|A) = \frac{P(A|E_i)P(E_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{\sum_{j=1}^{n} P(A|E_j)P(E_j)}.$$

Random Variables



- A random variable is a *mapping* from the sample space Ω to the set of real numbers.
- We shall denote random variables by boldface, i.e., $\mathbf{x}$, $\mathbf{y}$, etc., while individual or specific values of the mapping $\mathbf{x}$ are denoted by $\mathbf{x}(\omega)$.

Cumulative Distribution Function (cdf)

- cdf gives a complete description of the random variable. It is defined as:

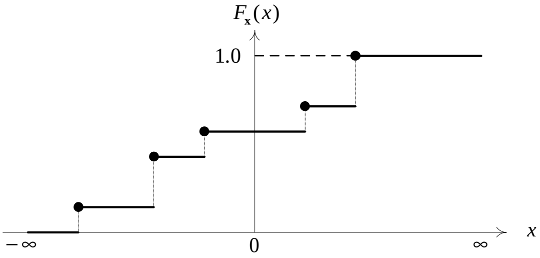
$$F_{\mathbf{x}}(x) = P(\omega \in \Omega : \mathbf{x}(\omega) \leq x) = P(\mathbf{x} \leq x).$$

- The cdf has the following properties:
  1. $0 \leq F_{\mathbf{x}}(x) \leq 1$ (this follows from Axiom 1 of the probability measure).
  2. $F_{\mathbf{x}}(x)$ is nondecreasing: $F_{\mathbf{x}}(x_1) \leq F_{\mathbf{x}}(x_2)$ if $x_1 \leq x_2$ (this is because event $\mathbf{x}(\omega) \leq x_1$ is contained in event $\mathbf{x}(\omega) \leq x_2$).
  3. $F_{\mathbf{x}}(-\infty) = 0$ and $F_{\mathbf{x}}(+\infty) = 1$ ($\mathbf{x}(\omega) \leq -\infty$ is the empty set, hence an impossible event, while $\mathbf{x}(\omega) \leq \infty$ is the whole sample space, i.e., a certain event).
  4. $P(a < \mathbf{x} \leq b) = F_{\mathbf{x}}(b) - F_{\mathbf{x}}(a)$.

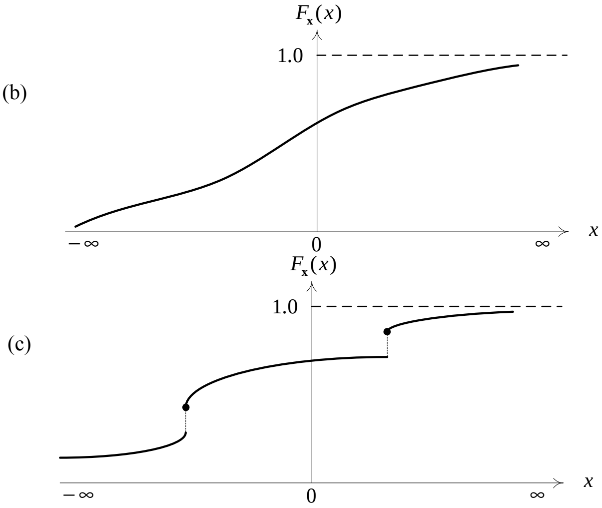
Typical Plots of cdf I

A random variable can be *discrete*, *continuous* or *mixed*.

(a)



Typical Plots of cdf II



Probability Density Function (pdf)

- The pdf is defined as the derivative of the cdf:

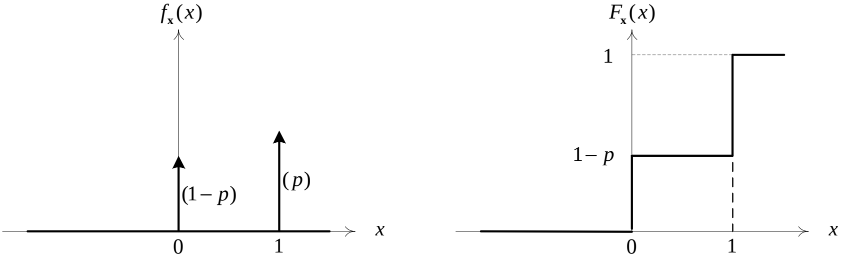
$$f_{\mathbf{x}}(x) = \frac{\mathrm{d}F_{\mathbf{x}}(x)}{\mathrm{d}x}.$$

- It follows that:

$$\begin{aligned} P(x_1 \leq \mathbf{x} \leq x_2) &= P(\mathbf{x} \leq x_2) - P(\mathbf{x} \leq x_1) \\ &= F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1) = \int_{x_1}^{x_2} f_{\mathbf{x}}(x)\mathrm{d}x. \end{aligned}$$

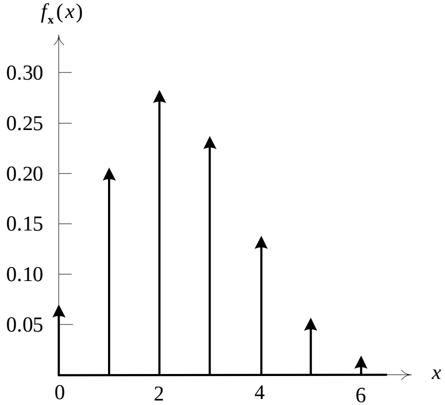
- Basic properties of pdf:
  1. $f_{\mathbf{x}}(x) \geq 0$.
  2. $\int_{-\infty}^{\infty} f_{\mathbf{x}}(x)\mathrm{d}x = 1$.
  3. In general, $P(\mathbf{x} \in \mathcal{A}) = \int_{\mathcal{A}} f_{\mathbf{x}}(x)\mathrm{d}x$.
- For discrete random variables, it is more common to define the *probability mass function* (pmf): $p_i = P(\mathbf{x} = x_i)$.
- Note that, for all i, one has $p_i \geq 0$ and $\sum_i p_i = 1$.

Bernoulli Random Variable



- A discrete random variable that takes two values 1 and 0 with probabilities p and $1-p$.
- Good model for a binary data source whose output is 1 or 0.
- Can also be used to model the channel errors.

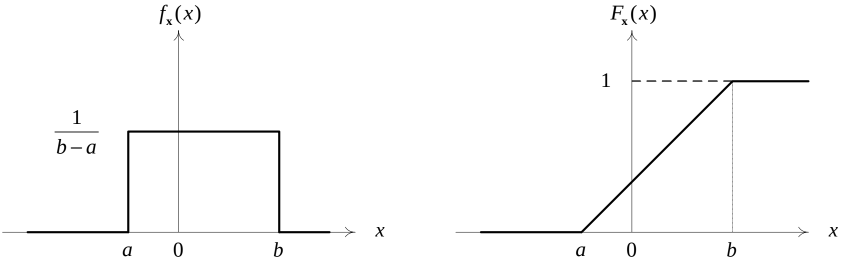
Binomial Random Variable



- A discrete random variable that gives the number of 1's in a sequence of n *independent* Bernoulli trials.

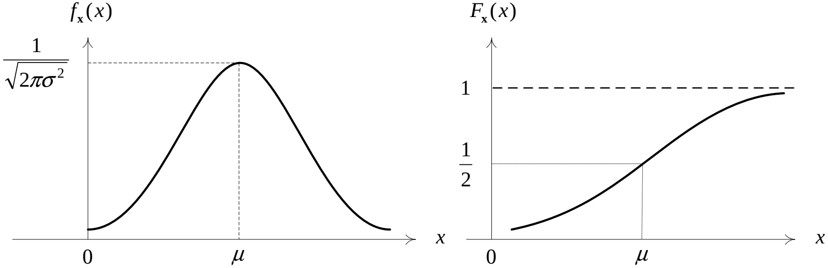
$$f_{\mathbf{x}}(x) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta(x-k), \text{ where } \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Uniform Random Variable



- A continuous random variable that takes values between a and b with equal probabilities over intervals of equal length.
- The phase of a received sinusoidal carrier is usually modeled as a uniform random variable between 0 and 2π. Quantization error is also typically modeled as uniform.

Gaussian (or Normal) Random Variable



- A continuous random variable whose pdf is:

$$f_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$

  μ and σ^2 are parameters. Usually denoted as $\mathcal{N}(\mu, \sigma^2)$.
- Most important and frequently encountered random variable in communications.

Functions of A Random Variable

- The function $\mathbf{y} = g(\mathbf{x})$ is itself a random variable.
- From the definition, the cdf of $\mathbf{y}$ can be written as

$$F_{\mathbf{y}}(y) = P(\omega \in \Omega : g(\mathbf{x}(\omega)) \leq y).$$

- Assume that for all y, the equation $g(x) = y$ has a countable number of solutions and at each solution point, $\mathrm{d}g(x)/\mathrm{d}x$ exists and is nonzero. Then the pdf of $\mathbf{y} = g(\mathbf{x})$ is:

$$f_{\mathbf{y}}(y) = \sum_i \frac{f_{\mathbf{x}}(x_i)}{\left| \left. \frac{\mathrm{d}g(x)}{\mathrm{d}x} \right|_{x=x_i} \right|},$$

where $\{x_i\}$ are the solutions of $g(x) = y$.

- *A linear function of a Gaussian random variable is itself a Gaussian random variable.*

Expectation of Random Variables I

Statistical averages, or *moments*, play an important role in the characterization of the random variable.

- The *expected value* (also called the mean value, first moment) of the random variable $\mathbf{x}$ is defined as

$$m_{\mathbf{x}} = E\{\mathbf{x}\} \equiv \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)\mathrm{d}x,$$

  where E denotes the *statistical expectation operator*.

- In general, the nth moment of $\mathbf{x}$ is defined as

$$E\{\mathbf{x}^n\} \equiv \int_{-\infty}^{\infty} x^n f_{\mathbf{x}}(x)\mathrm{d}x.$$

- For $n = 2$, $E\{\mathbf{x}^2\}$ is known as the mean-squared value of the random variable.

Expectation of Random Variables II

- The nth *central moment* of the random variable $\mathbf{x}$ is:

$$E\{\mathbf{y}\} = E\{(\mathbf{x} - m_{\mathbf{x}})^n\} = \int_{-\infty}^{\infty} (x - m_{\mathbf{x}})^n f_{\mathbf{x}}(x) \mathrm{d}x.$$

- When $n = 2$ the central moment is called the *variance*, commonly denoted as $\sigma_{\mathbf{x}}^2$:

$$\sigma_{\mathbf{x}}^2 = \mathrm{var}(\mathbf{x}) = E\{(\mathbf{x} - m_{\mathbf{x}})^2\} = \int_{-\infty}^{\infty} (x - m_{\mathbf{x}})^2 f_{\mathbf{x}}(x) \mathrm{d}x.$$

- The variance provides a measure of the variable's "randomness".
- The mean and variance of a random variable give a *partial description* of its pdf.

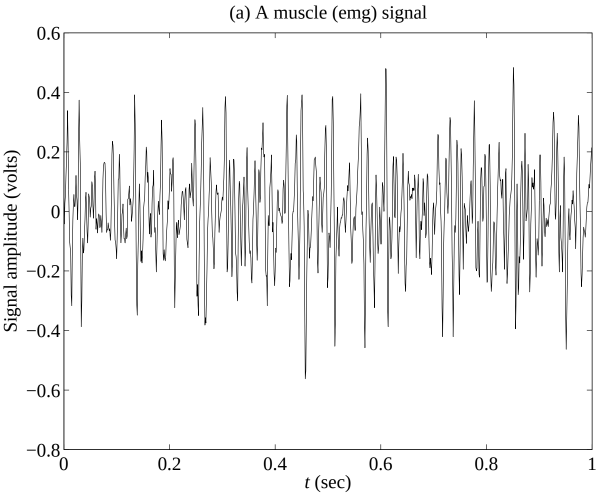
Expectation of Random Variables III

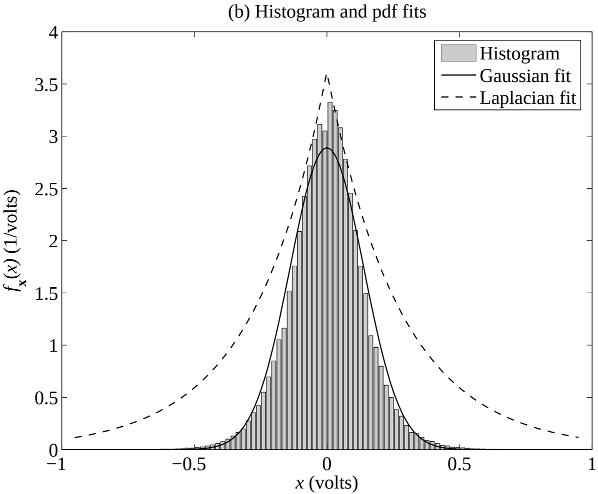
- Relationship between the variance, the first and second moments:

$$\sigma_{\mathbf{x}}^2 = E\{\mathbf{x}^2\} - [E\{\mathbf{x}\}]^2 = E\{\mathbf{x}^2\} - m_{\mathbf{x}}^2.$$

- An electrical engineering interpretation: *The AC power equals total power minus DC power.*
- The square-root of the variance is known as the *standard deviation*, and can be interpreted as the root-mean-squared (RMS) value of the AC component.

The Gaussian Random Variable

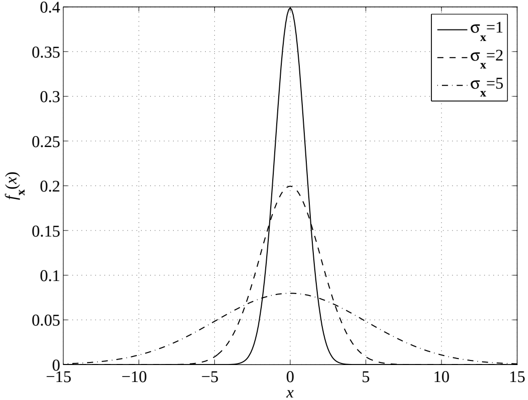




$$f_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi\sigma_{\mathbf{x}}^2}} \mathrm{e}^{-\frac{(x-m_{\mathbf{x}})^2}{2\sigma_{\mathbf{x}}^2}} \quad \text{(Gaussian)}$$

$$f_{\mathbf{x}}(x) = \frac{a}{2} \mathrm{e}^{-a|x|} \quad \text{(Laplacian)}$$

Gaussian Distribution (Univariate)



Range ($\pm k\sigma_{\mathbf{x}}$)	$k=1$	$k=2$	$k=3$	$k=4$
$P(m_{\mathbf{x}} - k\sigma_{\mathbf{x}} < \mathbf{x} \leq m_{\mathbf{x}} - k\sigma_{\mathbf{x}})$	0.683	0.955	0.997	0.999
Error probability	10^{-3}	10^{-4}	10^{-6}	10^{-8}
Distance from the mean	3.09	3.72	4.75	5.61

Multiple Random Variables I

- Often encountered when dealing with combined experiments or repeated trials of a single experiment.
- Multiple random variables are basically multidimensional functions defined on a sample space of a combined experiment.
- Let $\mathbf{x}$ and $\mathbf{y}$ be the two random variables defined on the same sample space Ω. The *joint cumulative distribution function* is defined as

$$F_{\mathbf{x},\mathbf{y}}(x,y) = P(\mathbf{x} \leq x, \mathbf{y} \leq y).$$

- Similarly, the *joint probability density function* is:

$$f_{\mathbf{x},\mathbf{y}}(x,y) = \frac{\partial^2 F_{\mathbf{x},\mathbf{y}}(x,y)}{\partial x \partial y}.$$

Multiple Random Variables II

- When the joint pdf is integrated over one of the variables, one obtains the pdf of other variable, called the *marginal pdf*:

$$\int_{-\infty}^{\infty} f_{\mathbf{x},\mathbf{y}}(x,y)\mathrm{d}x = f_{\mathbf{y}}(y),$$
$$\int_{-\infty}^{\infty} f_{\mathbf{x},\mathbf{y}}(x,y)\mathrm{d}y = f_{\mathbf{x}}(x).$$

- Note that:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{x},\mathbf{y}}(x,y)\mathrm{d}x\mathrm{d}y = F(\infty,\infty) = 1$$
$$F_{\mathbf{x},\mathbf{y}}(-\infty,-\infty) = F_{\mathbf{x},\mathbf{y}}(-\infty,y) = F_{\mathbf{x},\mathbf{y}}(x,-\infty) = 0.$$

Multiple Random Variables III

- The conditional pdf of the random variable $\mathbf{y}$, given that the value of the random variable $\mathbf{x}$ is equal to x, is defined as

$$f_{\mathbf{y}}(y|x) = \begin{cases} \dfrac{f_{\mathbf{x},\mathbf{y}}(x,y)}{f_{\mathbf{x}}(x)}, & f_{\mathbf{x}}(x) \neq 0 \\ 0, & \text{otherwise} \end{cases}.$$

- Two random variables $\mathbf{x}$ and $\mathbf{y}$ are statistically independent if and only if

$$f_{\mathbf{y}}(y|x) = f_{\mathbf{y}}(y) \quad \text{or equivalently} \quad f_{\mathbf{x},\mathbf{y}}(x,y) = f_{\mathbf{x}}(x) f_{\mathbf{y}}(y).$$

- The *joint moment* is defined as

$$E\{\mathbf{x}^j \mathbf{y}^k\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^j y^k f_{\mathbf{x},\mathbf{y}}(x,y)\mathrm{d}x\mathrm{d}y.$$

Multiple Random Variables IV

- The *joint central moment* is

$$E\{(\mathbf{x}-m_{\mathbf{x}})^j(\mathbf{y}-m_{\mathbf{y}})^k\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-m_{\mathbf{x}})^j(y-m_{\mathbf{y}})^k f_{\mathbf{x},\mathbf{y}}(x,y)\mathrm{d}x\mathrm{d}y$$

  where $m_{\mathbf{x}} = E\{\mathbf{x}\}$ and $m_{\mathbf{y}} = E\{\mathbf{y}\}$.

- The most important moments are

$$E\{\mathbf{xy}\} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{\mathbf{x},\mathbf{y}}(x,y)\mathrm{d}x\mathrm{d}y \quad \text{(correlation)}$$

$$\begin{aligned} \operatorname{cov}\{\mathbf{x},\mathbf{y}\} &\equiv E\{(\mathbf{x}-m_{\mathbf{x}})(\mathbf{y}-m_{\mathbf{y}})\} \\ &= E\{\mathbf{xy}\} - m_{\mathbf{x}}m_{\mathbf{y}} \quad \text{(covariance)}. \end{aligned}$$

Multiple Random Variables V

- Let $\sigma_{\mathbf{x}}^2$ and $\sigma_{\mathbf{y}}^2$ be the variance of $\mathbf{x}$ and $\mathbf{y}$. The covariance normalized w.r.t. $\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}$ is called the *correlation coefficient*:

$$\rho_{\mathbf{x},\mathbf{y}} = \frac{\text{cov}\{\mathbf{x}, \mathbf{y}\}}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}}.$$

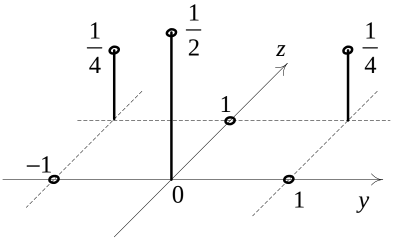
- $\rho_{\mathbf{x},\mathbf{y}}$ indicates the degree of *linear dependence* between two random variables.
- It can be shown that $|\rho_{\mathbf{x},\mathbf{y}}| \leq 1$.
- $\rho_{\mathbf{x},\mathbf{y}} = \pm 1$ implies an increasing/decreasing linear relationship.
- If $\rho_{\mathbf{x},\mathbf{y}} = 0$, $\mathbf{x}$ and $\mathbf{y}$ are said to be *uncorrelated*.
- It is easy to verify that if $\mathbf{x}$ and $\mathbf{y}$ are independent, then $\rho_{\mathbf{x},\mathbf{y}} = 0$: *Independence implies lack of correlation*.
- However, lack of correlation (no linear relationship) does not in general imply statistical independence.

Examples of Uncorrelated Dependent Random Variables

- Example 1: Let $\mathbf{x}$ be a discrete random variable that takes on $\{-1, 0, 1\}$ with probabilities $\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\}$, respectively. The random variables $\mathbf{y} = \mathbf{x}^3$ and $\mathbf{z} = \mathbf{x}^2$ are uncorrelated but dependent.
- Example 2: Let $\mathbf{x}$ be an uniformly random variable over $[-1, 1]$. Then the random variables $\mathbf{y} = \mathbf{x}$ and $\mathbf{z} = \mathbf{x}^2$ are uncorrelated but dependent.
- Example 3: Let $\mathbf{x}$ be a Gaussian random variable with zero mean and unit variance (*standard normal distribution*). The random variables $\mathbf{y} = \mathbf{x}$ and $\mathbf{z} = |\mathbf{x}|$ are uncorrelated but dependent.
- Example 4: Let $\mathbf{u}$ and $\mathbf{v}$ be two random variables (discrete or continuous) with the same probability density function. Then $\mathbf{x} = \mathbf{u} - \mathbf{v}$ and $\mathbf{y} = \mathbf{u} + \mathbf{v}$ are uncorrelated dependent random variables.

Example 1

$\mathbf{x} \in \{-1, 0, 1\}$ with probabilities $\{1/4, 1/2, 1/4\}$
$\Rightarrow \mathbf{y} = \mathbf{x}^3 \in \{-1, 0, 1\}$ with probabilities $\{1/4, 1/2, 1/4\}$
$\Rightarrow \mathbf{z} = \mathbf{x}^2 \in \{0, 1\}$ with probabilities $\{1/2, 1/2\}$
$m_{\mathbf{y}} = (-1)\frac{1}{4} + (0)\frac{1}{2} + (1)\frac{1}{4} = 0$; $m_{\mathbf{z}} = (0)\frac{1}{2} + (1)\frac{1}{2} = \frac{1}{2}$.
The joint pmf (similar to pdf) of $\mathbf{y}$ and $\mathbf{z}$:



$$P(\mathbf{y} = -1, \mathbf{z} = 0) = 0$$
$$P(\mathbf{y} = -1, \mathbf{z} = 1) = P(\mathbf{x} = -1) = 1/4$$
$$P(\mathbf{y} = 0, \mathbf{z} = 0) = P(\mathbf{x} = 0) = 1/2$$
$$P(\mathbf{y} = 0, \mathbf{z} = 1) = 0$$
$$P(\mathbf{y} = 1, \mathbf{z} = 0) = 0$$
$$P(\mathbf{y} = 1, \mathbf{z} = 1) = P(\mathbf{x} = 1) = 1/4$$

Therefore, $E\{\mathbf{yz}\} = (-1)(1)\frac{1}{4} + (0)(0)\frac{1}{2} + (1)(1)\frac{1}{4} = 0$
$\Rightarrow \operatorname{cov}\{\mathbf{y}, \mathbf{z}\} = E\{\mathbf{yz}\} - m_{\mathbf{y}} m_{\mathbf{z}} = 0 - (0)1/2 = 0!$

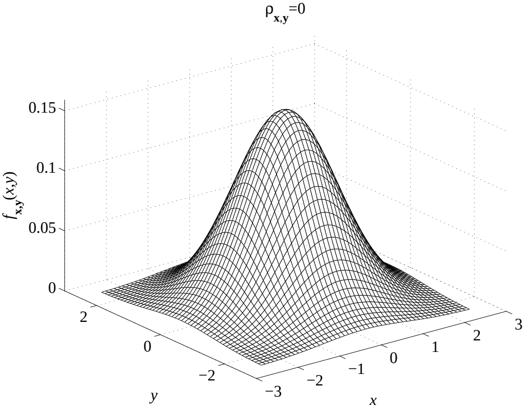
Jointly Gaussian Distribution (Bivariate)

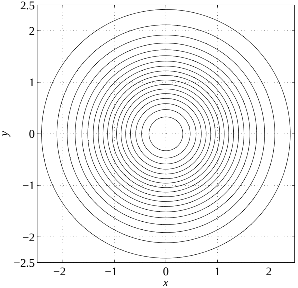
$$f_{\mathbf{x},\mathbf{y}}(x,y) = \frac{1}{2\pi\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}\sqrt{1-\rho_{\mathbf{x},\mathbf{y}}^2}} \exp\left\{ -\frac{1}{2(1-\rho_{\mathbf{x},\mathbf{y}}^2)} \times \left[\frac{(x-m_{\mathbf{x}})^2}{\sigma_{\mathbf{x}}^2} - \frac{2\rho_{\mathbf{x},\mathbf{y}}(x-m_{\mathbf{x}})(y-m_{\mathbf{y}})}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} + \frac{(y-m_{\mathbf{y}})^2}{\sigma_{\mathbf{y}}^2} \right] \right\},$$

where $m_{\mathbf{x}}$, $m_{\mathbf{y}}$, $\sigma_{\mathbf{x}}$, $\sigma_{\mathbf{y}}$ are the means and variances.

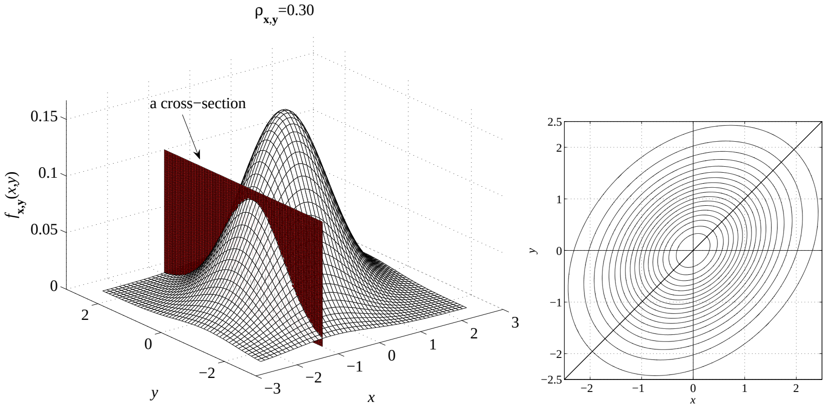
- $\rho_{\mathbf{x},\mathbf{y}}$ is indeed the correlation coefficient.
- Marginal density is Gaussian: $f_{\mathbf{x}}(x) \sim \mathcal{N}(m_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)$ and $f_{\mathbf{y}}(y) \sim \mathcal{N}(m_{\mathbf{y}}, \sigma_{\mathbf{y}}^2)$.
- When $\rho_{\mathbf{x},\mathbf{y}} = 0 \rightarrow f_{\mathbf{x},\mathbf{y}}(x,y) = f_{\mathbf{x}}(x) f_{\mathbf{y}}(y) \rightarrow$ random variables $\mathbf{x}$ and $\mathbf{y}$ are statistically independent.
- *Uncorrelatedness means that joint Gaussian random variables are statistically independent*. The converse is not true.
- Weighted sum of two jointly Gaussian random variables is also Gaussian.

Joint pdf and Contours for $\sigma_{\mathbf{x}} = \sigma_{\mathbf{y}} = 1$ and $\rho_{\mathbf{x},\mathbf{y}} = 0$

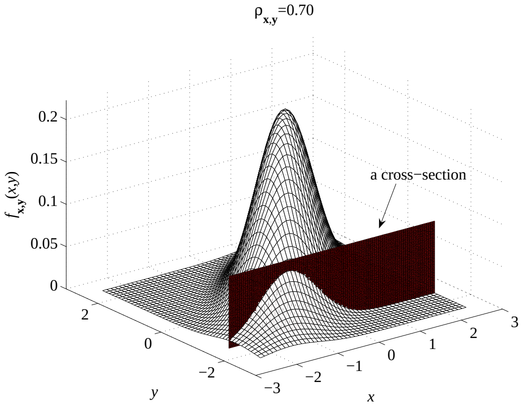


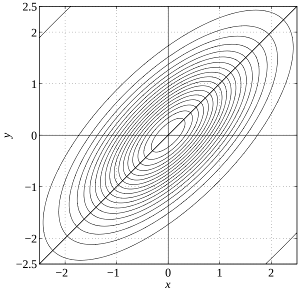


Joint pdf and Contours for $\sigma_{\mathbf{x}} = \sigma_{\mathbf{y}} = 1$ and $\rho_{\mathbf{x},\mathbf{y}} = 0.3$

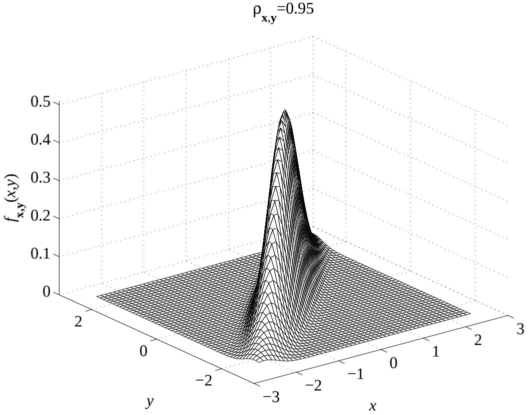


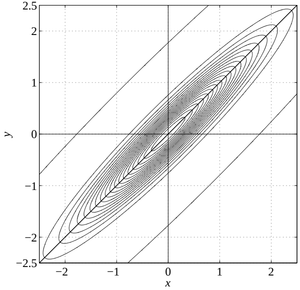
Joint pdf and Contours for $\sigma_{\mathbf{x}} = \sigma_{\mathbf{y}} = 1$ and $\rho_{\mathbf{x},\mathbf{y}} = 0.7$





Joint pdf and Contours for $\sigma_{\mathbf{x}} = \sigma_{\mathbf{y}} = 1$ and $\rho_{\mathbf{x},\mathbf{y}} = 0.95$





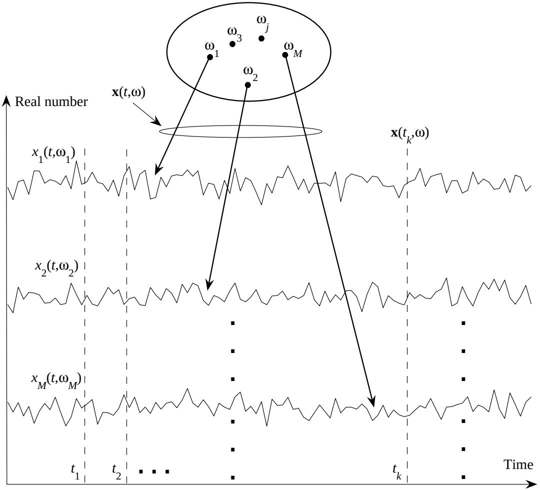
Multivariate Gaussian pdf

- Define $\overrightarrow{\mathbf{x}} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$, a vector of the means $\overrightarrow{m} = [m_1, m_2, \ldots, m_n]$, and the $n \times n$ covariance matrix C with $C_{i,j} = \mathrm{cov}(\mathbf{x}_i, \mathbf{x}_j) = E\{(\mathbf{x}_i - m_i)(\mathbf{x}_j - m_j)\}$.
- The random variables $\{\mathbf{x}_i\}_{i=1}^n$ are jointly Gaussian if:

$$f_{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n}(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det(C)}} \times \exp\left\{-\frac{1}{2}(\overrightarrow{x} - \overrightarrow{m})C^{-1}(\overrightarrow{x} - \overrightarrow{m})^\top\right\}.$$

- If C is *diagonal* (i.e., the random variables $\{\mathbf{x}_i\}_{i=1}^n$ are all uncorrelated), the joint pdf is a product of the marginal pdfs: Uncorrelatedness implies statistical independent for multiple Gaussian random variables.

Random Processes I



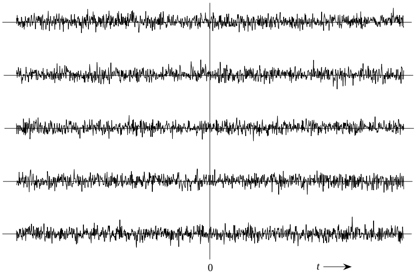
A mapping from a sample space to *a set of time functions*.

Random Processes II

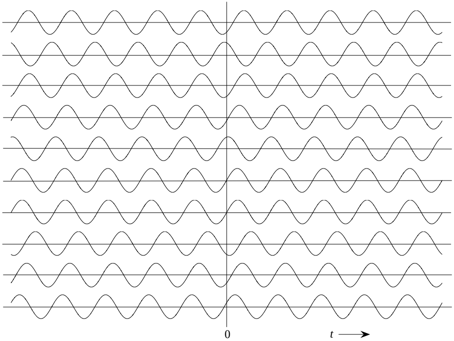
- *Ensemble*: The set of possible time functions that one sees.
- Denote this set by $\mathbf{x}(t)$, where the time functions $x_1(t, \omega_1)$, $x_2(t, \omega_2)$, $x_3(t, \omega_3)$, ... are specific members of the ensemble.
- At any time instant, $t = t_k$, we have random variable $\mathbf{x}(t_k)$.
- At any two time instants, say t_1 and t_2, we have two different random variables $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$.
- Any relationship between them is described by the joint pdf $f_{\mathbf{x}(t_1),\mathbf{x}(t_2)}(x_1, x_2; t_1, t_2)$.
- A complete description of the random process is determined by the joint pdf $f_{\mathbf{x}(t_1),\mathbf{x}(t_2),\ldots,\mathbf{x}(t_N)}(x_1, x_2, \ldots, x_N; t_1, t_2, \ldots, t_N)$.
- The most important joint pdfs are the first-order pdf $f_{\mathbf{x}(t)}(x; t)$ and the second-order pdf $f_{\mathbf{x}(t_1)\mathbf{x}(t_2)}(x_1, x_2; t_1, t_2)$.

Examples of Random Processes I

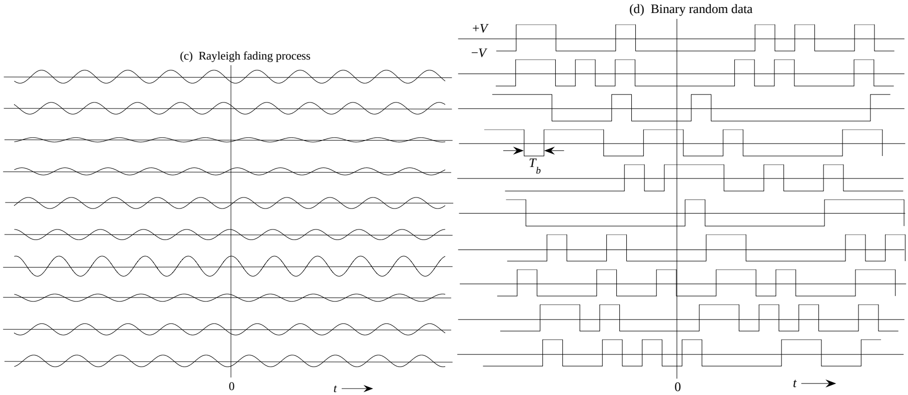
(a) Thermal noise



(b) Uniform phase



Examples of Random Processes II



Classification of Random Processes

- Based on whether its statistics change with time: the process is *non-stationary* or *stationary*.
- Different levels of stationarity:
  - Strictly stationary: the joint pdf of any order is independent of a shift in time.
  - Nth-order stationarity: the joint pdf does not depend on the time shift, but depends on time spacings:

$$f_{\mathbf{x}(t_1),\mathbf{x}(t_2),\dots\mathbf{x}(t_N)}(x_1, x_2, \dots, x_N; t_1, t_2, \dots, t_N) =$$
$$f_{\mathbf{x}(t_1+t),\mathbf{x}(t_2+t),\dots\mathbf{x}(t_N+t)}(x_1, x_2, \dots, x_N; t_1 + t, t_2 + t, \dots, t_N + t).$$

- The first- and second-order stationarity:

$$f_{\mathbf{x}(t_1)}(x, t_1) = f_{\mathbf{x}(t_1+t)}(x; t_1 + t) = f_{\mathbf{x}(t)}(x)$$

$$\begin{aligned} f_{\mathbf{x}(t_1),\mathbf{x}(t_2)}(x_1, x_2; t_1, t_2) &= f_{\mathbf{x}(t_1+t),\mathbf{x}(t_2+t)}(x_1, x_2; t_1 + t, t_2 + t) \\ &= f_{\mathbf{x}(t_1),\mathbf{x}(t_2)}(x_1, x_2; \tau), \quad \tau = t_2 - t_1. \end{aligned}$$

Statistical Averages or Joint Moments

- Consider N random variables $\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots \mathbf{x}(t_N)$. The joint moments of these random variables is

$$E\{\mathbf{x}^{k_1}(t_1), \mathbf{x}^{k_2}(t_2), \ldots \mathbf{x}^{k_N}(t_N)\} = \int_{x_1=-\infty}^{\infty} \cdots \int_{x_N=-\infty}^{\infty} x_1^{k_1} x_2^{k_2} \cdots x_N^{k_N} f_{\mathbf{x}(t_1),\mathbf{x}(t_2),\ldots\mathbf{x}(t_N)}(x_1, x_2, \ldots, x_N; t_1, t_2, \ldots, t_N) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_N,$$

  for all integers $k_j \geq 1$ and $N \geq 1$.
- Shall only consider the first- and second-order moments, i.e., $E\{\mathbf{x}(t)\}$, $E\{\mathbf{x}^2(t)\}$ and $E\{\mathbf{x}(t_1)\mathbf{x}(t_2)\}$. They are the *mean value*, *mean-squared value* and *(auto)correlation*.

Mean Value or the First Moment

- The mean value of the process at time t is

$$m_{\mathbf{x}}(t) = E\{\mathbf{x}(t)\} = \int_{-\infty}^{\infty} x f_{\mathbf{x}(t)}(x;t)\mathrm{d}x.$$

- The average is *across the ensemble* and if the pdf varies with time then the mean value is a (deterministic) function of time.
- If the process is stationary then the mean is independent of t or a constant:

$$m_{\mathbf{x}} = E\{\mathbf{x}(t)\} = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)\mathrm{d}x.$$

Mean-Squared Value or the Second Moment

- This is defined as

$$\text{MSV}_{\mathbf{x}}(t) = E\{\mathbf{x}^2(t)\} = \int_{-\infty}^{\infty} x^2 f_{\mathbf{x}(t)}(x;t)\mathrm{d}x \text{ (non-stationary)},$$

$$\text{MSV}_{\mathbf{x}} = E\{\mathbf{x}^2(t)\} = \int_{-\infty}^{\infty} x^2 f_{\mathbf{x}}(x)\mathrm{d}x \text{ (stationary)}.$$

- The second central moment (or the variance) is:

$$\begin{aligned}\sigma_{\mathbf{x}}^2(t) &= E\left\{[\mathbf{x}(t) - m_{\mathbf{x}}(t)]^2\right\} = \text{MSV}_{\mathbf{x}}(t) - m_{\mathbf{x}}^2(t) \text{ (non-stationary)},\\ \sigma_{\mathbf{x}}^2 &= E\left\{[\mathbf{x}(t) - m_{\mathbf{x}}]^2\right\} = \text{MSV}_{\mathbf{x}} - m_{\mathbf{x}}^2 \text{ (stationary)}.\end{aligned}$$

Correlation

- The autocorrelation function completely describes the power spectral density of the random process.
- Defined as the correlation between the two random variables $\mathbf{x}_1 = \mathbf{x}(t_1)$ and $\mathbf{x}_2 = \mathbf{x}(t_2)$:

$$\begin{aligned} R_{\mathbf{x}}(t_1, t_2) &= E\{\mathbf{x}(t_1)\mathbf{x}(t_2)\} \\ &= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} x_1 x_2 f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2; t_1, t_2) \mathrm{d}x_1 \mathrm{d}x_2. \end{aligned}$$

- For a stationary process:

$$\begin{aligned} R_{\mathbf{x}}(\tau) &= E\{\mathbf{x}(t)\mathbf{x}(t+\tau)\} \\ &= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} x_1 x_2 f_{\mathbf{x}_1,\mathbf{x}_2}(x_1, x_2; \tau) \mathrm{d}x_1 \mathrm{d}x_2. \end{aligned}$$

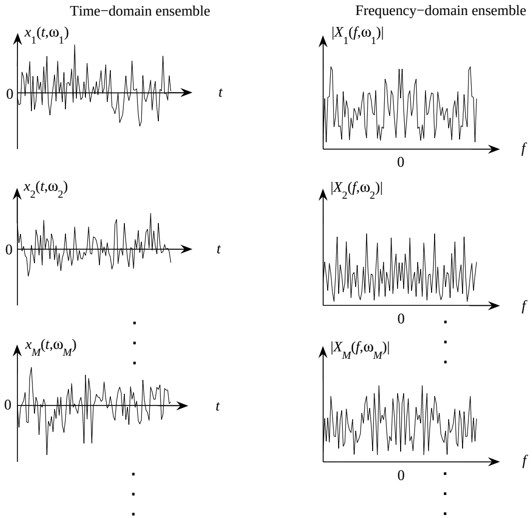
- *Wide-sense stationarity* (WSS) process: $E\{\mathbf{x}(t)\} = m_{\mathbf{x}}$ for any t, and $R_{\mathbf{x}}(t_1, t_2) = R_{\mathbf{x}}(\tau)$ for $\tau = t_2 - t_1$.

Properties of the Autocorrelation Function

1. $R_{\mathbf{x}}(\tau) = R_{\mathbf{x}}(-\tau)$. It is an even function of τ because the same set of product values is averaged across the ensemble, regardless of the direction of translation.
2. $|R_{\mathbf{x}}(\tau)| \leq R_{\mathbf{x}}(0)$. The maximum always occurs at $\tau = 0$, though there maybe other values of τ for which it is as big. Further $R_{\mathbf{x}}(0)$ is the mean-squared value of the random process.
3. If for some τ_0 we have $R_{\mathbf{x}}(\tau_0) = R_{\mathbf{x}}(0)$, then for all integers k, $R_{\mathbf{x}}(k\tau_0) = R_{\mathbf{x}}(0)$.
4. If $m_{\mathbf{x}} \neq 0$ then $R_{\mathbf{x}}(\tau)$ will have a constant component equal to $m_{\mathbf{x}}^2$.
5. Autocorrelation functions cannot have an arbitrary shape. The restriction on the shape arises from the fact that the Fourier transform of an autocorrelation function must be greater than or equal to zero, i.e., $\mathcal{F}\{R_{\mathbf{x}}(\tau)\} \geq 0$.

Power Spectral Density of a Random Process (I)

Taking the Fourier transform of the random process does not work.



Power Spectral Density of a Random Process (II)

- Need to determine how the average power of the process is distributed in frequency.
- Define a truncated process:

$$\mathbf{x}_T(t) = \begin{cases} \mathbf{x}(t), & -T \leq t \leq T \\ 0, & \text{otherwise} \end{cases}.$$

- Consider the Fourier transform of this truncated process:

$$\mathbf{X}_T(f) = \int_{-\infty}^{\infty} \mathbf{x}_T(t) \mathrm{e}^{-j2\pi f t} \mathrm{d}t.$$

- Average the energy over the total time, $2T$:

$$\mathbf{P} = \frac{1}{2T} \int_{-T}^{T} \mathbf{x}_T^2(t) \mathrm{d}t = \frac{1}{2T} \int_{-\infty}^{\infty} |\mathbf{X}_T(f)|^2 \,\mathrm{d}f \quad \text{(watts)}.$$

Power Spectral Density of a Random Process (III)

- Find the average value of $\mathbf{P}$:

$$E\{\mathbf{P}\} = E\left\{\frac{1}{2T}\int_{-T}^{T}\mathbf{x}_T^2(t)\mathrm{d}t\right\} = E\left\{\frac{1}{2T}\int_{-\infty}^{\infty}\left|\mathbf{X}_T(f)\right|^2\mathrm{d}f\right\}.$$

- Take the limit as $T \to \infty$:

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}E\left\{\mathbf{x}_T^2(t)\right\}\mathrm{d}t = \lim_{T\to\infty}\frac{1}{2T}\int_{-\infty}^{\infty}E\left\{\left|\mathbf{X}_T(f)\right|^2\right\}\mathrm{d}f,$$

- It follows that

$$\begin{aligned}\mathsf{MSV}_{\mathbf{x}} &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}E\left\{\mathbf{x}_T^2(t)\right\}\mathrm{d}t\\ &= \int_{-\infty}^{\infty}\lim_{T\to\infty}\frac{E\left\{\left|\mathbf{X}_T(f)\right|^2\right\}}{2T}\mathrm{d}f \quad \text{(watts)}.\end{aligned}$$

Power Spectral Density of a Random Process (IV)

- Finally,

$$S_{\mathbf{x}}(f) = \lim_{T\to\infty} \frac{E\left\{\left|\mathbf{X}_T(f)\right|^2\right\}}{2T} \quad \text{(watts/Hz)},$$

  is the *power spectral density* of the process.

- It can be shown that *the power spectral density and the autocorrelation function are a Fourier transform pair*:

$$R_{\mathbf{x}}(\tau) \longleftrightarrow S_{\mathbf{x}}(f) = \int_{\tau=-\infty}^{\infty} R_{\mathbf{x}}(\tau)\mathrm{e}^{-j2\pi f\tau}\mathrm{d}\tau.$$

Time Averaging and Ergodicity

- A process where any member of the ensemble exhibits the same statistical behavior as that of the whole ensemble.
- All time averages on a single ensemble member are equal to the corresponding ensemble average:

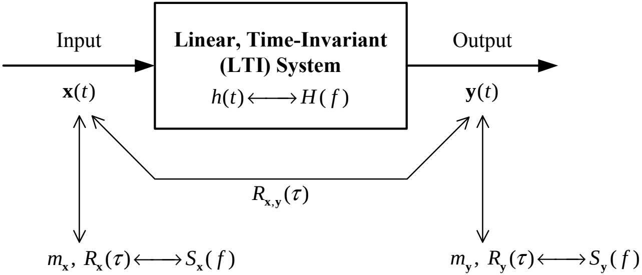
$$\begin{aligned} E\{\mathbf{x}^n(t))\} &= \int_{-\infty}^{\infty} x^n f_{\mathbf{x}}(x)\mathrm{d}x \\ &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} [\mathbf{x}_k(t,\omega_k)]^n \mathrm{d}t,\ \forall\ n,\ k. \end{aligned}$$

- For an ergodic process: To measure various statistical averages, it is sufficient to look at only one realization of the process and find the corresponding time average.
- *For a process to be ergodic it must be stationary*. The converse is not true.

Examples of Random Processes

- (Example 3.4) $\mathbf{x}(t) = A\cos(2\pi f_0 t + \mathbf{\Theta})$, where $\mathbf{\Theta}$ is a random variable uniformly distributed on $[0, 2\pi]$. This process is both stationary and ergodic.
- (Example 3.5) $\mathbf{x}(t) = \mathbf{x}$, where $\mathbf{x}$ is a random variable uniformly distributed on $[-A, A]$, where $A > 0$. This process is WSS, but not ergodic.
- (Example 3.6) $\mathbf{x}(t) = \mathbf{A}\cos(2\pi f_0 t + \mathbf{\Theta})$ where $\mathbf{A}$ is a zero-mean random variable with variance, $\sigma^2_{\mathbf{A}}$, and $\mathbf{\Theta}$ is uniform in $[0, 2\pi]$. Furthermore, $\mathbf{A}$ and $\mathbf{\Theta}$ are statistically independent. This process is not ergodic, but strictly stationary.

Random Processes and LTI Systems



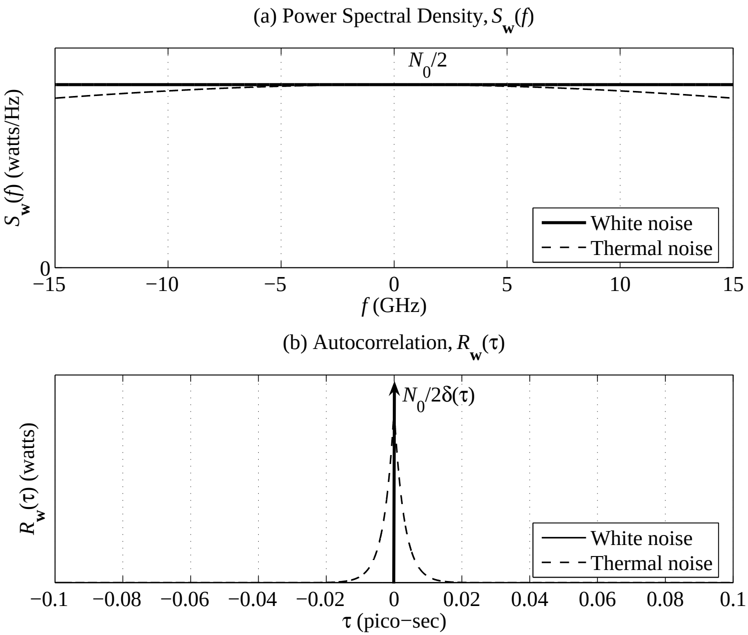
$$
\begin{aligned}
m_{\mathbf{y}} &= E\{\mathbf{y}(t)\} = E\left\{\int_{-\infty}^{\infty} h(\lambda)\mathbf{x}(t-\lambda)\mathrm{d}\lambda\right\} = m_{\mathbf{x}}H(0) \\
S_{\mathbf{y}}(f) &= |H(f)|^2\, S_{\mathbf{x}}(f) \\
R_{\mathbf{y}}(\tau) &= h(\tau) * h(-\tau) * R_{\mathbf{x}}(\tau).
\end{aligned}
$$

Thermal Noise in Communication Systems

- A natural noise source is thermal noise, whose amplitude statistics are well modeled to be Gaussian with zero mean.
- The autocorrelation and PSD are well modeled as:

$$\begin{aligned} R_{\mathbf{w}}(\tau) &= k\theta G \frac{\mathrm{e}^{-|\tau|/t_0}}{t_0} \quad \text{(watts)}, \\ S_{\mathbf{w}}(f) &= \frac{2k\theta G}{1+(2\pi f t_0)^2} \quad \text{(watts/Hz)}. \end{aligned}$$

where $k = 1.38 \times 10^{-23}$ joule/^{0}K is Boltzmann's constant, G is conductance of the resistor (mhos); θ is temperature in degrees Kelvin; and t_0 is the statistical average of time intervals between collisions of free electrons in the resistor (on the order of 10^{-12} sec).


(a) Power Spectral Density, $S_{\mathbf{w}}(f)$
$N_0/2$
$S_{\mathbf{w}}(f)$ (watts/Hz)
0
−15
−10
−5
0
5
10
15
f (GHz)
White noise
Thermal noise
(b) Autocorrelation, $R_{\mathbf{w}}(\tau)$
$N_0/2\delta(\tau)$
$R_{\mathbf{w}}(\tau)$ (watts)
−0.1
−0.08
−0.06
−0.04
−0.02
0
0.02
0.04
0.06
0.08
0.1
τ (pico−sec)
White noise
Thermal noise

- The noise PSD is approximately flat over the frequency range of 0 to 10 GHz $\Rightarrow$ let the spectrum be flat from 0 to ∞:

$$S_{\mathbf{w}}(f) = \frac{N_0}{2} \quad \text{(watts/Hz)},$$

where $N_0 = 4k\theta G$ is a constant.
- Noise that has a uniform spectrum over the entire frequency range is referred to as *white noise*
- The autocorrelation of white noise is

$$R_{\mathbf{w}}(\tau) = \frac{N_0}{2}\delta(\tau) \quad \text{(watts)}.$$

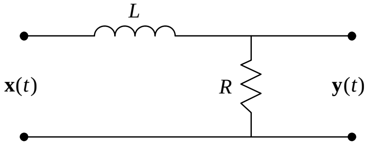
- Since $R_{\mathbf{w}}(\tau) = 0$ for $\tau \neq 0$, any two different samples of white noise, no matter how close in time they are taken, are *uncorrelated*.
- Since the noise samples of white noise are uncorrelated, if the noise is both white and Gaussian (for example, thermal noise) then the noise samples are also *independent*.

Example

Suppose that a (WSS) white noise process, $\mathbf{x}(t)$, of zero-mean and power spectral density $N_0/2$ is applied to the input of the filter.

(a) Find and sketch the power spectral density and autocorrelation function of the random process $\mathbf{y}(t)$ at the output of the filter.

(b) What are the mean and variance of the output process $\mathbf{y}(t)$?



$$H(f) = \frac{R}{R + j2\pi f L} = \frac{1}{1 + j2\pi f L/R}.$$

$$S_{\mathbf{y}}(f) = \frac{N_0}{2} \frac{1}{1 + \left(\frac{2\pi L}{R}\right)^2 f^2} \longleftrightarrow R_{\mathbf{y}}(\tau) = \frac{N_0 R}{4L} e^{-(R/L)|\tau|}.$$