CT516 Advanced Digital Communications
Lecture 2: Review of Probability and Random Processes

Yash M. Vasavada
Associate Professor, DA-IICT, Gandhinagar

4th January 2017
Overview of Today’s Talk

1. Introduction
2. Random Variables (RVs)
   - Definition of RV
   - Characterization of RVs
3. Random Processes
Overview of Today’s Talk

1. Introduction
2. Random Variables (RVs)
   - Definition of RV
   - Characterization of RVs
3. Random Processes
Overview of Today’s Talk

1. Introduction
2. Random Variables (RVs)
   - Definition of RV
   - Characterization of RVs
3. Random Processes
Why to Study Random Processes?

Random variables and processes let us talk about quantities and signals which are *not* known in advance:

- Data sent through the communication channel is best modeled as a random sequence.
- Noise, interference and fading affecting this data transmission are all random processes.
- Receiver performance is measured probabilistically.
Random Events

- Outcome of any random event can be viewed as a member of a set.
- This set is a set of all possible outcomes of this random event or random experiment.
  - Roll a six-sided die:
    - Set of all possible outcomes: \( S = \{1, 2, \ldots, 6\} \)
  - Toss a coin:
    - Set of all possible outcomes: \( S = \{\text{Head, Tail}\} \)
  - Transmit one bit in a noisy channel:
    - Set of all possible outcomes:
      \( S = \{\text{Bit is received correctly, Bit is received incorrectly}\} \)
Random Events

- An event is a subset of set $S$.
  - Roll a six-sided die:
    - An event: $A = \{1, 2\}$
  - Toss a coin:
    - An event: head turns up $A = \{\text{Head}\}$
  - Transmit one bit in a noisy channel:
    - An event: $A = \{\text{Bit is received correctly}\}$

- Set $S$ of all possible outcomes is a certain event. Its probability is 1.
- Set $\emptyset$ is a null event. Its probability is zero.
- Subset $A$ of set $S$ denotes a probable event. Its probability is a variable between 0 and 1.
Axioms of Probability

- Probability $P(A)$ is a number which measures the likelihood of event $A$.

- Following are three axioms of probability:
  1. $P(A) \geq 0$ (i.e., no event has a probability less than zero).
  2. $P(A) \leq 1$, and $P(A) = 1$ only if $A = S$, i.e., if $A$ is a certain event.
  3. Let $A$ and $B$ be two events such that $A \cap B = \emptyset$. In this case, $P(A \cup B) = P(A) + P(B)$ (i.e., probabilities of mutually exclusive events add).

- All other theorems of probability follow from these three axioms.
Rules of Probability

- **Joint probability** $P(A, B) = P(A \cap B)$ is the probability that both $A$ and $B$ occur.
- **Conditional probability** $P(A|B) = \frac{P(A, B)}{P(B)}$ is the probability that $A$ will occur given $B$ has occurred.
- **Statistical independence**: events $A$ and $B$ are statistically independent if $P(A, B) = P(A) \times P(B)$.
  - If $A$ and $B$ are statistically independent, $P(A|B) = P(A)$ and $P(B|A) = P(B)$. 
Introduction

Random Variables (RVs)

• Definition of RV

Random Variables

- A random variable (RV) $X(s)$ is a real-valued function of the underlying event space $s \in S$
- Typically we omit the notation $s$ and just denote the RV as $X$
- A random variable may be:
  → Discrete-valued with either finite range (e.g., $[0, 1]$) or infinite range
  → Continuous-valued (e.g., range can be the set $\mathcal{R}$ of real numbers)
- A random variable is described by its name, its range and its distribution
Cumulative Distribution Function (CDF)

- Definition: $F_X(x) = F(x) = P(X \leq x)$
- Properties:
  1. $F(x)$ is monotonically nondecreasing
  2. $F(-\infty) = 0$
  3. $F(\infty) = 1$
  4. $P(a < X \leq b) = F(b) - F(a)$

- CDF completely defines the probability distribution of a RV
- Alternate specifications are called PDF (Probability Density Function - for continuous variables) or PMF (Probability Mass Function - for discrete variables)
Characterization of RVs

Probability Density Function (PDF)

- **Definition**: \( p_X(x) = \frac{dF(x)}{dx} \)

- **Interpretations**: PDF measures
  - how fast CDF is increasing
  - how likely a random variable is to lie at a particular value

- **Properties**
  1. \( p(x) \geq 0 \)
  2. \( \int_{-\infty}^{\infty} p(x) \, dx = 1 \)
  3. \( P(a < X \leq b) = \int_{a}^{b} p(x) \, dx \)
Expected Values

- Sometimes the PDF is cumbersome to specify, or it may not be known.
- Expected values are shorthand ways of describing the behavior of RVs.
- Most important examples are:
  - Mean: $E(x) = m_x = \int_{-\infty}^{\infty} x \cdot p(x) \, dx$
  - Variance: $E((x - m_x)^2) = \int_{-\infty}^{\infty} (x - m_x)^2 \cdot p(x) \, dx$

- Expectation operator works with any function $Y = g(X)$.
  - $E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot p(x) \, dx$
Characterization of RVs

Example RVs

Uniform PDF

\[ p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{else} \end{cases} \]
Characterization of RVs

Example PDFs

Uniform CDF

\[ F(x) = \begin{cases} 
0, & x \leq 0 \\
\frac{x - a}{b - a}, & a \leq x \leq b \\
1, & x \geq b 
\end{cases} \]
Characterization of RVs

Example PDFs

Uniform RV

- Mean: \( m_x = \int_a^b x \, p(x) \, dx = \frac{1}{b-a} \int_a^b x \, dx = \frac{a+b}{2} \)

- Variance: \( \sigma_x^2 = \int_a^b (x - m_x)^2 \, p(x) \, dx = \frac{(b-a)^2}{12} \)

- Probability:
  \[
P(a_1 \leq x < b_1) = \int_{a_1}^{b_1} p(x) \, dx = \frac{b_1 - a_1}{b - a}, \quad a < a_1, b_1 < b
\]
Characterization of RVs

Example RVs

Gaussian PDF

\[ p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left( -\frac{(x - m_x)^2}{2\sigma_x^2} \right) \]
Characterization of RVs

Example RVs

Gaussian CDF

\[ m_x = 0, \sigma_x^2 = 1 \]
\[ m_x = 3, \sigma_x^2 = 0.4 \]
Characterization of RVs

Example RVs

Gaussian PDF

http://turtleinvestor888.blogspot.in/2012_01_01_archive.html
Let $X_1, X_2, \ldots, X_N$ be $N$ independent RVs with identical PDFs.

Let $Y = \sum_{i=1}^{N} X_i$.

A theorem of probability theory called Central Limit Theorem or CLT: as $N \to \infty$, distribution of $Y$ tends to a Gaussian distribution.

In practice, $N = 10$ is sufficient to see the tendency of $Y$ to follow the Gaussian PDF.

Importance of CLT:

- Thermal noise results from random movements of many electrons, and it is well modeled by the Gaussian PDF.
- Interference from many identically distributed interferers in a CDMA system tends toward the Gaussian PDF.
Characterization of RVs

Example RVs

Central Limit Theorem: A Uniform Distribution
Characterization of RVs

Example RVs

Central Limit Theorem. Average of $N = 2$ identically distributed Uniform RVs

Blue: PDF of $Y = \frac{1}{N} \sum_{i=1}^{N} X_i$, $N = 2$

Red: Gaussian PDF with Same Variance
Characterization of RVs

Example RVs

Central Limit Theorem. Average of $N = 3$ identically distributed Uniform RVs
Characterization of RVs

Example RVs

Central Limit Theorem. Average of $N = 4$ identically distributed Uniform RVs

Blue: PDF of $Y = (1/N) \times \sum_{i=1}^{N} X_i$. $N = 4$

Red: Gaussian PDF with Same Variance
Characterization of RVs

Example RVs
Central Limit Theorem. Average of $N = 10$ identically distributed Uniform RVs

Blue: PDF of $Y = \frac{1}{N} \sum_{i=1}^{N} X_i$, $N = 10$
Red: Gaussian PDF with Same Variance
Example RVs

Central Limit Theorem. Average of $N = 19$ identically distributed Uniform RVs

Blue: PDF of $Y = (1/N) \times \sum_{i=1}^{N} X_i$, $N = 19$

Red: Gaussian PDF with Same Variance
An application of Gaussian PDFs: signal level at the output of a digital communications receiver can often be given as $r = s + n$, where

- $r$ is the received signal level,
- $s = -a$ is the transmitted signal level, and
- $n$ is the Gaussian noise with mean 0 and variance $\sigma_n^2$.

Probability that the signal level $-a$ can be mistaken by the receiver as the signal level $+a$ is given as:

$$P(r > 0) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp \left( -\frac{(x + a)^2}{2\sigma_x^2} \right) \, dx = Q \left( \frac{a}{\sigma_n} \right)$$

Definition: $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp \left( -\frac{v^2}{2} \right) \, dv$
Suppose \( r = \sqrt{x_1^2 + x_2^2} \), where \( x_1 \) and \( x_2 \) are Gaussian with zero mean and variance \( \sigma^2 \).

PDF \( p(r) = \frac{r}{\sigma^2} \exp \left(- \frac{r^2}{2\sigma^2} \right) \) is the Rayleigh PDF.

Used to model fading when no line of sight is present.
Probability Mass Functions or PMFs

- For discrete random variables, the concept analogous to probability density function or PDF is PMF; \( P(X = x) = p(x) \)
- Properties are analogous to those of the PDFs:
  1. \( p(x) \geq 0 \)
  2. \( \sum_x p(x) = 1 \)
  3. \( P(a < X \leq b) = \sum_{x=a}^{b} p(x) \)
Example PMFs

**Binary Distribution**

- **Outcome of the toss of a fair coin:** \( p(x) = \begin{cases} 
\frac{1}{2}, & x = 0 \text{ (head)}, \\
\frac{1}{2}, & x = 1 \text{ (tail)} 
\end{cases} \)

  \[ m_x = \sum x p(x) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = 1/2 \]

  \[ \sigma_x^2 = 1/4 \]

- **If** \( X_1 \) and \( X_2 \) are independent binary random variables,

  \[ P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0) \]
Characterization of RVs

Example PMFs

Binomial Distribution

- Let \( Y = \sum_{i=1}^{n} X_i \), where \( \{X_i\}; i = 1, \ldots, N \) are independent binary RVs with \( p(x) = \begin{cases} 1 - p, & x = 0 \text{(head)}, \\ p, & x = 1 \text{(tail)} \end{cases} \)

- In this case, RV \( Y \) follows the Binomial Distribution given as

\[
P_Y(y) = \binom{N}{y} p^y (1 - p)^{N-y}
\]

- Mean \( m_y = N \times p \)

- Variance \( \sigma^2_y = N \times p \times (1 - p) \)
Suppose we transmit a 31 bit sequence (i.e., a codeword) with a channel code that is capable of correcting upto 3 bits in error.

If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?

$$P(\text{codeword in error}) = 1 - P(\text{correct codeword}) = 1 - \sum_{i=0}^{3} \binom{31}{i} (0.999)^{31-i} (0.001)^i \approx 3 \times 10^{-8}$$

What is the error probability if no channel coding is used?
Binomial Distribution

An Application

- Suppose we transmit a 31 bit sequence (i.e., a codeword) with a channel code that is capable of correcting up to 3 bits in error.
- If the probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?
- $P(\text{codeword in error}) = 1 - P(\text{correct codeword}) = 1 - \sum_{i=0}^{3} \binom{31}{i} (0.999)^{31-i} (0.001)^{i} \approx 3 \times 10^{-8}$
- What is the error probability if no channel coding is used?
Suppose we transmit a 31 bit sequence (i.e., a codeword) with a channel code that is capable of correcting upto 3 bits in error.

If probability of an individual bit in error is \( p = 0.01 \), what is the probability that the codeword is received in error?

\[
P(\text{codeword in error}) = 1 - P(\text{correct codeword}) = 1 - \sum_{i=0}^{3} \binom{31}{i} (0.999)^{31-i} (0.001)^i \approx 3 \times 10^{-8}
\]

What is the error probability if no channel coding is used?
Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding.

If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?

Answer:

$\rightarrow$ Probability of receiving one bit correctly: $1 - p = 0.999$

$\rightarrow$ Recall: If $X_1$ and $X_2$ are independent binary random variables,

$$P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$$

$\rightarrow$ Therefore, probability of receiving two bits correctly:

$$(1 - p)^2 = (0.999)^2$$

$\rightarrow$ Probability of receiving a total of $i$ bits correctly:

$$(1 - p)^i = (0.999)^i$$

$\rightarrow$ Thus, probability of receiving all 31 bits correctly:

$$(1 - p)^{31} = (0.999)^{31}$$

$\rightarrow$ Thus, probability of making a codeword error:

$$1 - (0.999)^{31} = 0.0305$$

Compare this against error probability of $3 \times 10^{-8}$ obtained with the channel coding of prior slide.
Characterization of RVs

Binomial Distribution
An Application

- Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding
- If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?
- Answer:
  - Probability of receiving one bit correctly: $1 - p = 0.999$
  - Recall: If $X_1$ and $X_2$ are independent binary random variables, $P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$
  - Therefore, probability of receiving two bits correctly: $(1 - p)^2 = (0.999)^2$
  - Probability of receiving a total of $i$ bits correctly: $(1 - p)^i = (0.999)^i$
  - Thus, probability of receiving all 31 bits correctly: $(1 - p)^{31} = (0.999)^{31}$
  - Thus, probability of making a codeword error: $1 - (0.999)^{31} = 0.0305$
- Compare this against error probability of $3 \times 10^{-8}$ obtained with the channel coding of prior slide
Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding.

If probability of an individual bit in error is \( p = 0.01 \), what is the probability that the codeword is received in error?

Answer:

- Probability of receiving one bit correctly: \( 1 - p = 0.999 \)
- Recall: If \( X_1 \) and \( X_2 \) are independent binary random variables, \( P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0) \)
- Therefore, probability of receiving two bits correctly: \( (1 - p)^2 = (0.999)^2 \)
- Probability of receiving a total of \( i \) bits correctly: \( (1 - p)^i = (0.999)^i \)
- Thus, probability of receiving all 31 bits correctly: \( (1 - p)^{31} = (0.999)^{31} \)
- Thus, probability of making a codeword error: \( 1 - (0.999)^{31} = 0.0305 \)

Compare this against error probability of \( 3 \times 10^{-8} \) obtained with the channel coding of prior slide.
Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding.

If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?

Answer:

- Probability of receiving one bit correctly: $1 - p = 0.999$
- Recall: If $X_1$ and $X_2$ are independent binary random variables, $P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$
- Therefore, probability of receiving two bits correctly: $(1 - p)^2 = (0.999)^2$
- Probability of receiving a total of $i$ bits correctly: $(1 - p)^i = (0.999)^i$
- Thus, probability of receiving all 31 bits correctly: $(1 - p)^{31} = (0.999)^{31}$
- Thus, probability of making a codeword error: $1 - (0.999)^{31} = 0.0305$

Compare this against error probability of $3 \times 10^{-8}$ obtained with the channel coding of prior slide.
Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding
If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?

Answer:

→ Probability of receiving one bit correctly: $1 - p = 0.999$
→ Recall: If $X_1$ and $X_2$ are independent binary random variables, $P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$
→ Therefore, probability of receiving two bits correctly: $(1 - p)^2 = (0.999)^2$
→ Probability of receiving a total of $i$ bits correctly: $(1 - p)^i = (0.999)^i$
→ Thus, probability of receiving all 31 bits correctly: $(1 - p)^{31} = (0.999)^{31}$
→ Thus, probability of making a codeword error: $1 - (0.999)^{31} = 0.0305$

Compare this against error probability of $3 \times 10^{-8}$ obtained with the channel coding of prior slide
Characterization of RVs

Binomial Distribution

An Application

- Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding
- If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?
- Answer:
  - Probability of receiving one bit correctly: $1 - p = 0.999$
  - Recall: If $X_1$ and $X_2$ are independent binary random variables, $P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$
  - Therefore, probability of receiving two bits correctly: $(1 - p)^2 = (0.999)^2$
  - Probability of receiving a total of $i$ bits correctly: $(1 - p)^i = (0.999)^i$
  - Thus, probability of receiving all 31 bits correctly: $(1 - p)^{31} = (0.999)^{31}$
  - Thus, probability of making a codeword error: $1 - (0.999)^{31} = 0.0305$
- Compare this against error probability of $3 \times 10^{-8}$ obtained with the channel coding of prior slide
Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding.

If probability of an individual bit in error is $p = 0.01$, what is the probability that the codeword is received in error?

**Answer:**

- Probability of receiving one bit correctly: $1 - p = 0.999$
- Recall: If $X_1$ and $X_2$ are independent binary random variables, $P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0)$
- Therefore, probability of receiving two bits correctly: $(1 - p)^2 = (0.999)^2$
- Probability of receiving a total of $i$ bits correctly: $(1 - p)^i = (0.999)^i$
- Thus, probability of receiving all 31 bits correctly: $(1 - p)^{31} = (0.999)^{31}$
- Thus, probability of making a codeword error: $1 - (0.999)^{31} = 0.0305$

Compare this against error probability of $3 \times 10^{-8}$ obtained with the channel coding of prior slide.
Suppose we transmit a 31 bit sequence (i.e., a codeword) without any channel coding.

If probability of an individual bit in error is \( p = 0.01 \), what is the probability that the codeword is received in error?

**Answer:**

- Probability of receiving one bit correctly: \( 1 - p = 0.999 \)
- Recall: If \( X_1 \) and \( X_2 \) are independent binary random variables,
  \[ P_{X_1X_2}(x_1 = 0, x_2 = 0) = P_{X_1}(x_1 = 0)P_{X_2}(x_2 = 0) \]
- Therefore, probability of receiving two bits correctly:
  \[ (1 - p)^2 = (0.999)^2 \]
- Probability of receiving a total of \( i \) bits correctly: \( (1 - p)^i = (0.999)^i \)
- Thus, probability of receiving all 31 bits correctly:
  \[ (1 - p)^{31} = (0.999)^{31} \]
- Thus, probability of making a codeword error:
  \[ 1 - (0.999)^{31} = 0.0305 \]

Compare this against error probability of \( 3 \times 10^{-8} \) obtained with the channel coding of prior slide.
Suppose we want to simulate Bit Error Rate (BER) of a digital communication system on a computer.

In a typical approach, we would run, say, \( N = 100000 \) independent simulation experiments or trials on the computer and count the total number of bits in error in \( N \) simulation trial. Say, this number \( K \) is equal to 100.

Therefore, our estimate of BER is \( \frac{K}{N} = 10^{-3} \). We want to know how accurate this estimate is.

We model estimated BER as \( Y = \frac{1}{N} \sum_{i=1}^{N} X_i \), where \( \{X_i\}; i = 1, \ldots, N \) are independent binary RVs (outcome of each simulation trial) with

\[
p(x) = \begin{cases} 
1 - p, & x = 0; \text{bit correctly received}, \\
p, & x = 1; \text{received bit in error}
\end{cases}
\]

It is seen that the BER is a Binomial RV. We determine its mean and variance. This tells us whether the mean is close to the actual bit error probability of \( p \) and how much deviation is likely from this mean.
Chebyshev’s Inequality

- For any RV $X$ with mean $m_x$ and variance $\sigma_x^2$, following holds for any given $\delta$:
  
  $$P(|X - m_x| \geq \delta) \leq \frac{\sigma_x^2}{\delta^2}$$

- This is Chebyshev’s Inequality. It shows that the variance is a key parameter that determines a random variable’s likelihood to remain close to its mean value.
Chebyshev’s Inequality

An Application

Earlier we saw that we can model the estimated BER (in an simulation based experiment of digital communication system) as a Binomial RV. Given the number $N$ of simulation trials and the actual probability $p$ of bit in error, we can determine the mean and the variance of the estimated BER.

- Mean of estimated BER: $p$
- Variance of estimated BER: $\frac{p(1-p)}{N}$

Chebyshev’s Inequality can be used to determine the likelihood that the estimated BER is close to the mean value.
Random Process

- An RV has a single value. However, in many cases, the random phenomenon changes with time.
- RVs model unknown events.
- Random processes model unknown signals.
- A random process can be thought of as a collection of RVs.
- If $X(t)$ is a random process, $X(1), X(1.5), X(3.72)$ are all random variables obtained at different values of time instant $t$. 

Random Process
Characterization

- Knowing the PDFs of individual samples of a random process is not sufficient to characterize it fully. We need also to know how these samples are related to each other.
- Two mathematical concepts toward this purpose:
  - Autocorrelation Function or ACF
  - Power Spectral Density or PSD
Random Process

Autocorrelation Function

- Autocorrelation Function measures how a random process changes with time.
- Intuitively, $X(1)$ and $X(1.1)$ are likely to be more strongly related compared to $X(1)$ and $X(1000)$.
- Definition of ACF: $\phi(t, \tau) = E[X(t)X(t + \tau)]$
- Note that the power is given as $\phi(t, \tau = 0)$.
Random Process

Different Types

- Stationary Random Process has statistical properties which do not change with time.
- Stationarity is a mathematical idealization. Can be tough to prove that a given random process is stationary.
- A Wide Sense Stationary (WSS) process has mean and autocorrelation functions which do not change with time.
- An ergodic random process has the time average which is identical to the statistical average.
- Typically the random process is assumed to be WSS and ergodic.
- For WSS process, the dependence of ACF $\phi(t, \tau)$ on time variable $t$ is not present and ACF is given as $\phi(\tau)$. 
Random Process
Power Spectral Density

- $\Phi(f)$ tells us how much power is present at a given frequency $f$
- Wiener-Khinchine Theorem: $\Phi(f) = \mathcal{F}\{\phi(\tau)\}$
  - Applies to WSS random processes. Shows that the PSD and ACF for a WSS random process are Fourier Transform pairs
- Properties of PSD:
  1. $\Phi(f) \geq 0$
  2. $\Phi(f) \geq \Phi(-f)$
  3. Power $= \int_{-\infty}^{\infty} \Phi(f) \, df$
Consider a linear system with impulse response $h(t)$ that is excited by a deterministic (non-random) input $x(t)$. Let the deterministic output of the system be denoted as $y(t)$. Since the system is linear,

→ In time domain, the output is the convolution of impulse response with the input: $y(t) = h(t) * x(t)$
→ In frequency domain, the output is the product of the Fourier Transforms of the impulse response and the input: $Y(f) = H(f) \times X(f)$

When the input is stochastic, the above deterministic relations are not readily applicable. Instead we need a mathematical relation between the statistical properties of the input and the output. This is given as follows:

$$\phi_Y(\tau) = h(-\tau) \ast \phi_X(\tau) \ast h(\tau)$$
$$\Phi_Y(f) = |H(f)|^2 \Phi_X(f)$$
A type of random processes, called Gaussian Random Process, is such that any sample point from this random process is a Gaussian RV. It has two characteristics that make it special:

1. If a Gaussian random process is WSS, it is also stationary.
2. If the input to a linear system is a Gaussian random process, the output is also a Gaussian random process.