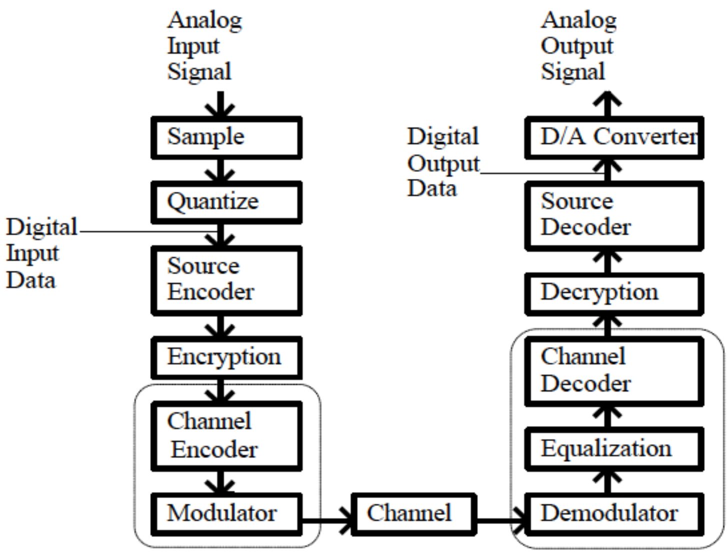
EE571 Digital Communications I

Proakis Chapter 2
Deterministic and Random Signal Analysis
Dr. Samir Alghadhban

1

Digital Communications System



2

Why are Random Processes important?

- Random Variables and Processes let us talk about quantities and signals which are unknown in advance:
- The data sent through a communication system is modeled as random
- The noise, interference, and fading introduced by the channel can all be modeled as random processes
- Even the measure of performance (Probability of Bit Error) is expressed in terms of a probability.

3

Random Events

- When we conduct a random experiment, we can use set notation to describe possible outcomes.
- Example: Roll a six-sided die.
  Possible Outcomes: *S*={1,2,3,4,5,6}
- An event is any subset of possible outcomes: *A*={1,2}
- The complementary event: $\bar{A} = S - A = \{3, 4, 5, 6\}$
- The set of all outcomes is the certain event: *S*
- The null event: ϕ
- Transmitting a data bit is also an experiment

4

Probability

- The probability P(A) is a number which measures the likelihood of the event A.

Axioms of Probability:

- No event has probability less than zero: $P(A) \geq 0$
- Also $P(A) \leq 1$ and $P(A) = 1 \Leftrightarrow A = S$
- Let A and B be two events such that: $A \cap B = \phi$
  Then: $P(A \cup B) = P(A) + P(B)$
- All other laws of probability follow from these axioms

5

Relationships Between Random Events

- Joint Probability: $P(A, B) = P(A \cap B)$
  - Probability that both A and B occur
- Conditional Probability: $P(A|B) = \dfrac{P(A, B)}{P(B)}$
  - Probability that A will occur given that B has occurred
- Statistical Independence:
  - Events A and B are statistically independent if:
    $P(A, B) = P(A) \cdot P(B)$
  - If A and B are independent then:
    $P(A|B) = P(A)$ and $P(B|A) = P(B)$

6

Random Variables

- A random variable *X(s)* is a real-valued function of the underlying event space: s □ S
- A random variable may be:
  - Discrete-valued: range is finite (e.g. {0,1}) or countably infinite (e.g., {1,2,3,...})
  - Continuous-valued - range is uncountably infinite (e.g.) □
- A random variable may be described by:
  - A name: *X*
  - It's range: X □□
  - A description of its distribution

7

Probability Distribution Function (PDF)

- Also called Cumulative Distribution Function (CDF)
- Definition: $F_X(x) = F(x) = P(X \le x)$
- Properties:

  $F(x)$ is monotonically nondecreasing

  $F(-\infty) = 0$

  $F(\infty) = 1$

  $P(a < X \le b) = F(b) - F(a)$
- While the PDF completely defines the distribution of a random variable, we will usually work with the pdf or pmf

8

Probability Density Function (pdf)

- Defn: $p_X(x) = \frac{dF_X(x)}{dx}$ or $p(x) = \frac{dF(x)}{dx}$
- Interpretations:
  - pdf measures how fast PDF is increasing or how likely a random variable is to lie at a particular value
- Properties:

$$p(x) \geq 0$$

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

$$P(a < X \leq b) = \int_a^b p(x)dx$$

9

Expected Values

- Expected values are a shorthand way of describing a random variable
- The most important examples are:
  - Mean: $E(X) = m_X = \int_{-\infty}^{\infty} xp(x)dx$
  - Variance: $E\left([X - m_X]^2\right) = \int_{-\infty}^{\infty} (x - m_X)^2 p(x)dx$
- The expectation operator works with any function:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)p(x)dx$$

10

Chebyshev Inequality

- Let X
  be a random variable with mean: m_X
  and variance: σ_X^2
- Then for any δ , $P(|X - m_X| \geq \delta) \leq \frac{\sigma_X^2}{\delta^2}$

- The size of the variance determines how a random variable is to lie close to it's mean value

11

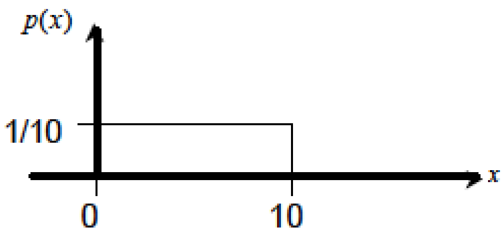
Chernoff Bound

- Let Y be a random variable
- Then, for any value of ν>0 and δ>0:
  $$\Pr[Y \geq \delta] \leq E\left(e^{\nu(Y-\delta)}\right)$$
- Very useful for upper bounding low probability events on the tails of distributions
- Example:
  $$p_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \Rightarrow \Pr[Y \geq \delta] \leq e^{-\delta^2/2}$$

12

Example #1: Uniform pdf

- $p(x) = \begin{cases} 1/10, & 0 \le x \le 10 \\ 0, & \text{else} \end{cases}$



Example #1 (Find the mean and the variance)

- Mean:

$$m_x = \int_{-\infty}^{\infty} x \cdot p(x)dx = \int_0^{10} x \cdot \frac{1}{10}dx = \left[\frac{x^2}{20}\right]_0^{10} = 5$$

- Variance:

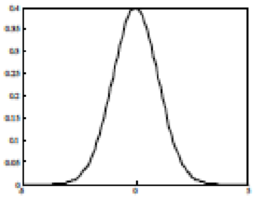
$$\sigma_x^2 = \int_{-\infty}^{\infty} (x-5)^2 \cdot p(x)dx = \int_0^{10} (x-5)^2 \cdot \frac{1}{10}dx = \frac{5}{6}$$

- Probability Calculation:

$$P(6 \le x \le 9) = \int_6^9 p(x)dx = \int_6^9 \frac{1}{10}dx = 0.3$$

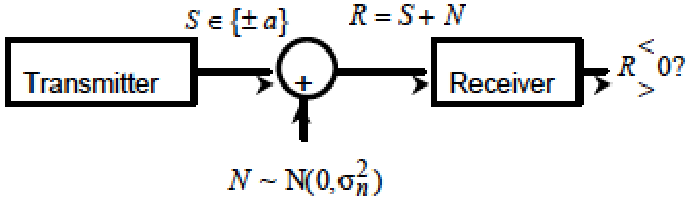
Example #2: Gaussian pdf

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



- A Gaussian random variable is completely determined by its mean and variance

A Communication System with Gaussian Noise



- The probability that the receiver will make an error is:

$$P(R > 0 | S = -a) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(x+a)^2}{2\sigma_n^2}} = Q\left(\frac{a}{\sigma_n}\right)$$

The Q-function

- The Q-function is a standard form for expressing error probalities without a closed form:

$$Q(x)=\int_x^\infty \frac{1}{\sqrt{2\pi}}e^{-u^2/2}du=\frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)$$

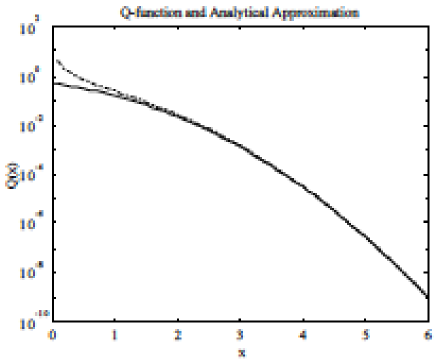
- Numerical Calculation of Q-function:

$$Q(x)=\frac{1}{x\sqrt{2\pi}}e^{-x^2/2}\left[1-\frac{1}{x^2}+\frac{1\cdot 3}{x^4}+\cdots+\frac{(-1)^n\cdot 1\cdot 3\cdots(2n-1)}{x^{2n}}\right]$$

$$\approx \frac{1}{x\sqrt{2\pi}}e^{-x^2/2},\ \text{for } x\geq 3$$

17

The Q-function and its Approximation



18

Example #3 - Rayleigh pdf

- Let: $R = \sqrt{X_1^2 + X_2^2}$

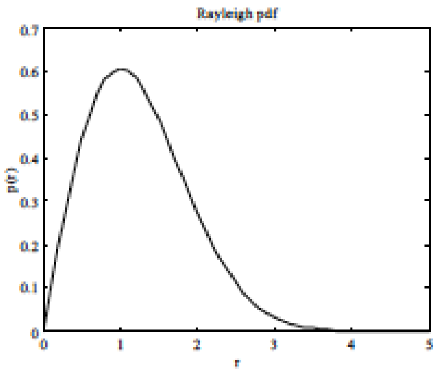
  where X_1 and X_2 are Gaussian with mean 0 and variance σ^2
- Then R is a Rayleigh random variable with pdf:

$$p_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}$$

- Rayleigh pdf's are frequently used to model fading when no line of site signal is present

19

Rayleigh pdf



20

Probability Mass Functions (pmf)

- A discrete random variable can be described a pdf if we allow impulse functions
- We usually use probability mass functions (pmf):
  $p(x) = P(X = x)$
- Properties are analogous to pdf:
  $p(x) \geq 0$
  $\sum_X p(x) = 1$
  - $P(a \leq X \leq b) = \sum_{x=a}^{b} p(x)$

21

Example #1: Binary Distribution

- $p(x) = \begin{cases} 1/2, & x = 0 \\ 1/2, & x = 1 \end{cases}$
- This is frequently used to model binary data
- Mean: $m_X = \sum_x x \cdot p(x) = 0 \cdot 1/2 + 1 \cdot 1/2 = 1/2$
- Variance:
  $\sigma_X^2 = \sum_x (x - m_X)^2 \cdot p(x) = (1/2)^2 \cdot 1/2 + (1/2)^2 \cdot 1/2 = 1/4$
- If X_1 and X_2 are independent binary random variables, then $p_{X_1 X_2}(0,0) = p_{X_1}(0) \cdot p_{X_2}(0) = 1/2 \cdot 1/2 = 1/4$

22

Example #2: Binomial Distribution

- Let $Y = \sum_{i=1}^{n} X_i$ where $\{X_i, i = 1, \ldots, n\}$ are independent binary RVs with: $p_X(x) = \begin{cases} 1-p, & x=0 \\ p, & x=1 \end{cases}$
- Then $p_Y(y) = \binom{n}{y} p^y (1-p)^{n-y}$, $\binom{n}{y} = \frac{n!}{y! \cdot (n-y)!}$
- Mean: $m_X = n \cdot p$
- Variance: $\sigma_X^2 = n \cdot p \cdot (1-p)$

23

Example #2 (continued)

- Suppose that we transmit a 31 bit sequence with error correction capable of correcting up to 3 errors.
- If the probability of a bit error is p=0.001, what is the probability that the codeword is received in error?

$$P(\text{codeword error}) = 1 - P(\text{correct codeword})$$

$$= 1 - \sum_{i=0}^{3} \binom{31}{i} (0.999)^{31-i} (0.001)^i \approx 3 \times 10^{-8}$$

- If no error correction is used, the error probability is:

$$1 - (1 - 0.001)^{31} = 0.0305$$

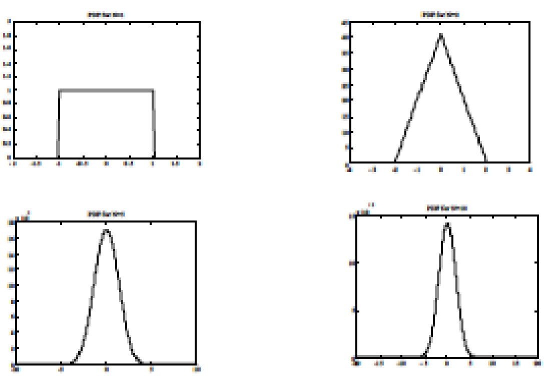
24

Central Limit Theorem

- Let $X_1, X_2, \ldots, X_N$ be a set of independent random variables with identical pdfs
- Let: $Y = \sum_{i=1}^{N} X_i$
- Then as $N \to \infty$, the distribution of will tend towards a Gaussian distribution
- In practice, N=10 is usually enough to see this effect
- Thermal noise results from the random movement of many electrons - it is well modeled by a Gaussian distribution

25

Example of Central Limit Theorem:



26

Random Processes

- A random variable has a single value. However, actual signals change with time.
- Random variables model unknown events.
- Random processes model unknown signals.
- A random process is just a collection of random variables.
- If *X(t)* is a random process then *X(1)*, *X(1.5)*, and *X(37.5)* are all random variables for any specific time *t*

27

Terminology Describing Random Processes

- A stationary random process has statistical properties which do not change at all with time (i.e., all joint pdfs do not change)
- A wide sense stationary (WSS) process has a mean and autocorrelation function which do not change with time (this is usually sufficient)
- A random process is ergodic if the time average always converges to the statistical average.
- Unless specified, we will assume that all random processes are WSS and ergodic.

28

Description of Random Processes

- Knowing the pdf of individual samples of the random process is not sufficient. We also need to know how how individual samples are related to each other.
- Two tools are available to describe this relationship:
  - Autocorrelation function
  - Power spectral density function

29

Autocorrelation

- Autocorrelation measures how a random process changes with time.
- Intuitively, *X(1)* and *X(1.1)* will be more strongly related than *X(1)* and *X(100000)* (although it is possible to construct counterexamples). The autocorrelation function quantifies this.
- Defn (for WSS random processes):
$$\phi_X(\tau) = E[X(t)X(t+\tau)]$$
- Note that Power $= \phi_X(0)$

30

Power Spectral Density

- $\Phi(f)$ tells us how much power is at each frequency
- Wiener-Khinchine Theorem: $\Phi(f) = F\{\phi(\tau)\}$

  Power spectral density and autocorrelation are a Fourier Transform pair!
- Properties of Power Spectral Density:

  $\Phi(f) \geq 0$

  $\Phi(f) = \Phi(-f)$

  ◆ Power $= \int_{-\infty}^{\infty} \Phi(f) df$

31

Gaussian Random Processes

Gaussian Random Processes have several special properties:

- If a Gaussian random process is wide-sense stationary, then it is also stationary.
- Any sample point from a Gaussian random process is a Gaussian random variable
- If the input to a linear system is a Gaussian random process, then the output is also a Gaussian random process

32

Linear Systems

- Input: $x(t)$
- Impulse Response: $h(t)$
- Output: $y(t)$

33

Computing the Output of Linear Systems

- Deterministic Signals:
  - Time Domain: $y(t) = h(t) * x(t)$
  - Frequency Domain: $Y(f) = F\{y(t)\} = X(f)H(f)$
- For a random process, we can still relate the statistical properties of the input and output signal
  - Time Domain: $\phi_Y(\tau) = \phi_X(\tau) * h(\tau) * h(-\tau)$
  - Frequency Domain: $\Phi_Y(f) = \Phi_X(f) \cdot |H(f)|^2$

34

Gaussian Process

- Suppose that we observe a random process X(t) for an interval that starts at time t=0 and lasts until t=T.
- Suppose also that we weight the random process X(t) by some function g(t) and then integrate the product g(t)X(t) over this observation interval
- Thereby obtaining a random variable Y defined by:

$$Y = \int_{o}^{T} g(t)X(t)dt$$

- Y is a linear function of X(t)

35

Gaussian Process

- If the mean-square value of the random variable Y is finite and if the random variable Y is a Gaussian-distributed random variable for every g(t) in this class of functions,
- Then the process X(t) is a Gaussian process
- In other words, the process X(t) is Gaussian process if every linear function of X(t) is a Gaussian random variable.
- For example, $Y = \sum_{i=1}^{N} a_i X_i$

36

Gaussian Process

- The random variable Y has a Gaussian distribution if its pdf is:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left[-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right]$$

Where μ_Y is the mean and σ^2_Y is the variance

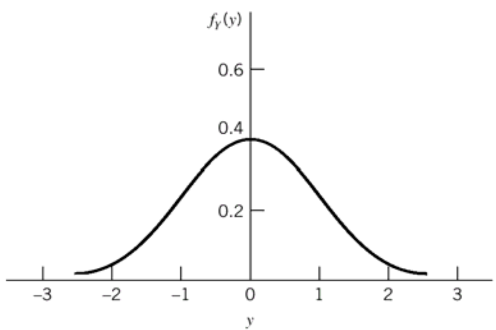
- The normalized Gaussian random variable Y has a zero mean ($\mu_Y = 0$) and unit variance ($\sigma^2_Y = 1$) :

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)$$

37

Normalized Gaussian distribution

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)$$



38

Gaussian Process

- A Gaussian process has two main virtues:

1- the Gaussian process has many properties that make analytic results possible

2- By experimental verifications, physical phenomena usually follow a Gaussian model.

Central Limit Theorem

- The central limit theorem provides the mathematical justification for using a Gaussian process as a model for a large number of different physical phenomena in which the observed random variable, at a particular instant of time. is the result of a large number of individual random events.
- To formulate this important theorem, let X_i , i = 1, 2, . . . , N be a set of random variables that satisfies the following requirements:
  1. The X_i are statistically independent.
  2. The X_i have the same probability distribution with mean μ_X and variance σ^2_X.

Central Limit Theorem

- The X_i so described are said to constitute a set of independently and identically distributed (i.i.d.) random variables.
- Let these random variables be normalized as follows:

$$Y_i = \frac{1}{\sigma_X}(X_i - \mu_X), \qquad i = 1, 2, \ldots, N$$

So that we have $E[Y_i] = 0$ And $\text{var}[Y_i] = 1$

Define the new random variable $V_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y_i$

41

Central Limit Theorem

- The central limit theorem states that the probability distribution of V_N approaches a normalized Gaussian distribution *N*(0, 1) in the limit as the number of random variables N approaches infinity.

42

Properties of a Gaussian Process

<u>Property 1:</u>

- If a Gaussian process X(t) is applied to a stable linear filter, then the random process Y(t) developed at the output of the filet is also Gaussian.

Properties of a Gaussian Process

<u>Property 2:</u>

- Consider the set of random variables or samples $X(t_1)$, $X(t_2)$, . . ., $X(t_n)$, obtained by observing a random process X(t) at times $t_1, t_2, \ldots, t_n$. If the process X(t) is Gaussian, then this set of random variables is jointly Gaussian for any *n*, with their *n*-fold joint probability density function being completely determined by specifying
  1- the set of means $\mu_{X(t_i)} = E[X(t_i)], \quad i = 1, 2, \ldots, n$
  2- and the set of covariance functions

$$C_X(t_k, t_i) = E[(X(t_k) - \mu_{X(t_k)})(X(t_i) - \mu_{X(t_i)})], \quad k, i = 1, 2, \ldots, n$$

Multivariate Gaussian distribution

- Let the *n*-by-1 vector X denote the set of random variables $X(t_1), \ldots, X(t_n)$ derived from the Gaussian process X(t) by sampling it at times $t_1 \ldots, t_n$. Let x denote a value of X. According to Property 2, the random vector X has a multivariate Gaussian distribution defined in matrix form as:

$$f_{X(t_1),\ldots,X(t_2)}(x_1, \ldots, x_2) = \frac{1}{(2\pi)^{n/2}\Delta^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

Where the superscript T denotes transposition and

$$\begin{aligned} \boldsymbol{\mu} &= \text{mean vector} \\ &= [\mu_1, \mu_2, \ldots, \mu_n]^T \\ \boldsymbol{\Sigma} &= \text{covariance matrix} \\ &= \{C_X(t_k, t_i)\}_{k,i=1}^n \\ \boldsymbol{\Sigma}^{-1} &= \text{inverse of covariance matrix} \\ \Delta &= \text{determinant of covariance matrix } \boldsymbol{\Sigma} \end{aligned}$$

45

Properties of a Gaussian Process

<u>Property 3</u>

- If a Gaussian process is stationary, then the process is also strictly stationary.

<u>Property 4</u>

If the random variable $X(t_1), \ldots, X(t_n)$ obtained by sampling a Gaussian process X(t) at times $t_1, t_2, \ldots, t_n$, are uncorrelated, that is,

$$E[(X(t_k) - \mu_{X(t_k)})(X(t_i) - \mu_{X(t_i)})] = 0, \qquad i \neq k$$

Then these random variables are statistically independent.

46

Properties of a Gaussian Process

Property 4 (continue)

That means that the covariance matrix of X(t) is diagonal:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & & & \mathbf{O} \\ & \sigma_2^2 & & \\ & & \ddots & \\ \mathbf{O} & & & \sigma_n^2 \end{bmatrix}$$

Where

$$\sigma_i^2 = E[(X(t_i) - E[X(t_i)])^2], \qquad i = 1, 2, \dots, n$$

47

Properties of a Gaussian Process

Property 4 (continue)

Therefore, the multivariate Gaussian distribution will be:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

Where

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - \mu_{X_i})^2}{2\sigma_i^2}\right)$$

48

Properties of a Gaussian Process

Property 4 (continue)

In words, if the Gaussian random variables $X(t_l)$, . . ., $X(t_n)$ are uncorrelated., then they are statistically independent, which, in turn, means that the joint probability density function of this set of random variables can be expressed as the product of the probability density functions of the individual random variables in the set.

49

Noise

What do we mean by noise? From where does it come from?

- The term noise is used customarily to designate unwanted signals that tend to disturb the transmission and processing of signals in communication systems and over which we have incomplete control.
- There are many potential sources of noise in a communication system.
  - **external to the system** (e.g., atmospheric noise, galactic noise, man-made noise)
  - **internal to the system**, such as the noise that arises from spontaneous fluctuations of current or voltage in electrical circuits. This type of noise represents a basic limitation on the transmission or detection of signals in communication systems involving the use of electronic devices.
- The two most common examples of spontaneous fluctuations in electrical circuits are shot noise anti thermal noise.
- White noise is an idealized form of noise used in communication systems analysis.

50

Shot Noise

- Shot noise arises in electronic devices such as diodes and transistors because of the discrete nature of current flow in these devices.
- For example, in a *photodetector* circuit a current pulse is generated every time an electron is emitted by the cathode due to incident fight from a source of constant intensity.
- The electrons are naturally emitted at random times denoted by τ_k where $-\infty<k<\infty$

51

Shot Noise

- Thus, the total current flowing through the photo-detector may be modeled as an infinite sum of current pulses as:

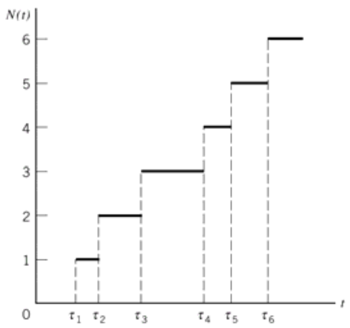
$$X(t) = \sum_{k=-\infty}^{\infty} h(t - \tau_k)$$

Where $h(t-\tau_k)$ is the current pulse generated at time τ_k

52

Shot Noise

- The number of electrons, *N(t),* emitted in the time interval [0, *t*] constitutes a discrete stochastic process, the value of which increases by one each time an electron is emitted.



the mean value of the number of electrons , *ν,* emitted between times *t* and $t + t_a$ is:

$$E[\nu] = \lambda t_0$$

The parameter **λ** is a constant called the ***rate of the process***. The total number of electrons emitted in the interval $[t, t + t_0]$ is:

$$\nu = N(t + t_0) - N(t)$$

Which follows a *Poisson distribution* with a mean value equal to

53

Shot Noise

The probability that *k* electrons are emitted in the interval $[t, t + t_0]$ is defined by:

$$P(\nu = k) = \frac{(\lambda t_0)^k}{k!} e^{-\lambda t_0} \qquad k = 0, 1, \ldots$$

54

Thermal Noise

Thermal noise is the name given to the electrical noise arising from the random motion of electrons in a conductor.

The mean-square value of the thermal noise voltage V_{TN} appearing across the terminals of a resistor, measured in a bandwidth of Δf Hertz, is given by:

$$E[V_{TN}^2] = 4kTR\ \Delta f \text{ volts}^2$$

← Noise Power

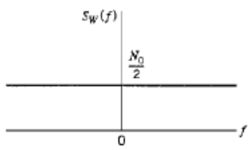
where *k* is *Boltzmann's constant* equal to 1.38 X 10^{-23} joules per degree Kelvin
T is the *absolute temperature* in degrees Kelvin
and R is the resistance in ohms

55

White Noise

- The noise analysis of communication systems is customarily based on an idealized form of noise called *white noise,* the power spectral density of which is independent of the operating frequency.
- The adjective *white* is used in the sense that white light contains equal amounts of all frequencies within the visible band of electromagnetic radiation.
- We express the **power spectral density** of white noise, with a sample function denoted by *w(t),* as

$$S_W(f) = \frac{N_0}{2}$$



56

White Noise

- The dimensions of N_0 are in watts per Hertz.
- The parameter N_0 is usually referenced to the input stage of the receiver of a communication system. It may be expressed as:

$$N_0 = kT_e$$

where k is Boltzmann's constant and T_e is the *equivalent noise temperature* of the receiver

- The equivalent noise temperature of a system is defined as

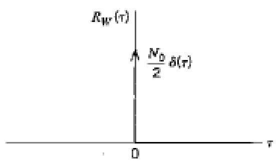
the temperature at which a noisy resistor has to be maintained such that, by connecting the resistor to the input of a noiseless version of the system, it produces the same available noise power at the output of the system as that produced by all the sources of noise in the actual system

57

White Noise

Since the **autocorrelation function** is the inverse Fourier transform of the power spectral density, it follows that for white noise:

$$R_W(\tau) = \frac{N_0}{2}\,\delta(\tau)$$



Strictly speaking, white noise has infinite average power and, as such, it is not physically realizable. Nevertheless, white noise has simple mathematical properties which make it useful in statistical system analysis.

58