Probability, Random Variables, and Random Vectors

Outcomes of an "experiment": Outcomes are the observable (measurable) results or objects of an a random experiment. There basically three types depending on the ~~n~~ number of outcomes: (i) finite, (ii) countably infinite, and (iii) uncountable. The first ~~to~~ two cases are discrete and the third case is continuous.

All outcomes of an experiment form the

Sample space Ω

Events of the experiment. Event E corresponds to a feature or property that multiple outcomes share. An event coinciding with a single outcome is called an elementary event. Events are subsets of the sample space Ω

An algebra for events called a sigma field is defined. Since set operations on events and outcomes arise in many problems, it will be necessary to define a σ-field which places a restriction on the allowable events that are assigned probabilities. The sample space and the σ-field $\mathcal{F}$ together comprise the event space $\{\Omega, \mathcal{F}\}$.

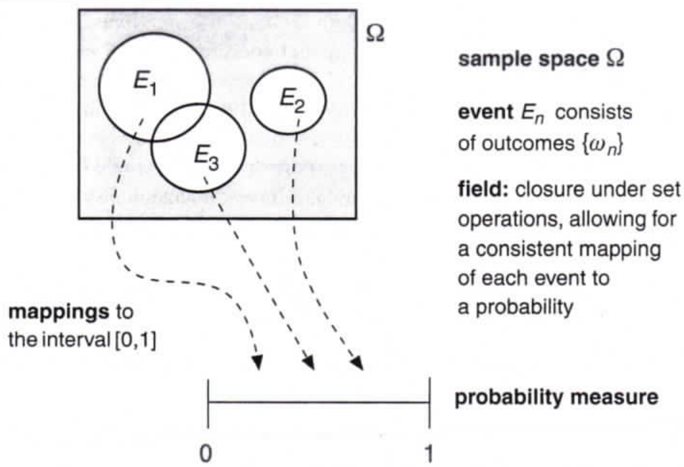
Probabilities are assigned to events according to three axioms. These axioms allow for any probability of an event in the event space to be calculated in a consistent manner. The sample space, the σ-field, and the probability measure P together comprise the probability space $\{\Omega, \mathcal{F}, P\}$.

This framework is pictorially shown ~~the~~ in the figure below. A rectangle is used to denote the sample space Ω, and the events $\{E_n\}$ are depicted by circles. This is the so-called Venn Diagram. Events that have some common outcomes $\{\omega_n\}$ are shown to be overlapping.

A σ-field $\mathcal{F}$ is defined so that all operations of events in $\mathcal{F}$ yield an event that is also in $\mathcal{F}$.

Event operations include $\cup$ (or) $\cap$ (and), ^{c} (complement), $-$ (difference), and $\oplus$ (exclusive or)

A σ-field and the three axioms allow us to define a consistent probability measure P with values in the closed interval $[0, 1]$



Pictorial representation of sample space Ω, events $\{E_n\}$ in Ω, and mappings to probabilities in the interval [0, 1]. A σ-field $\mathcal{F}$ is specified for Ω to describe events for which probabilities are assigned.

The previous description is suitable for relatively small sample spaces. For more complicated situations, which basically comprise most problems of interest, it is necessary that the probability space be mapped to a random variable defined on the real line $\mathbb{R}$. This allows us to perform complicated operations on the random variable such as transformation from one random variable to another. We can also compute various quantities that characterize a random experiment such as moments (mean, variance, and so on). This mapping is summarized below

- Outcomes in the sample space Ω are mapped to numbers on the real line R generating random variable X. Since engineering problems are formulated using variables and numbers, it will be necessary to map outcomes in the original sample space Ω to R (or integers $\mathbb{Z}$ or complex numbers $\mathbb{C}$) The original probability space carries over to the real line and provided the mapping is measurable, the resulting random variable also has a consistent probability space

- Random variables can be defined directly without resorting to an underlying abstract probability space. The probability space for random variable X is $\{\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X\}$ where $\mathcal{B}(\mathbb{R})$ is the Borel σ-field defined by all open intervals on $\mathbb{R}$. With this probability, we can utilize standard families of random variables such as Gaussian, binomial and so on to model various physical phenomena. The mapping of events to intervals of random variable X is shown ~~the~~ in the figure below. The goal is to describe events in a problem of interest using the real line $\mathbb{R}$. For discrete random variables, a probability mass function (pmf) describes the probability of each possible outcome, which are usually denoted by some subset of the integers $\mathbb{Z}$. The probability of an event is obtained by summing over the integer values corresponding to the event

$$P(a \leq X \leq b) = \sum_{x=a}^{b} p_X[x]$$

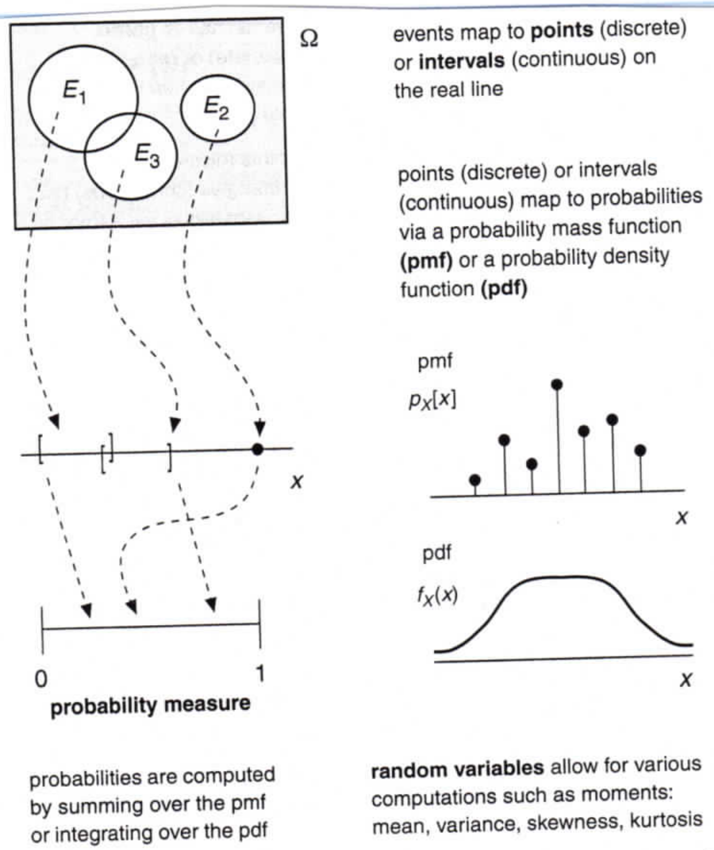
where $p_X[x] = P[X = x]$ is the pmf and is represented by the solid circles in the figure.

For continuous random variables a probability density function (pdf) describes the probability density on $\mathbb{R}$: probabilities are obtained by integrating the pdf over an interval of interest.

$$P[a \leq X \leq b] = \int_a^b f_X(x)\,dx$$

where $f_X(x)$ is the pdf. and has a continuous form as shown in the figure. It is also possible to have a mixed random variable with discrete and continuous components. As mentioned above in the last item, once we are familiar with random variables, there is usually no need to consider the underlying abstract sample space Ω: we can operate directly on the pmf (discrete) ~~of~~ or the pdf (continuous)

This notion of random variable can be extended to the notion random vectors but that will be ~~dissecuse~~ discussed later



Pictorial representation of the mapping of events in the sample space Ω to random variable X. The probability measure of a random variable is defined by a probability mass function (pmf) (for discrete outcomes) or a probability density function (pdf) (for continuous outcomes).

~~Random sequences and Rand~~

<u>Example of pmf</u>

Suppose we observe three calls at a telephone switch where voice calls (v) and data calls (d) are equally likely. Let X denote the number of voice calls, Y the

~~*~~ number of data calls, and let $R = XY$

Find the pmf $P_X[x]$, $P_Y(y]$, $P_R(r)$

What ~~are~~ ~~the~~ is the sample space?

Outcomes	ddd	ddv	dvd	dvv	vdd	vdv	vvd	vvv
P	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$	$1/8$
Random variables								
X	0	1	1	2	1	2	2	3
Y	3	2	2	1	2	1	1	0
R	0	2	2	2	2	2	2	0

$P_X(x) = P[X = x]$

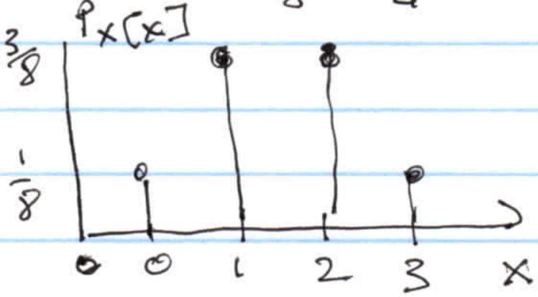
$P_X[0] = \frac{1}{8}$, $P_X[1] = \frac{3}{8}$, $P_X[2] = \frac{3}{8}$, $P_X[3] = \frac{1}{8}$

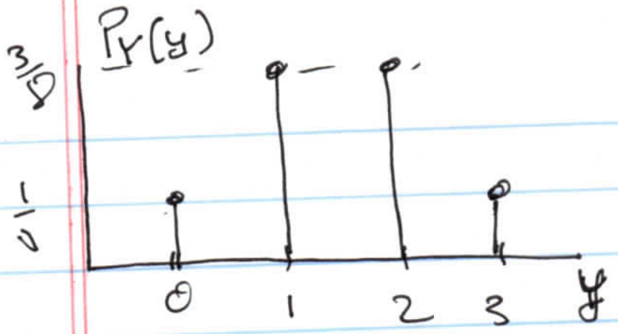
$P_Y(y) = P[Y = y]$

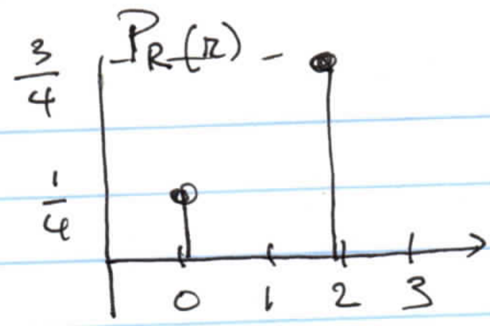
$P_Y(0) = \frac{1}{8}$, $P_Y(1) = \frac{3}{8}$, $P_Y(2) = \frac{3}{8}$ $P_Y(3) = \frac{1}{8}$

$P_R(r) = P[R = r]$

$P_R(0) = \frac{2}{8} = \frac{1}{4}$ $P_R(2) = \frac{6}{8} = \frac{3}{4}$







$$P[1 \le X \le 2] = \sum_{x=1}^{2} P_X[x] = \frac{3}{8} + \frac{3}{8} = \frac{3}{4}$$

Random Sequences and Random Processes

An example of a random sequence is several consecutive tosses of a fair coin. Each possible sequence of outcomes is called a realization. For example, if a fair coin is tossed 10 times $\{H, H, H, H, H, H, H, H, H, H\}$ is one possible realization: in fact, there are 2^{10} realizations for this experiment. We summarize the approach that will be taken to describe random sequences and random processes:

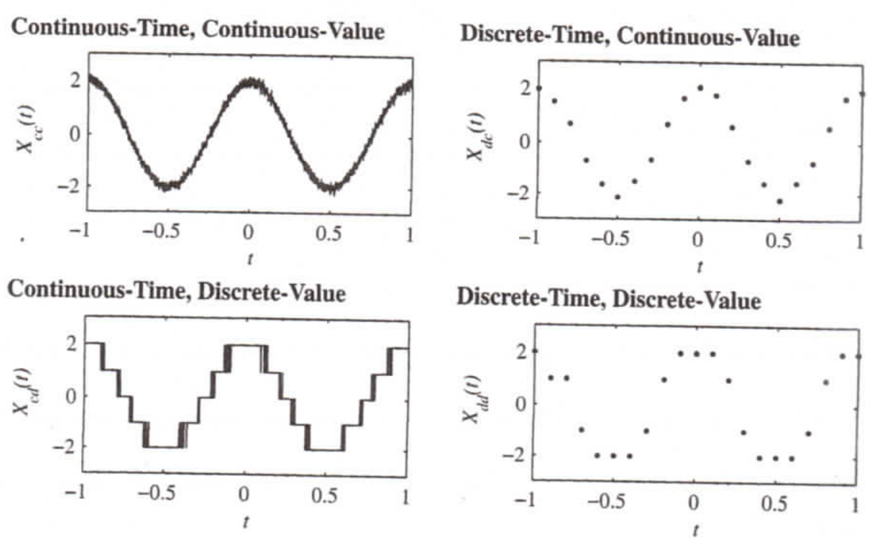
Random sequences: A collection of random variables indexed by discrete time k

They may have discrete or countable values or continuous.

Random process: A random process ~~is~~ $X(t)$ is specified on real line $\mathbb{R}$ representing continuous time. the outcome may be discrete or continuous

TABLE 1.1 Types of Random Processes with Examples

Random Sequence $X[k]$	Random Process $X(t)$
Discrete time, discrete outcomes (*Bernoulli sequence*)	Continuous time, discrete outcomes (*Poisson process*)
Discrete time, continuous outcomes (*Gaussian sequence*)	Continuous time, continuous outcomes (*Wiener process*)



Probability theory

Three basic types of experiments

(i) a finite number of outcomes
ex: single coin toss

(ii) an infinite but countable number of outcomes (such as tossing a single coin ad infinitum) and (iii) a continuum (uncountable number) of outcomes (such as temperature measurements). More complicated experiments are combinations of these three types of outcomes

It is easy to assign probabilities to events from experiments in category (i). Extending it to category (ii) requires some work but is also

~~stret~~ straightforward because the outcomes are countable: there is a one-to-one mapping of outcomes to the natural numbers $N = \{1, 2, 3 \cdots\}$. Experiments in category (iii) are the most difficult to quantify because there is a continuum of outcomes: for any two values $a < b$ there exists $\varepsilon > 0$ such that $a < a + \varepsilon < b$. Thus it is not immediately obvious how to assign probabilities to an uncountable number of outcomes; in fact, it turns out that the probability of a single outcome in a continuous experiment is zero.

Summary of how probabilities are generally computed for each of the three types of experiments:

- Finite: Probabilities are computed by counting the outcomes comprising an event of interest relative to the total of outcomes in the sample space Ω. Basic rules of combinatorics are used for finite problems

- Countably infinite: Probabilities are computed by summing over a discrete probability mass function that characterizes the experiment. Techniques for computing finite and infinite sums are used for countably infinite ~~sum~~ problems

- Uncountable. Probabilities are computed by integrating over a continuous probability density function that characterizes the experiment. Techniques from calculus are used for uncountable problems.

TABLE 2.1 Basic Types of Random Experiments with Examples

Outcomes	Finite	Infinite
Discrete	Tossing two coins (*finite*)	Repeated tosses of a single coin (*countably infinite*, denumerable)
Continuous	None	Temperature (*uncountable*, nondenumerable)

Finite experiments can also be solved using technique for countably infinite problems, but it is often easier to simply count outcomes as mentioned above. Countably infinite and uncountable experiments problems are not easily examined without a more rigorous characterization as provided by the probability mass function (pmf) and the probability density function (pdf).

Example in the two coin toss. There are only four possible outcomes {HH, TT, TH, HT}. If the probability of at least one H is of interest, we simply count the number of such outcomes (which is three) and divide that by the total number of outcomes. Thus the probability of at least one H is 0.75, it is denoted by P[at least one H) = P(~~HH or HT or TH~~)

P(HH or HT or TH) = P(HH) + P(HT) + P(TH)

For the countably infinite example in the table we might be interested in the probability that H first appears after the second toss. Clearly this more difficult to compute than the previous finite outcome example: we need to consider all sequences of the form TTTHH...., TTHH....., and so on. As we will see the pmf allows for a straightforward probability computation which is easier than attempting to directly count all outcomes of ~~the~~ interest. Likewise for the uncountable (continuum) case where the pdf is used to compute probabilities via an integration over the outcomes of interest which are described by intervals on $\mathbb{R}$.

Although we use the term experiment to describe how random outcomes are generated, in real applications the notion of experiment is artificial because events "just occur" (such as temperature fluctuations, radioactive decay, and so on). However, we use the term experiment and trial not only for synthetic experiment such as tossing a die or using a computer program to generate a random number but also for natural events. In the latter case, the experiment can be viewed as ~~as~~ a model of the underlying mechanisms for such events.

A probabilistic modeling of events in the physical world is advantageous because most phenomena are too complicated to accurately represent using physical models (coin toss on a table)

By using a probabilistic model, we bypass the need for an accurate physical model, and can make predictions about the occurrence of an event with a high degree of accuracy, and which is consistent with our intuition about randomness (the frequency of outcomes).

A probabilistic model is a powerful representation that allows us to "say something" about events that arise due to the complicated interactions of many underlying physical mechanism.

Good example: use of ~~statist~~ probability and statistics in solid-state physics and thermodynamics

Sets and Sample Spaces

Definition: Set: A set is a collection of objects or numbers that represent those objects. These objects are called elements or points of the set

sets will be denoted by uppercase letters usually at the beginning of the Latin alphabet

Example 1

set $A = \{2, 4, 6, 8\}$, and set $B = \{\ldots, -1, 0, 1, 2 \ldots\}$ consists of all integers

The set of all integers is represented by $\mathbb{Z}$ and positive integers by $\mathbb{Z}^+$ which includes zero

Example 2

Set $C = (1, 5)$ consists of all real numbers between 1 and 5 ~~ex~~ excluding 1 and 5, whereas set $D = [1, 5]$ includes the end points

The set of all real numbers is represented by $\mathbb{R}$ and positive real numbers by $\mathbb{R}^{+}$ which includes zero. We often use the symbols $\pm\infty$ which technically are not real numbers. Although complex quantities are not extensively here, they do appear in the application ~~par~~ part. the symbol for all ~~co~~ complex numbers is $\mathbb{C}$

Element a in set A is written as $a \in A$, element b that is not in set A is written as $b \notin A$. The type of set can be characterized according to the number of

<u>Definition</u>: Cardinality: The cardinality of a set is the number of elements in that set. The cardinality of set E is denoted $|E|$

A countable set can have either a finite or an infinite number of elements, the former is called finite (cardinality $< |\mathbb{N}|$) and the latter is ~~co~~ countably infinite or denumerable (cardinality $= |\mathbb{N}|$)

$\mathbb{N}$: set of natural numbers $\{1, 2, 3, \ldots\}$
does not include zero

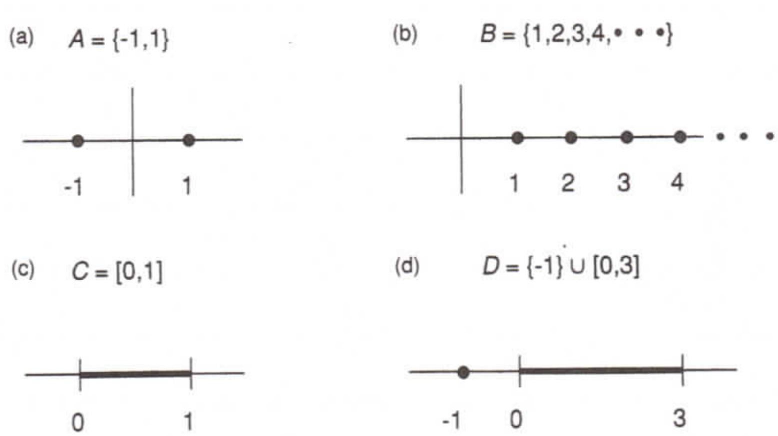
The elements of a countably infinite are in one-to-one correspondence with the natural numbers $\mathbb{N} \triangleq \{1, 2, 3 \ldots\}$

Example 3 $\mathbb{Z}$ an $\mathbb{Z}^+$ have the same cardinality as $\mathbb{N}$ and thus are countably infinite. The rational numbers denoted by $\mathbb{Q}$ and having the form a/b with $a, b \in \mathbb{Z}$ $(b \neq 0)$ also have cardinality $|\mathbb{N}|$

Example 4

The irrational numbers have ~~a~~ cardinality $> |\mathbb{N}|$ and are uncountable. Likewise $\mathbb{R}$ and any interval on $\mathbb{R}$ have a cardinality $>$ $|\mathbb{N}|$, and both are uncountable. "Almost all" irrational numbers are transcendental of which π and e are perhaps the most well-known of the set.

The figure below shows the different types of sets



Examples of the three basic types of sets and a mixed set. (a) Finite. (b) Countably infinite. (c) Uncoun (an interval). (d) Mixed (an interval and one point).

These set distinctions are useful later when developing a probability space for events of an experiment. It is ~~necess~~ necessary that we define precisely events of an experiment and the corresponding algebra of operations which allow for consistent probability measure

Example 5

Discrete set: $A = \{x^3 - 1 : x = 0, 2, 4\}$

$$= \{-1, 7, 63\}$$

Continuous set $B = \{x : x \geq -1\} = [-1, \infty)$

A set defined ~~by~~ in terms of other sets

$$C = \{x : x \in A \text{ or } x \in B\} = A \cup B$$

Definition: Intervals

$(a,b) \triangleq \{x : a < x < b\}$

$(a,b] \triangleq \{x : a < x \leq b\}$

$[a,b) \triangleq \{x : a \leq x < b\}$

$[a,b] \triangleq \{x : a \leq x \leq b\}$

if $x \in (a,b)$ there is an $\varepsilon > 0$ such that $x+\varepsilon \in (a,b)$ and $x-\varepsilon \in (a,b)$: open interval

Singletons : boundary points and points in $\mathbb{R}$

<u>Definition</u> Sample Space Ω : The sample space Ω is the set of all possible outcomes in an experiment. It is also called the universe or the universal set

Example 6

$\Omega = \{0, 1\}$ two possible symbols

used in digital communication system.

Coin toss $H = 1$ $T = 0$

Example 7

$$\Omega = [0, \infty) = \mathbb{R}^+$$

Definition

Subset : if $x \in A$ then $x \in B$

$A \subset B$ sometimes $A \subseteq B$ possibility

of equality

Definition $A = B$ $A \subset B$ and $B \subset A$

Definition A^c : all elements in Ω that

are not in A $a \notin A$ implies $a \in A^c$

$$A^c = \{x : x \notin A\}$$

Example 9

if $\Omega = [0, 100]$ and ~~A = 25~~

$A = [0, 25)$ then $A^c = [25, 100]$

if $\Omega = \mathbb{R}$ and $B = (0, \infty)$ then

$B^c = (-\infty, 0]$

Ω must be defined in order for A^c to have a meaning.

Definition: Empty set : The empty set is the set containing no element

$\phi = \Omega^c$ The empty set is needed when defining an algebra of operations for sets

Definition: Power Set $\mathcal{P}(\Omega)$: the power set contains all possible subsets of the sample space Ω. It includes ϕ and Ω itself.

The power set can also be defined for any subset $E \subset \Omega$, it is denoted by $\mathcal{P}(E)$

Example 10

$$\Omega = \{2, 4, 6\}$$

$$\mathcal{P}(\Omega) = \{\phi, \Omega, 2, 4, 6, \{2,4\}, \{2,6\}, \{4,6\}\}$$ has eight element

$$E = \{2, 4\} \subset \Omega$$

$$\mathcal{P}[E] = \{\phi, E, 2, 4\}$$ has four elements

if E is finite then $|P[E]|$ is necessarily finite