Contemporary Communication Systems

Chapter 6

Probability and Random Processes

M.F. Mesiya

Probability Concepts

- The fundamental concept in probability theory is the concept of **random experiment**, which is any experiment whose outcome cannot be predicted with certainty
  - A simple example is coin tossing experiment. We know that heads and tails are possible outcomes, although the outcome (head or tail?) of a particular experiment (toss) is uncertain

Probability Concepts (cont.)

- Let us define the following concepts associated with a random experiment:
  - **Outcome** (ξ) – the result of a random experiment
  - **Sample space** (Ω) – the set of all possible outcomes of a random experiment
  - **Event** (A) – any collection of outcomes, in other words, a subset of Ω
  - The empty subset ϕ, is called the **null** or **impossible event**, and the whole set Ω is called the **whole** or **sure event**

Probability Axioms

- In the axiomatic approach, the probability is defined as a function that assigns a real number, denoted by $P(A)$, to every event A in the sample space Ω such that:

P1 $0 \le P(A) \le 1$

P2 The whole event Ω will occur each time we perform the random experiment

$$P(\Omega) = 1$$

P3 If the events are **mutually exclusive** (i.e., can not occur at the same time), the probability of their union is the sum of their probabilities

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$$

Probability Axioms (contd)

- (P1) – (P3) are called the **axioms** of probability
- By using the above axioms, we can derive following important properties of the probability function:

P4 The probability of the null event is zero.

$$P(\phi) = 0$$

P5 $P(\bar{A}) = 1 - P(A)$, $\quad \bar{A}$ = complement of A

- If the events $A_1, A_2, \ldots$ are not mutually exclusive, the probability of their union is upper-bounded by the sum of probabilities of the constituent events. That is,

$$P(A_1 \cup A_2 \cup \ldots\ldots) \le P(A_1) + P(A_2) + \ldots\ldots$$ **Union Bound**

Conditional Probability

- The probability $P(A)$ is *a priori* probability of the occurrence of an event A
  - Reflects our knowledge of A before the random experiment takes place
- The **conditional probability** $P(A|B)$ is the *a posteriori* probability of event A knowing that event B has already occurred
- It is defined as

$$P(A \mid B) = \frac{P(AB)}{P(B)}, \qquad \text{provided } P(B) > 0$$

- Conditioning by event B has the effect of restricting the universe of outcomes for the event A to the subset B of Ω
  - Definition satisfies all probability axioms

Independent Events

- A and B are said to be independent events if

  $P(AB) = P(A)P(B)$

- One should not confuse independent events with mutually exclusive or disjoint events
  - Mutually exclusive events have no outcome in common, i.e., $AB = \phi$ implying that $P(AB) = 0$
  - Independent events in most cases are not disjoint
- Substituting into the definition of conditional probability yields

  $$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

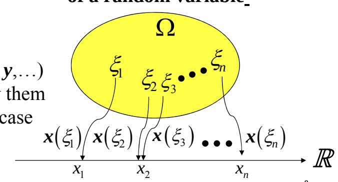
- $\Rightarrow$ that the occurrence of B does not provide any more information about the event A

Random Variable

- A **random variable** is defined as a rule that assigns a real number to each possible outcome $\xi \in \Omega$ of a random experiment
  - Thus, random variable is a function that maps every outcome $\xi \in \Omega$ to a real number x as illustrated in Figure

Conceptual model of a random variable

We will denote random variables in a bold font ($\mathbf{x}, \mathbf{y}, \ldots$) and the values assumed by them are displayed by the lowercase letters ($x, y, \ldots$).



Discrete Random Variables

- Random variables may be discrete, continuous or mixed depending upon the range of values they assume
- A **discrete** random variable $\mathbf{x}$ can take on a **countable** number of values $x_1, x_2, x_3, \ldots$ with probabilities

  $P\{\mathbf{x} = x_i\},\ i = 0, 1, 2, \ldots$

  - e.g., # of defective chips from a semiconductor wafer
- A **probability mass function (PMF)** $p_{\mathbf{x}}(x_i)$ completely characterizes a discrete random variable. It is defined as

  $p_{\mathbf{x}}(x_i) = P\{\mathbf{x} = x_i\}$

- Since $p_{\mathbf{x}}(x_i)$ is a probability, it satisfies following properties
  $0 \le p_{\mathbf{x}}(x_i) \le 1, \qquad \sum_i p_{\mathbf{x}}(x_i) = \sum_i P\{\mathbf{x}(\xi_i) = x_i \mid \xi_i \in \Omega\} = 1$

Continuous Random Variables

- A **continuous** random variable $\mathbf{x}$ takes values in a continuous set of numbers. The range of $\mathbf{x}$ may include the whole real line $\mathbb{R}$ or an interval thereof
- Continuous random variables model many real life phenomena that include file download time on Internet, voltage across a resistor, and phase of a carrier signal produced by a radio transmitter
- One characteristic that distinguishes a continuous random variable from the discrete one is that the probability of an individual outcome is zero. That is, $P\{\mathbf{x} = x\} = 0$, where x is any number in the range of $\mathbf{x}$
- Therefore, we can not use the PMF for a continuous random variable. Instead we shall use the cumulative distribution function which serves as an appropriate probability measure for any random variable

Cumulative Distribution Function (CDF)

- The **cumulative distribution function (CDF)**, $F_{\mathbf{x}}(x)$, of a random variable $\mathbf{x}$ is defined as

  $F_{\mathbf{x}}(x) = P\{\mathbf{x} \le x\}$

- For any real number x, the CDF measures the probability that the random variable $\mathbf{x}$ is no larger than x
  - (a) $0 \le F_{\mathbf{x}}(x) \le 1$
  - (b) $\lim_{x \to -\infty} F_{\mathbf{x}}(x) = 0$ and $\lim_{x \to \infty} F_{\mathbf{x}}(x) = 1$
  - (c) $P\{a < \mathbf{x} \le b\} = F_{\mathbf{x}}(b) - F_{\mathbf{x}}(a)$
  - (d) $F_{\mathbf{x}}(x)$ is nondecreasing

Probability Density Function (PDF)

- A **probability density function (PDF)**, $f_{\mathbf{x}}(x)$, of a continuous random variable $\mathbf{x}$ is derivative of its CDF. That is.

  $$f_{\mathbf{x}}(x) = \frac{dF_{\mathbf{x}}(x)}{dx}$$

- The CDF of a continuous random variable $\mathbf{x}$ is integral of its PDF

  $$F_{\mathbf{x}}(a) = \int_{-\infty}^{a} f_{\mathbf{x}}(x)dx$$

  The PDF indicates the probability that $\mathbf{x}$ is in the close vicinity of x

  - (a) $f_{\mathbf{x}}(x) \ge 0$
  - (b) $\int_{-\infty}^{\infty} f_{\mathbf{x}}(x)dx = 1$
  - (c) $\int_{a}^{b} f_{\mathbf{x}}(x)dx = P\{a < \mathbf{x} \le b\}$

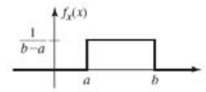
Some Common Continuous Random Variables

- Here we introduce three important continuous random variables:
  - Uniform
  - Gaussian
  - Exponential
  - Poisson
  - Rayleigh

Uniform Random Variable

- $\boldsymbol{x}$ is a uniform random variable if its PDF is given by

$$f_x(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b, \\ 0, & \text{otherwise.} \end{cases}$$



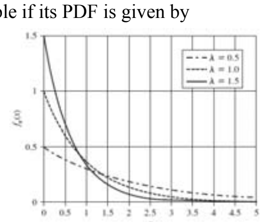
- The uniform random variable is a good model when each outcome of a random experiment is equally likely, and constrained to lie in the interval $[b, a]$, $b > a$.

Exponential Random Variable

- $\boldsymbol{x}$ is an exponential random variable if its PDF is given by

$$f_x(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0, \\ 0, & \text{otherwise.} \end{cases}$$

where $\lambda > 0$



- For $x \ge 0$,

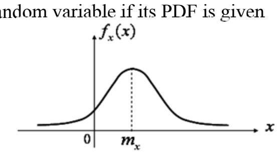
$$F_x(x) = P\{\boldsymbol{x} \le x\} = \int_0^x \lambda e^{-\lambda t} dt = \int_0^x \lambda e^{-\lambda x} dx = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$$

- The exponential random variable is frequently used to model lifetimes (e.g., duration of a phone call) or waiting times (e.g. until some event happens)

Gaussian or Normal Random Variable

- $\boldsymbol{x}$ is a normal or Gaussian random variable if its PDF is given by

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(x-m_x)^2/2\sigma_x^2}$$



- Characterized by **mean** m_x **and variance** σ_x^2
  - σ_x called the **standard deviation**
- A Gaussian random variable with mean m_x and variance σ_x^2 is denoted by $\mathcal{N}(m_x, \sigma_x^2)$
- It is most frequently used random variable in the analysis and modeling of communication systems. Thermal noise, which is ubiquitously present in communication systems, has a Gaussian PDF

Gaussian or Normal Random Variable (contd)

- The CDF $F_x(x)$ of the Gaussian random variable $\boldsymbol{x}$ is given by

$$F_x(x) = P\{\boldsymbol{x} \le x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(t-m_x)^2/2\sigma_x^2} dt$$

- There is no closed form solution for the integral on the right hand side. However, it can be written in terms of the *Q*-function as

$$F_x(x) = 1 - Q\left(\frac{x - m_x}{\sigma_x}\right) = Q\left(\frac{m_x - x}{\sigma_x}\right)$$

where

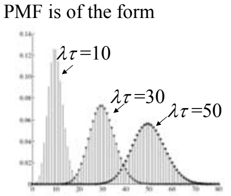
$$Q(a) = P\{\boldsymbol{x} > a\} = \frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-y^2/2} dy$$

Poisson Random Variable

- The Poisson random variable $\boldsymbol{x}$ models the number of events (k) occurring in any interval $(t_o, t_o + \tau)$ if the occurrence of these events, at an average rate λ, is independent of t_o and depends only on the length of interval τ
- It is common in the literature to refer to the occurrence of a Poisson event as an arrival
- $\boldsymbol{x}$ is a Poisson random variable if its PMF is of the form

$$\begin{aligned} p_x(k) &= P(\boldsymbol{x} = k) \\ &= P\{k \text{ arrivals in interval } \tau\} \\ &= e^{-\lambda\tau} \frac{(\lambda\tau)^k}{k!}, \quad k = 0, 1, 2, \cdots, \infty \end{aligned}$$

where λ = average arrival rate



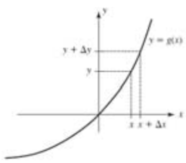
Functions of a Random Variable

- We are frequently interested in statistics of a random signal after passage through a system
- Let $\boldsymbol{x}$ be a random variable whose PDF is known and suppose that $g(.)$ denotes the transfer characteristic of a linear or nonlinear system
- We want to determine the PDF $f_y(y)$ of the new random variable $\boldsymbol{y}$ related to $\boldsymbol{x}$ by

$$\boldsymbol{y} = g(\boldsymbol{x})$$

- For a monotone increasing or decreasing function, the PDF of $\boldsymbol{y}$ is given by

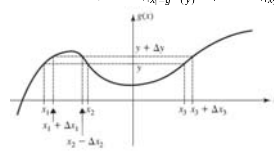
$$f_y(y) = \left.\frac{f_x(x)}{|g'(x)|}\right|_{x=g^{-1}(y)}$$



Functions of a Random Variable (contd)

- An arbitrary function $g(x)$ can be viewed as consisting of several piecewise monotonic segments over the range of $\boldsymbol{x}$
- Figure shows $g(x)$ with three piecewise monontonic segments. As a result the intervals $(x_1, x_1+\Delta x_1)$, $(x_2, x_2+\Delta x_2)$, and $(x_3, x_3+\Delta x_3)$ are mapped by into the same interval $(y, y+\Delta y)$
- The PDF of $\boldsymbol{y}$ is given by

$$f_y(y) = \left.\frac{f_x(x_1)}{|g'(x_1)|}\right|_{x_1=g^{-1}(y)} + \left.\frac{f_x(x_2)}{|g'(x_2)|}\right|_{x_2=g^{-1}(y)} + \left.\frac{f_x(x_3)}{|g'(x_3)|}\right|_{x_3=g^{-1}(y)}$$



where x_1, x_2 and x_3 are roots of the equation $g(x) = y$

Joint Cumulative Distribution Functions

- Consider a random experiment with sample space Ω. We are interested in a function $(\boldsymbol{x}(\xi), \boldsymbol{y}(\xi))$ that assigns a pair of real numbers to each outcome $\xi \in \Omega$ of the random experiment
- That is, we are dealing with a vector function that maps Ω into the real plane $\mathbb{R}^2$ or a subset thereof
- The **joint cumulative distribution function (CDF)** of two random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

$$F_{xy}(x,y) = P\{\boldsymbol{x} \le x, \boldsymbol{y} \le y\}$$

- Note that $F_{xy}(x,y)$ measures the probability of event

$$A = \{\xi \in \Omega : \boldsymbol{x}(\xi) \le x, \boldsymbol{y}(\xi) \le y\}$$

Properties of Joint CDFs

- For any pair of random variables $\boldsymbol{x}$ and $\boldsymbol{y}$,

(a) $0 \le F_{xy}(x,y) \le 1$

(b) $F_{xy}(\infty,\infty) = 1$

(c) $F_{xy}(x,-\infty) = F_{xy}(-\infty,y) = 0$

(d) $F_{xy}(x,y)$ is nondecreasing

Joint Probability Density Function

- The **joint probability density function,** $f_{xy}(x,y)$**,** of two random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

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

$$\Rightarrow F_{xy}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{xy}(u,v)\, dudv$$

(a) $f_{xy}(x,y) \ge 0$ for all (x,y)

(b) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{xy}(x,y)\, dxdy = F_{xy}(\infty,\infty) = 1$

(c) For a rectangle $\{a < \boldsymbol{x} \le b, c < \boldsymbol{y} \le d\}$ in x-y plane,

$$P\{a < \boldsymbol{x} \le b, c < \boldsymbol{y} \le d\} = \int_a^b\int_c^d f_{xy}(x,y)\, dxdy$$

Conditional Distributions (contd)

- The conditional PDF of random variable $\boldsymbol{x}$ given $\{\boldsymbol{y} = y\}$, denoted by $f_x(x|y)$, is defined as

$$f_x(x|y) = f_x(x|\boldsymbol{y} = y) = \frac{f_{xy}(x,y)}{f_y(y)}, \quad f_y(y) > 0$$

- Note that for each y with $f_y(y) > 0$, the conditional PDF $f_x(x|y)$ provides a new probabilistic description of the random variable $\boldsymbol{x}$
- Similarly, we can define

$$f_y(y|x) = f_y(y|\boldsymbol{x} = x) = \frac{f_{xy}(x,y)}{f_x(x)}, \quad f_x(x) > 0$$

Independent Random Variables

- Two random variables $\mathbf{x}$ and $\mathbf{y}$ are said to be **statistically independent** if

$$F_{xy}(x,y) = P\{\mathbf{x} \le x, \mathbf{y} \le y\}$$
$$= P\{\mathbf{x} \le x\} P\{\mathbf{y} \le y\} = F_x(x) F_y(y)$$

- Equivalently, for independent random variables

$$f_{xy}(x,y) = f_x(x) f_y(y)$$

- For independent random variables,

The PDF of $\mathbf{x}$ after knowledge of the event $\{\mathbf{y} = y\}$ same as its PDF before the knowledge

$$f_x(x|y) = \frac{f_{xy}(x,y)}{f_y(y)} = \frac{f_x(x) f_y(y)}{f_y(y)} = f_x(x)$$

$$f_y(y|x) = f_y(y)$$

Statistics of a Random Variable

- The *expected value* or *mean* of a continuous random variable $\mathbf{x}$ is defined as

$$m_x = \overline{\mathbf{x}} = E\{\mathbf{x}\} = \int_{-\infty}^{+\infty} x\, f_x(x) dx$$

- The expected value of a random variable represents its average value in a very large number of trials
- The mean of the function $\mathbf{y} = g(\mathbf{x})$ is

$$\overline{g(\mathbf{x})} = E\{g(\mathbf{x})\} = \int_{-\infty}^{+\infty} g(x) f_x(x) dx$$

- The variance $Var(\mathbf{x})$ of a random variable $\mathbf{x}$ is defined as

$$Var(\mathbf{x}) = \sigma_x^2 = E\{(\mathbf{x} - m_x)^2\} = \int_{-\infty}^{+\infty} (x - m_x)^2 f_x(x) dx \ge 0$$

  - Describes the spread of its PDF around the expected value

Statistics of Pair of Random Variables

- **Expected value** of $\mathbf{x} + \mathbf{y}$

$$E\{\mathbf{x} + \mathbf{y}\} = E\{\mathbf{x}\} + E\{\mathbf{y}\}$$

  - More generally, expectation is a **linear operator**

$$E\left\{\sum_i \alpha_i \mathbf{x}_i\right\} = \sum_i \alpha_i E\{\mathbf{x}_i\}$$

- **Variance** of $\mathbf{x} + \mathbf{y}$

$$Var(\mathbf{x} + \mathbf{y}) = Var(\mathbf{x}) + Var(\mathbf{y}) + 2E\{(\mathbf{x} - m_x)(\mathbf{y} - m_y)\}$$

- **Covariance** of $\mathbf{x}$ and $\mathbf{y}$

$$Cov(\mathbf{x}, \mathbf{y}) = E\{(\mathbf{x} - m_x)(\mathbf{y} - m_y)\}$$
$$\Rightarrow Var(\mathbf{x} + \mathbf{y}) = Var(\mathbf{x}) + Var(\mathbf{y}) + 2Cov(\mathbf{x}, \mathbf{y})$$

Correlation and Covariance

- The **correlation** of two random variables $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$R_{xy} = E\{\mathbf{xy}\}$$

- It is very simple exercise to prove that

$$Cov(\mathbf{x}, \mathbf{y}) = E\{\mathbf{xy}\} - E\{\mathbf{x}\}E\{\mathbf{y}\} = R_{xy} - m_x m_y$$

- $\mathbf{x}$ and $\mathbf{y}$ are called **uncorrelated** random variables if

$$Cov(\mathbf{x}, \mathbf{y}) = 0$$
$$\Rightarrow E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$$

- The **correlation coefficient** of two random variables $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$\rho_{xy} = \frac{Cov(\mathbf{x}, \mathbf{y})}{\sigma_x \sigma_y}$$

Pair of Gaussian Random Variables

- The joint PDF of two Gaussian random variables $\mathbf{x} \sim \mathcal{N}(m_x, \sigma_x^2)$ and $\mathbf{y} \sim \mathcal{N}(m_y, \sigma_y^2)$ is given by

$$f_{xy}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} e^{\frac{-1}{2(1-\rho_{xy}^2)}\left(\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2\rho_{xy}(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right)} \quad (*)$$

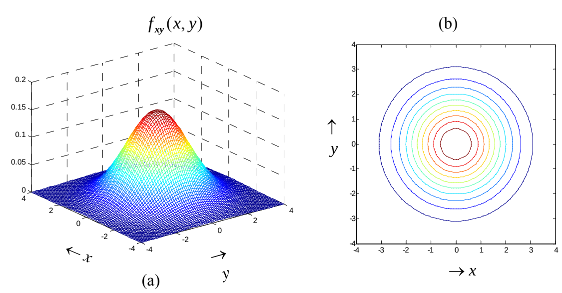
where $|\rho_{xy}| < 1$

- (*) is also called the **bivariate** Gaussian density
- For uncorrelated Gaussian random variables, $\rho_{xy} = 0$

$$f_{xy}(x,y)\Big|_{\rho_{xy}=0} = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(x-m_x)^2/2\sigma_x^2} \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-(y-m_y)^2/2\sigma_y^2}$$

Uncorrelated Gaussian rvs $\Rightarrow$ statistically independent

Bivariate Gaussian Density for $\rho_{xy} = 0$



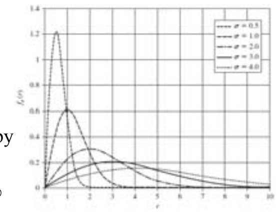
(a)

Rayleigh PDF

- Assume that the random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ are $\mathcal{N}(0,\sigma^2)$ and statistically independent. Let $\boldsymbol{r}$ and $\boldsymbol{\phi}$ be the polar coordinate representation of $\boldsymbol{x}$ and $\boldsymbol{y}$

$$\boldsymbol{r} = g(\boldsymbol{x},\boldsymbol{y}) = \sqrt{\boldsymbol{x}^2+\boldsymbol{y}^2}$$

$$\boldsymbol{\phi} = h(\boldsymbol{x},\boldsymbol{y}) = \tan^{-1}\frac{\boldsymbol{y}}{\boldsymbol{x}}$$



- The PDFs of $\boldsymbol{r}$ and $\boldsymbol{\phi}$ are given by

$$f_{\boldsymbol{r}}(r) = \frac{r}{\sigma^2}e^{-r^2/2\sigma^2}, \quad 0 \le r < \infty$$

- $f_{\boldsymbol{r}}(r)$ is referred to as the **Rayleigh's** PDF
- $\boldsymbol{\phi}$ is uniformly distributed over $[-\pi, \pi]$

Central Limit Theorem

- Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ be n independent, identically distributed random variables with finite mean m and variance σ^2
- Consider their scaled sum

$$\boldsymbol{z}_n = \frac{\sum_{i=1}^{n}(\boldsymbol{x}_i - m)}{\sigma\sqrt{n}} = \frac{\boldsymbol{s}_n - nm}{\sigma\sqrt{n}}$$

- Then

$$\lim_{n\to\infty} P\{\boldsymbol{z}_n \le z\} = \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}du \Rightarrow \boldsymbol{z}_n \sim \mathcal{N}(0,1)$$

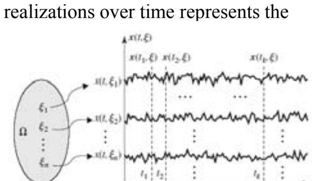
- That is, the CDF of $\boldsymbol{z}_n$ converges to a Gaussian CDF $\mathcal{N}(0,1)$ as n approaches ∞, independent of the distribution of random variables $\boldsymbol{x}_n$

Central Limit Theorem (contd)

- This result is known as the **central limit theorem**
- In a nutshell, the central limit theorem, states that the sum of almost any set of independent and randomly generated random variables rapidly converges to the Gaussian distribution
- This explains why the Gaussian distribution arises so commonly in practice to reflect the additive effect of multiple random occurrences

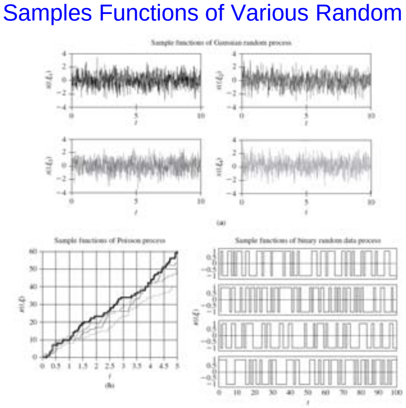
Random Process – Basic Concept

- Random variables assign one or more numbers to each outcome ξ of a random experiment
- In the case of random process $\boldsymbol{x}(t)$, every such outcome is assigned a waveform $x(t,\xi)$
- $x(t,\xi)$ is called the **sample function** and the **ensemble** of all such sample functions or realizations over time represents the random process $\boldsymbol{x}(t)$
- Note that various sample functions themselves are deterministic



- The randomness is associated with the occurrence of a particular outcome which in turn determines the sample function observed

Samples Functions of Various Random Processes



Characterization of a Random Process

- A random process is described by its ensemble of sample functions (waveforms) and the PDF over the ensemble
- Each time we perform the random experiment, a number $x(t_1,\xi)$ is observed at $t = t_1$ depending upon the particular outcome ξ and the corresponding realization $x(t,\xi)$
  - The amplitudes of sample functions at any fixed instant instant t_1, therefore, represent the random variable $\boldsymbol{x}(t_1,\xi)$
- We will use the notation $\boldsymbol{x}_1$ to represent the random variable at instant t_1. Thus, random variables $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ represent amplitudes of sample functions at $t = t_1, t_2, \ldots, t_n$
- A random process can, therefore, be viewed as a collection of infinite number of random variables
  - The joint PDF $f_{\boldsymbol{x}}(x_1, x_2, \ldots, x_n, t_1, t_2, \ldots, t_n)$ for all n and for any choice of $t_1, t_2, \ldots, t_n$ completely describes it

Expected Value and Correlation

- The expected value of the random process $\boldsymbol{x}(t)$ is defined as

$$m_x(t) = \overline{\boldsymbol{x}(t)} = E\{\boldsymbol{x}(t)\} = \int_{-\infty}^{+\infty} x\, f_{\boldsymbol{x}}(x,t)dx$$

- In general, $m_x(t)$ is a function of time
- Similarly, the mean-square value of the random process $\boldsymbol{x}(t)$ is given by

$$\overline{\boldsymbol{x}^2(t)} = E\{\boldsymbol{x}^2(t)\} = \int_{-\infty}^{+\infty} x^2\, f_{\boldsymbol{x}}(x,t)dx$$

- The **autocorrelation function** of the random process $\boldsymbol{x}(t)$ is defined as

$$R_x(t_1,t_2) = E\left\{\boldsymbol{x}(t_1)\boldsymbol{x}(t_2)\right\}$$

- It is a measure of correlation between sample function values of the random process $\boldsymbol{x}(t)$ at time instants t_1 and t_2

Autocovariance Function

- The **autocovariance** $C_x(t_1,t_2)$ of the random process $\boldsymbol{x}(t)$ is defined as

$$C_x(t_1,t_2) = E\left\{\left[\boldsymbol{x}(t_1) - m_x(t_1)\right]\left[\boldsymbol{x}(t_2) - m_x(t_2)\right]\right\}$$
$$= R_x(t_1,t_2) - m_x(t_1)m_x(t_2)$$

- Note that the variance of $\boldsymbol{x}(t)$ can be obtained from $C_x(t_1,t_2)$ as

$$Var\left(\boldsymbol{x}(t)\right) = E\left\{\left[\boldsymbol{x}(t) - m_x(t)\right]^2\right\} = C_x(t,t)$$

- The **correlation coefficient** $\rho_x(t_1,t_2)$ of the random process $\boldsymbol{x}(t)$ is defined as

$$\rho_x(t_1,t_2) = \frac{C_x(t_1,t_2)}{\sqrt{C_x(t_1,t_1)}\sqrt{C_x(t_2,t_2)}}$$

Wide-Sense Stationary Random Processes

- A process whose statistical properties do not change with time is called a **strictly stationary** random process.
- A random process $\boldsymbol{x}(t)$ is said to be **wide-sense stationary (WSS)** if

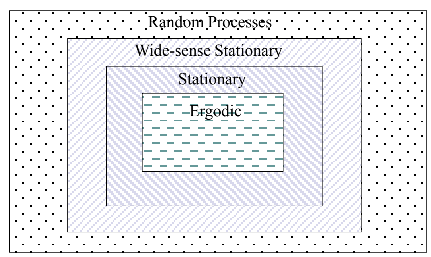
$$m_x(t) = \overline{\boldsymbol{x}(t)} = E\{\boldsymbol{x}(t)\} = \text{constant}$$

$$R_x(t, t+\tau) = R_x(\tau)$$

- Thus in order for a random process to be WSS, we only require that its mean is a constant and that the autocorrelation function depends only on the time difference
- A strict-sense stationary random process is always wide-sense stationary
  - However, the converse is not true in general, except for the Gaussian random process

Ergodic Random Processes

- A random process is **ergodic** if all time averages of sample functions equal corresponding ensemble averages
- Since time averages by definition are independent of time variable, it follows that an ergodic process is always stationary



Cross-correlation Function

- The **cross-correlation** function provides a measure of correlation between sample function amplitudes of processes $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ at time instants t_1 and t_2, respectively. It is defined as

$$R_{xy}(t_1,t_2) = E\left\{\boldsymbol{x}(t_1)\boldsymbol{y}(t_2)\right\}$$

- The random processes $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ are said to be
  - **Uncorrelated**

$$R_{xy}(t,t+\tau) = E\left\{\boldsymbol{x}(t)\boldsymbol{y}(t+\tau)\right\} = E\left\{\boldsymbol{x}(t)\right\}E\left\{\boldsymbol{y}(t+\tau)\right\}$$

  - **Orthogonal**

$$R_{xy}(\tau) = 0$$

Orthogonal and Independent Random Processes

- Note that if either of the processes $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ has a zero mean, uncorrelatedness implies orthogonality and vice versa
- $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ are said to be **independent** random processes if the set of random variables $\boldsymbol{x}(t_1), \boldsymbol{x}(t_2), \ldots, \boldsymbol{x}(t_n)$ is statistically independent of the set of random variables $\boldsymbol{y}(t_1'), \boldsymbol{y}(t_2'), \ldots, \boldsymbol{y}(t_n')$ for any choice of $t_1, t_2, \ldots, t_n$ and $t_1', t_2', \ldots, t_n'$
- Independence implies that the joint PDF of the random variables is the product of the PDFs of the individual variables
- The random processes $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ are said to **jointly stationary in wide-sense** if (a) $\boldsymbol{x}(t)$ is WSS; (b) $\boldsymbol{y}(t)$ is WSS; and (c) their cross-correlation is invariant under the shift of time origin. That is,

$$R_{xy}(t,t+\tau) = R_{xy}(\tau)$$

Gaussian Random Process

- A random process $\mathbf{x}(t)$ is said to be a Gaussian process if the random variables $\mathbf{x}(t_1), \mathbf{x}(t_2), ..., \mathbf{x}(t_n)$ are jointly Gaussian for any n and for any choice of $t_1, t_2, ..., t_n$
- Many processes that arise from natural phenomena are approximated well by Gaussian processes, using central limit theorem arguments
  - Examples include thermal noise in resistors and diffusion noise in semiconductors.
- Gaussian processes are also relatively easy to handle analytically. That is why they are so important in communication and control systems.

Important properties of Gaussian processes

- A Gaussian process $\mathbf{x}(t)$ is completely specified by the set of means

  $m_i = E\{\mathbf{x}(t_i)\}$

  and the set of autocorrelation functions

  $R_x(t_i, t_j) = E\{\mathbf{x}(t_i)\mathbf{x}(t_j)\}$

- For a Gaussian random process $\mathbf{x}(t)$, if $\mathbf{x}(t_1), \mathbf{x}(t_2), ..., \mathbf{x}(t_n)$ for any set of distinct time instants $t_1, t_2, ..., t_n$ are uncorrelated, then they are statistically independent
- If $\mathbf{x}(t)$ is a wide-sense stationary Gaussian process, then $\mathbf{x}(t)$ is a strictly stationary Gaussian process
- For an LTI system with Gaussian input process $\mathbf{x}(t)$, the output process $\mathbf{y}(t)$ is also Gaussian. Moreover, $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly Gaussian processes

Power Spectrum of a Random Process

- The **power spectral density (PSD)** $\mathcal{G}_x(f)$ of power signal $x(t)$ from (2.171) is given by

  $$\mathcal{G}_x(f) = \lim_{T\to\infty} \frac{|X_T(f)|^2}{T}$$

- For each sample function $x(t,\xi_i)$, we can write its PSD as

  $$\mathcal{G}_x(f, \xi_i) = \lim_{T\to\infty} \frac{|X_T(f,\xi_i)|^2}{T}$$

  where $X_T(f,\xi_i)$ is the FT of the truncated sample function $x_T(t,\xi_i)$
- A meaningful definition for the PSD of a random process would be the ensemble average of PSDs of all the sample functions.

  $$G_x(f) \triangleq E\{\mathcal{G}_x(f,\xi_i)\} = \lim_{T\to\infty} \frac{E\left\{|X_T(f,\xi_i)|^2\right\}}{T}$$

Power Spectrum of a Random Process (contd)

- The PSD describes power distribution of a of a random signal as a function of frequency
- For a WSS random process $\mathbf{x}(t)$, the PSD $G_x(f)$ is the Fourier transform of its autocorrelation function $R_x(\tau)$.

  $$G_x(f) = \Im[R_x(\tau)] = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau$$ ← **Wiener-Khinchin Theorem**

- Conversely,

  $$R_x(\tau) = \Im^{-1}[G_x(f)] = \int_{-\infty}^{\infty} G_x(f) e^{j2\pi f\tau} df$$

- It follows that

  $$\overline{\mathbf{x}^2(t)} = R_x(0) = \int_{-\infty}^{\infty} G_x(f) df$$

- That is, the area under $G_x(f)$ represents the total power of the random process $\mathbf{x}(t)$

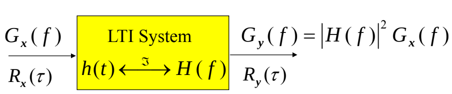
Response of Linear Systems to Random Signals

- The output of an LTI system with impulse response function $h(t)$ to the wide-sense stationary random signal $\mathbf{x}(t)$ is given by

  $$\mathbf{y}(t) = \mathbf{x}(t) \otimes h(t) = \int_{-\infty}^{+\infty} \mathbf{x}(t-u) h(u) du$$

- The PSD of output random process $\mathbf{y}(t)$ is related to the input process $\mathbf{x}(t)$ is related by

  $$G_y(f) = |H(f)|^2 \; G_x(f)$$

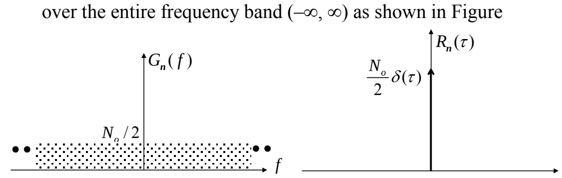


- Apply inverse FT to both sides, we obtain

  $$R_y(\tau) = h(\tau) \otimes h(-\tau) \otimes R_x(\tau)$$

White Gaussian Noise (WGN)

- The term **white noise** $\boldsymbol{n}(t)$ is used to describe a wide-sense stationary random process whose power spectral density is flat over the entire frequency band $(-\infty, \infty)$ as shown in Figure



- The constant spectral density, by convention, is denoted by $N_o/2$

  $$G_n(f) = \frac{N_o}{2}$$

WGN (contd)

- Taking the inverse Fourier transform yields

$$R_n(\tau) = \frac{N_o}{2}\delta(\tau) \qquad (*)$$

- White noise represents the ultimate in randomness since (*) implies instantaneous decorrelation
  - That is, any two samples of WGN are uncorrelated no matter how closely spaced they are
- White noise processes that are also Gaussian are called **white Gaussian noise (WGN)**
  - The samples of WGN $\boldsymbol{n}(t_1), \boldsymbol{n}(t_2), \ldots, \boldsymbol{n}(t_n)$ for any set of distinct time instants $t_1, t_2, \ldots, t_n$ are jointly Gaussian random variables
  - It follows from (*) and the Gaussian property that they are statistically independent

WGN (contd)

- WGN is an idealization of the noise observed in electronic components. This noise is caused by the chaotic motion of electrons in these components, and is commonly referred to as **thermal** noise
- Since the random motion of a large number of electrons contributes to this noise, we can apply the central limit theorem to conclude that this noise is a Gaussian random process
- Experiments conducted by Johnson (and verified analytically by Nyquist) in the 1920s showed that the power spectral density of thermal noise was constant for frequencies as high as 1000 GHz
- Although WGN is a useful mathematical abstraction, it does not conform to any random signal or noise observed in real life

Filtered White Gaussian Noise

- If WGN is passed through a nonideal filter with transfer function $H(f)$, the output noise spectral density is given by

$$G_y(f) = \frac{N_o}{2}|H(f)|^2$$

- The mean-square output power is given by

$$\overline{y^2(t)} = \int_{-\infty}^{\infty} G_y(f)df = N_o\int_0^{\infty}|H(f)|^2\, df \qquad (*)$$

- If we have an ideal filter with bandwidth B_N and gain equal to the maximum gain of the nonideal filter, $|H(f)|^2_{\max}$, as shown in Figure, the mean-square output power is given by

$$\overline{y^2(t)} = \frac{N_o}{2}\times 2B_N \times |H(f)|^2_{\max} = N_o B_N |H(f)|^2_{\max} \qquad (**)$$

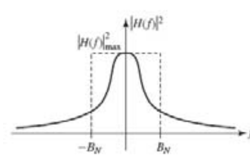
Noise-Equivalent Bandwidth

- We would now like to determine the equivalent bandwidth that passes the same amount of noise power as the nonideal filter. Comparing (*) and (**), we obtain

$$N_o B_N |H(f)|^2_{\max} = N_o\int_0^{\infty}|H(f)|^2\, df$$

- Solving for yields

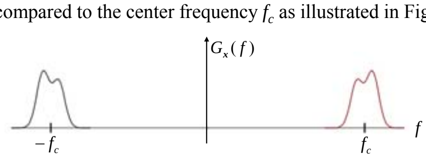
$$B_N = \frac{\int_0^{\infty}|H(f)|^2\, df}{|H(f)|^2_{\max}}$$



- B_N is called the **noise-equivalent bandwidth** of the nonideal filter $H(f)$

Narrowband Noise

- The power spectral density of a narrowband random process is nonzero only in a narrow frequency band which is very small compared to the center frequency f_c as illustrated in Figure



- It is convenient to represent the narrowband random process $\boldsymbol{x}(t)$ in terms of *in-phase* and *quadrature* components, $\boldsymbol{x}_c(t)$ and $\boldsymbol{x}_s(t)$, respectively

$$\boldsymbol{x}(t) = \boldsymbol{x}_c(t)\cos(2\pi f_c t) - \boldsymbol{x}_s(t)\sin(2\pi f_c t)$$

Spectral Densities of Quadrature Components

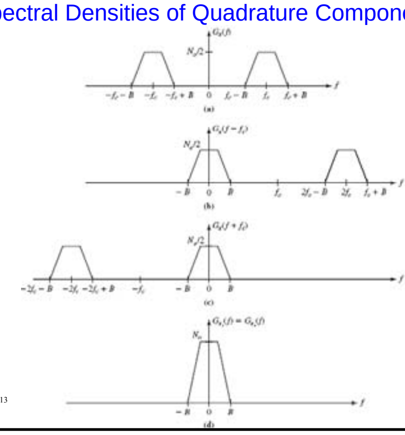
- The PSDs of quadrature noise components $\boldsymbol{x}_c(t)$ and $\boldsymbol{x}_s(t)$ are given by

$$G_{x_c}(f) = G_{x_s}(f) = \begin{cases} G_x(f+f_c) + G_x(f-f_c), & |f| \le B \\ 0, & \text{otherwise} \end{cases}$$

  where $G_x(f)$ is PSD of narrowband noise $\boldsymbol{x}(t)$
- Consider the noise $\boldsymbol{n}(t)$ obtained by passing white Gaussian noise with spectral density $N_o/2$ through an ideal bandpass filter centered at frequency f_c. Assume $f_c >> 2B$
- $\boldsymbol{n}(t)$ can be expressed in terms of its quadrature components $\boldsymbol{n}_c(t)$ and $\boldsymbol{n}_s(t)$ as

$$\boldsymbol{n}(t) = \boldsymbol{n}_c(t)\cos(2\pi f_c t) - \boldsymbol{n}_s(t)\sin(2\pi f_c t)$$

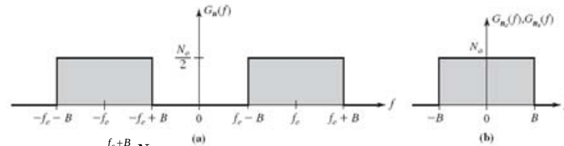
Spectral Densities of Quadrature Components



Narrowband White Gaussian Noise (contd)

- The spectral densities of $\boldsymbol{n}_c(t)$ and $\boldsymbol{n}_s(t)$ are obtained as

$$G_{\boldsymbol{n}_c}(f) = G_{\boldsymbol{n}_c}(f) = \begin{cases} N_o & |f| \le B \\ 0 & \text{otherwise} \end{cases}$$



$$\overline{\boldsymbol{n}^2(t)} = 2\int_{f_c-B}^{f_c+B} \frac{N_o}{2} df = 2N_oB$$

$$\overline{\boldsymbol{n}_c^2(t)} = \overline{\boldsymbol{n}_s^2(t)} = \int_{-B}^{B} N_o df = 2N_oB$$