

# Introduction to Communication Systems

## The Theory of Probability

### Probability Theory

#### Why Probability

Probability theory is applied to situations where *uncertainty exists*. These situations include

1. The characterization of traffic at the intersection US 460 and Peppers Ferry Road (i.e., the number of cars that cross the intersection as a function of time)
2. The prediction of the weather in Blacksburg
3. The number of students traversing the Drill Field between 8:50 and 9:00 on Mondays
4. The thermally induced (Brownian) motion of molecules in a
  - (a) copper wire
  - (b) a JFET amplifier

Note that the latter two situations would resort in a phenomenon known as noise whose power would be a function of the Temperature of the molecules.

In all of these situations, we could develop an excellent approximation or prediction for each of these *experiments*. However, we would never be able to characterize them with absolute certainty (i.e., *deterministically*). However, we could characterize them in a probabilistic fashion (or *probabilistically*) via the elements of *probability theory*. This is what weather forecasters are doing when they say

The chance of rain tomorrow is 40%

In such cases they are saying that the probability of a *rain even* is 0.4. Other probabilistic characterizations include

- One the average 300 cars per minute cross US 460 and Peppers Ferry Road at Noon on Saturdays. The chances of so many cars crossing that intersection at 8:00 AM on Sunday is very small
- The average noise power produced by this amplifier is  $1\mu W$

Scientists and Engineers apply the theories of Probability and Random Processes to those repeating situations in nature where

1. We can *roughly predict* what may happen.
2. We cannot *exactly determine* what may happen

Whenever we cannot exactly predict an occurrence, we say that such an occurrences is *random*. Random occurrences occur for the following reasons

- All the causal forces at work are unknown.
- Insufficient data for the conditions of the problem do not exist
- The physical mechanisms driving the problem are so complicated that the direct calculation of the problem is not feasible
- There exists some basic indeterminacy in the physical world

#### The Notions of Probability

One can approach probability through an abstract mathematical concept called *measure theory*, which results in the *axiomatic theory of probability*, or through heuristic approach called *relative frequency*, which is a less complete (and slightly flawed) definition of probability. However, it suits our need for this course. Student's that continue on to graduate studies will be introduced to the more abstract but powerful axiomatic theory. Before continuing it is necessary to define the following important terms

**Definition 1** An *experiment* is a set of rules that governs a specific operation that is being performed

**Definition 2** A *trial* is the performance or exercise of that *experiment*

**Definition 3** An *outcome* is the result of a given *trial*

**Definition 4** An *event* is an outcome or any combination of outcomes

**Example 1** Consider the experiment of selecting at random one card from a deck of 52 playing cards and writing down the number and suit of that card. Notice that the rules are well defined

1. We have a deck of cards
2. Somebody selects the card
3. The records the result

Suppose somebody decided to perform the experiment. We would then say that he/she conducted a trial. The result of that trial could have been the 3 of spades (or  $3\spadesuit$ ). Hence the outcome would be:  $3\spadesuit$ . Another outcome could have been  $10\heartsuit$  or  $J\clubsuit$ . Indeed there are as many as **52 possible outcomes**.

An event is a collection of possible outcomes. So  $3\spadesuit$  is an event. However,  $10\heartsuit$  is also an event, and  $J\clubsuit$  is an event as well. Further  $\{3\spadesuit, J\clubsuit\}$ ,  $\{3\spadesuit, 10\heartsuit\}$ , and  $\{2\heartsuit, 5\clubsuit, K\heartsuit, A\clubsuit\}$  are also events. **Any combination of possible outcomes is an event**. Note that for this experiment, there are  $2^{52} = 4.50359962737 \times 10^{15}$  different events

## The Axiomatic Theory of Probability

This is actually an application of a mathematical theory called **Measure Theory**. Both theories apply basic concepts from **set theory**.

The **axiomatic theory of probability** is based on a triplet

$$(\Omega, \mathfrak{S}, P)$$

where

- $\Omega$  is the *sample space*, which is the set of all possible outcomes
- $\mathfrak{S}$  is the *sigma algebra* or *sigma field*, which is the set of all possible events (or combination of outcomes)
- $P$  is the *probability function*, which can be **any set function**, whose **domain** is  $\Omega$  and the **range** is the closed unit interval  $[0, 1]$ , i.e., any number between 0 and 1 including 0 and 1. The only requirement for this function is that it must obey the following three rules
  - (a)  $P[\Omega] = 1$
  - (b) Let  $A$  be any event in  $\mathfrak{S}$ , then  $P[A] \geq 0$
  - (c) Let  $A$  and  $B$  be two events in  $\mathfrak{S}$  such that  $A \cap B = \phi$ , then  $P[A \cup B] = P[A] + P[B]$

## Relative Frequency Definition.

The relative frequency approach is based on the following definition: Suppose we conduct a large number of trials of this a given experiment. The probability that a given event, say  $\mathcal{A}$ , is the following limit

$$P[\mathcal{A}] = \lim_{n \rightarrow \infty} \frac{n_{\mathcal{A}}}{n} \tag{1}$$

where  $n_{\mathcal{A}}$  is the number of occurrences and  $n$  is the number of trials.

For example suppose we conduct the above experiment 10,000 times ( $n = 10000$ ). Further suppose the event  $\mathcal{A} = \{3\spadesuit\}$  occurred 188 times ( $n_{\mathcal{A}} = 188$ ). Then

$$\frac{n_{\mathcal{A}}}{n} = \frac{188}{10000} = 0.0188$$

As  $n$  increased to infinity (and assuming that the cards are fair), then the ratio would approach the probability of  $\mathcal{A}$

$$P[\mathcal{A}] = \lim_{n \rightarrow \infty} \frac{n_{\mathcal{A}}}{n} = \frac{1}{52} = 0.0196$$

Suppose the event  $\mathcal{B}$  was the set of all spades. Then

$$\mathcal{B} = \{A\spadesuit, 2\spadesuit, 3\spadesuit, 4\spadesuit, 5\spadesuit, 6\spadesuit, 7\spadesuit, 8\spadesuit, 9\spadesuit, 10\spadesuit, J\spadesuit, Q\spadesuit, K\spadesuit\}$$

Now we say that **The event  $\mathcal{B}$  occurs whenever an outcome is of a trial is contained in  $\mathcal{B}$ .** In this case, whenever the outcome of a trial contains a  $\spadesuit$ , the outcome belongs to  $\mathcal{B}$  and we say that  $\mathcal{B}$  occurred. So, suppose we conducted the above experiment 10,000 times, and the event  $\mathcal{B}$  occurred 3000 times. Then

$$\frac{n_{\mathcal{B}}}{n} = \frac{3000}{10000} = 0.3$$

As  $n$  increased to infinity (and assuming that the cards are fair), then the ratio would approach the probability of  $\mathcal{B}$

$$P[\mathcal{B}] = \lim_{n \rightarrow \infty} \frac{n_{\mathcal{B}}}{n} = \frac{1}{4} = 0.25$$

Note that this definition makes sense. Since  $P[\mathcal{B}] = \frac{1}{4}$ , it follows that the chance of  $\mathcal{B}$  occurring in any given trial is 25%. Similarly, the chance of a  $4\clubsuit$  occurring from a particular trial is  $(1/52) * 100$  or 1.9%. However, note that the **probability of any event will always be a number between 0 and 1.**

Now, suppose the event  $\mathcal{C}$  consists of the set of all diamonds. Then

$$\mathcal{C} = \{A\diamond, 2\diamond, 3\diamond, 4\diamond, 5\diamond, 6\diamond, 7\diamond, 8\diamond, 9\diamond, 10\diamond, J\diamond, Q\diamond, K\diamond\}$$

Let's look at both  $\mathcal{B}$  and  $\mathcal{C}$ . Note that none of the members in  $\mathcal{B}$  belong to  $\mathcal{C}$  and none of the members in  $\mathcal{C}$  belong to  $\mathcal{B}$ . We say that  $\mathcal{B}$  and  $\mathcal{C}$  are *disjoint* and that their *intersection* is the *empty set*

$$\mathcal{B} \cap \mathcal{C} = \phi$$

In this case the probability of either  $\mathcal{B}$  or  $\mathcal{C}$  occurring is equal to

$$P[\mathcal{B} \text{ or } \mathcal{C}] = \lim_{n \rightarrow \infty} \frac{n_{\mathcal{B}} + n_{\mathcal{C}}}{n} = P[\mathcal{B} \cup \mathcal{C}] \quad (2)$$

where the  $\cup$  symbol stands for the *union* of two sets. Note that (2) describes the additive concept of probability: **If two events are disjoint, then the probability of their sum is the sum of their probabilities.**

Finally, let the event  $\Omega$  equal all possible outcome. We call  $\Omega$  the *certain event* or the *sample space*. Since every possible outcome is contained in this event,  $n_{\Omega} = n$ . Therefore

$$P[\Omega] = \lim_{n \rightarrow \infty} \frac{n_{\Omega}}{n} = \lim_{n \rightarrow \infty} \frac{n}{n} = 1 \quad (3)$$

**Example 2** Consider the experiment of flipping a fair coin twice and observing the output. What are the possible outcomes? List all possible events. Assume that all outcomes are equally likely, assign probabilities to each event.

The set of all possible outcomes is

$$\Omega = \{HH, HT, TH, TT\}$$

There are  $2^4 = 16$  events. They are

$$\begin{aligned} &\{HH\}, \{HT\}, \{TH\}, \{TT\} \\ &\{HH, HT\}, \{HH, TH\}, \{HH, TT\}, \{HT, TH\}, \{HT, TT\}, \{TH, TT\} \\ &\{HH, HT, TH\}, \{HH, HT, TT\}, \{HH, TH, TT\}, \{HT, TH, TT\} \\ &\{HH, HT, TH, TT\}, \phi \end{aligned}$$

Note that  $\{HH, HT, TH, TT\}$  is the certain event. Therefore

$$P[\{HH, HT, TH, TT\}] = 1 \quad (4)$$

However, from the additive property, it follows that

$$P[\{HH, HT, TH, TT\}] = P[\{HH\}] + P[\{HT\}] + P[\{TH\}] + P[\{TT\}] \quad (5)$$

but since each outcome is equally likely we have that

$$P[\{HH\}] = P[\{HT\}] = P[\{TH\}] = P[\{TT\}] \quad (6)$$

Therefore from (5) and (6), it follows that

$$P[\{HH\}] = P[\{HT\}] = P[\{TH\}] = P[\{TT\}] = \frac{1}{4} \quad (7)$$

Using the additive concept of probability we can see that

$$\begin{aligned} P[\{HH, HT\}] &= P[\{HH, TH\}] = P[\{HH, TT\}] = \\ &= P[\{HT, TH\}] = P[\{HT, TT\}] = P[\{TH, TT\}] = \frac{1}{2} \end{aligned} \quad (8)$$

$$P[\{HH, HT, TH\}] = P[\{HH, HT, TT\}] = P[\{HH, TH, TT\}] = P[\{HT, TH, TT\}] = \frac{3}{4} \quad (9)$$

$$P[\phi] = 0 \quad (10)$$

## Joint Probability

Let the events  $\mathcal{D}$  and  $\mathcal{E}$  be defined as follows

$$\begin{aligned} \mathcal{D} &= \{A\spadesuit, 2\clubsuit, 8\diamond, 4\heartsuit\} \\ \mathcal{E} &= \{4\heartsuit, K\spadesuit, 4\heartsuit, 2\clubsuit\} \end{aligned}$$

Note that

$$\begin{aligned} P[\mathcal{D}] &= \frac{4}{52} = \frac{1}{13} \\ P[\mathcal{E}] &= \frac{4}{52} = \frac{1}{13} \end{aligned}$$

Then the *intersection* of  $\mathcal{D}$  and  $\mathcal{E}$  is **another** event that contains the elements of  $\mathcal{D}$  and  $\mathcal{E}$

$$\mathcal{D} \cap \mathcal{E} = \{4\heartsuit, 2\clubsuit\}$$

and

$$P[\mathcal{D} \cap \mathcal{E}] = P[\mathcal{D}, \mathcal{E}] = P[\mathcal{D} \text{ and } \mathcal{E}] = \frac{1}{26} \quad (11)$$

We call  $P[\mathcal{D} \cap \mathcal{E}] = P[\mathcal{D}, \mathcal{E}] = P[\mathcal{D} \text{ and } \mathcal{E}]$  the *joint probability* of  $\mathcal{D}$  and  $\mathcal{E}$ . It describes the probability of **both** events occurring

## Conditional Probability

In a number of cases, knowledge about one event provides us additional information about the occurrence of another event. Suppose that we conduct an experiment and we find out that  $\mathcal{D}$  has occurred. Thus we know that the outcome was either a  $A\spadesuit$ ,  $2\clubsuit$ ,  $8\diamond$ , or  $4\heartsuit$ . Does this tell us anything about the occurrence of  $\mathcal{E}$ ? The answer is yes.

Given that  $\mathcal{D}$  has occurred, we know that the outcome was either  $A\spadesuit$ ,  $2\clubsuit$ ,  $8\diamond$ , or  $4\heartsuit$ . Since each outcome is equally likely, and we know that these are now the only possibilities, we have that the probability of each of these outcomes **given** that  $\mathcal{D}$  has occurred is  $\frac{1}{4}$  in other words

$$P[A\spadesuit|\mathcal{D}] = P[2\clubsuit|\mathcal{D}] = P[8\diamond|\mathcal{D}] = P[4\heartsuit|\mathcal{D}] = \frac{1}{4} \quad (12)$$

Note that  $P[A|B]$  means the probability of  $A$  given the  $B$  has occurred

Now look at the event  $\mathcal{E} = \{4\heartsuit, K\spadesuit, 4\heartsuit, 2\clubsuit\}$ . Given that  $\mathcal{D}$  has occurred and that the possible outcomes were only  $A\spadesuit$ ,  $2\clubsuit$ ,  $8\diamond$ , or  $4\heartsuit$ , we can conclude that the outcomes  $4\heartsuit$  and  $K\spadesuit$  could not have happened because  $4\heartsuit$  and  $K\spadesuit$  are **not in**  $\mathcal{D}$ . Therefore

$$P[4\heartsuit|\mathcal{D}] = P[K\spadesuit|\mathcal{D}] = 0 \quad (13)$$

Therefore, given that we know  $\mathcal{D}$  has occurred, we can write the probability of  $\mathcal{E}$  as follows

$$\begin{aligned} P[\mathcal{E}|\mathcal{D}] &= P[4\heartsuit|\mathcal{D}] + P[K\spadesuit|\mathcal{D}] + P[8\diamond|\mathcal{D}] + P[4\heartsuit|\mathcal{D}] \\ &= 0 + 0 + \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned} \quad (14)$$

We can also compute the probability of  $\mathcal{E}$  given that  $\mathcal{D}$  has occurred using the following definition.

**Definition 5** Let  $A$  and  $B$  be two events from the same experiment. Then the *conditional probability of  $A$  given  $B$* ,  $P[A|B]$  is defined as follows

$$P[A|B] \triangleq \frac{P[A, B]}{P[B]} \tag{15}$$

If  $P[B] = 0$ , then  $P[A|B]$  is *undefined*.

So, for the case of  $\mathcal{E}$  given that  $\mathcal{D}$ , we have that

$$P[\mathcal{E}|\mathcal{D}] = \frac{P[\mathcal{E}, \mathcal{D}]}{P[\mathcal{D}]} = \frac{1/26}{1/13} = \frac{13}{26} = \frac{1}{2}$$

Similarly for the case of  $\mathcal{A}$  given  $\mathcal{C}$  above we have that

$$P[\mathcal{A}|\mathcal{C}] = \frac{P[\mathcal{A}, \mathcal{C}]}{P[\mathcal{C}]} = \frac{0}{0.25} = 0$$

Another important concept in probability is *independence*.

**Definition 6** Let  $A$  and  $B$  be two events from the same experiment. Then  $A$  and  $B$  are said to be *independent* if the joint probability of  $A$  and  $B$  is equal to the **product** of their two individual probabilities.

$$P[A, B] = P[A] P[B] \tag{16}$$

From this it follows that

$$P[A|B] = \frac{P[A, B]}{P[B]} = \frac{P[A] P[B]}{P[B]} = P[A] \tag{17}$$

This implies if two events are *independent*, the **knowledge about one event provides absolutely know information about the other event**.

## Random Variables

You may have noticed thus far that we have working with sets and events that contain odd symbols such as  $3\spadesuit, 4\clubsuit, 5\diamond, 6\heartsuit, \dots$ . Other examples, both in this note set and the book, have worked with events with other symbols such as  $\{HH\}$  for the coin problem and  $\{\cdot\cdot\}$  for the dice problem. We would like to perform some mathematical analysis and manipulation on these events. However, such a task would be difficult and would not provide any insight.

It would be better if we could *assign* (or *map*) each outcome a number. Then we could work with these numbers using the standard mathematical techniques that we have learned over the years.

**Example 3** Consider our card experiment. Suppose we assign each card an integer as follows

$$\begin{aligned} A\spadesuit &\rightarrow 1 \\ 2\spadesuit &\rightarrow 2 \\ &\vdots \\ K\spadesuit &\rightarrow 13 \\ A\clubsuit &\rightarrow 14 \\ 2\clubsuit &\rightarrow 15 \\ &\vdots \\ K\clubsuit &\rightarrow 26 \\ A\diamond &\rightarrow 27 \\ 2\diamond &\rightarrow 28 \\ &\vdots \\ K\diamond &\rightarrow 39 \\ A\heartsuit &\rightarrow 40 \\ 2\heartsuit &\rightarrow 41 \\ &\vdots \\ K\heartsuit &\rightarrow 52 \end{aligned}$$

Then each card is assigned an integer on the real line. Suppose we assign this mapping the function  $X(\zeta)$ , where  $\zeta$  is any outcome

Such a mapping is called a *random variable*.

**Definition 7** A random variable,  $X(\zeta)$  is a **deterministic function** that assigns each outcome,  $\zeta \in \Omega$ , in an experiment, a number.

$$X : \Omega \rightarrow \mathfrak{R} \quad (18)$$

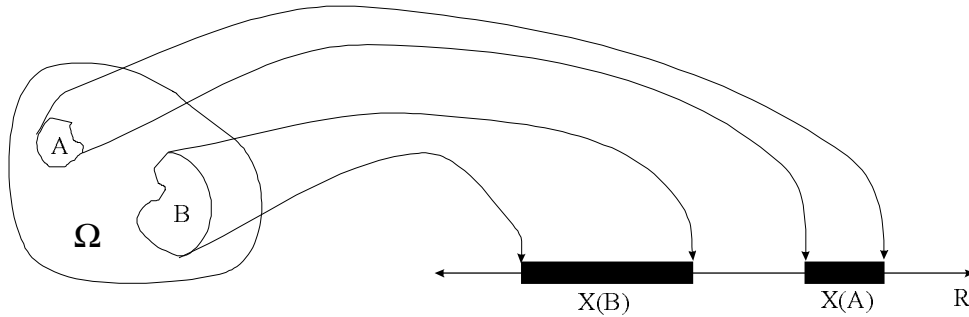
So that  $P[X(\zeta) = \infty] = P[X(\zeta) = -\infty] = 0$

The first thing you can notice about a random variable is that

1. There is nothing random about it
2. It is not a variable. Rather it is a set function.

However, in performing analysis, it is convenient to treat  $X(\zeta)$  as a variable. Consider a second example of a random variable. since the random variable is a function of an outcome, which has an associated probability, the random variable also has an associated probability.

$$P[X(\zeta) = x] = P[\zeta = X^{-1}(x)] \quad (19)$$



**Example 4** Consider the same card experiment. Suppose we define a random variable  $S(\zeta)$  such that

$$\begin{aligned} \text{if } \zeta \text{ contains a } \spadesuit : S(\zeta) &= 1 \\ \text{if } \zeta \text{ contains a } \clubsuit : S(\zeta) &= 2 \\ \text{if } \zeta \text{ contains a } \diamondsuit : S(\zeta) &= 3 \\ \text{if } \zeta \text{ contains a } \heartsuit : S(\zeta) &= 4 \end{aligned}$$

Note that the outcome can be assigned the same number. The function defined by the random variable can be **any** set function provided that the associated probabilities as the  $\pm\infty$  is zero

What about the probabilities

A random variable is a function of the outcomes of an experiment. Therefore, since each outcome has a probability, the number *assigned to that function by a random variable*, also has a probability. The probability of a given random variable is determined as follows.

$$P[X(\zeta) = x] = P[\zeta = X^{-1}(x)] \quad (20)$$

where  $X^{-1}(x)$  is a mapping from the real line,  $\mathfrak{R}$ , to the set of all outcomes,  $\Omega$ .

**Example 5** Consider the random variable  $X(\zeta)$  defined in Example 3

$$\begin{aligned} P[X(\zeta) = 1] &= P[\zeta = X^{-1}(1)] = P[\zeta = A\spadesuit] = \frac{1}{52} \\ P[X(\zeta) = 26] &= P[\zeta = X^{-1}(26)] = P[\zeta = K\clubsuit] = \frac{1}{52} \\ P[X(\zeta) = 28.9] &= P[\zeta = X^{-1}(28.9)] = P[\phi] = 0 \\ P[X(\zeta) = 0] &= P[\zeta = X^{-1}(0)] = P[\phi] = 0 \\ P[X(\zeta) \leq 2] &= P\{\zeta = X^{-1}(-\infty, 2]\} = P[\zeta = \{A\spadesuit, 2\spadesuit\}] = \frac{2}{52} \end{aligned}$$

**Example 6** Consider the random variable  $S(\zeta)$ , defined in Example 4

$$\begin{aligned}
 P[S(\zeta) = 1] &= P[\zeta = S^{-1}(1)] \\
 &= P[\zeta = \{A\spadesuit, 2\spadesuit, 3\spadesuit, 4\spadesuit, 5\spadesuit, 6\spadesuit, 7\spadesuit, 8\spadesuit, 9\spadesuit, 10\spadesuit, J\spadesuit, Q\spadesuit, K\spadesuit\}] = \frac{1}{4} \\
 P[S(\zeta) = 2] &= P[\zeta = S^{-1}(2)] = \frac{1}{4} \\
 P[S(\zeta) = 2.5] &= P[\zeta = S^{-1}(2.5)] = P[\phi] = 0 \\
 P[2 \leq S(\zeta) \leq 3] &= P\{\zeta = S^{-1}[2, 3]\} = P[\zeta = \{\text{all } \clubsuit\text{'s and all } \heartsuit\}] = \frac{1}{2}
 \end{aligned}$$

## Cumulative Distribution Functions and Probability Density Functions

Two ways to analyze experiments via random variables is through the Cumulative Distribution Function (CDF) and probability density function (pdf)

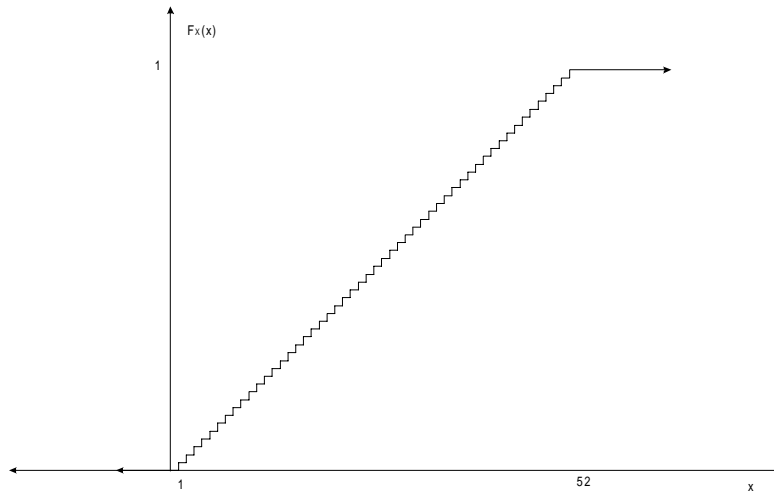
**Definition 8** Let  $X(\zeta)$  be a random variable defined on a sample space  $\Omega$ . Then the *Cumulative Distribution Function (CDF)* of  $X$  is denoted by the function  $F_X(x_0)$  and is defined as follows

$$F_X(x_0) = P[X \leq x_0] \quad (21)$$

where the set  $[X \leq x_0]$  defines an event or a collection of outcomes.

**Example 7** Consider the random variable in Example 3. The CDF of  $X$  is

$$F_X(x_0) = P[X \leq x_0] = \frac{1}{52} \sum_{k=1}^{52} u(x_0 - k)$$

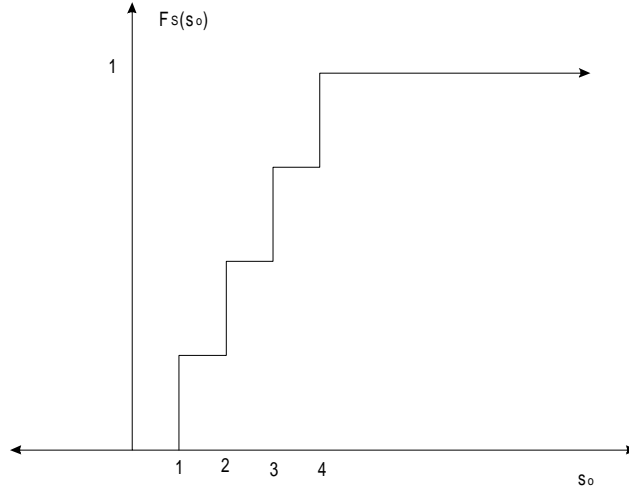


**Example 8** Consider the random variable in Example 4. The CDF of  $S$  is

$$F_S(s_0) = \frac{1}{4}u(s_0 - 1) + \frac{1}{4}u(s_0 - 2) + \frac{1}{4}u(s_0 - 3) + \frac{1}{4}u(s_0 - 4) \quad (22)$$

So

$$F_S(s_0) = \begin{cases} 0 & s_0 < 1 \\ \frac{1}{4} & 1 \leq s_0 < 2 \\ \frac{2}{4} & 2 \leq s_0 < 3 \\ \frac{3}{4} & 3 \leq s_0 < 4 \\ 1 & s_0 \geq 4 \end{cases} \quad (23)$$



**Definition 9** Let  $X(\zeta)$  be a random variable defined on a sample space  $\Omega$ . Then the *Probability Density Function (pdf)* of  $X$  is denoted by the symbol  $f_X(x)$  and is defined as the derivative of its CDF

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} P[X(\zeta) \leq x] \quad (24)$$

Note that we can recover the CDF from the PDF through integration

$$F_X(x) = \int_{-\infty}^x f_X(\alpha) d\alpha \quad (25)$$

Therefore  $f_X(x)$  will always integrate to 1

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (26)$$

Furthermore, the probability that  $X(\zeta)$  is between  $x_1$  and  $x_2$ ,  $x_2 > x_1$  and  $x_2$  inclusive, can be computed as follows

$$P[x_1 < X \leq x_2] = P[X \leq x_2] - P[X \leq x_1] = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x) dx \quad (27)$$

Note that the argument,  $\zeta$ , has been dropped from the random variable,  $X$ . For the remainder of this discussion, we will assume that the dependence of a random variable  $X$  on  $\zeta$  is implicit and omit the argument for convenience

## Moments of Random Variables

Moments of random variables include a number of quantities such as *averages*, *variances*, *standard deviations*, etc. They are particularly useful in communications because they provide valuable information about a random variable with having to know its statistics (i.e., the CDF or pdf)

**Definition 10** Let  $X$  be a random variable with the pdf  $f_X(x)$ . Then the *expected value* or *average* of  $X$ , is

$$X_{avg} = E[X] \triangleq \int_{-\infty}^{\infty} x f_X(x) \quad (28)$$

**Definition 11** Let  $X$  be a random variable with pdf  $f_X(x)$ . Then, the *variance* of  $X$  is

$$\sigma_X^2 = E[(X - X_{avg})^2] \triangleq \int_{-\infty}^{\infty} (x - X_{avg})^2 f_X(x) dx \quad (29)$$

and the *standard deviation* of  $X$  is

$$\sigma_X = \sqrt{E[(X - X_{avg})^2]} \triangleq \sqrt{\int_{-\infty}^{\infty} (x - X_{avg})^2 f_X(x) dx} \quad (30)$$



**Definition 12** Let  $X$  be a random variable with pdf  $f_X(x)$ . Then, the  $n^{\text{th}}$  moment of  $X$  is

$$m_X^n = \int_{-\infty}^{\infty} x^n f_X(x) dx \quad (31)$$

and the  $n^{\text{th}}$  central moment of  $X$  is

$$\sigma_X^n = \int_{-\infty}^{\infty} (x - X_{avg})^n f_X(x) dx \quad (32)$$

Finally, let  $X$  be a random variable and let  $g(X)$  be some function of that random variable. In most cases this  $g(X)$  will transform  $X$  into a new random variable, which we will call  $Y$

$$Y = g(X) \quad (33)$$

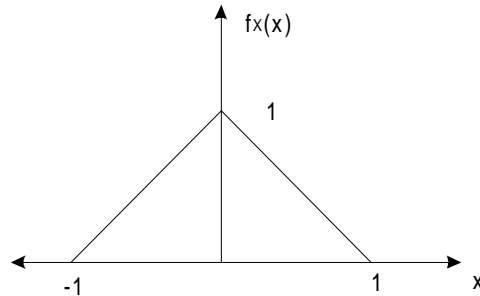
which has its own CDF and pdf, which must be found through the use of random variable transformation techniques. However, if we are interested in the *moments* of  $Y$ , such as its *average*, then knowledge about  $Y$ 's CDF or pdf are not required. Rather we can just use the already known CDF and pdf of  $X$ .

**Theorem 1** Let  $X$  be a random variable with pdf  $f_X(x)$ . Let  $Y = g(X)$  be some transformation of  $X$ . Then the expected value or average of  $Y$  is

$$Y_{ave} = E[Y] = E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (34)$$

**Example 9** Let  $X$  be a random variable with the following pdf

$$f_X(x) = \begin{cases} x+1 & -1 \leq x \leq 0 \\ 1-x & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (35)$$



$$X_{avg} = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-1}^0 x(x+1) dx + \int_0^1 x(1-x) dx = 0$$

$$\sigma_X^2 = E[(X - X_{avg})^2] = E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{-1}^0 x^2(x+1) dx + \int_0^1 x^2(1-x) dx = \frac{2}{3}$$

$$E[X + \pi] = \int_{-\infty}^{\infty} (x + \pi) f_X(x) dx = \int_{-1}^0 (x + \pi)(x+1) dx + \int_0^1 (x + \pi)(1-x) dx = \pi$$

## Functions of a Random Variable

**Definition 1** Consider a function of a random variable,  $g(X)$ . The result is another random variable, which we could call  $Y$ ,  $Y = g(X)$ . However the CDF and pdf of  $Y$  will be different than the CDF and pdf of  $X$ . In fact we will use information about the  $F_X(x)$ ,  $f_X(x)$ , and  $g(X)$  to determine  $F_Y(y)$  and  $f_Y(y)$

The distribution of  $Y = g(X)$

The PDF of the random variable,  $Y$ , is nothing more than the probability of the event that  $\{\mathbf{Y} \leq y\}$ . This consists of all outcomes  $\zeta \in \Omega$  such that  $\mathbf{Y}(\zeta) \leq y$ . This is **equivalent** to the set of all outcomes  $\zeta \in \Omega$  such that  $g[X(\zeta)] \leq y$ , and this is **equivalent** to the set of all outcome  $\zeta \in \Omega$  such that  $X(\zeta) \leq g^{-1}(y)$ . So

$$F(y) = P(\{\mathbf{Y} \leq y\}) = P(\{g(\mathbf{X}) \leq y\}) = P(\{\mathbf{X} \leq g^{-1}(y)\}) \quad (36)$$

The Probability Density Function (pdf) of  $Y = g(X)$

Once the PDF of  $Y$  is computed, the computation of the pdf of  $Y$  is straightforward

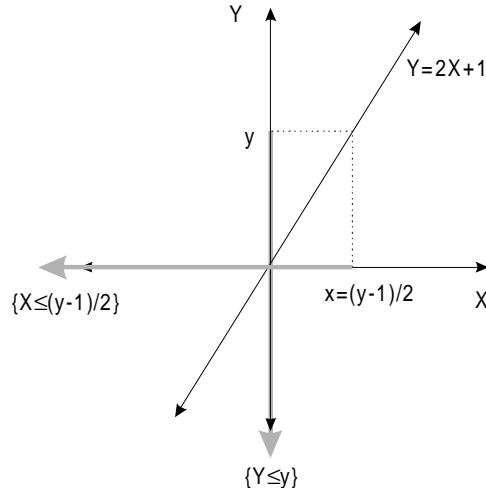
$$f_Y(y) = \frac{d}{dy} F_Y(y) \quad (37)$$

Examples

**Example 10** Let  $X$  be a random variable. Let

$$Y = g(X) = 2X + 1$$

Find  $F_Y(y)$  and  $f_Y(y)$ . To find  $F_Y(y)$ , we must evaluate the event  $\{Y \leq y\}$  where  $Y$  is the random variable and  $y$  is simply a number. Consider the following graph



The shaded vertical line corresponds to the event  $\{\zeta : Y(\zeta) \leq y\}$  and the horizontal line corresponds to the event  $\{\zeta : X(\zeta) \leq \frac{1}{2}(y-1)\}$ . Both events have the exact same outcomes. Therefore

$$\{\zeta : Y(\zeta) \leq y\} = \left\{ \zeta : X(\zeta) \leq \frac{1}{2}(y-1) \right\}$$

Therefore

$$\{Y \leq y\} = \{2X + 1 \leq y\} = \left\{ X \leq \frac{1}{2}(y-1) \right\} = \{X \leq g^{-1}(y)\}$$

Note that the inverse function  $g^{-1}(Y)$  is

$$g^{-1}(Y) = \frac{1}{2}(X-1)$$

So

$$F_Y(y) = P[Y \leq y] = P[2X + 1 \leq y] = P\left[X \leq \frac{1}{2}(y-1)\right] = F_X\left(\frac{1}{2}(y-1)\right)$$

Now suppose  $X$  was uniformly distributed between 0 and 1. Then

$$f_X(x) = u(x) - u(x-1) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (38)$$

$$F_X(x) = x[u(x) - u(x-1)] = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases} \quad (39)$$

We can now plot  $F_Y(y)$ . Let's plug in some values

$$\begin{aligned} y = 0 & F_Y(0) = F_X\left(\frac{1}{2}(0-1)\right) = F_X\left(-\frac{1}{2}\right) = 0 \\ y = 1 & F_Y(1) = F_X\left(\frac{1}{2}(1-1)\right) = F_X(0) = 0 \\ y = 2 & F_Y(2) = F_X\left(\frac{1}{2}(2-1)\right) = F_X\left(\frac{1}{2}\right) = \frac{1}{2} \\ y = 3 & F_Y(3) = F_X\left(\frac{1}{2}(3-1)\right) = F_X(1) = 1 \end{aligned} \quad (40)$$

So

$$F_Y(y) = \begin{cases} 0 & y < 1 \\ \frac{1}{2}(y-1) & 1 \leq y \leq 3 \\ 1 & y > 3 \end{cases} \quad (41)$$

The density of  $Y$ ,  $f_Y(y)$ , is simply

$$f_Y(y) = \frac{d}{dy} F_Y(y) \quad (42)$$

and from (41) we have that

$$f_Y(y) = \begin{cases} \frac{1}{2} & 1 \leq y \leq 3 \\ 0 & \text{otherwise} \end{cases} \quad (43)$$

We can also compute the  $f_Y(y)$  as follows

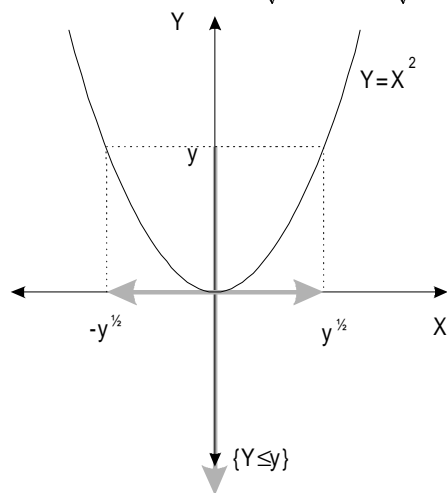
$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{1}{2}(y-1)\right) = \frac{1}{2} f_X\left(\frac{1}{2}(y-1)\right) \quad (44)$$

Substituting (38) into the above equation, we get the same result as (43).

**Example 11** Let  $X$  be a random variable. Let

$$Y = g(X) = X^2$$

Then for  $y \geq 0$ , we have that  $Y \leq y$ , when  $X^2 \leq y$  for  $-\sqrt{y} \leq X \leq \sqrt{y}$ .



Note that in this case there are two inverse functions for  $g(X)$

$$g_1^{-1}(Y) = \sqrt{Y} \quad (45)$$

$$g_2^{-1}(Y) = -\sqrt{Y} \quad (46)$$

So

$$F_Y(y) = P(Y < y) = P(X^2 = y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

and

$$f_Y(y) = \frac{d}{dx} F_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})$$

Now for  $y < 0$ , there are no values of  $x$  such that  $x^2 < y$ . therefore

$$F_Y(y) = 0 \quad y < 0$$

Again, suppose  $X$  is uniform between the interval  $[0,1]$ . Then we have that

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ \sqrt{y} & 0 < y \leq 1 \\ 1 & y > 1 \end{cases}$$

and

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & 0 \leq y \leq 1 \\ 0 & \text{Otherwise} \end{cases} \quad (47)$$

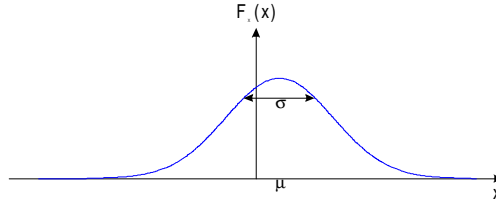


Figure 1: pdf of a Gaussian Random Variable

## Common Random Variables

### The Gaussian Random Variable

The Gaussian random variable is the most common of all random variables. It is used to characterize a number of random phenomenon such as noise most communication system to the random fluctuations of the desired received voltage in non-line-of-sight wireless communications systems such as cellular systems and personal communication systems. The Gaussian random variable has a pdf which is *completely defined* by its mean  $m$  and variance  $\sigma^2$ . Let  $X$  be a Gaussian random variable with a mean of  $m_X$  and a variance of  $\sigma_X^2$ . Then,  $f_X(x)$  is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} \quad (48)$$

and is plotted as follows. The CDF of  $X$  is found in the usual way

$$F_X(x) = \int_{-\infty}^x f_X(\alpha) d\alpha = \frac{1}{\sqrt{2\pi\sigma_X^2}} \int_{-\infty}^x e^{-\frac{(\alpha-m_X)^2}{2\sigma_X^2}} d\alpha \quad (49)$$

This integral cannot be evaluated in closed form. However the integral is well tabulated. A tabulation common to communication engineers is the Q-function where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{\alpha^2}{2}} d\alpha \quad (50)$$

Therefore,

$$F_X(x) = 1 - Q\left[\frac{x - m_X}{\sigma_X}\right] \quad (51)$$

Other well tabulated integrals include the *error function* and the *complementary error function*. The error function is defined as follows

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\alpha^2} d\alpha \quad (52)$$

The complementary error function is

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-\alpha^2} d\alpha \quad (53)$$

Note that the complementary error function and the Q-function are related as follows

$$Q(x) = \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right) \quad (54)$$

Note that  $P[x_1 < X \leq x_2]$  is computed as follows

$$P[x_1 < X \leq x_2] = \frac{1}{\sqrt{2\pi\sigma_X^2}} \int_{x_1}^{x_2} e^{-\frac{(\alpha-m_X)^2}{2\sigma_X^2}} d\alpha = Q\left(\frac{x_2 - m_X}{\sigma_X}\right) - Q\left(\frac{x_1 - m_X}{\sigma_X}\right) \quad (55)$$

## The Rayleigh Random Variable

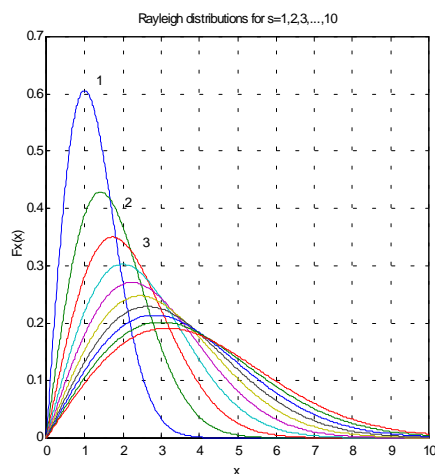
Let  $X$  and  $Y$  be two independent Gaussian random variables with  $m_X = m_Y = 0$  and  $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ . Suppose we define a new random variable  $Z$  from the transformation

$$Z = \sqrt{X^2 + Y^2}$$

Now if  $X$  and  $Y$  represent the received voltages from two signal components in a wireless channel, then  $Z$  represents the *received envelope*, namely  $\sqrt{V_X^2 + V_Y^2}$ .  $Z$  is called a *Rayleigh random variable*. The pdf of  $Z$  is

$$f_Z(z) = \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}} u(z) \quad (56)$$

and is plotted below for various values of  $\sigma_X^2 = \sigma_Y^2 = \sigma^2$



The CDF of  $Z$  is found by integrating the pdf. Therefore

$$F_Z(z) = \int_{-\infty}^z f_Z(z) dz = \int_0^z f_Z(z) dz = 1 - e^{-\frac{z^2}{2\sigma^2}} \quad (57)$$

The received signal envelop in a cellular or PCS system, where line-of-sight is not established, is typically modeled as a Rayleigh random variable

## The Exponential Random Variable

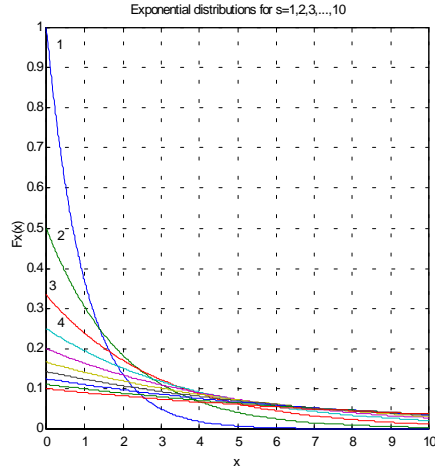
Let  $Z$  be a Rayleigh random variable. Suppose we define a new random variable,  $W$ , such that

$$W = Z^2 \quad (58)$$

If  $Z$  represents the voltage envelope of a signal, the  $W$  represents the *power* in a signal, to within a constant. Then  $W$  has the following pdf

$$f_W(w) = \frac{1}{\sigma^2} e^{-\frac{w}{\sigma^2}} u(w) \quad (59)$$

$W$  is called an exponential random variable, and the pdf of  $w$  looks as follows



The CDF of  $W$  is

$$F_W(w) = \left(1 - e^{-\frac{w}{\sigma^2}}\right) u(w) \quad (60)$$

Exponential random variables are used to model the power of cellular and PCS systems where line-of-sight propagation is not established

## Uniform Random Variable

In systems or signals where the phase is not known, it is generally modeled as a random variable  $\Theta$ , which is uniformly distributed between 0 and  $2\pi$ . The pdf of such a random variable is

$$f_{\Theta}(\theta) = \frac{1}{2\pi} [u(\theta) - u(\theta - 2\pi)] = \begin{cases} \frac{1}{2\pi} & 0 \leq \theta \leq 2\pi \\ 0 & \text{otherwise} \end{cases} \quad (61)$$

The CDF is

$$F_{\Theta}(\theta) = \frac{\theta}{2\pi} [u(\theta) - u(\theta - 2\pi)] + u(\theta + 2\pi) = \begin{cases} 0 & \theta < 0 \\ \frac{\theta}{2\pi} & 0 \leq \theta \leq 2\pi \\ 1 & \theta > 2\pi \end{cases} \quad (62)$$

## The Central Limit Theorem

Suppose we take a large number of random variables and sum them together. The *central limit theorem* states that the resultant random variable will have a Gaussian distribution. This is one of the reasons why Gaussian random variables are so common. A more formal statement of the theorem follows

**Theorem 2** Let  $X_i, i = 1, 2, \dots, N$  be a collection of independent random variables with a mean of 0 and a variance of 1. Define a new random variable,  $Y$ , as follows

$$Y = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N X_i \quad (63)$$

Then  $Y$  is a Gaussian random variable with a mean of 0 and a variance of 1, i.e.,  $m_Y = 0, \sigma_Y^2 = 1$ .

## Rician Random Variables

We just learned that when line-of-sight does not exist in a cellular system, that the received envelope follows a Rayleigh distribution. However, if line-of-sight is established, then the signal follows a *Rician distribution*, which has the following pdf

$$f_Z(z) = \frac{z}{\sigma^2} e^{-\frac{z^2+v^2}{2\sigma^2}} I_0\left(\frac{zv}{\sigma^2}\right) u(z) \quad (64)$$

where  $I_0(\cdot)$  is the *modified Bessel function of the 0<sup>th</sup> order*. Note that when  $v = 0$ ,  $I_0(0) = 1$  and  $Z$  degenerates to a Rayleigh distribution.