PROBABILITY THEORY AND STOCHASTIC PROCESS

OBJECTIVES:

1. To provide mathematical background and sufficient experience so that student can read, write and understand sentences in the language of probability theory.
2. To introduce students to the basic methodology of ―probabilistic thinking‖ and apply it to problems.
3. To understand basic concepts of Probability theory and Random Variables, how to deal with multiple Random Variables.
4. To understand the difference between time averages statistical averages.
5. To teach students how to apply sums and integrals to compute probabilities, and expectations.

Course Outcomes: Upon completing this course, the students will be able to:

1. Apply operations on single random variables
2. Apply operations on multiple random variables
3. Analyze spectral and temporal characteristics
4. Characterize LTI systems using ACFs and PSDs
5. Explain Noise and Information theory

UNIT-I

Probability and Random Variable

Probability: Set theory, Experiments and Sample Spaces, Discrete and Continuous Sample Spaces, Events, Probability Definitions and Axioms, Mathematical Model of Experiments, Joint Probability, Conditional Probability, Total Probability, Bayes‘ Theorem, and Independent Events, Bernoulli‘s trials.

The Random Variable: Definition of a Random Variable, Conditions for a Function to be a Random Variable, Discrete and Continuous, Mixed Random Variable

UNIT II:

Distribution and density functions and Operations on One Random Variable

Distribution and density functions: Distribution and Density functions, Properties, Binomial, Poisson, Uniform, Exponential Gaussian, Rayleigh and Conditional Distribution, Methods of defining Conditioning Event, Conditional Density function and its properties, problems.

Operation on One Random Variable: Expected value of a random variable, function of a random variable, moments about the origin, central moments, variance and skew, characteristic function, moment generating function, transformations of a random variable, monotonic transformations for a continuous random variable, non monotonic transformations of continuous random variable, transformations of Discrete random variable

Multiple Random Variables and Operations on Multiple Random Variables

Multiple Random Variables: Vector Random Variables, Joint Distribution Function and Properties, Joint density Function and Properties, Marginal Distribution and density Functions, conditional Distribution and density Functions, Statistical Independence, Distribution and density functions of Sum of Two Random Variables and Sum of Several Random Variables, Central Limit Theorem - Unequal Distribution, Equal Distributions

Operations on Multiple Random Variables: Expected Value of a Function of Random Variables, Joint Moments about the Origin, Joint Central Moments, Joint Characteristic Functions, and Jointly Gaussian Random Variables: Two Random Variables case and N Random Variable case, Properties, Transformations of Multiple Random Variables

UNIT III:

Stochastic Processes-Temporal Characteristics: The Stochastic process Concept, Classification of Processes, Deterministic and Nondeterministic Processes, Distribution and Density Functions, Statistical Independence and concept of Stationarity: First-Order Stationary Processes, Second- Order and Wide-Sense Stationarity, Nth-Order and Strict-Sense Stationarity, Time Averages and

Ergodicity, Mean-Ergodic Processes, Correlation-Ergodic Processes Autocorrelation Function and Its Properties, Cross-Correlation Function and Its Properties, Covariance Functions and its properties, Gaussian Random Processes. Linear system Response: Mean and Mean-squared value, Autocorrelation, Cross-Correlation Functions.

UNIT I V:

Stochastic Processes-Spectral Characteristics: The Power Spectrum and its Properties, Relationship between Power Spectrum and Autocorrelation Function, the Cross-Power Density Spectrum and Properties, Relationship between Cross-Power Spectrum and Cross-Correlation Function.

Spectral characteristics of system response: power density spectrum of response, cross power spectral density of input and output of a linear system

UNIT-V

Resistive noise / Thermal noise sources Arbitrary noise sources ,effective noise temperature ,noise equivalent bandwidth, Average noise figures,Averages noise figure of cascaded networks narrow band noise quradrature noise representation of narrowband noise and it's properties

TEXT BOOKS:

Probability, Random Variables & Random Signal Principles -Peyton Z. Peebles, TMH, 4th Edition, 2001.

Probability and Random Processes-Scott Miller, Donald Childers,2Ed,Elsevier,2012

REFERENCE BOOKS:

Theory of probability and Stochastic Processes-Pradip Kumar Gosh, University Press

Probability and Random Processes with Application to Signal Processing - Henry Stark and John W. Woods, Pearson Education, 3rd Edition.

Probability Methods of Signal and System Analysis- George R. Cooper, Clave D. MC Gillem, Oxford, 3rd Edition, 1999.

Statistical Theory of Communication -S.P. Eugene Xavier, New Age Publications 2003

Probability, Random Variables and Stochastic Processes Athanasios Papoulis and S.Unnikrishna Pillai, PHI, 4th Edition, 2002.

UNIT – 1
PROBABILITY AND RANDOM VARIABLE

PROBABILITY

Introduction

It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge.

A brief history

Probability has an amazing history. A practical gambling problem faced by the French nobleman Chevalier de Méré sparked the idea of probability in the mind of Blaise Pascal (1623-1662), the famous French mathematician. Pascal's correspondence with Pierre de Fermat (1601-1665), another French Mathematician in the form of seven letters in 1654 is regarded as the genesis of probability. Early mathematicians like Jacob Bernoulli (1654-1705), Abraham de Moivre (1667-1754), Thomas Bayes (1702-1761) and Pierre Simon De Laplace (1749-1827) contributed to the development of probability. Laplace's Theory Analytique des Probabilities gave comprehensive tools to calculate probabilities based on the principles of permutations and combinations. Laplace also said, "Probability theory is nothing but common sense reduced to calculation."

Later mathematicians like Chebyshev (1821-1894), Markov (1856-1922), von Mises (1883- 1953), Norbert Wiener (1894-1964) and Kolmogorov (1903-1987) contributed to new developments. Over the last four centuries and a half, probability has grown to be one of the most essential mathematical tools applied in diverse fields like economics, commerce, physical sciences, biological sciences and engineering. It is particularly important for solving practical electrical-engineering problems in communication, signal processing and computers. Notwithstanding the above developments, a precise definition of probability eluded the mathematicians for centuries. Kolmogorov in 1933 gave the axiomatic definition of probability and resolved the problem.

Randomness arises because of

random nature of the generation mechanism
Limited understanding of the signal dynamics inherent imprecision in measurement, observation, etc.
For example, thermal noise appearing in an electronic device is generated due to random motion of electrons. We have deterministic model for weather prediction; it takes into account of the factors affecting weather. We can locally predict the temperature or the rainfall of a place on the basis of previous data. Probabilistic models are established from observation of a random phenomenon. While probability is concerned with analysis of a random phenomenon, statistics help in building such models from data.

Deterministic versus probabilistic models

A deterministic model can be used for a physical quantity and the process generating it provided sufficient information is available about the initial state and the dynamics of the process generating the physical quantity. For example,

We can determine the position of a particle moving under a constant force if we know the initial position of the particle and the magnitude and the direction of the force.

We can determine the current in a circuit consisting of resistance, inductance and capacitance for a known voltage source applying Kirchoff's laws.

Many of the physical quantities are random in the sense that these quantities cannot be predicted with certainty and can be described in terms of probabilistic models only. For example,

The outcome of the tossing of a coin cannot be predicted with certainty. Thus the outcome of tossing a coin is random.

The number of ones and zeros in a packet of binary data arriving through a communication channel cannot be precisely predicted is random.

The ubiquitous noise corrupting the signal during acquisition, storage and transmission can be modelled only through statistical analysis.

How to Interpret Probability

Mathematically, the probability that an event will occur is expressed as a number between 0 and 1. Notationally, the probability of event A is represented by P (A).

If P (A) equals zero, event A will almost definitely not occur.
If P (A) is close to zero, there is only a small chance that event A will occur.
If P (A) equals 0.5, there is a 50-50 chance that event A will occur.
If P(A) is close to one, there is a strong chance that event A will occur.
If P(A) equals one, event A will almost definitely occur.

In a statistical experiment, the sum of probabilities for all possible outcomes is equal to one. This means, for example, that if an experiment can have three possible outcomes (A, B, and C), then P(A) + P(B) + P(C) = 1.

Applications

Probability theory is applied in everyday life in risk assessment and in trade on financial markets. Governments apply probabilistic methods in environmental regulation, where it is called pathway analysis

Another significant application of probability theory in everyday life is reliability. Many consumer products, such as automobiles and consumer electronics, use reliability theory in product design to reduce the probability of failure. Failure probability may influence a manufacturer's decisions on a product's warranty.

THE BASIC CONCEPTS OF SET THEORY

Some of the basic concepts of set theory are:

Set: A set is a well defined collection of objects. These objects are called elements or members of the set. Usually uppercase letters are used to denote sets.

The set theory was developed by George Cantor in 1845-1918. Today, it is used in almost every branch of mathematics and serves as a fundamental part of present-day mathematics.
In everyday life, we often talk of the collection of objects such as a bunch of keys, flock of birds, pack of cards, etc. In mathematics, we come across collections like natural numbers, whole numbers, prime and composite numbers.

We assume that,

the word set is synonymous with the word collection, aggregate, class and comprises of elements.

Objects, elements and members of a set are synonymous terms.

Sets are usually denoted by capital letters A, B, C, , etc.

Elements of the set are represented by small letters a, b, c, , etc.

If _a‘ is an element of set A, then we say that _a‘ belongs to A. We denote the phrase _belongs to‘ by the Greek symbol _∈_ (epsilon). Thus, we say that a ∈ A.

If _b‘ is an element which does not belong to A, we represent this as b ∉ A.

Examples of sets:

Describe the set of vowels.

If A is the set of vowels, then A could be described as A = {a, e, i, o, u}.

Describe the set of positive integers.

Since it would be impossible to list all of the positive integers, we need to use a rule to describe this set. We might say A consists of all integers greater than zero.
Set A = {1, 2, 3} and Set B = {3, 2, 1}. Is Set A equal to Set B?

Yes. Two sets are equal if they have the same elements. The order in which the elements are listed does not matter. What is the set of men with four arms?

Since all men have two arms at most, the set of men with four arms contains no elements. It is the null set (or empty set).

Set A = {1, 2, 3} and Set B = {1, 2, 4, 5, 6}. Is Set A a subset of Set B?

Set A would be a subset of Set B if every element from Set A were also in Set B. However, this is not the case. The number 3 is in Set A, but not in Set B. Therefore, Set A is not a subset of Set B

Some important sets used in mathematics are

N: the set of all natural numbers = {1, 2, 3, 4, }

Z: the set of all integers = {....., -3, -2, -1, 0, 1, 2, 3, }

Q: the set of all rational numbers

R: the set of all real numbers

Z+: the set of all positive integers

W: the set of all whole numbers

The different types of sets are explained below with examples.

Empty Set or Null Set:

A set which does not contain any element is called an empty set, or the null set or the void set and it is denoted by ∅ and is read as phi. In roster form, ∅ is denoted by {}. An empty set is a finite set, since the number of elements in an empty set is finite, i.e., 0.

For example: (a) the set of whole numbers less than 0.

(b) Clearly there is no whole number less than 0.

Therefore, it is an empty set.

(c) N = {x : x ∈ N, 3 < x < 4}

Let A = {x : 2 < x < 3, x is a natural number}

Here A is an empty set because there is no natural number between 2 and 3.

Let B = {x : x is a composite number less than 4}.

Here B is an empty set because there is no composite number less than 4.

Note:

∅ ≠ {0} ∴ has no element.

{0} is a set which has one element 0.

The cardinal number of an empty set, i.e., n(∅) = 0

Singleton Set:

A set which contains only one element is called a singleton set.

For example:

A = {x : x is neither prime nor composite}

It is a singleton set containing one element, i.e., 1.

B = {x : x is a whole number, x < 1}

This set contains only one element 0 and is a singleton set.

Let A = {x : x ∈ N and x^2 = 4}

Here A is a singleton set because there is only one element 2 whose square is 4.

Let B = {x : x is a even prime number}

Here B is a singleton set because there is only one prime number which is even, i.e., 2.

Finite Set:

A set which contains a definite number of elements is called a finite set. Empty set is also called a finite set.

For example:

The set of all colors in the rainbow.

N = {x : x ∈ N, x < 7}

P = {2, 3, 5, 7, 11, 13, 17, 97}

Infinite Set:

The set whose elements cannot be listed, i.e., set containing never-ending elements is called an infinite set.

For example:

Set of all points in a plane

A = {x : x ∈ N, x > 1}

Set of all prime numbers

B = {x : x ∈ W, x = 2n}

Note:

All infinite sets cannot be expressed in roster form.

For example:

The set of real numbers since the elements of this set do not follow any particular pattern.

Cardinal Number of a Set:

The number of distinct elements in a given set A is called the cardinal number of A. It is denoted by n(A). And read as _the number of elements of the set‘.

For example:

A {x : x ∈ N, x < 5}

A = {1, 2, 3, 4}

Therefore, n(A) = 4

B = set of letters in the word ALGEBRA B = {A, L, G, E, B, R}
Therefore, n(B) = 6

Equivalent Sets:

Two sets A and B are said to be equivalent if their cardinal number is same, i.e., n(A) = n(B). The symbol for denoting an equivalent set is _↔‘.

For example:

A = {1, 2, 3} Here n(A) = 3

B = {p, q, r} Here n(B) = 3 Therefore, A ↔ B

Equal sets:

Two sets A and B are said to be equal if they contain the same elements. Every element of A is an element of B and every element of B is an element of A.

For example:

A = {p, q, r, s}

B = {p, s, r, q} Therefore, A = B
Disjoint Sets:

Two sets A and B are said to be disjoint, if they do not have any element in common.

For example;

A = {x : x is a prime number}

B = {x : x is a composite number}.

Clearly, A and B do not have any element in common and are disjoint sets.

Overlapping sets:

Two sets A and B are said to be overlapping if they contain at least one element in common.

For example;

A = {a, b, c, d}

B = {a, e, i, o, u}

X = {x : x ∈ N, x < 4} Y = {x : x ∈ I, -1 < x < 4}
Here, the two sets contain three elements in common, i.e., (1, 2, 3)

Definition of Subset:

If A and B are two sets, and every element of set A is also an element of set B, then A is called a subset of B and we write it as A ⊆ B or B ⊇ A

The symbol ⊂ stands for _is a subset of' or _is contained in'

Every set is a subset of itself, i.e., A ⊂ A, B ⊂ B.

Empty set is a subset of every set.

Symbol _⊆' is used to denote _is a subset of' or _is contained in'.

A ⊆ B means A is a subset of B or A is contained in B.

B ⊆ A means B contains A.

Examples;

1. Let A = {2, 4, 6}

B = {6, 4, 8, 2}

Here A is a subset of B

Since, all the elements of set A are contained in set B. But B is not the subset of A
Since, all the elements of set B are not contained in set A.

Notes:

If ACB and BCA, then A = B, i.e., they are equal sets. Every set is a subset of itself.
Null set or ∅ is a subset of every set.

The set N of natural numbers is a subset of the set Z of integers and we write $N \subset Z$.

3. Let A = {2, 4, 6}

B = {x : x is an even natural number less than 8} Here $A \subset B$ and $B \subset A$.
Hence, we can say A = B

4. Let A = {1, 2, 3, 4}

B = {4, 5, 6, 7}

Here $A \not\subset B$ and also $B \not\subset C$ [$\not\subset$ denotes _not a subset of‟]

Super Set:

Whenever a set A is a subset of set B, we say the B is a superset of A and we write, $B \supseteq A$. Symbol $\supseteq$ is used to denote _is a super set of‟
For example;

A = {a, e, i, o, u}

B = {a, b, c, , z}

Here A ⊆ B i.e., A is a subset of B but B ⊇ A i.e., B is a super set of A

Proper Subset:

If A and B are two sets, then A is called the proper subset of B if A ⊆ B but B ⊇ A i.e., A ≠ B. The symbol ‗⊂' is used to denote proper subset. Symbolically, we write A ⊂ B.

For example;

1. A = {1, 2, 3, 4}

Here n(A) = 4

B = {1, 2, 3, 4, 5}

Here n(B) = 5
We observe that, all the elements of A are present in B but the element ‗5' of B is not present in A. So, we say that A is a proper subset of B.
Symbolically, we write it as A ⊂ B

Notes:

No set is a proper subset of itself.

Null set or ∅ is a proper subset of every set.

2. A = {p, q, r}

B = {p, q, r, s, t}

Here A is a proper subset of B as all the elements of set A are in set B and also A ≠ B.

Notes:

No set is a proper subset of itself.

Empty set is a proper subset of every set.

Power Set:

The collection of all subsets of set A is called the power set of A. It is denoted by P(A). In P(A), every element is a set.

For example;

If A = {p, q} then all the subsets of A will be P(A) = {∅, {p}, {q}, {p, q}}
Number of elements of P(A) = n[P(A)] = 4 = 22

In general, n[P(A)] = 2m where m is the number of elements in set A.

Universal Set

A set which contains all the elements of other given sets is called a universal set. The symbol for denoting a universal set is ∪ or ξ.

For example;

1. If A = {1, 2, 3} B = {2, 3, 4} C = {3, 5, 7}

then U = {1, 2, 3, 4, 5, 7}

[Here A ⊆ U, B ⊆ U, C ⊆ U and U ⊇ A, U ⊇ B, U ⊇ C]

If P is a set of all whole numbers and Q is a set of all negative numbers then the universal set is a set of all integers.

3. If A = {a, b, c} B = {d, e} C = {f, g, h, i}

then U = {a, b, c, d, e, f, g, h, i} can be taken as universal set.

Operations on sets:

Definition of Union of Sets:

Union of two given sets is the smallest set which contains all the elements of both the sets.

To find the union of two given sets A and B is a set which consists of all the elements of A and all the elements of B such that no element is repeated.

The symbol for denoting union of sets is ‗U‘. Some properties of the operation of union:
AUB = BUA (Commutative law)

AU(BUC) = (AUB)UC (Associative law)

A U Φ = A (Law of identity element, is the identity of U)

AUA = A (Idempotent law)

UUA = U (Law of U) U is the universal set.

Notes:

A U Φ = ΦU A = A i.e. union of any set with the empty set is always the set itself. Examples:
If A = {1, 3, 7, 5} and B = {3, 7, 8, 9}. Find union of two set A and B.

Solution:

A U B = {1, 3, 5, 7, 8, 9}
No element is repeated in the union of two sets. The common elements 3, 7 are taken only once.

Let X = {a, e, i, o, u} and Y = {ϕ}. Find union of two given sets X and Y.

Solution:

X U Y = {a, e, i, o, u}

Therefore, union of any set with an empty set is the set itself.

2. Definition of Intersection of Sets:

Intersection of two given sets is the largest set which contains all the elements that are common to both the sets.

To find the intersection of two given sets A and B is a set which consists of all the elements which are common to both A and B.

The symbol for denoting intersection of sets is ‗∩‗. Some properties of the operation of intersection

$A \cap B = B \cap A$ (Commutative law)

$(A \cap B) \cap C = A \cap (B \cap C)$ (Associative law)

$\Phi \cap A = \Phi$ (Law of Φ)

$U \cap A = A$ (Law of ∪)

$A \cap A = A$ (Idempotent law)

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributive law) Here ∩ distributes over ∪ Also, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributive law) Here ∪ distributes over ∩ Notes:
$A \cap \Phi = \Phi \cap A = \Phi$ i.e. intersection of any set with the empty set is always the empty set. Solved examples :
If A = {2, 4, 6, 8, 10} and B = {1, 3, 8, 4, 6}. Find intersection of two set A and B.

Solution:

$A \cap B = \{4, 6, 8\}$

Therefore, 4, 6 and 8 are the common elements in both the sets.

If X = {a, b, c} and Y = {ф}. Find intersection of two given sets X and Y.

Solution:

$X \cap Y = \{ \}$

Difference of two sets

If A and B are two sets, then their difference is given by A - B or B - A.

If A = {2, 3, 4} and B = {4, 5, 6}

A - B means elements of A which are not the elements of B. i.e., in the above example A - B = {2, 3}
In general, $B - A = \{x : x \in B, \text{ and } x \notin A\}$

If A and B are disjoint sets, then A – B = A and B – A = B Solved examples to find the difference of two sets:
1. A = {1, 2, 3} and B = {4, 5, 6}.

Find the difference between the two sets:

A and B

B and A

Solution:

The two sets are disjoint as they do not have any elements in common. (i) A - B = {1, 2, 3} = A
(ii) B - A = {4, 5, 6} = B

2. Let A = {a, b, c, d, e, f} and B = {b, d, f, g}.

Find the difference between the two sets:

A and B

B and A

Solution:

A - B = {a, c, e}

Therefore, the elements a, c, e belong to A but not to B

B - A = {g)

Therefore, the element g belongs to B but not A.

4. Complement of a Set

In complement of a set if S be the universal set and A a subset of S then the complement of A is the set of all elements of S which are not the elements of A.

Symbolically, we denote the complement of A with respect to S as A‘.

Some properties of complement sets

A ∪ A' = A' ∪ A = U (Complement law)

(A ∩ B') = ϕ (Complement law) - The set and its complement are disjoint sets.

(A ∪ B) = A' ∩ B' (De Morgan's law)

(A ∩ B)' = A' ∪ B' (De Morgan's law)

(A')' = A (Law of complementation)

Φ' = U (Law of empty set - The complement of an empty set is a universal set.

U' = Φ and universal set) - The complement of a universal set is an empty set.

For Example; If S = {1, 2, 3, 4, 5, 6, 7}

A = {1, 3, 7} find A'.

Solution:

We observe that 2, 4, 5, 6 are the only elements of S which do not belong to A. Therefore, A' = {2, 4, 5, 6}
Algebraic laws on sets:

Commutative Laws:

For any two finite sets A and B;

A U B = B U A

A ∩ B = B ∩ A

Associative Laws:

For any three finite sets A, B and C;

(A U B) U C = A U (B U C)

(A ∩ B) ∩ C = A ∩ (B ∩ C)

Thus, union and intersection are associative.

Idempotent Laws:

For any finite set A;

A U A = A

A ∩ A = A

Distributive Laws:

For any three finite sets A, B and C;

A U (B ∩ C) = (A U B) ∩ (A U C)

A ∩ (B U C) = (A ∩ B) U (A ∩ C)

Thus, union and intersection are distributive over intersection and union respectively.

De Morgan's Laws:

For any two finite sets A and B;

(i) A – (B U C) = (A – B) ∩ (A – C)

(ii) A - (B ∩ C) = (A – B) U (A – C)

De Morgan's Laws can also we written as:

(A U B)' = A' ∩ B'

(A ∩ B)' = A' U B'

More laws of algebra of sets:

For any two finite sets A and B;

A – B = A ∩ B'

B – A = B ∩ A'

A – B = A ⇔ A ∩ B = ∅

(A – B) U B = A U B

(v) (A – B) ∩ B = ∅

(vi) (A – B) U (B – A) = (A U B) – (A ∩ B)

Definition of De Morgan's law:

The complement of the union of two sets is equal to the intersection of their complements and the complement of the intersection of two sets is equal to the union of their complements. These are called De Morgan's laws.

For any two finite sets A and B;

(A U B)' = A' ∩ B' (which is a De Morgan's law of union).

(A ∩ B)' = A' U B' (which is a De Morgan's law of intersection).

Venn Diagrams:

Pictorial representations of sets represented by closed figures are called set diagrams or Venn diagrams.

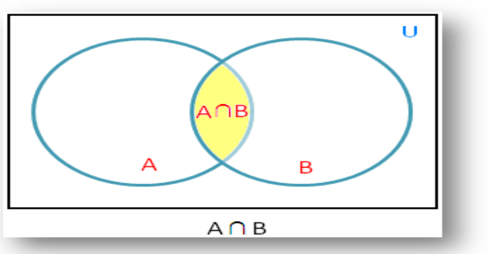
Venn diagrams are used to illustrate various operations like union, intersection and difference. We can express the relationship among sets through this in a more significant way.
In this,

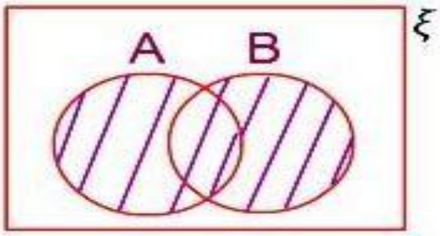
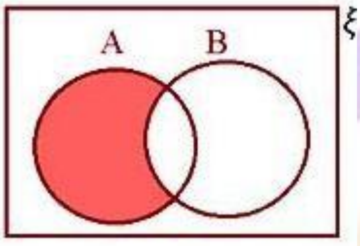
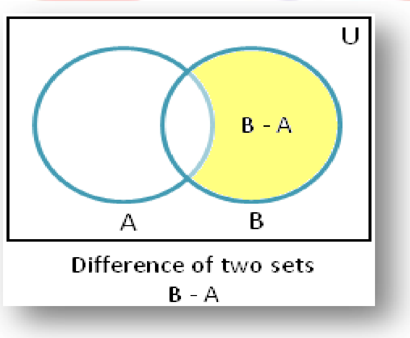
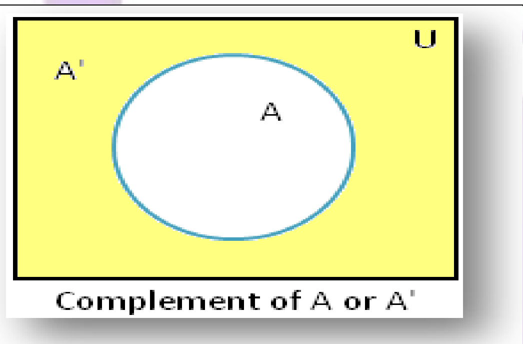
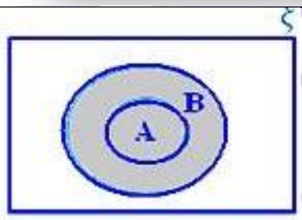
A rectangle is used to represent a universal set.

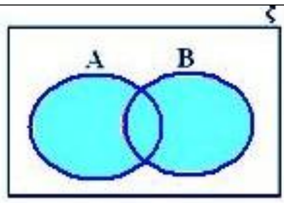
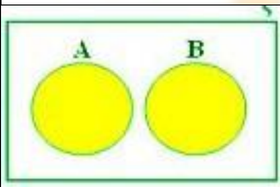
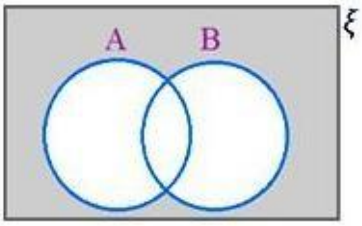
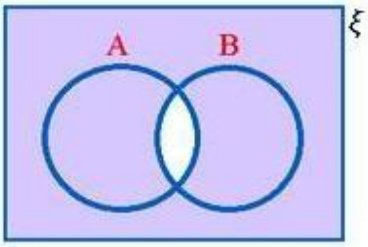
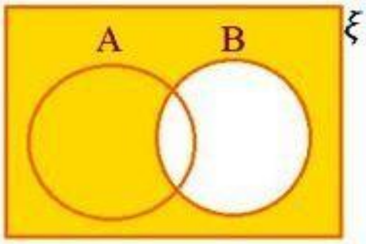
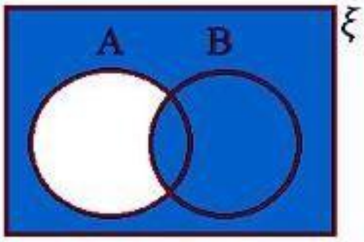
Circles or ovals are used to represent other subsets of the universal set.

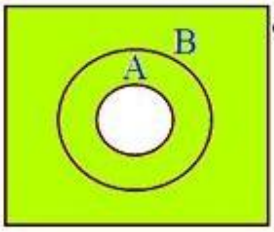
Venn diagrams in different situations

In these diagrams, the universal set is represented by a rectangular region and its subsets by circles inside the rectangle. We represented disjoint set by disjoint circles and intersecting sets by intersecting circles.

S.No	Set &Its relation	Venn Diagram
1	Intersection of A and B	

2	Union of A and B	
3	Difference : A-B	
4	Difference : B-A	
5	Complement of set A	
6	A U B when A ⊂ B	

7	A ∪ B when neither A ⊂ B nor B ⊂ A	
8	A ∪ B when A and B are disjoint sets	
9	(A ∪ B)’ (A union B dash)	
10	(A ∩ B)’ (A intersection B dash)	
11	B’ (B dash)	
12	(A - B)’ (Dash of sets A minus B)	

13	$(A \subset B)'$ (Dash of A subset B)	

Problems of set theory:

Let A and B be two finite sets such that $n(A) = 20$, $n(B) = 28$ and $n(A \cup B) = 36$, find $n(A \cap B)$.

Solution:

Using the formula $n(A \cup B) = n(A) + n(B) - n(A \cap B)$. then $n(A \cap B) = n(A) + n(B) - n(A \cup B)$
$= 20 + 28 - 36$

$= 48 - 36$

$= 12$

If $n(A - B) = 18$, $n(A \cup B) = 70$ and $n(A \cap B) = 25$, then find $n(B)$.

Solution:

Using the formula $n(AUB) = n(A - B) + n(A \cap B) + n(B - A)$ $70 = 18 + 25 + n(B - A)$
$70 = 43 + n(B - A)$ $n(B - A) = 70 - 43$
$n(B - A) = 27$

Now $n(B) = n(A \cap B) + n(B - A)$

$= 25 + 27$

$= 52$

In a group of 60 people, 27 like cold drinks and 42 like hot drinks and each person likes at least one of the two drinks. How many like both coffee and tea?

Solution:

Let A = Set of people who like cold drinks. B = Set of people who like hot drinks.
Given

$(A \cup B) = 60$ $n(A) = 27$ $n(B) = 42$ then; $n(A \cap B) = n(A) + n(B) - n(A \cup B)$
$= 27 + 42 - 60$

$= 69 - 60 = 9$

$= 9$

Therefore, 9 people like both tea and coffee.

There are 35 students in art class and 57 students in dance class. Find the number of students who are either in art class or in dance class.

When two classes meet at different hours and 12 students are enrolled in both activities.

When two classes meet at the same hour.

Solution:

$n(A) = 35,$ $n(B) = 57,$ $n(A \cap B) = 12$

(Let A be the set of students in art class. B be the set of students in dance class.)

When 2 classes meet at different hours $n(A \cup B) = n(A) + n(B) - n(A \cap B)$

$= 35 + 57 - 12$

$= 92 - 12$

$= 80$

When two classes meet at the same hour, $A \cap B = \emptyset$ $n(A \cup B) = n(A) + n(B) - n(A \cap B)$

$= n(A) + n(B)$

$= 35 + 57$

$= 92$

In a group of 100 persons, 72 people can speak English and 43 can speak French. How many can speak English only? How many can speak French only and how many can speak both English and French?

Solution:

Let A be the set of people who speak English. B be the set of people who speak French.
A - B be the set of people who speak English and not French. B - A be the set of people who speak French and not English. $A \cap B$ be the set of people who speak both French and English. Given,
$n(A) = 72$ $n(B) = 43$ $n(A \cup B) = 100$

Now, $n(A \cap B) = n(A) + n(B) - n(A \cup B)$

$= 72 + 43 - 100$

$= 115 - 100$

$= 15$

Therefore, Number of persons who speak both French and English = 15 $n(A) = n(A - B) + n(A \cap B)$
$\Rightarrow n(A - B) = n(A) - n(A \cap B)$

$= 72 - 15$

$= 57$

and $n(B - A) = n(B) - n(A \cap B)$

$= 43 - 15$

$= 28$

Therefore, Number of people speaking English only = 57 Number of people speaking French only = 28

Probability Concepts

Before we give a definition of probability, let us examine the following concepts:

Experiment:

In probability theory, an experiment or trial (see below) is any procedure that can be infinitely repeated and has a well-defined set of possible outcomes, known as the sample space. An experiment is said to be random if it has more than one possible outcome, and deterministic if it has only one. A random experiment that has exactly two (mutually exclusive) possible outcomes is known as a Bernoulli trial.

Random Experiment:

An experiment is a random experiment if its outcome cannot be predicted precisely. One out of a number of outcomes is possible in a random experiment. A single performance of the random experiment is called a trial.

Random experiments are often conducted repeatedly, so that the collective results may be subjected to statistical analysis. A fixed number of repetitions of the same experiment can be thought of as a composed experiment, in which case the individual repetitions are called trials. For example, if one were to toss the same coin one hundred times and record each result, each toss would be considered a trial within the experiment composed of all hundred tosses.

Mathematical description of an experiment:
A random experiment is described or modeled by a mathematical construct known as a probability space. A probability space is constructed and defined with a specific kind of experiment or trial in mind.
A mathematical description of an experiment consists of three parts:

A sample space, Ω (or S), which is the set of all possible outcomes.
A set of events , where each event is a set containing zero or more outcomes.
The assignment of probabilities to the events—that is, a function P mapping from events to probabilities.
An outcome is the result of a single execution of the model. Since individual outcomes might be of little practical use, more complicated events are used to characterize groups of outcomes. The collection of all such events is a sigma-algebra . Finally, there is a need to specify each event's likelihood of happening; this is done using the probability measure function,P.

Sample Space: The sample space S is the collection of all possible outcomes of a random experiment. The elements of S are called sample points.
A sample space may be finite, countably infinite or uncountable.
A finite or countably infinite sample space is called a discrete sample space.

An uncountable sample space is called a continuous sample space

Ex:1. For the coin-toss experiment would be the results ―Head‖and ―Tail‖, which we may represent by S={H T}.

Ex. 2. If we toss a die, one sample space or the set of all possible outcomes is S = { 1, 2, 3, 4, 5, 6}
The other sample space can be S = {odd, even}
Types of Sample Space:

Finite/Discrete Sample Space:

Consider the experiment of tossing a coin twice. The sample space can be
S = {HH, HT, T H , TT} the above sample space has a finite number of sample points. It is called a finite sample space.

Countably Infinite Sample Space:

Consider that a light bulb is manufactured. It is then tested for its life length by inserting it into a socket and the time elapsed (in hours) until it burns out is recorded. Let the measuring instrument is capable of recording time to two decimal places, for example 8.32 hours.

Now, the sample space becomes count ably infinite i.e.

S = {0.0, 0.01, 0.02}

The above sample space is called a countable infinite sample space.

Un Countable/ Infinite Sample Space:

If the sample space consists of unaccountably infinite number of elements then it is called Un Countable/ Infinite Sample Space.

Event: An event is simply a set of possible outcomes. To be more specific, an event is a subset A of the sample space S.

$A \subseteq S$

For a discrete sample space, all subsets are events.

Ex: For instance, in the coin-toss experiment the events A={Heads} and B={Tails} would be mutually exclusive.

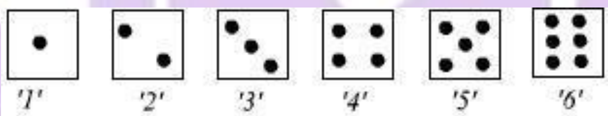
An event consisting of a single point of the sample space 'S' is called a simple event or elementary event.

Some examples of event sets:

Example 1: tossing a fair coin

The possible outcomes are H (head) and T (tail). The associated sample space is $S = \{H, T\}$ It is a finite sample space. The events associated with the sample space are: $S, \{H\}, \{T\}$ and ϕ. S

Example 2: Throwing a fair die:
The possible 6 outcomes are:



. . .

The associated finite sample space is $S = \{'1', '2', '3', '4', '5', '6'\}$.Some events are

A = The event of getting an odd face={'1', '3', '5'}.

B = The event of getting a six={'6'}

And so on.

Example 3: Tossing a fair coin until a head is obtained

We may have to toss the coin any number of times before a head is obtained. Thus the possible outcomes are:
H, TH, TTH, TTTH,
How many outcomes are there? The outcomes are countable but infinite in number. The countably infinite sample space is $S = \{H, TH, TTH, ...\}$.

Example 4 : Picking a real number at random between -1 and +1

The associated Sample space is

$$S = \{s \mid s \in \mathbb{R}, -1 \leq s \leq 1\} = [-1, 1]$$

Clearly S is a continuous sample space.

Example 5: Drawing cards

Drawing 4 cards from a deck: Events include all spades, sum of the 4 cards is (assuming face cards have a value of zero), a sequence of integers, a hand with a 2, 3, 4 and 5. There are many more events.

Types of Events:

Exhaustive Events:

A set of events is said to be exhaustive, if it includes all the possible events.

Ex. In tossing a coin, the outcome can be either Head or Tail and there is no other possible outcome. So, the set of events { H , T } is exhaustive.

Mutually Exclusive Events:

Two events, A and B are said to be mutually exclusive if they cannot occur together.

i.e. if the occurrence of one of the events precludes the occurrence of all others, then such a set of events is said to be mutually exclusive.

If two events are mutually exclusive then the probability of either occurring is

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B).$$

Ex. In tossing a die, both head and tail cannot happen at the same time.

Equally Likely Events:
If one of the events cannot be expected to happen in preference to another, then such events are said to be Equally Likely Events.(Or) Each outcome of the random experiment has an equal chance of occuring.
Ex. In tossing a coin, the coming of the head or the tail is equally likely.

Independent Events:

Two events are said to be independent, if happening or failure of one does not affect the happening or failure of the other. Otherwise, the events are said to be dependent.

If two events, A and B are independent then the joint probability is

$$P(A \text{ and } B) = P(A \cap B) = P(A)P(B),$$

Non-. Mutually Exclusive Events:

If the events are not mutually exclusive then

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

Probability Definitions and Axioms:
Relative frequency Definition:

Consider that an experiment E is repeated n times, and let A and B be two events associated w i t h E. Let nA and nB be the number of times that the event A and the event B occurred among the n repetitions respectively.

The relative frequency of the event A in the 'n' repetitions of E is defined as f(A) = nA /n
The Relative frequency has the following properties:

$\boxed{f(A)=n_A/n}$

1.0 ≤f(A) ≤ 1

f(A) =1 if and only if A occurs every time among the n repetitions.

If an experiment is repeated n times under similar conditions and the event A occurs in n_A times, then the probability of the event A is defined as

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

Limitation:
Since we can never repeat an experiment or process indefinitely, we can never know the probability of any event from the relative frequency definition. In many cases we can't even obtain a long series of repetitions due to time, cost, or other limitations. For example, the probability of rain today can't really be obtained by the relative frequency definition since today can't be repeated again.

2. .The classical definition:
Let the sample space (denoted by S) be the set of all possible distinct outcomes to an experiment. The probability of some event is

$$\frac{\text{number of ways the event can occur}}{\text{number of outcomes in } S},$$

provided all points in S are equally likely. For example, when a die is rolled the probability of getting a 2 is $\frac{1}{6}$ because one of the six faces is a 2.

Limitation:

What does "equally likely" mean? This appears to use the concept of probability while trying to define it! We could remove the phrase "provided all outcomes are equally likely", but then the definition would clearly be unusable in many settings where the outcomes in S did not tend to occur equally often.

Example1:A fair die is rolled once. What is the probability of getting a _6' ?

Here $S = \{'1', '2', '3', '4', '5', '6'\}$ and $A = \{'6'\}$

$$\therefore N = 6 \quad \text{and} \quad N_A = 1$$
$$\therefore P(A) = \frac{1}{6}$$

Example2:A fair coin is tossed twice. What is the probability of getting two _heads'?

Here $S = \{HH, TH, HT, TT\}$ and $A = \{HH\}$.
Total number of outcomes is 4 and all four outcomes are equally likely. Only outcome favourable to A is {HH}

$\therefore P(A) = \frac{1}{4}$

Probability axioms:

Given an event E in a sample space S which is either finite with N elements or countably infinite with $N = \infty$ elements, then we can write

$$S \equiv \left(\bigcup_{i=1}^{N} E_i\right),$$

and a quantity $P(E_i)$, called the probability of event E_i, is defined such that

Axiom1: The probability of any event A is positive or zero. Namely P(A)≥0. The probability measures, in a certain way, the difficulty of event A happening: the smaller the probability, the more difficult it is to happen. i.e

$0 \leq P(E_i) \leq 1$.

Axiom2: The probability of the sure event is 1. Namely P(Ω)=1. And so, the probability is always greater than 0 and smaller than 1: probability zero means that there is no possibility for it to happen (it is an impossible event), and probability 1 means that it will always happen (it is a sure event).i.e

$P(S) = 1$.

Axiom3: The probability of the union of any set of two by two incompatible events is the sum of the probabilities of the events. That is, if we have, for example, events A,B,C, and these are two by two incompatible, then P(AUBUC)=P(A)+P(B)+P(C). i.e Additivity:

$P(E_1 \cup E_2) = P(E_1) + P(E_2)$, where E_1 and E_2 are mutually exclusive.

$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$ for $n = 1$, 2, ..., N where E_1, E_2, ... are mutually exclusive (i.e., $E_1 \cap E_2 = \emptyset$).

Main properties of probability: If A is any event of sample space S then

1. P(A)+P($\bar{A}$)=1. Or P($\bar{A}$)=1-P(A)
Since A ∪ $\bar{A}$ = S , P(A ∪ $\bar{A}$)=1
The probability of the impossible event is 0, i.e P(Ø)=0
If A⊂B, then P(A)≤P(B).
If A and B are two incompatible events, and therefore, P(A−B)=P(A)−P(A∩B).and P(B−A)=P(B)−P(A∩B).
Addition Law of probability:

$P(A \cup B)=P(A)+P(B)-P(A \cap B)$

Rules of Probability:

Rule of Subtraction:

Rule of Subtraction The probability that event A will occur is equal to 1 minus the probability that event A will not occur.

$$P(A) = 1 - P(A')$$

Rule of Multiplication:

Rule of Multiplication The probability that Events A and B both occur is equal to the probability that Event A occurs times the probability that Event B occurs, given that A has occurred.

$$P(A \cap B) = P(A)\ P(B|A)$$

Rule of addition:

Rule of Addition The probability that Event A or Event B occurs is equal to the probability that Event A occurs plus the probability that Event B occurs minus the probability that both Events A and B occur.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B))$$

Note: Invoking the fact that $P(A \cap B) = P(A)P(B \mid A)$, the Addition Rule can also be expressed as

$$P(A \cup B) = P(A) + P(B) - P(A)P(B \mid A)$$

PERMUTATIONS and COMBINATIONS:

S.No.	PERMUTATIONS	COMBINATIONS:
1	Arrangement of things in a specified order is called permutation. Here all things are taken at a time	In permutations, the order of arrangement of objects is important. But, in combinations, order is not important, but only selection of objects.
2	Arrangement of _r' things taken at a time from _n' things ,where r < n in a specified order in called r-permutation.	
3	Consider the letters a,b and c . Considering all the three letters at a time, the possible permutations are ABC ,a c b , b c a , b a c , c b a and c a b .	
4	The number of permutations taking r things at a time from _n' available things is denoted as p (n , r) or n pr	The number of combinations taking r things at a time from _n' available things is denoted as C(n , r) or n Cr
5	nPr= r!/nCr= n!/(n-r)!	n C r = P (n , r) / r ! = n ! / r ! (n - r) !

Example 1: An urn contains 6 red balls, 5 green balls and 4 blue balls. 9 balls were picked at random from the urn without replacement. What is the probability that out of the balls 4 are red, 3 are green and 2 are blue?
Sol:

9 balls can be picked from a population of 15 balls in $^{15}C_9 = \frac{15!}{9!6!}$.

Therefore the required probability is $\frac{^6C_4 \times {^5C_3} \times {^4C_2}}{^{15}C_9}$

Example2: What is the probability that in a throw of 12 dice each face occurs twice.

Solution: The total number of elements in the sample space of the outcomes of a single throw of 12 dice is $= 6^{12}$

The number of favourable outcomes is the number of ways in which 12 dice can be arranged in six groups of size 2 each – group 1 consisting of two dice each showing 1, group 2 consisting of two dice each showing 2 and so on. Therefore, the total number distinct groups

$$= \frac{12!}{2!2!2!2!2!2!}$$

Hence the required probability is $= \frac{12!}{(2)^6 6^{12}}$

Conditional probability

The answer is the conditional probability of B given A denoted by $P(B/A)$. We shall develop the concept of the conditional probability and explain under what condition this conditional probability is same as $P(B)$.

Notation
P (B/A) = Conditional probability of B given A

Let us consider the case of equiprobable events discussed earlier. Let N_{AB} sample points be favourable for the joint event A∩B

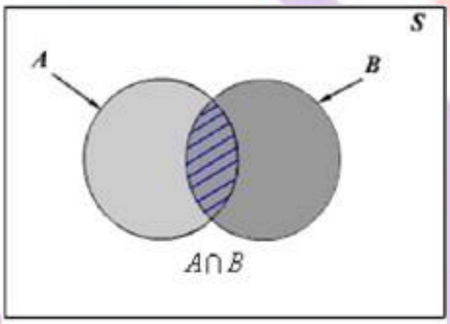


Figure 1

$$P(B/A) = \frac{\text{Number of outcomes favourable to A and B}}{\text{Number of outcomes in A}}$$

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

This concept suggests us to define conditional probability. The probability of an event B under the condition that another event A has occurred is called the conditional probability of B given A and defined by

$$P(B/A)=\frac{P(A\cap B)}{P(A)},\ \ P(A)\neq 0$$

We can similarly define the conditional probability of A given B , denoted by $P(A/B)$.
From the definition of conditional probability, we have the joint probability $P(A\cap B)$ of two events A and B as follows

$$\boxed{P(A\cap B)=P(A)P(B/A)=P(B)P(A/B)}$$

Problems:

Example 1 Consider the example tossing the fair die. Suppose

$$\begin{aligned}&A=\text{event of getting an even number }=\{2,4,6\}\\&B=\text{event of getting a number less than 4 }=\{1,2,3\}\\&\therefore A\cap B=\{2\}\\&\therefore P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{1/6}{3/6}=\frac{1}{3}\end{aligned}$$

Example 2 A family has two children. It is known that at least one of the children is a girl. What is the probability that both the children are girls?

A = event of at least one girl

B = event of two girls

$$\begin{aligned}&S\ =\ \{gg,\ gb,\ bg,\ bb\},\ A=\{gg,\ gb,\ bg\}\text{ and }B=\{gg\}\\&A\cap B=\{gg\}\\&\therefore P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{1/4}{3/4}=\frac{1}{3}\end{aligned}$$

Properties of Conditional probability:

If $B \subseteq A$, then $P(B/A) = 1$ and $P(A/B) \geq P(A)$ We have, $A \cap B = B$

$$\therefore P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

and

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Since

$$= \frac{P(A)P(B/A)}{P(B)}$$

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

We have,

$$\therefore P(S/A) = \frac{P(S \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

$$P(A \cap B) \geq 0, P(A) > 0$$

$$\therefore P(B/A) = \frac{P(A \cap B)}{P(A)} \geq 0$$

Chain Rule of Probability/Multiplication theorem:

$$P(A_1 \cap A_2 ... A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)...P(A_n/A_1 \cap A_2 .... \cap A_{n-1})$$

We have,

$$(A \cap B \cap C) = (A \cap B) \cap C$$

$$P(A \cap B \cap C) = P(A \cap B)P(C/A \cap B)$$

$$= P(A)P(B/A)P(C/A \cap B)$$

$$\therefore P(A \cap B \cap C) = P(A)P(B/A)P(C/A \cap B)$$

We can generalize the above to get the chain rule of probability for n events as

$$P(A_1 \cap A_2 ... A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)...P(A_n/A_1 \cap A_2 .... \cap A_{n-1})$$

Joint probability

Joint probability is defined as the probability of both A and B taking place, and is denoted by P (AB) or P(A∩B)

Joint probability is not the same as conditional probability, though the two concepts are often confused. Conditional probability assumes that one event has taken place or will take place, and then asks for the probability of the other (A, given B). Joint probability does not have such conditions; it simply asks for the chances of both happening (A and B). In a problem, to help distinguish between the two, look for qualifiers that one event is conditional on the other (conditional) or whether they will happen concurrently (joint).

Probability definitions can find their way into CFA exam questions. Naturally, there may also be questions that test the ability to calculate joint probabilities. Such computations require use of the multiplication rule, which states that the joint probability of A and B is the product of the conditional probability of A given B, times the probability of B. In probability notation:

P(AB) = P(A | B) * P(B)

Given a conditional probability P(A | B) = 40%, and a probability of B = 60%, the joint probability P(AB) = 0.6*0.4 or 24%, found by applying the multiplication rule.

P(AUB)=P(A)+P(B)-P(A∩B)

For independent events: P(AB) = P(A) * P(B)

Moreover, the rule generalizes for more than two events provided they are all independent of one another, so the joint probability of three events P(ABC) = P(A) * (P(B) *P(C), again assuming independence.

Summary of probabilities

Event	Probability
A	$P(A) \in [0,1]$
not A	$P(A^c) = 1 - P(A)$
A or B	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $P(A \cup B) = P(A) + P(B)$ if A and B are mutually exclusive
A and B	$P(A \cap B) = P(A\|B)P(B) = P(B\|A)P(A)$ $P(A \cap B) = P(A)P(B)$ if A and B are independent
A given B	$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B\|A)P(A)}{P(B)}$

Total Probability theorem:

Let $A_1, A_2, \ldots, A_n$ be n events such that

$$S = A_1 \cup A_2 \ldots\ldots \cup A_n \text{ and } A_i \cap A_j = \phi \text{ for } i \neq j.$$

Then for any event B,

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B/A_i)$$

Proof: We have

$$\therefore P(B) = P(\bigcup_{i=1}^{n} B \cap A_i)$$

$$= \sum_{i=1}^{n} P(B \cap A_i)$$

$$= \sum_{i=1}^{n} P(A_i)P(B/A_i)$$

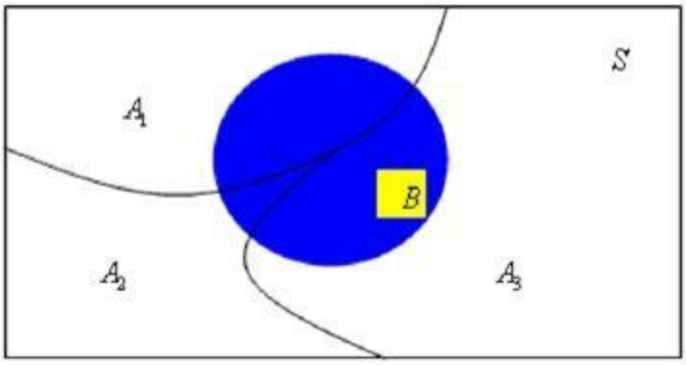


Figure 3

A decomposition of a set S into 2 or more disjoint nonempty subsets is called a partition of S.The subsets $A_1, A_2, \ldots, A_n$ form a partition of S if

$$S = A_1 \cup A_2 \ldots\ldots \cup A_n \text{ and } A_i \cap A_j = \phi \text{ for i} \neq \text{j}.$$

The theorem of total probability can be used to determine the probability of a complex event in terms of related simpler events. This result will be used in Bays' theorem to be discussed to the end of the lecture.

Bayes' Theorem:

Suppose $A_1, A_2, \ldots A_n$ *are partitions on S such that* $S = A_1 \cup A_2 \ldots\ldots \cup A_n$ and $A_i \cap A_j = \phi$ for i ≠ j.
Suppose the event B occurs if one of the events $A_1, A_2, \ldots A_n$ occurs. Thus we have the information of the probabilities $P(A_i)$ and $P(B/A_i), i = 1, 2.., n$. We ask the following question:
Given that B has occured what is the probability that a particular event A_k *has occured*? In other words *what is* $P(A_k/B)$?

We have $P(B) = \sum_{i=1}^{n} P(A_i)\, P(B|A_i)$ (Using the theorem of total probability)

$$\therefore\ P(A_k|B) = \frac{P(A_k)\, P(B/A_k)}{P(B)}$$

$$= \frac{P(A_k)\, P(B/A_k)}{\sum_{i=1}^{n} P(A_i) P(B/A_i)}$$

This result is known as the Baye's theorem. The probability $P(A_k)$ is called the a priori probability and $P(A_k/B)$ is called the a posteriori probability. Thus the Bays' theorem enables us to determine the a posteriori probability $P(A_k/B)$ from the observation that B has occurred. This result is of practical importance and is the heart of Baysean classification, Baysean estimation etc.

Example1:
In a binary communication system a zero and a one is transmitted with probability 0.6 and 0.4 respectively. Due to error in the communication system a zero becomes a one with a probability 0.1 and a one becomes a zero with a probability 0.08. Determine the probability (i) of receiving a one and (ii) that a one was transmitted when the received message is one
Solution:

Let S is the sample space corresponding to binary communication. Suppose T_0 be event of Transmitting 0 and T_1 be the event of transmitting 1 and R_0 and R_1 be corresponding events of receiving 0 and 1 respectively.

Given and $P(T_0) = 0.6,\ P(T_1) = 0.4,\ P(R_1/T_0) = 0.1$ $P(R_0/T_1) = 0.08.$

(i) $P(R_1)$ = Probabilty of receiving 'one'

$$= P(T_1)P(R_1/T_1) + P(T_0)P(R_1/T_0)$$
$$= 0.4\times 0.92 + 0.6\times 0.1$$
$$= 0.448$$

(ii) Using the Baye's rule

$$P(T_1/R_1) = \frac{P(T_1)P(R_1/T_1)}{P(R_1)}$$
$$= \frac{P(T_1)P(R_1/T_1)}{P(T_1)P(R_1/T_1) + P(T_0)P(R_1/T_0)}$$
$$= \frac{0.4\times 0.92}{0.4\times 0.92 + 0.6\times 0.1}$$
$$= 0.8214$$

Example 7: In an electronics laboratory, there are identically looking capacitors of three makes A_1, A_2 and A_3 in the ratio 2:3:4. It is known that 1% of A_1, 1.5% of A_2 and 2% of A_3 are defective. What percentages of capacitors in the laboratory are defective? If a capacitor picked at defective is found to be defective, what is the probability it is of make A_3 ?

Let D be the event that the item is defective. Here we have to find $P(D)$ and $P(A_3/D)$.

$$P(A_1) = \frac{2}{9},\ P(A_2) = \frac{1}{3} \text{ and } P(A_3) = \frac{4}{9}$$
$$P(D/A_1) = 0.01,\ P(D/A_2) = 0.015 \text{ and } P(D/A_3) = 0.02$$

$$\therefore P(D) = P(A_1)P(D/A_1) + P(A_2)P(D/A_2) + P(A_3)\ P(D/A_3)$$
$$= \frac{2}{9}\times 0.01 + \frac{1}{3}\times 0.015 + \frac{4}{9}\times 0.02$$
$$= 0.0167$$

and

$$P(A_3/D) = \frac{P(A_3)P(D/A_3)}{P(D)}$$
$$= \frac{\frac{4}{9}\times 0.02}{0.0167}$$
$$= 0.533$$

Independent events

Two events are called independent if the probability of occurrence of one event does not affect the probability of occurrence of the other. Thus the events A and B are independent if

$P(B/A)=P(B)$ and $P(A/B)=P(A)$ where $P(A)$ and $P(B)$ are assumed to be non-zero.
Equivalently if A and B are independent, we have

$$\frac{P(A\cap B)}{P(A)}=P(B)$$

or --------------------

$$P(A\cap B)=P(A)P(B)$$

Joint probability is the product of individual probabilities.

Two events A and B are called statistically dependent if they are not independent. Similarly, we can define the independence of n events. The events $A_1, A_2, \ldots, A_n$ are called independent if and only if
Example: Consider the example of tossing a fair coin twice. The resulting sample space is given by

$$P(A_i\cap A_j)=P(A_i)P(A_j)$$
$$P(A_i\cap A_j\cap A_k)=P(A_i)P(A_j)P(A_k)$$
$$P(A_i\cap A_j\cap A_k\cap \ldots A_n)=P(A_i)P(A_j)P(A_k)\ldots P(A_n)$$

$S=\{HH,HT,TH,TT\}$ and all the outcomes are equiprobable.

Let $A=\{TH,TT\}$ be the event of getting _tail' in the first toss and $B=\{TH,HH\}$ be the event of getting _head' in the second toss. Then

$P(A)=\frac{1}{2}$ and $P(B)=\frac{1}{2}$.
Again, so that

$$(A\cap B)=\{TH\}$$

$$P(A\cap B)=\frac{1}{4}=P(A)P(B)$$

Hence the events A and B are independent.

Problems:

Example1.A dice of six faces is tailored so that the probability of getting every face is proportional to the number depicted on it.

What is the probability of extracting a 6?
In this case, we say that the probability of each face turning up is not the same, therefore we cannot simply apply the rule of Laplace. If we follow the statement, it says that the probability of each face turning up is proportional to the number of the face itself, and this means that, if we say that the probability of face 1 being turned up is k which we do not know, then:
$P(\{1\})=k$, $P(\{2\})=2k$, $P(\{3\})=3k$, $P(\{4\})=4k$, $P(\{5\})=5k$, $P(\{6\})=6k$.
Now, since $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$ form an events complete system , necessarily
$P(\{1\})+P(\{2\})+P(\{3\})+P(\{4\})+P(\{5\})+P(\{6\})=1$

Therefore

$k+2k+3k+4k+5k+6k=1$

which is an equation that we can already solve:
$21k=1$

thus
$k=1/21$

And so, the probability of extracting 6 is $P(\{6\})=6k=6\cdot(1/21)=6/21$.

What is the probability of extracting an odd number?

The cases favourable to event A= "to extract an odd number" are: $\{1\},\{3\},\{5\}$. Therefore, since they are incompatible events,
$P(A)=P(\{1\})+P(\{3\})+P(\{5\})=k+3k+5k=9k=9\cdot(1/21)=9/21$

Example2: Roll a red die and a green die. Find the probability the total is 5. Solution: Let (x, y) represent getting x on the red die and y on the green die.
Then, with these as simple events, the sample space is

$$
\begin{aligned}
S = \{ \quad & (1,1) \quad (1,2) \quad (1,3) \quad \cdots \quad (1,6) \\
& (2,1) \quad (2,2) \quad (2,3) \quad \cdots \quad (2,6) \\
& (3,1) \quad (3,2) \quad (3,3) \quad \cdots \quad (3,6) \\
& \qquad\quad -- \quad\; -- \quad\; -- \\
& (6,1) \quad (6,2) \quad (6,3) \quad \cdots \quad (6,6)\}
\end{aligned}
$$

The sample points giving a total of 5 are (1,4) (2,3) (3,2), and (4,1). Therefore P(total is 5) = $\frac{4}{36}$

Example3: Suppose the 2 dice were now identical red dice. Find the probability the total is 5.

Solution : Since we can no longer distinguish between (x, y) and (y, x), the only distinguishable points in S are

$$
\begin{aligned}
S = \{ \quad & (1,1) \quad (1,2) \quad (1,3) \quad \cdots \quad (1,6) \\
& \qquad\quad\;\; (2,2) \quad (2,3) \quad \cdots \quad (2,6) \\
& \qquad\qquad\qquad\;\; (3,3) \quad \cdots \quad (3,6) \\
& \qquad\qquad\qquad\qquad\quad\;\; .. \qquad .. \\
: & \qquad\qquad\qquad\qquad\qquad\quad\;\; (6,6)\}
\end{aligned}
$$

Using this sample space, we get a total of 5 from points $(1,4)$ and $(2,3)$ only. If we assign equal probability $\frac{1}{21}$ to each point (simple event) then we get P(total is 5) = $\frac{2}{21}$.

Example4: Draw 1 card from a standard well-shuffled deck (13 cards of each of 4 suits - spades, hearts, diamonds, and clubs). Find the probability the card is a club.

Solution 1: Let S = { spade, heart, diamond, club}. (The points of S are generally listed between brackets {}.) Then S has 4 points, with 1 of them being "club", so P(club) = $\frac{1}{4}$.

Solution 2: Let S = {each of the 52 cards}. Then 13 of the 52 cards are clubs, so

$$P(\text{club}) = \frac{13}{52} = \frac{1}{4}.$$

Example 5: Suppose we draw a card from a deck of playing cards. What is the probability that we draw a spade?

Solution: The sample space of this experiment consists of 52 cards, and the probability of each sample point is 1/52. Since there are 13 spades in the deck, the probability of drawing a spade is

P(Spade) = (13)(1/52) = 1/4

Example 6: Suppose a coin is flipped 3 times. What is the probability of getting two tails and one head?

Solution: For this experiment, the sample space consists of 8 sample points. S = {TTT, TTH, THT, THH, HTT, HTH, HHT, HHH}
Each sample point is equally likely to occur, so the probability of getting any particular sample point is 1/8. The event "getting two tails and one head" consists of the following subset of the sample space.

A = {TTH, THT, HTT}

The probability of Event A is the sum of the probabilities of the sample points in A. Therefore,

P(A) = 1/8 + 1/8 + 1/8 = 3/8

Example7: An urn contains 6 red marbles and 4 black marbles. Two marbles are drawn without replacement from the urn. What is the probability that both of the marbles are black?

Solution: Let A = the event that the first marble is black; and let B = the event that the second marble is black. We know the following:

In the beginning, there are 10 marbles in the urn, 4 of which are black. Therefore, P(A) = 4/10.
After the first selection, there are 9 marbles in the urn, 3 of which are black. Therefore, P(B|A) = 3/9.
Therefore, based on the rule of multiplication:

$P(A \cap B) = P(A)\, P(B|A)$
$P(A \cap B) = (4/10) * (3/9) = 12/90 = 2/15$

RANDOM VARIABLE

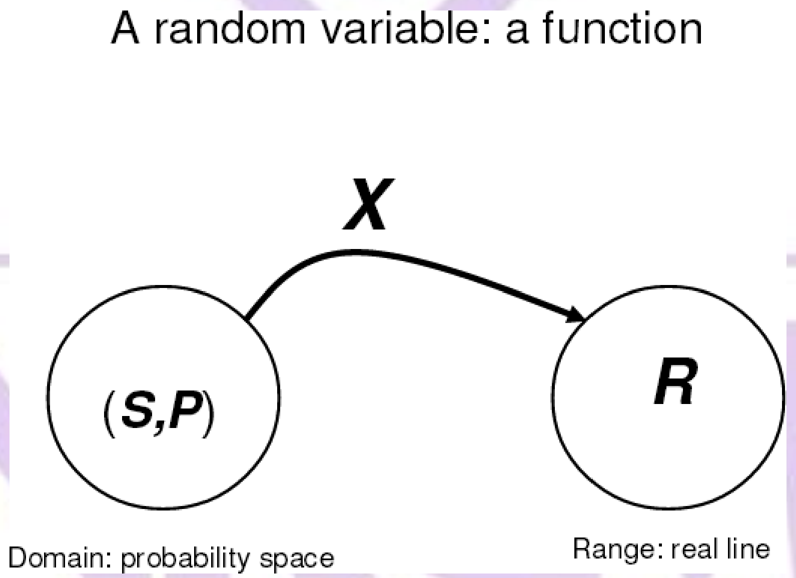
INTRODUCTION

In many situations, we are interested in numbers associated with the outcomes of a random experiment. In application of probabilities, we are often concerned with numerical values which are random in nature. For example, we may consider the number of customers arriving at a service station at a particular interval of time or the transmission time of a message in a communication system. These random quantities may be considered as real-valued function on the sample space. Such a real-valued function is called real random variable and plays an important role in describing random data. We shall introduce the concept of random variables in the following sections.

Random Variable Definition

A random variable is a function that maps outcomes of a random experiment to real numbers. (or)
A random variable associates the points in the sample space with real numbers

A (real-valued) random variable, often denoted by X(or some other capital letter), is a function mapping a probability space (S; P) into the real line R. This is shown in Figure 1.Associated with each point s in the domain S the function X assigns one and only one value X(s) in the range R. (The set of possible values of X(s) is usually a proper subset of the real line; i.e., not all real numbers need occur. If S is a finite set with m elements, then X(s) can assume at most an m different value as s varies in S.)



Example1
A fair coin is tossed 6 times. The number of heads that come up is an example of a random variable.
HHTTHT – 3, THHTTT -- 2.
These random variables can only take values between 0 and 6. The Set of possible values of random variables is known as its Range. Example2
A box of 6 eggs is rejected once it contains one or more broken eggs. If we examine 10 boxes of eggs and define the randomvariablesX1, X2 as
1 X1- the number of broken eggs in the 10 boxes 2 X2- the number of boxes rejected
Then the range of X1 is {0, 1,2,3,4-------------- 60} and X2 is {0,1,2 10}
Figure 2: A (real-valued) function of a random variable is itself a random variable, i.e., a function mapping a probability space into the real line.

Example 3 Consider the example of tossing a fair coin twice. The sample space is S={ HH,HT,TH,TT} and all four outcomes are equally likely. Then we can define a random variable X as follows

Sample Point	Value of the random Variable
HH	0
HT	1
TH	2
TT	3

Here $R_X = \{0,1,2,3\}$.
Example 4 Consider the sample space associated with the single toss of a fair die. The sample space is given by $S = \{1,2,3,4,5,6\}$.

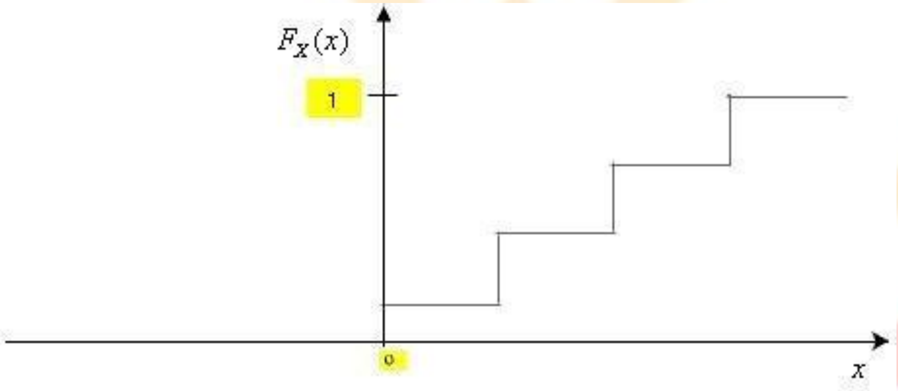
If we define the random variable X that associates a real number equal to the number on the face of the die, then $X = \{1,2,3,4,5,6\}$.
Types of random variables:

There are two types of random variables, discrete and continuous.

Discrete random variable:

A discrete random variable is one which may take on only a countable number of distinct values such as 0, 1,2,3,4,.Discrete random variables are usually (but not necessarily) counts. If
a random variable can take only a finite number of distinct values, then it must be discrete
(Or)

A random variable X is called a discrete random variable if $F_X(x)$ is piece-wise constant. Thus $F_X(x)$ is flat except at the points of jump discontinuity. If the sample space is discrete the random variable X defined on it is always discrete.



Plot of Distribution function of discrete random variable

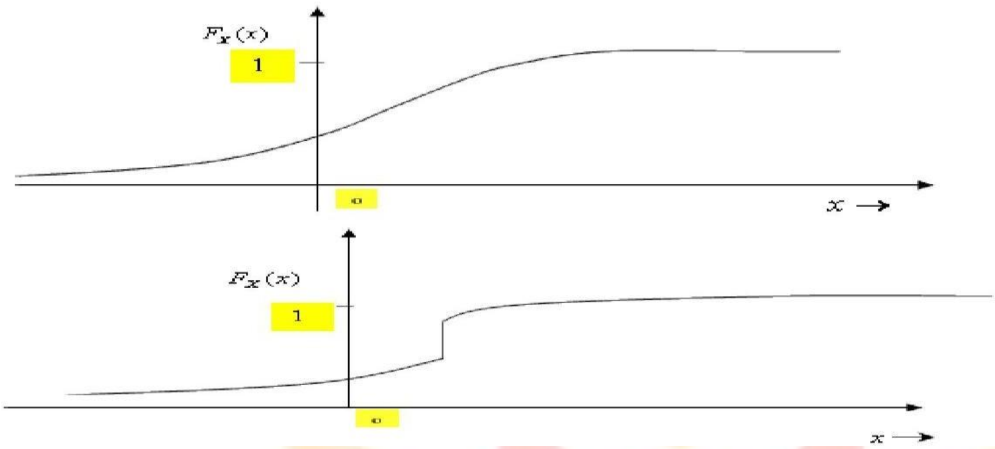
•A discrete random variable has a finite number of possible values or an infinite sequence of countable real numbers.
–X: number of hits when trying 20 free throws.
–X: number of customers who arrive at the bank from 8:30 – 9:30AM Mon--Fri.
–E.g. Binomial, Poisson...

Continuous random variable:

A continuous random variable is one which takes an infinite number of possible values. Continuous rando variables are usually measurements. E
A continuous random variable takes all values in an interval of real numbers. (or)
X is called a continuous random variable if is a n absolutely continuous function of x . Thus $F_X(x)$ is continuous everywhere on $\mathbb{R}$ and exists everywhere except at finite or countably infinite points

Mixed random variable:

X is called a mixed random variable if $F_X(x)$ has jump discontinuity at countable number of points and increases continuously at least in one interval of X. For a such type RV X.



Plot of Distribution function of continuous and mixed random variables

DISTRIBUTION AND DENSITY FUNCTIONS

Probability Distribution

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function.

More formally, the probability distribution of a discrete random variable X is a function which gives the probability p(xi) that the random variable equals xi, for each value xi:
p(xi) = P(X=xi)

It satisfies the following conditions:

a. $0 \leq p(xi) \leq 1$
b. $\sum p(xi) = 1$

Cumulative Distribution Function

All random variables (discrete and continuous) have a cumulative distribution function. It is a function giving the probability that the random variable X is less than or equal to x, for every value x.

Formally, the cumulative distribution function F(x) is defined to be:

$$F(x) = P(X \leq x)$$

for
$-\infty < x < \infty$

For a discrete random variable, the cumulative distribution function is found by summing up the probabilities as in the example below.

For a continuous random variable, the cumulative distribution function is the integral of its probability density function.

Example
Discrete case : Suppose a random variable X has the following probability distribution p(xi):

xi	0	1	2	3	4	5
p(xi)	1/32	5/32	10/32	10/32	5/32	1/32

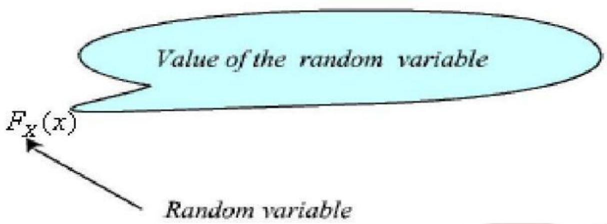
This is actually a binomial distribution: Bi(5, 0.5) or B(5, 0.5). The cumulative distribution function F(x) is then:

xi 0 1 2 3 4 5

F(xi) 1/32 6/32 16/32 26/32 31/32 32/32

F(x) does not change at intermediate values. For example: F(1.3) = F(1) = 6/32 and F(2.86) = F(2) = 16/32
Probability Distribution Function

The probability $P(\{X \le x\}) = P(\{s \mid X(s) \le x,\ s \in S\})$ is called the probability distribution function (also called the cumulative distribution function , abbreviated as CDF) of X and denoted by $F_X(x)$. Thus
$F_X(x) = P(\{X \le x\})$



Properties of the Distribution Function

1. $0 \le F_X(x) \le 1$

2. $F_X(x)$ is a non-decreasing function of X. Thus, if $x_1 < x_2$, then $F_X(x_1) < F_X(x_2)$

$$\begin{aligned}
& x_1 < x_2 \\
& \Rightarrow \{X(s) \le x_1\} \subseteq \{X(s) \le x_2\} \\
& \Rightarrow P\{X(s) \le x_1\} \le P\{X(s) \le x_2\} \\
& \therefore F_X(x_1) < F_X(x_2)
\end{aligned}$$

- $F_X(x)$ Is right continuous.

$$F_X(x^+) = \lim_{\substack{h \to 0 \\ h > 0}} F_X(x+h) = F_X(x)$$

$$\begin{aligned}
\text{Because, } \lim_{\substack{h \to 0 \\ h > 0}} F_X(x+h) &= \lim_{\substack{h \to 0 \\ h > 0}} P\{X(s) \le x+h\} \\
&= P\{X(s) \le x\} \\
&= F_X(x)
\end{aligned}$$

3. $F_X(-\infty) = 0$

. Because, $F_X(-\infty) = P\{s \mid X(s) \leq -\infty\} = P(\phi) = 0$

4. $F_X(\infty) = 1$

5. $P(\{x_1 < X \leq x_2\}) = F_X(x_2) - F_X(x_1) \leq \infty\} = P(S) = 1$

We have,

$$F_X(x^-) = \lim_{\substack{h \to 0 \\ h > 0}} F_X(x-h)$$

$$= \lim_{h \to 0} P\{X(s) \leq x-h\}$$

6. $P(\{X > x\}) = P(\{x < X < \infty\}) = 1 - F_X(x)$

$$= P\{X(s) \leq x\} - P(X(s) = x)$$

$$= F_X(x) - P(X = x)$$

Example: Consider the random variable X in the above example. We have

Value of the random Variable $X = x$	$P(\{X = x\})$
0	1/4
1	1/4
2	1/4
3	1/4

For $x < 0$,

$$F_X(x) = P(\{X \le x\}) = 0$$

For $0 \le x < 1$,

$$F_X(x) = P(\{X \le x\}) = P(\{X = 0\}) = \frac{1}{4}$$

For $1 \le x < 2$,

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(\{X = 0\} \cup \{X = 1\}) \\ &= P(\{X = 0\}) + P(\{X = 1\}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

For $2 \le x < 3$,

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\}) \\ &= P(\{X = 0\}) + P(\{X = 1\}) + P(\{X = 2\}) \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} \end{aligned}$$

For $x \ge 3$,

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(S) \\ &= 1 \end{aligned}$$

For $2 \le x < 3$,

$$F_X(x) = P(\{X \le x\})$$
$$= P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\})$$
$$= P(\{X = 0\}) + P(\{X = 1\}) + P(\{X = 2\})$$
$$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

Thus we have seen that given $F_X(x), \ -\infty < x < \infty$, we can determine the probability of any event involving values of the random variable X .Thus $F_X(x)\ \forall x \in X$ is a complete description of the random variable X .

Example 5 Consider the random variable X defined by

$$F_X(x) = 0, \qquad x < -2$$
$$= \frac{1}{8}x + \frac{1}{4}, \quad -2 \le x < 0$$
$$= 1, \qquad x \ge 0$$

Find a) $P(X = 0)$.

. $P\{X \le 0\}$

. $P\{X > 2\}$

. $P\{-1 < X \le 1\}$

Solutio

a) $P(X = 0) = F_X(0^+) - F_X(0^-)$

$$= 1 - \frac{1}{4} = \frac{3}{4}$$

b) $P\{X \le 0\} = F_X(0)$

$$= 1$$

c) $P\{X > 2\} = 1 - F_X(2)$

$$= 1 - 1 = 0$$

Probability Density Function

The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval.

More formally, the probability density function, f(x), of a continuous random variable X is the derivative of the cumulative distribution function F(x):

$$f(x) = \frac{d}{dx} F(x)$$

Since $F(x) = P(X < x)$ it follows that:

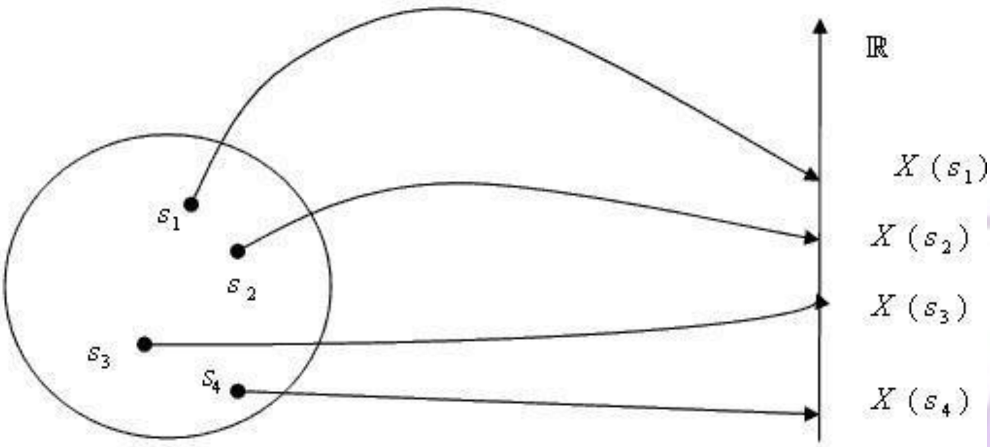
$$\int f(x)\, dx = F(b) - F(a) = P(a < X < b)$$

If f(x) is a probability density function then it must obey two conditions:

that the total probability for all possible values of the continuous random variable X is 1:

$$\int f(x)\, dx = 1$$

that the probability density function can never be negative: f(x) > 0 for all x.

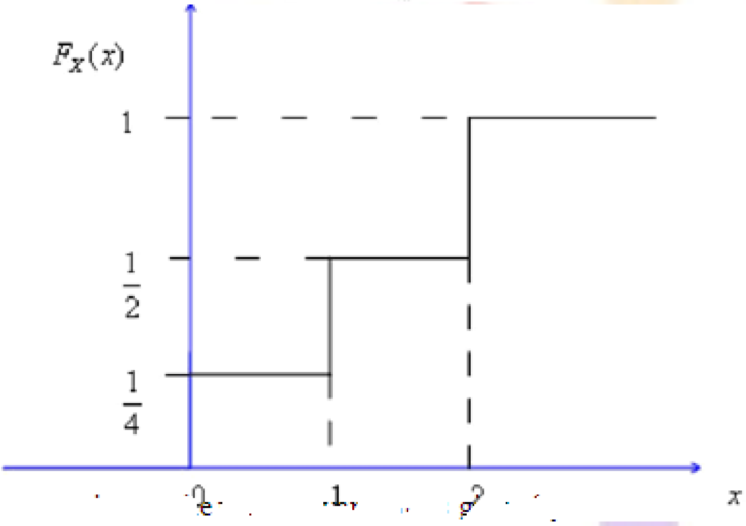


Example 1

Consider the random variable X with the distribution function

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{1}{2} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

The plot of the is shown in Figure 7 on next page.



The probability mass function of the random variable is given by

Value of the random variable $X = x$	$p_X(x)$
0	$\frac{1}{4}$
1	$\frac{1}{4}$
2	$\frac{1}{2}$

Properties of the Probability Density Function

1. $f_X(x) \geq 0$.------- This follows from the fact that $F_X(x)$ is a non-decreasing function

2. $F_X(x) = \int_{-\infty}^{x} f_X(u)du$

3. $\int_{-\infty}^{\infty} f_X(x)dx = 1$

4. $P(x_1 < X \leq x_2) = \int_{-x_1}^{x_2} f_X(x)dx$

Other Distribution and density functions of Random variable:

Bernoulli random variable:

Suppose X is a random variable that takes two values 0 and 1, with probability mass functions

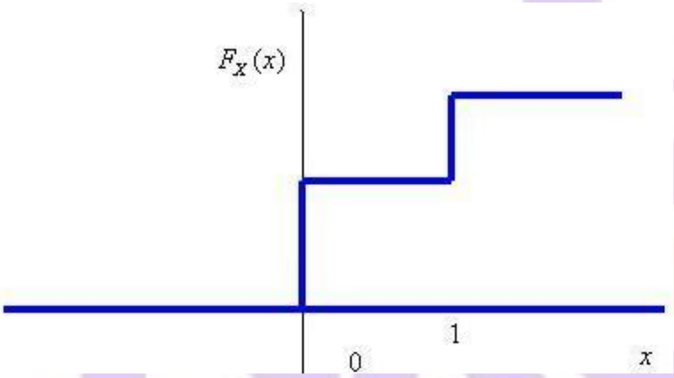
$p_X(1) = P\{X = 1\} = p$

And

$p_X(0) = 1 - p,$ $0 \leq p \leq 1$

Such a random variable X is called a Bernoulli random variable, because it describes the outcomes of a Bernoulli trial.

The typical CDF of the Bernoulli RV X is as shown in Figure 2



Mean and variance of the Bernoulli random:

$$\mu_X = EX = \sum_{k=0}^{1} k p_X(k) = 1\times p + 0\times(1-p) = p$$

$$EX^2 = \sum_{k=0}^{1} k^2 p_X(k) = 1\times p + 0\times(1-p) = p$$

$$\therefore \sigma_X^2 = EX^2 - \mu_X^2 = p(1-p)$$

Remark

The Bernoulli RV is the simplest discrete RV. It can be used as the building block for many discrete RVs.
For the Bernoulli RV,

$$EX^m = p \quad m = 1,2,3\ldots$$

Thus all the moments of the Bernoulli RV have the same value of p.

Binomial random variable
Suppose X is a discrete random variable taking values from the set $\{0,1,\ldots,n\}$. X is called a binomial random variable with parameters n and $0 \le p \le 1$ if

$$p_X(k) = {}^nC_k p^k (1-p)^{n-k} \quad k = 0,1,\ldots,n$$

where

$${}^nC_k = \frac{n!}{k!(n-k)!}$$

The trials must meet the following requirements:

the total number of trials is fixed in advance;
there are just two outcomes of each trial; success and failure;
the outcomes of all the trials are statistically independent;
all the trials have the same probability of success.

As we have seen, the probability of k successes in n independent repetitions of the Bernoulli trial is given by the binomial law. If X is a discrete random variable representing the number of successes in this case, then X is a binomial random variable. For example, the number of heads in _n
' independent tossing of a fair coin is a binomial random variable.

The notation $X \sim B(n,p)$ is used to represent a binomial RV with the parameters and n

p.

$$\sum_{k=0}^{n} p_X(k) = \sum_{k=0}^{n} {}^nC_k p^k (1-p)^{n-k} = [p+(1-p)]^n = 1.$$

The sum of n independent identically distributed Bernoulli random variables is a binomial random variable.
The binomial distribution is useful when there are two types of objects - good, bad; correct, erroneous; healthy, diseased etc

Example1:In a binary communication system, the probability of bit error is 0.01. If a block of 8 bits are transmitted, find the probability that

Exactly 2 bit errors will occur

At least 2 bit errors will occur

More than 2 bit errors will occur

All the bits will be erroneous

Suppose X is the random variable representing the number of bit errors in a block of 8 bits.
Then $X \sim B(8, 0.01)$.

Therefore,

(a) Probability that exactly 2 bit errors will occur

$$= p_X(2)$$
$$= {}^8C_2 \times 0.01^2 \times 0.99^6$$
$$= 0.0026$$

(b) Probability that at least 2 bit errors will occur

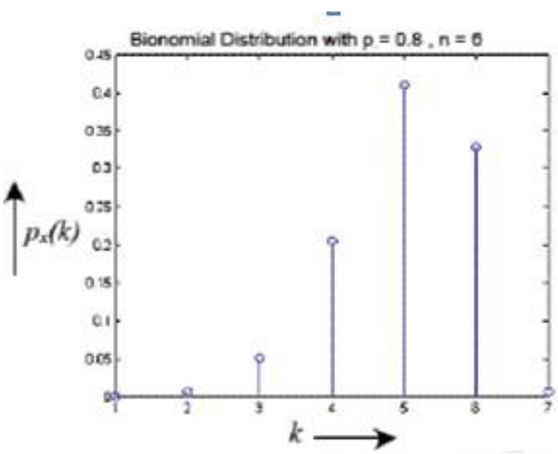
$$= p_X(0) + p_X(1) + p_X(2)$$
$$= 0.99^8 + {}^8C_1 \times 0.01^1 \times 0.99^7 + {}^8C_2 \times 0.01^2 \times 0.99^6$$
$$= 0.9999$$

(c) Probability that more than 2 bit errors will occur

$$= 1 - \sum_{k=0}^{2} p_X(k)$$
$$= 1 - 0.9999$$
$$= 0.0001$$

(d) Probability that all 8 bits will be erroneous

$$= p_X(8)$$
$$= 0.01^8 = 10^{-16}$$



The probability mass function for a binomial random variable with n = 6 and p =0.8

Mean and Variance of the Binomial Random Variable

We have

$$EX = \sum_{k=0}^{n} k p_X(k)$$

$$= \sum_{k=0}^{n} k\, {}^nC_k p^k (1-p)^{n-k}$$

$$= 0 \times q^n + \sum_{k=1}^{n} k\, {}^nC_k p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{n-1!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-1-k_1}$$

$$= np \sum_{k_1=0}^{n-1} \frac{n-1!}{k_1!(n-1-k_1)!} p^{k_1} (1-p)^{n-1-k_1} \quad \text{(Substituting } k_1 = k-1\text{)}$$

$$= np(p+1-p)^{n-1}$$

$$= np$$

Similarly

$$EX^2 = \sum_{k=0}^{n} k^2 p_X(k)$$
$$= \sum_{k=0}^{n} k^2\, {}^nC_k p^k (1-p)^{n-k}$$
$$= 0^2 \times q^n + \sum_{k=1}^{n} k\, {}^nC_k p^k (1-p)^{n-k}$$
$$= \sum_{k=1}^{n} k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$
$$= \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$
$$= np \sum_{k=1}^{n} (k-1+1) \frac{n-1!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-1-(k-1)}$$
$$= np \sum_{k=1}^{n} (k-1) \frac{n-1!}{(k-1)!(n-1-k+1)!} p^{k-1} (1-p)^{n-1-(k-1)} + np \sum_{k=1}^{n} \frac{n-1!}{(k-1)!(n-1-k+1)!} p^{k-1} (1-p)^{n-1-(k-1)}$$
$$= np \times (n-1)p + np$$
$$= n(n-1)p^2 + np$$

Where

$$\sum_{k=1}^{n} (k-1) \frac{(n-1)!}{(k-1)!(n-1-k+1)!} p^{k-1} (1-p)^{n-1-(k-1)}$$

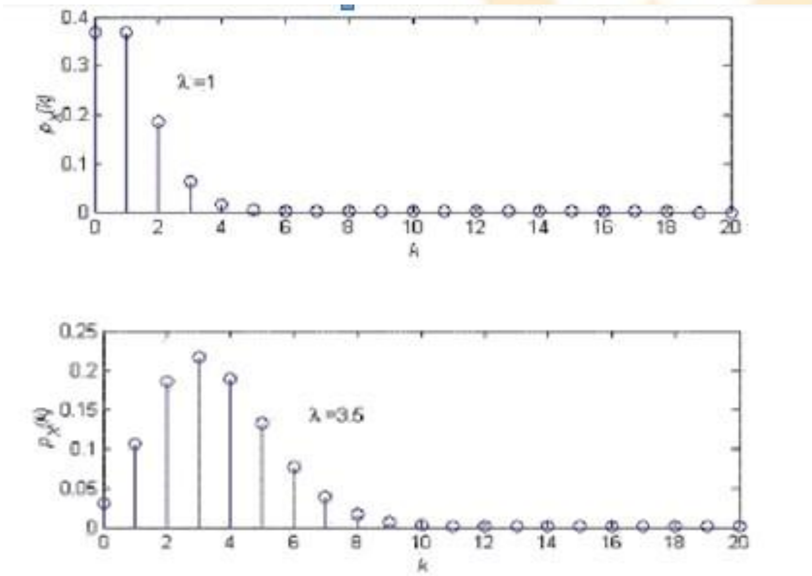
is the mean of $B(n-1,p)$

$$\therefore \sigma_X^2 = \text{variance of } X$$
$$= n(n-1)p^2 + np - n^2p^2$$
$$= np(1-p)$$

Poisson Random Variable

A discrete random variable X is called a Poisson random variable with the parameter λ if $\lambda > 0$ and
$P_X(k) = (e^{-\lambda}\lambda^k)/k!$

The plot of the pmf of the Poisson RV is shown



Mean and Variance of the Poisson RV

The mean of the Poisson RV X is given by

$$\mu_X = \sum_{k=0}^{\infty} k p_X(k)$$
$$= 0 + \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!}$$
$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{k-1!}$$
$$= \lambda$$

$$EX^2 = \sum_{k=0}^{\infty} k^2 p_X(k)$$
$$= 0 + \sum_{k=1}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} \frac{k\lambda^k}{k-1!}$$
$$= e^{-\lambda} \sum_{k=1}^{\infty} \frac{(k-1+1)\lambda^k}{k-1!}$$

$$= e^{-\lambda}\left(0 + \sum_{k=2}^{\infty} \frac{\lambda^k}{k-2!}\right) + e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k-1!}$$
$$= e^{-\lambda}\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{k-2!} + e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{k-1!}$$

$$= e^{-\lambda}\lambda^2 e^{\lambda} + e^{-\lambda}\lambda e^{\lambda}$$
$$= \lambda^2 + \lambda$$
$$\therefore \sigma_X^2 = EX^2 - \mu_X^2 = \lambda$$

Example:The number of calls received in a telephone exchange follows a Poisson distribution with an average of 10 calls per minute. What is the probability that in one-minute duration?

no call is received
exactly 5 calls are received
More than 3 calls are received.

Solution: Let X be the random variable representing the number of calls received. Given

Where Therefore, $\lambda = 10.$

probability that no call is received $= p_X(0) = e^{-10} =$ 0.000095

probability that exactly 5 calls are received $= p_X(5) = \frac{e^{-10} \times 10^5}{5!} =$ 0.0378

probability that more the 3 calls are received

$$= 1 - \sum_{k=0}^{5} p_X(k) = 1 - e^{-10}\left(1 + \frac{10}{1} + \frac{10^2}{2!} + \frac{10^3}{3!}\right) =$$ 0.9897

Poisson Approximation of the Binomial Random Variable

The Poisson distribution is also used to approximate the binomial distribution $B(n,p)$ when n is very large and p is small.

Consider binomial RV with $X \sim B(n,p)$ with

$n \to \infty,\ p \to 0$ so that $EX = np = \lambda$ remains constant.

Then

$$p_X(k) = {}^nC_k p^k (1-p)^{n-k}$$

$$= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} p^k (1-p)^{n-k}$$

I

$$= \frac{n^k(1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{k-1}{n})}{k!} p^k (1-p)^{n-k}$$

$$= \frac{(1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{k-1}{n})}{k!} (np)^k (1-p)^{n-k}$$

$$= \frac{n^k (1-\frac{1}{n})(1-\frac{2}{n})\dots(1-\frac{k-1}{n})}{k!} p^k (1-p)^{n-k}$$

$$= \frac{(1-\frac{1}{n})(1-\frac{2}{n})\dots(1-\frac{k-1}{n})}{k!} (np)^k (1-p)^{n-k}$$

$$= \frac{(1-\frac{1}{n})(1-\frac{2}{n})\dots(1-\frac{k-1}{n})(\lambda)^k (1-\frac{\lambda}{n})^n}{k!(1-\frac{\lambda}{n})^k}$$

Note that $\lim_{n\to\infty}(1-\frac{\lambda}{n})^n = e^{-\lambda}$.

$$\therefore p_X(k) = \lim_{n\to\infty} \frac{(1-\frac{1}{n})(1-\frac{2}{n})\dots(1-\frac{k-1}{n})(\lambda)^k (1-\frac{\lambda}{n})^n}{k!(1-\frac{\lambda}{n})^k} = \frac{e^{-\lambda}\lambda^k}{k!}$$

Thus the Poisson approximation can be used to compute binomial probabilities for large n. It also makes the analysis of such probabilities easier. Typical examples are:

number of bit errors in a received binary data file
Number of typographical errors in a printed page

Example 4 Suppose there is an error probability of 0.01 per word in typing. What is the probability that there will be more than 1 error in a page of 120 words?

Solution: Suppose X is the RV representing the number of errors per page of 120 words.

$X \sim B(120, p)$ Where $p = 0.01$. Therefore,

$$\therefore \lambda = 120 \times 0.01 = 0.12$$
$$\text{P(more than one errors)}$$
$$= 1 - p_X(0) - p_X(1)$$
$$\simeq 1 - e^{-\lambda} - \lambda e^{-\lambda}$$
$$= 0.0066$$

In the following we shall discuss some important continuous random variables.

Uniform Random Variable

A continuous random variable X is called uniformly distributed over the interval [a, b],

, if its probability density function is given by
$-\infty < a < b < \infty$

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

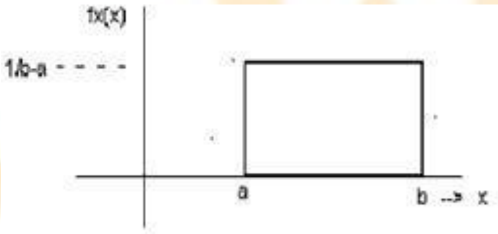


Figure 1

We use the notation $X \sim U(a,b)$ to denote a random variable X uniformly distributed over the interval [a,b]

Distribution function $F_X(x)$

For $x < a$

$F_X(x) = 0$

For $a \le x \le b$

$$\int_{-\infty}^{x} f_X(u)du$$

$$= \int_{a}^{x} \frac{du}{b-a}$$

$$= \frac{x-a}{b-a}$$

For $x > b$,

$F_X(x) = 1$

Figure 2 illustrates the CDF of a uniform random variable.

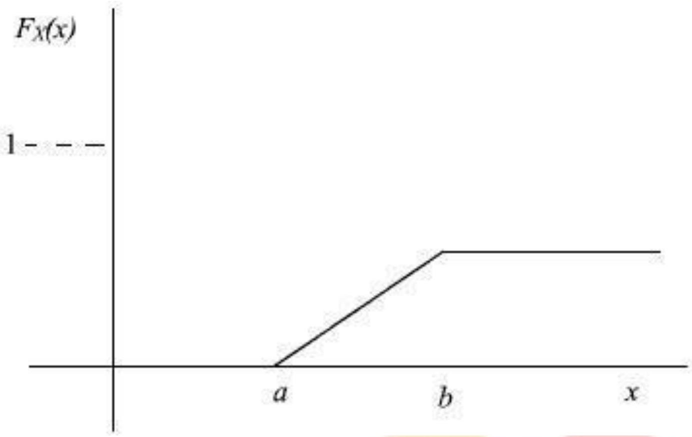


Figure 2: CDF of a uniform random variable

Mean and Variance of a Uniform Random Variable:

$$\mu_X = EX = \int_{-\infty}^{\infty} x f_X(x)dx = \int_a^b \frac{x}{b-a}dx$$

$$= \frac{a+b}{2}$$

$$EX^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_a^b \frac{x^2}{b-a}dx$$

$$= \frac{b^2+ab+a^2}{3}$$

$$\therefore \sigma_X^2 = EX^2 - \mu_X^2 = \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{(b-a)^2}{12}$$

The characteristic function of the random variable $X \sim U(a,b)$ is given by

$$\phi_X(w) = Ee^{jwx} = \int_a^b \frac{e^{jwx}}{b-a}dx$$

$$= \frac{e^{jwb} - e^{jwa}}{jw(b-a)}$$

Normal or Gaussian Random Variable

The normal distribution is the most important distribution used to model natural and man made phenomena. Particularly, when the random variable is the result of the addition of large number of independent random variables, it can be modeled as a normal random variable.

continuous random variable X is called a normal or a Gaussian random variable with parameters μ_X and σ_X^2 if its probability density function is given by,

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2}, \qquad -\infty < x < \infty$$

Where μ_X and $\sigma_X > 0$ are real numbers.
We write that X is $N(\mu_X, \sigma_X^2)$ distributed.
If $\mu_X = 0$ and $\sigma_X^2 = 1$, and the random variable X is called the standard normal variable.

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

Figure 3 illustrates two normal variables with the same mean but different variances.

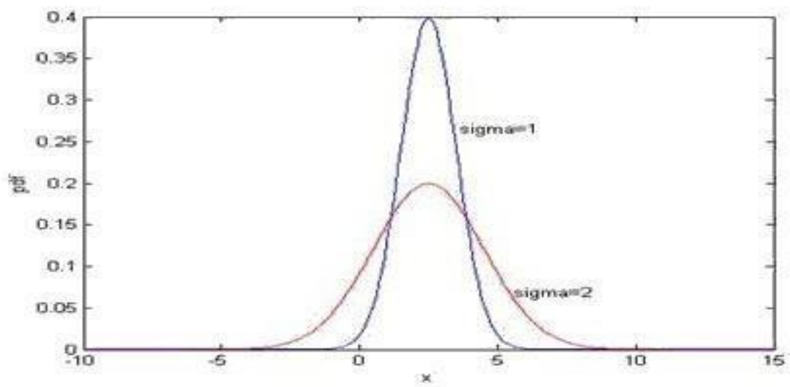


Figure 3

$f_X(x)$

σ_X^2

Is a bell-shaped function, symmetrical about $x = \mu_X$.

Determines the spread of the random variable X . If σ_X is small X is more concentrated around the mean μ_X.

Distribution function of a Gaussian variable is

$$F_X(x) = P(X \le x)$$
$$= \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu_X}{\sigma_X}\right)^2} dt$$

Substituting $u = \frac{t-\mu_X}{\sigma_X}$, we get

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu_X}{\sigma_X}} e^{-\frac{1}{2}u^2} du$$
$$= \Phi\left(\frac{x-\mu_X}{\sigma_X}\right)$$

where $\Phi(x)$ is the distribution function of the standard normal variable.

Thus $F_X(x)$ can be computed from tabulated values of $\Phi(x)$. The table $\Phi(x)$ was very useful in the pre-computer days.

In communication engineering, it is customary to work with the Q function defined by,

$$Q(x) = 1 - \Phi(x)$$
$$= \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\frac{u^2}{2}} du$$

Note that $Q(0) = \frac{1}{2}$, and $Q(-x) = Q(x)$

$$Q(x) = 1 - \phi(-x)$$

These results follow from the symmetry of the Gaussian pdf. The function $Q(x)$ is tabulated and the tabulated results are used to compute probability involving the Gaussian random variable.

Using the Error Function to compute Probabilities for Gaussian Random Variables

The function $Q(x)$ is closely related to the error function $erf(x)$ and the complementary error function .

Note that, $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du$

And the complementary error function $erfc(x)$ is given

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

Mean and Variance of a Gaussian Random Variable

If X is $N(\mu_X, \sigma_X^2)$ distributed, then

$$EX = \mu_X$$
$$\text{var}(X) = \sigma_X^2$$

Proof:

$$EX = \int_{-\infty}^{\infty} x f_X(x) dx = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} dx$$
$$= \frac{\sigma_X}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (u\sigma_X + \mu_X) e^{-\frac{1}{2}u^2} \sigma_X du$$
$$= \frac{1}{\sigma_X\sqrt{2\pi}} \int_{-\infty}^{\infty} u du + \frac{\mu_X}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du$$
$$= 0 + \mu_X \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du$$
$$= \frac{\mu_X}{\sqrt{2\pi}} 2\int_0^\infty e^{-\frac{u^2}{2}} du = \mu_X$$

Substituting $\frac{x-\mu_X}{\sigma_X} = u$

so that $x = u\sigma_X + \mu_X$

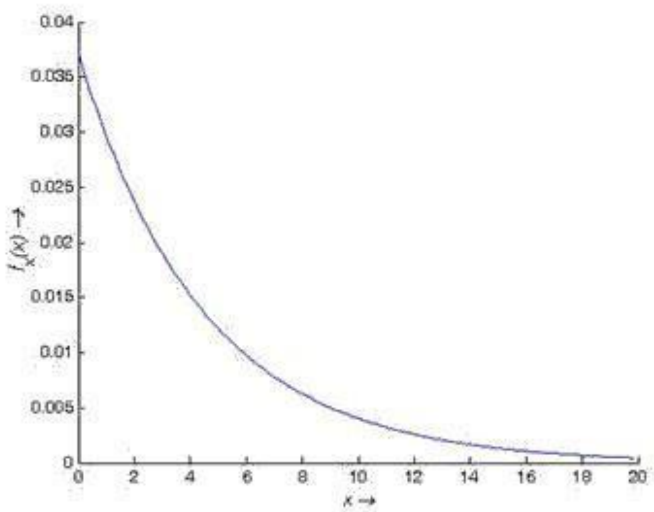
$$
\begin{aligned}
Var(X) &= E\left(X-\mu_X\right)^2 \\
&= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{\infty}(x-\mu_X)^2 e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2}dx \\
&= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{\infty}\sigma_X^2 u^2 e^{-\frac{1}{2}u^2}\sigma_X du \qquad (\text{substituting } u = \frac{x-\mu_X}{\sigma_X}) \\
&= 2\times\frac{\sigma_X^2}{\sqrt{2\pi}}\int_0^{\infty}u^2 e^{-\frac{1}{2}u^2}du \\
&= 2\times\frac{\sigma_X^2}{\sqrt{2\pi}}\sqrt{2}\int_0^{\infty}t^{\frac{1}{2}}e^{-t}dt \qquad (\text{substituting } t = \frac{u^2}{2}) \\
&= 2\times\frac{\sigma_X^2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) \\
&= 2\times\frac{\sigma_X^2}{\sqrt{\pi}}\frac{1}{2}\Gamma\left(\frac{1}{2}\right) \\
&= \frac{\sigma_X^2}{\sqrt{\pi}}\times\sqrt{\pi} \\
&= \sigma_X^2
\end{aligned}
$$

Exponential Random Variable

A continuous random variable X is called exponentially distributed with the parameter

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

if the probability density function is of the PDF of Exponential Random Variable is $\lambda > 0$



Example 1

Suppose the waiting time of packets in X in a computer network is an exponential RV with

$$f_X(x) = 0.5e^{-0.5x} \qquad x \geq 0$$

Then,

$$P(\{0.1 < X \leq 0.5\}) = \int_{0.1}^{0.5} 0.5e^{-0.5x}dx$$

$$= e^{-0.5\times0.5} - e^{-0.5\times0.1}$$

$$= 0.0241$$

Rayleigh Random Variable

A Rayleigh random variable X is characterized by the PDF

$$f_X(x) = \begin{cases} \dfrac{x}{\sigma^2}e^{-x^2/2\sigma^2}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where σ is the parameter of the random variable.

probability density functions for the Rayleigh RVs are illustrated in Figure

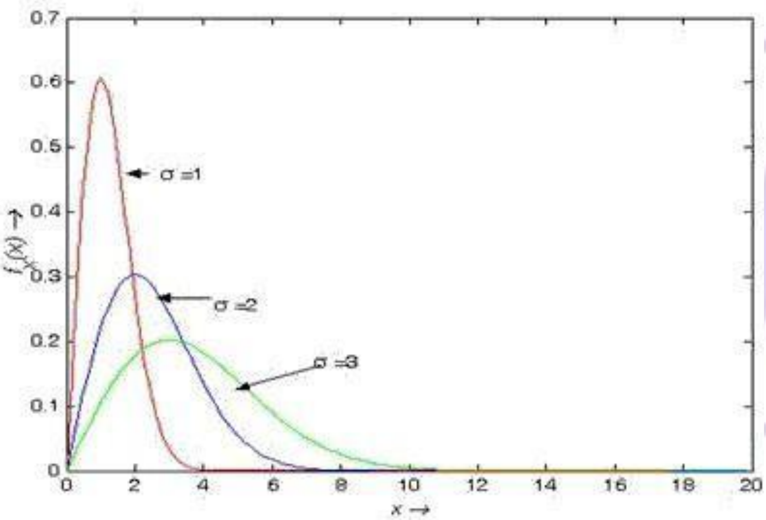


Figure 6

Mean and Variance of the Rayleigh Distribution

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x f_X(x)dx \\ &= \int_0^{\infty} x \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} dx \\ &= \frac{\sqrt{2\pi}}{\sigma} \int_0^{\infty} \frac{x^2}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx \\ &= \frac{\sqrt{2\pi}}{\sigma} \frac{\sigma^2}{2} \\ &= \sqrt{\frac{\pi}{2}}\sigma \end{aligned}$$

similarly

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 f_X(x)dx \\ &= \int_0^{\infty} x^2 \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} dx \\ &= 2\sigma^2 \int_0^{\infty} u e^{-u} du \quad (\textbf{Substituting } u = \frac{x^2}{2\sigma^2}) \\ &= 2\sigma^2 \quad (\textbf{Noting that } \int_0^{\infty} u e^{-u} du \textbf{ is the mean of the exponential RV with } \lambda = 1) \end{aligned}$$

$$\begin{aligned} \therefore \sigma_X^2 &= 2\sigma^2 - \left(\sqrt{\frac{\pi}{2}}\sigma\right)^2 \\ &= (2 - \frac{\pi}{2})\sigma^2 \end{aligned}$$

Relation between the Rayleigh Distribution and the Gaussian distribution

A Rayleigh RV is related to Gaussian RVs as follow: If $X_1 \sim N(0, \sigma^2)$ and $X_2 \sim N(0, \sigma^2)$ are independent, then the envelope $X = \sqrt{X_1^2 + X_2^2}$ has the Rayleigh distribution with the parameter σ.

We shall prove this result in a later lecture. This important result also suggests the cases where the Rayleigh RV can be used.

Application of the Rayleigh RV

- Modeling the root mean square error-
- Modeling the envelope of a signal with two orthogonal components as in the case of a signal of the following form:

Conditional Distribution and Density functions

We discussed conditional probability in an earlier lecture. For two events A and B with $P(B) \neq 0$, the conditional probability $P(A/B)$ was defined as

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Clearly, the conditional probability can be defined on events involving a Random Variable X

Conditional distribution function:

Consider the event $\{X \leq x\}$ and any event B involving the random variable X . The conditional distribution function of X given B is defined as

$$F_X(x/B) = P[\{X \leq x\}/B]$$
$$= \frac{P[\{X \leq x\} \cap B]}{P(B)} \qquad P(B) \neq 0$$

We can verify that $F_X(x/B)$ satisfies all the properties of the distribution function. Particularly.

- $F_X(-\infty/B) = 0$ And $F_X(\infty/B) = 1$.
- $0 \leq F_X(x/B) \leq 1$
- $F_X(x/B)$ Is a non-decreasing function of x.
- $$P(\{x_1 < X \leq x_2\}/B) = P(\{X \leq x_2\}/B) - P(\{X \leq x_1\}/B)$$
$$= F_X(x_2/B) - F_X(x_1/B)$$

Conditional Probability Density Function

In a similar manner, we can define the conditional density function $f_X(x/B)$ of the random variable X given the event B as

$$f_X(x/B) = \frac{d}{dx} F_X(x/B)$$

All the properties of the pdf applies to the conditional pdf and we can easily show that

- $f_X(x/B) \geq 0$
- $\int_{-\infty}^{\infty} f_X(x/B)dx = F_X(\infty/B) = 1$
- $F_X(x/B) = \int_{-\infty}^{x} f_X(u/B)du$

$$P(\{x_1 < X \leq x_2\}/B) = F_X(x_2/B) - F_X(x_1/B) = \int_{x_1}^{x_2} f_X(x/B)dx$$

Example 1 Suppose X is a random variable with the distribution function $F_X(x)$. Define $B = \{X \leq b\}$

$$F_X(x/B) = \frac{P(\{X \leq x\} \cap B)}{P(B)} = \frac{P(\{X \leq x\} \cap \{X \leq b\})}{P\{X \leq b\}} = \frac{P(\{X \leq x\} \cap \{X \leq b\})}{F_X(b)}$$

Case 1: $x < b$

Then

$$F_X(x/B) = \frac{P(\{X \leq x\} \cap \{X \leq b\})}{F_X(b)} = \frac{P(\{X \leq x\})}{F_X(b)} = \frac{F_X(x)}{F_X(b)}$$

And

$$f_X(x/B) = \frac{d}{dx}\frac{F_X(x)}{F_X(b)} = \frac{f_X(x)}{F_X(b)}$$

Case 2: $x \geq b$

$$F_X(x/B) = \frac{P(\{X \leq x\} \cap \{X \leq b\})}{F_X(b)}$$

$$= \frac{P(\{X \leq b\})}{F_X(b)} = \frac{F_X(b)}{F_X(b)} = 1$$

$F_X(x/B)$ and $f_X(x/B)$ are plotted in the following figures.

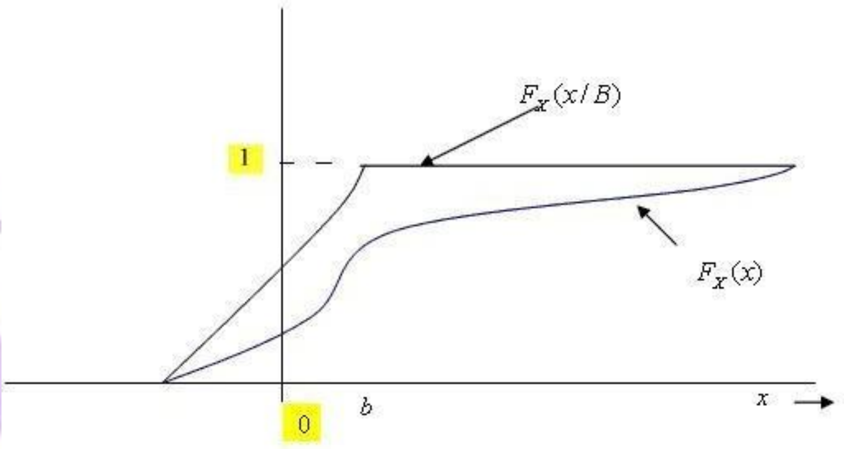
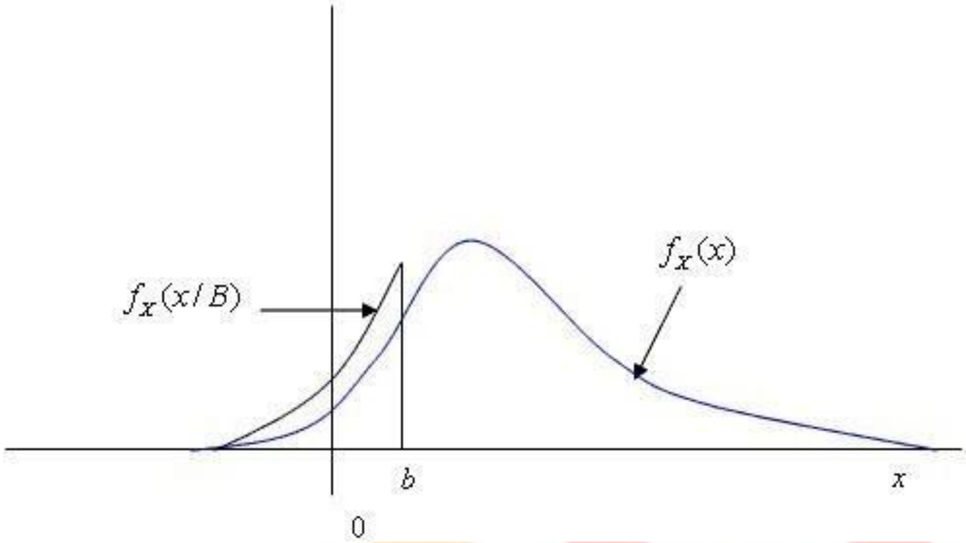


Figure 1


$f_X(x)$
$f_X(x/B)$
b
x
0

Example2: Suppose X is a random variable with the distribution function and $F_X(x)$

. $B = \{X > b\}$

$$F_X(x/B) = \frac{P(\{X \le x\} \cap B)}{P(B)}$$

$$= \frac{P(\{X \le x\} \cap \{X > b\})}{P\{X > b\}}$$

$$= \frac{P(\{X \le x\} \cap \{X > b\})}{1 - F_X(b)}$$

Then

For $x \le b$ $\{X \le x\} \cap \{X > b\} = \phi$ .Therefore,

$$F_X(x/B) = 0 \qquad x \le b$$

For $x > b$, $\{X \le x\} \cap \{X > b\} = \{b < X \le x\}$ .Therefore,

$$F_X(x/B) = \frac{P(\{b < X \le x\})}{1 - F_X(b)}$$

$$= \frac{F_X(x) - F_X(b)}{1 - F_X(b)}$$

Thus,

$$F_X(x/B) = \begin{cases} 0 \quad , & x \le b \\ \dfrac{F_X(x) - F_X(b)}{1 - F_X(b)} \quad , & \text{otherwise} \end{cases}$$

the corresponding pdf is given by

$$f_X(x/B) = \begin{cases} 0 \quad , & x \le b \\ \dfrac{f_X(x)}{1 - F_X(b)} & \text{otherwise} \end{cases}$$