Probability theory and stochastic process (AECB08)

B.Tech ECE III Semester

Prepared by

Dr. M V Krishna Rao

G. Ajitha

N.Nagaraju

Module-I

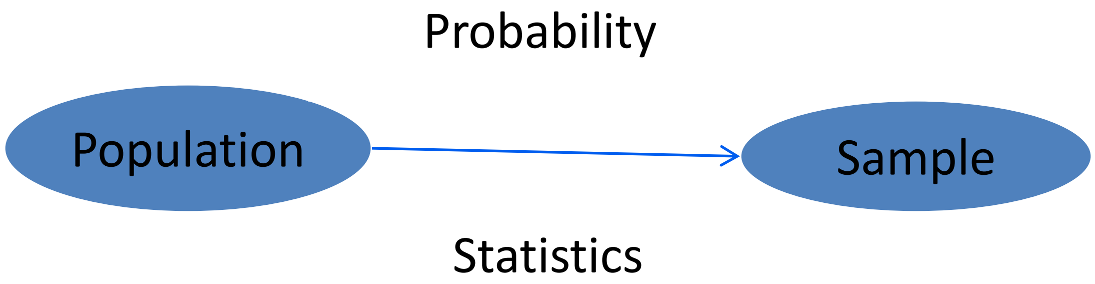
PROBABILITY, RANDOM VARIABLES AND OPERATIONS ON RANDOM VARIABLES

Index

- Why study probability
- What is probability
- Basic concepts
  - Relative frequency
  - Experiment
  - Sample space
  - Events
  - Probability definitions and axioms

Why study Probability?

- Nothing in life is certain. In everything we do, we gauge the chances of successful outcomes, from business to medicine
- A probability provides a quantitative description of the chances or likelihoods associated with various outcomes
- It provides a bridge between descriptive and inferential statistics



Probabilistic vs Statistical Reasoning

- Suppose I know exactly the proportions of car makes in India. Then I can find the probability that the first car I see in the street is a Maruthi. This is probabilistic reasoning as I know the population and predict the sample

- Now suppose that I do not know the proportions of car makes in India, but would like to estimate them. I observe a random sample of cars in the street and then I have an estimate of the proportions of the population. This is statistical reasoning

What is Probability?

- We use graphs and numerical measures to describe data sets which were usually called samples.
- We measured "how often" using

Relative frequency = *f/n*

- As *n* gets larger,

Sample → Population

And "How often" = Relative frequency → Probability

Basic Concepts

- An experiment is the process by which an observation (or measurement) is obtained.

- An event is an outcome of an experiment, usually denoted by a capital letter.

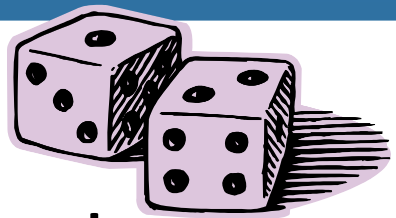
  - The basic element to which probability is applied
  - When an experiment is performed, a particular event either happens, or it doesn't!

Experiments and Events

- Experiment: Record an age
  - A: person is 28 years old
  - B: person is older than 55

- Experiment: Toss a die
  - A: observe an even number
  - B: observe a number greater than 4

Basic Concepts



- Two events are **mutually exclusive** if, when one event occurs, the other cannot, and vice versa.

- **Experiment: Toss a die**
  - A: observe an odd number
  - B: observe a number greater than 2
  - C: observe a 6
  - D: observe a 3

Not Mutually Exclusive

Mutually Exclusive

B and C?
B and D?

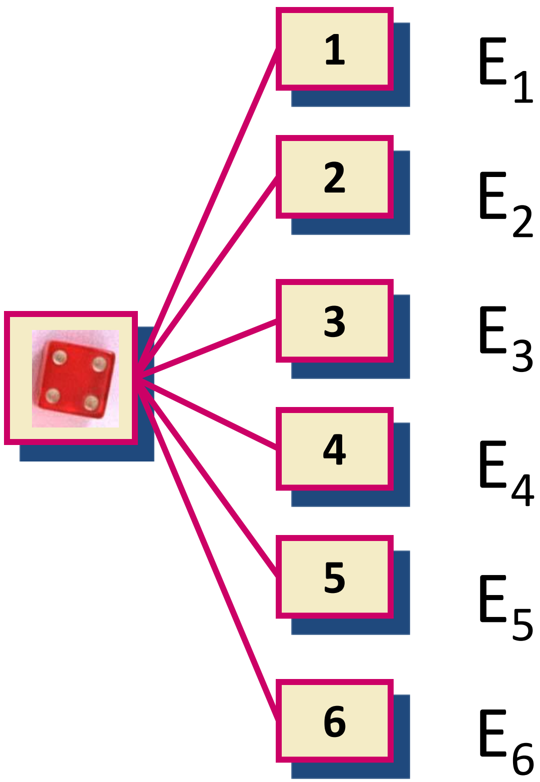
Basic Concepts

- An event that cannot be decomposed is called a simple event.
- Denoted by E with a subscript.
- Each simple event will be assigned a probability, measuring "how often" it occurs.
- The set of all simple events of an experiment is called the sample space, S.
- For a die tossing experiment the sample space consists of six events of discrete and finite, it is called discrete sample space (shown in next slide)

Example

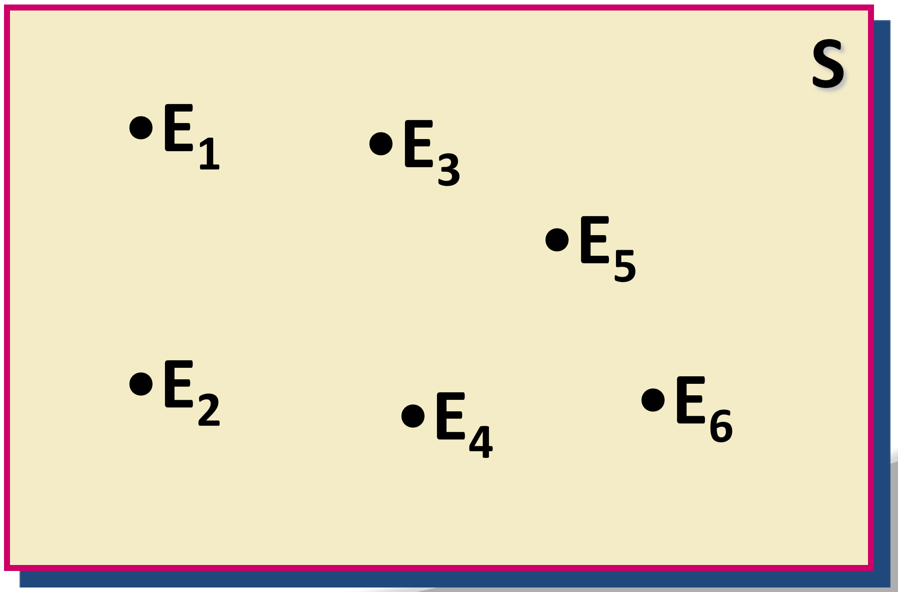
- **The die toss:**

Simple events:



Sample space:

$$S = \{E_1, E_2, E_3, E_4, E_5, E_6\}$$



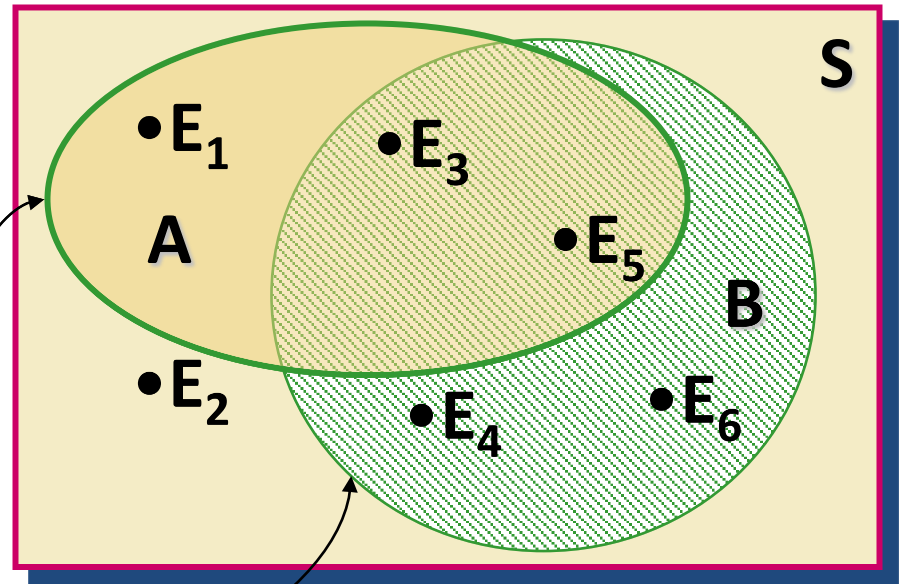
Basic Concepts

- **An event is a collection of one or more simple events.**

- **The die toss:**
  - –A: an odd number
  - –B: a number > 2

A ={E_1, E_3, E_5}

B ={E_3, E_4, E_5, E_6}



The Probability of an Event

- The probability of an event A measures "how often" A will occur. We write P(A).
- Suppose that an experiment is performed *n* times. The relative frequency for an event A is

$$\frac{\text{Number of times A occurs}}{n} = \frac{f}{n}$$

- If we let *n* get infinitely large,

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

The Probability of an Event

- P(A) must be between 0 and 1.
  - If event A can never occur, P(A) = 0. If event A always occurs when the experiment is performed, P(A) =1. (axiom-1)
- The sum of the probabilities for all simple events in S equals 1. (axiom-2)

• The **probability of an event A** is found by adding the probabilities of all the simple events contained in A. (axiom-3)

Finding Probabilities

- Probabilities can be found using
  - Estimates from empirical studies
  - Common sense estimates based on equally likely events.

- **Examples:**
  - –Toss a fair coin.

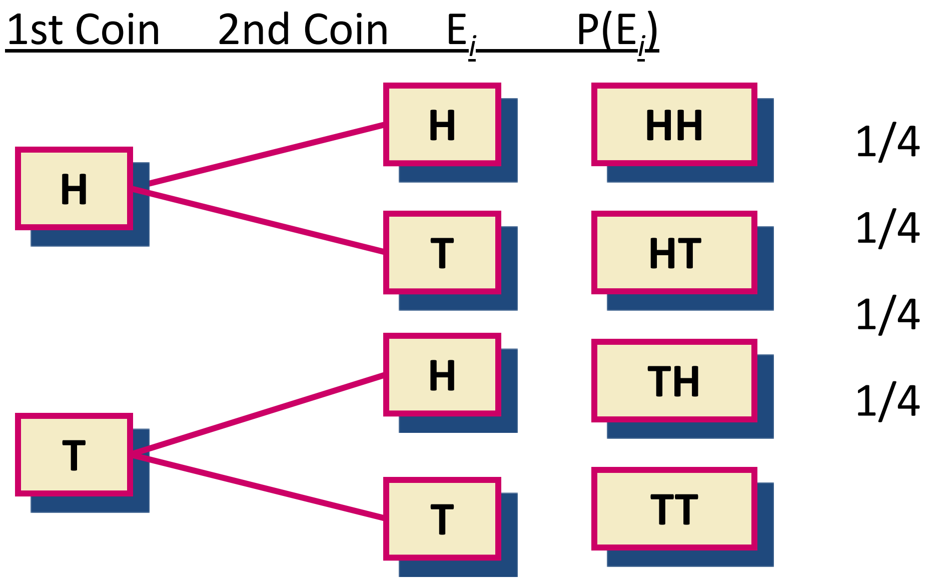
P(Head) = 1/2

  - – Suppose that 10% of a city population has red hair. Then for a person selected at random

P(Red hair) = .10

Example 1

Toss a fair coin twice. What is the probability of observing at least one head?



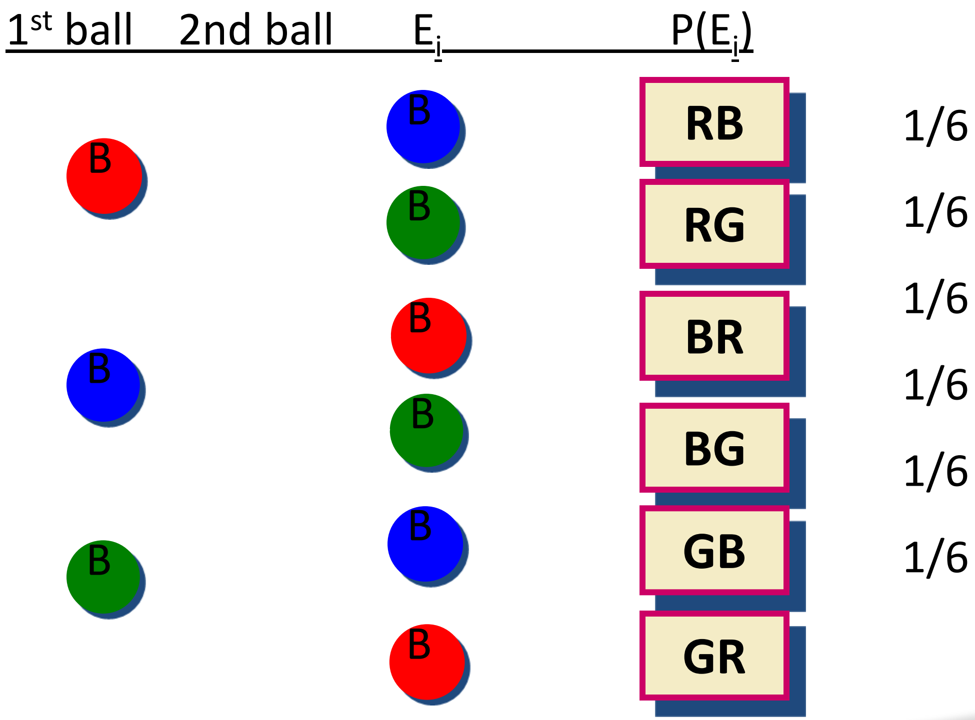
P(at least 1 head)

$= P(E_1) + P(E_2) + P(E_3)$

$= 1/4 + 1/4 + 1/4 = 3/4$

Example 2

A bowl contains three balls, one red, one blue and one green. A child selects two balls at random. What is the probability that at least one is red?

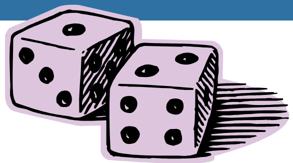


P(at least 1 red)

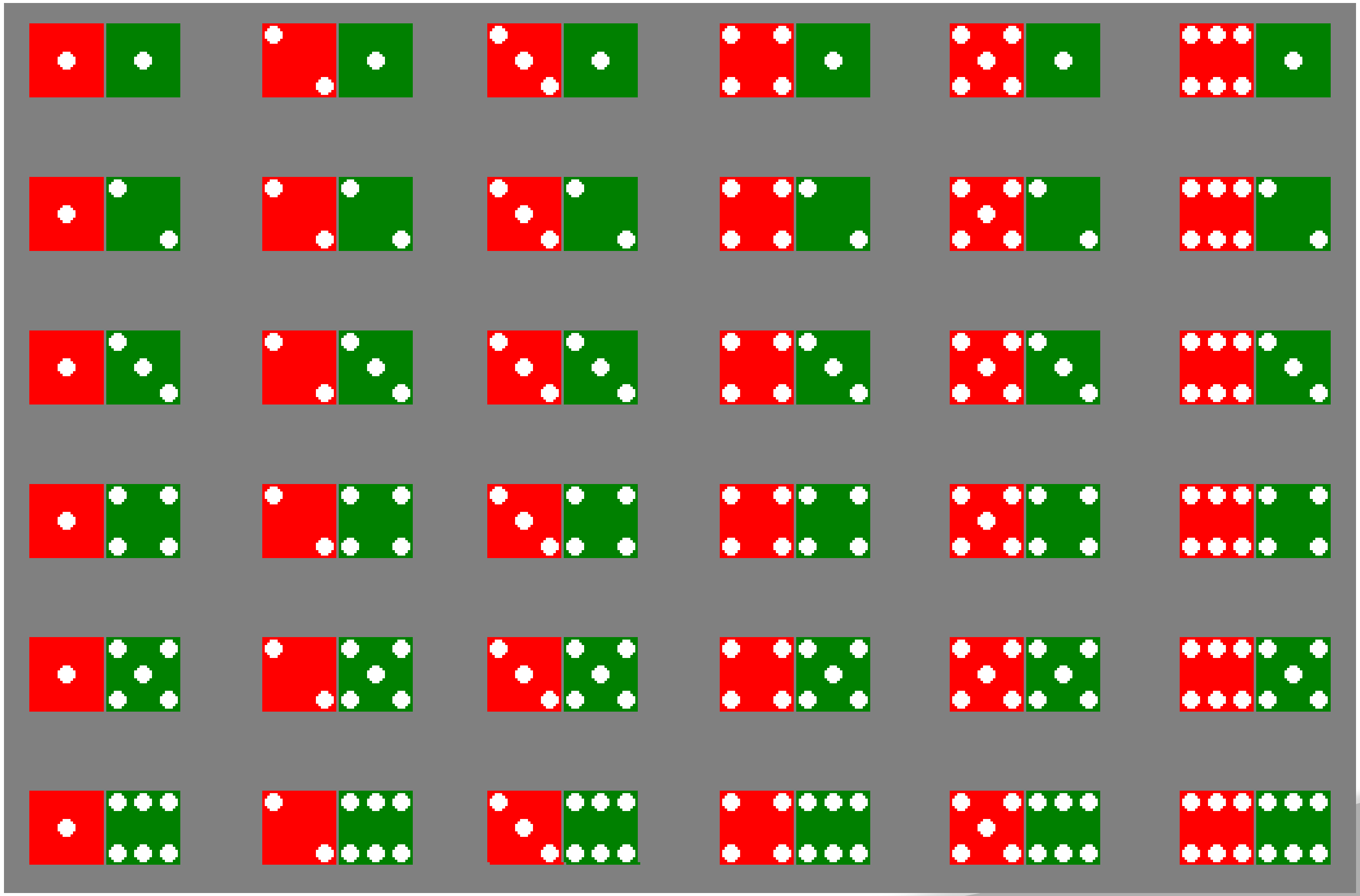
= P(RB) + P(BR)+ P(RG) + P(GR)

= 4/6 = 2/3

Example 3



The sample space of throwing a pair of dice is



Example 3 (contd..)

Event	Simple events	Probability
Dice add to 3	(1,2),(2,1)	2/36
Dice add to 6	(1,5),(2,4),(3,3), (4,2),(5,1)	5/36
Red die show 1	(1,1),(1,2),(1,3), (1,4),(1,5),(1,6)	6/36
Green die show 1	(1,1),(2,1),(3,1), (4,1),(5,1),(6,1)	6/36

Index

- Permutations and Combinations
- Event relations
- Conditional probability
- Total probability
- Independent events
- Bay’s theorem

Counting Rules

- Sample space of throwing 3 dice has 216 entries,
- Sample space of throwing 4 dice has 1296 entries, …
- At some point, we have to stop listing and start thinking …
- We need some counting rules

The *mn* Rule

- **If an experiment is performed in two stages, with *m* ways to accomplish the first stage and *n* ways to accomplish the second stage, then there are *mn* ways to accomplish the experiment.**
- **This rule is easily extended to *k* stages, with the number of ways equal to**

$$n_1 n_2 n_3 \ldots n_k$$

Example: Toss two coins. The total number of simple events is:

$$2 \times 2 = 4$$

Examples

Example: Toss three coins. The total number of simple events is:

$$2 \times 2 \times 2 = 8$$

Example: Toss two dice. The total number of simple events is:

$$6 \times 6 = 36$$

Example: Toss three dice. The total number of simple events is:

$$6 \times 6 \times 6 = 216$$

Example: Two balls are drawn from a dish containing two red and two blue balls. The total number of simple events is:

$$4 \times 3 = 12$$

Permutations

- **The number of ways you can arrange *n* distinct objects, taking them *r* at a time is**

$$P_r^n = \frac{n!}{(n-r)!}$$

where $n! = n(n-1)(n-2)...(2)(1)$ and $0! \equiv 1.$

Example: How many 3-digit lock combinations can we make from the numbers 1, 2, 3, and 4?

The order of the choice is important!

$$P_3^4 = \frac{4!}{1!} = 4(3)(2) = 24$$

Example

Example: A lock consists of five parts and can be assembled in any order. A quality control engineer wants to test each order for efficiency of assembly. How many orders are there?

The order of the choice is important!

$$P_5^5 = \frac{5!}{0!} = 5(4)(3)(2)(1) = 120$$

Combinations

- **The number of distinct combinations of *n* distinct objects that can be formed, taking them *r* at a time is**

$$C_r^n = \frac{n!}{r!(n-r)!}$$

Example: Three members of a 5-person committee must be chosen to form a subcommittee. How many different subcommittees could be formed?

The order of the choice is not important!

$$C_3^5 = \frac{5!}{3!(5-3)!} = \frac{5(4)(3)(2)1}{3(2)(1)(2)1} = \frac{5(4)}{(2)1} = 10$$

Example-1

- **A box contains six balls, four red and two green. A child selects two balls at random. What is the probability that exactly one is red?**

The order of the choice is not important!

$$C_2^6 = \frac{6!}{2!4!} = \frac{6(5)}{2(1)} = 15$$

ways to choose 2 balls.

$$C_1^2 = \frac{2!}{1!1!} = 2$$

ways to choose 1 green ball.

$$C_1^4 = \frac{4!}{1!3!} = 4$$

ways to choose 1 red ball

$4 \times 2 = 8$ ways to choose 1 red and 1 green ball.

P(exactly one red) = 8/15

Example-2

A deck of cards consists of 52 cards, 13 "kinds" each of four suits (spades, hearts, diamonds, and clubs). The 13 kinds are Ace (A), 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack (J), Queen (Q), King (K). In many poker games, each player is dealt five cards from a well shuffled deck.

There are $C_5^{52} = \dfrac{52!}{5!(52-5)!} = \dfrac{52(51)(50)(49)48}{5(4)(3)(2)1} = 2{,}598{,}960$ possible hands

Example-2 (Contd..)

- **Four of a kind: 4 of the 5 cards are the same “kind”. What is the probability of getting four of a kind in a five card hand?**

- There are 13 possible choices for the kind of which to have four, and 52-4=48 choices for the fifth card.
- Once the kind has been specified, the four are completely determined: you need all four cards of that kind.
- Thus there are 13×48=624 ways to get four of a kind.
- The probability=624/2598960=.000240096

Example-3

- One pair: two of the cards are of one kind, the other three are of three different kinds.
- What is the probability of getting one pair in a five card hand?

There are 13 possible choices for the kind of which to have a pair; given the choice, there are $C_2^4 = 6$ possible choices of two of the four cards of that kind

Example-3 (Contd..)

- There are 12 kinds remaining from which to select the other three cards in the hand.
- We must insist that the kinds be different from each other and from the kind of which we have a pair, or we could end up with a second pair, three or four of a kind, or a full house.

Example-3 (Contd..)

There are $C_3^{12} = 220$ ways to pick the kinds of the remaining three cards. There are 4 choices for the suit of each of those three cards, a total of $4^3 = 64$ choices for the suits of all three.

Therefore the number of "one pair" hands is $13 \times 6 \times 220 \times 64 = 1{,}098{,}240$.

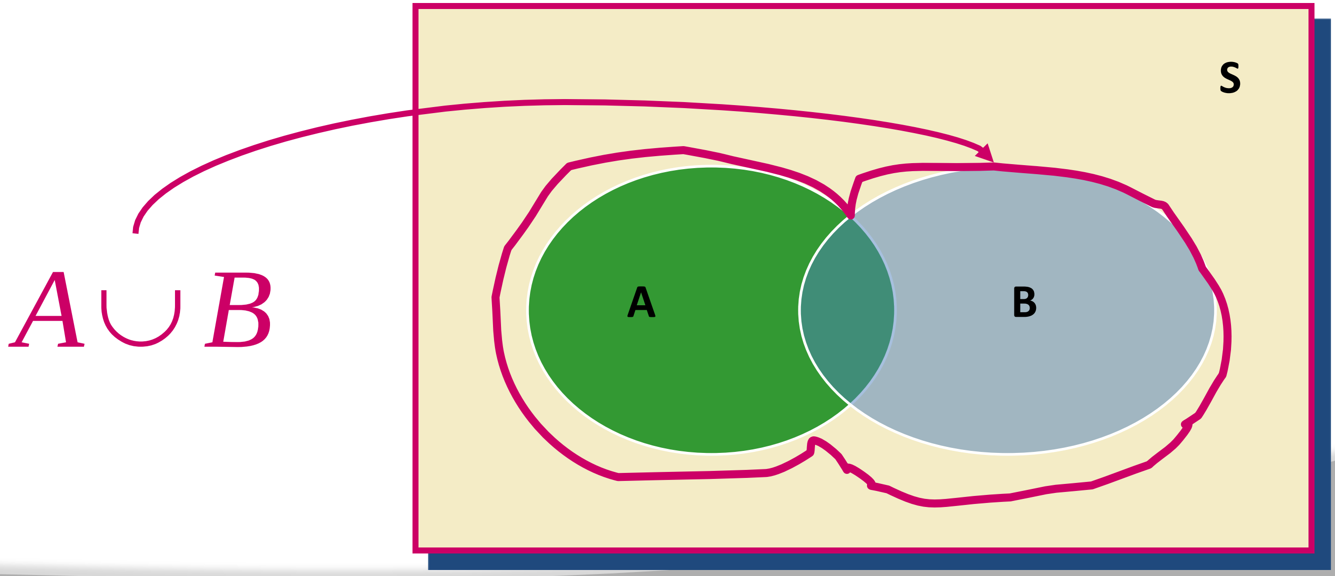
The probability $= 1098240/2598960 =$

$= .422569$

Event Relations

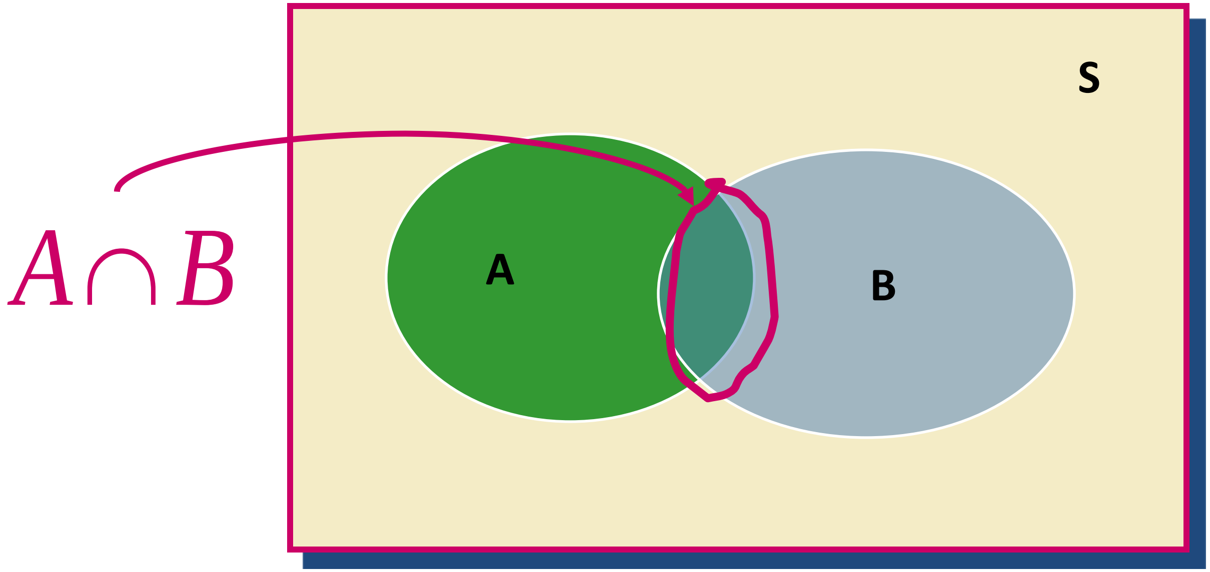
- The beauty of using events, rather than simple events, is that we can combine events to make other events using logical operations: and, or and not.
- The union of two events, A and B, is the event that either A or B or both occur when the experiment is performed. We write

$$A \cup B$$



Event Relations (Contd..)

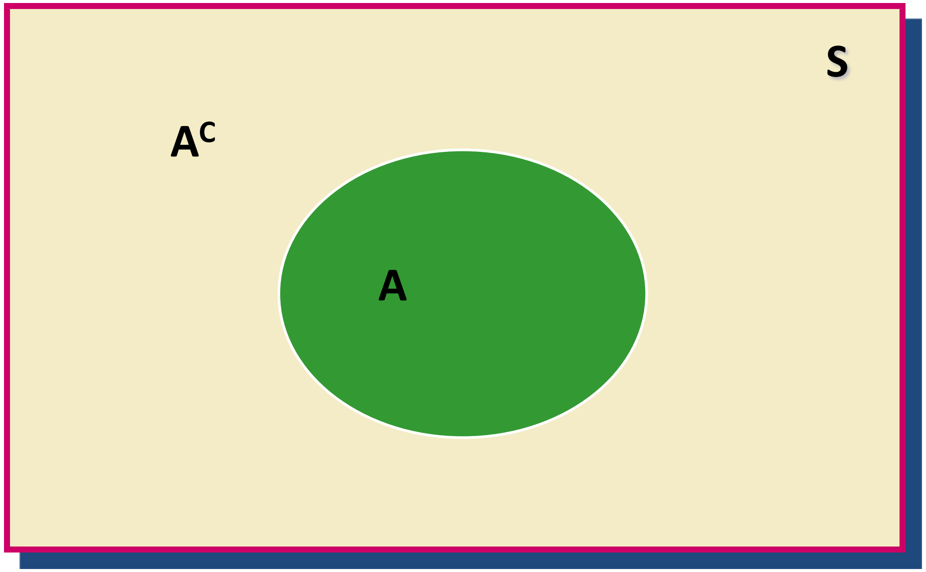
- **The intersection of two events, A and B, is the event that both A and B occur when the experiment is performed. We write A ∩ B.**



- If two events A and B are **mutually exclusive,** then P(**A ∩ B) = 0.**

Event Relations (Contd..)

- **The complement of an event A consists of all outcomes of the experiment that do not result in event A. We write A^C.**



Example

- **Select a student from the classroom and record his/her hair color and gender.**
  - **A: student has brown hair**
  - **B: student is female**
  - **C: student is male**

Mutually exclusive; $B = C^C$

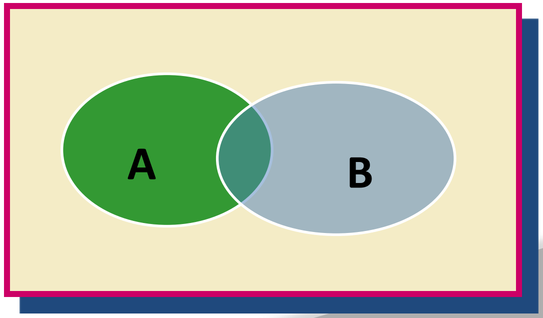
What is the relationship between events **B** and **C**?

- **A^C:** Student does not have brown hair
- **$B \cap C$:** Student is both male and female = $\varnothing$
- **$B \cup C$:** Student is either male and female = all students = S

Calculating Probabilities for Unions and Complements

- There are special rules that will allow you to calculate probabilities for composite events.
- **The Additive Rule for Unions:**
- For any two events, A and B, the probability of their union, $P(A \cup B)$, is

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



Example: Additive Rule

Example: Suppose that there were 120 students in the classroom, and that they could be classified as follows:

A: brown hair
P(A) = 50/120
B: female
P(B) = 60/120

	Brown	Not Brown
Male	20	40
Female	30	30

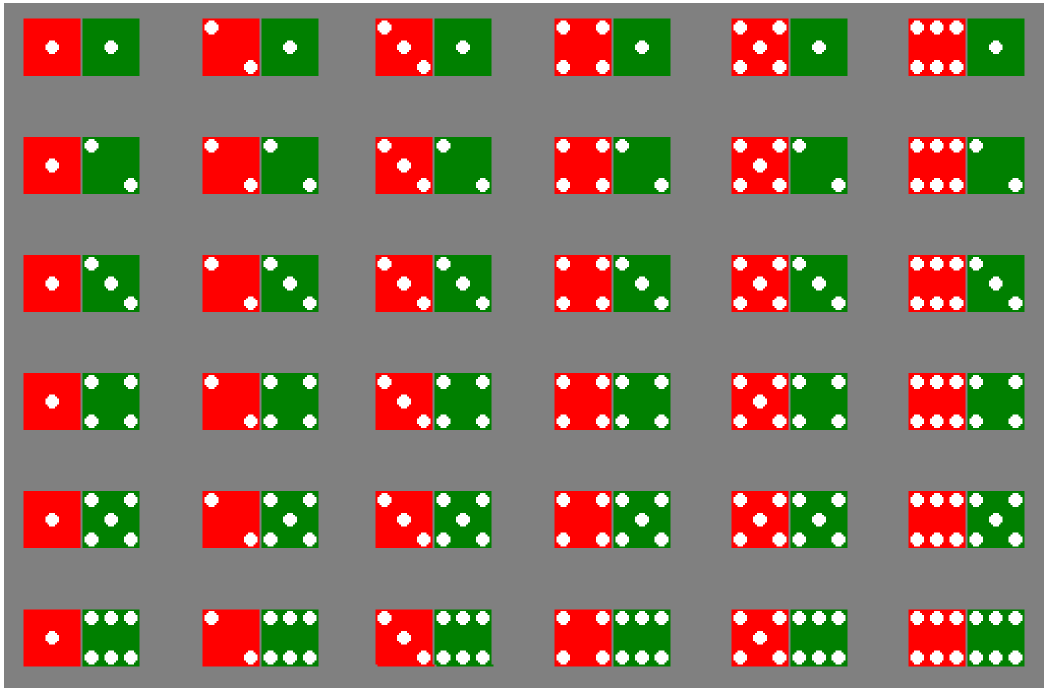
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$= 50/120 + 60/120 - 30/120$
$= 80/120 = 2/3$

Check: $P(A \cup B)$
$= (20 + 30 + 30)/120$

Example: Two Dice

A: red die show 1

B: green die show 1



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

$= 6/36 + 6/36 - 1/36$

$= 11/36$

A Special Case

- When two events A and B are **mutually exclusive,**
  $P(A \cap B) = 0$
  and $\mathbf{P(A \cup B) = P(A) + P(B).}$

A: male with brown hair
$P(A) = 20/120$
B: female with brown hair
$P(B) = 30/120$

	Brown	Not Brown
Male	20	40
Female	30	30

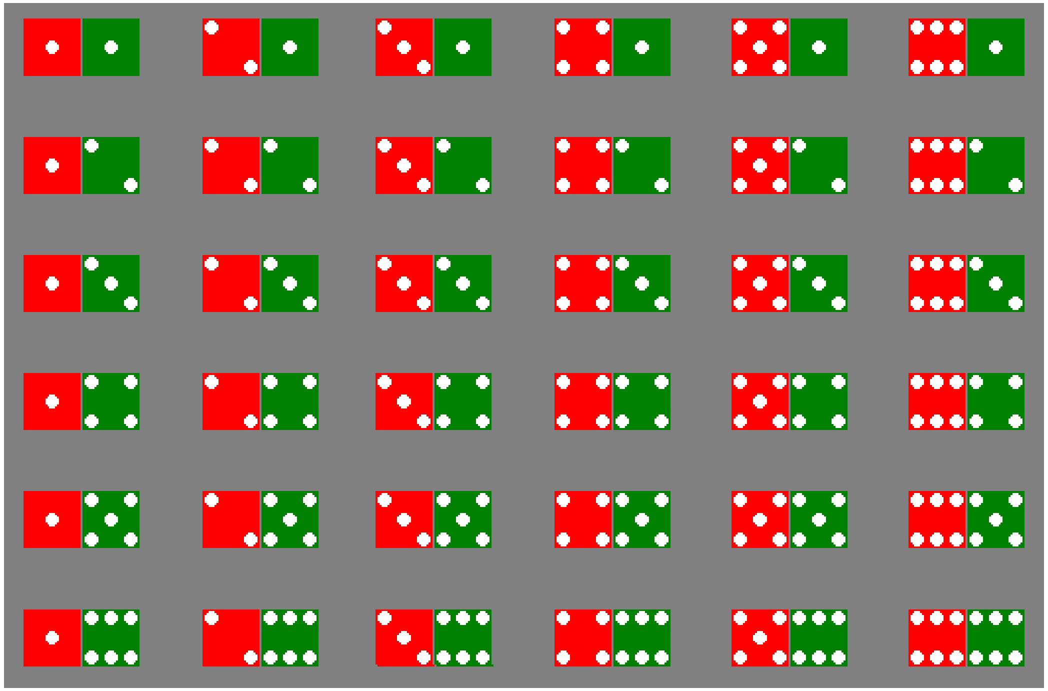
A and B are mutually exclusive, so that →

$P(A \cup B) = P(A) + P(B)$
$= 20/120 + 30/120$
$= 50/120$

Example: Two Dice

A: dice add to 3

B: dice add to 6



A and B are mutually exclusive, so that

$P(A \cup B) = P(A) + P(B)$
$= 2/36 + 5/36$
$= 7/36$

Calculating Probabilities for Complements

- We know that for any event A:
  - $P(A \cap A^C) = 0$
- Since either A or A^C must occur,

  $P(A \cup A^C) = 1$
- so that $P(A \cup A^C) = P(A) + P(A^C) = 1$

$$P(A^C) = 1 - P(A)$$

Example

- Select a student at random from the classroom. Define:

A: male

P(A) = 60/120

B: female

P(B) = ?

	Brown	Not Brown
Male	20	40
Female	30	30

A and B are complementary, so that

P(B) = 1- P(A)
= 1- 60/120 = 60/120

Calculating Probabilities for Intersections

- **In the previous example, we found P(A ∩ B) directly from the table. Sometimes this is impractical or impossible.**
- **The rule for calculating P(A ∩ B) depends on the idea of independent and dependent events.**

Two events, **A** and **B**, are said to be **independent** if the occurrence or nonoccurrence of one of the events does not change the probability of the occurrence of the other event.

Conditional Probabilities

- **The probability that A occurs, given that event B has occurred is called the conditional probability of A given B and is defined as**

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ if } P(B) \neq 0$$

"given"

Example-1

- Toss a fair coin twice. Define
  - A: head on second toss
  - B: head on first toss

Outcome	Probability
HH	1/4
HT	1/4
TH	1/4
TT	1/4

$P(A|B) = ½$

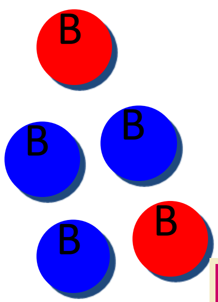
$P(A|\text{not } B) = ½$

P(A) does not change, whether B happens or not...

→ A and B are independent!

Example-2

- **A bowl contains five balls, two red and three blue. Randomly select two balls, and define**
  - **A: second ball is red.**
  - **B: first ball is blue.**



$P(A|B) = P(2^{nd}\ red|1^{st}\ blue) = 2/4 = 1/2$

$P(A|not\ B) = P(2^{nd}\ red|1^{st}\ red) = 1/4$

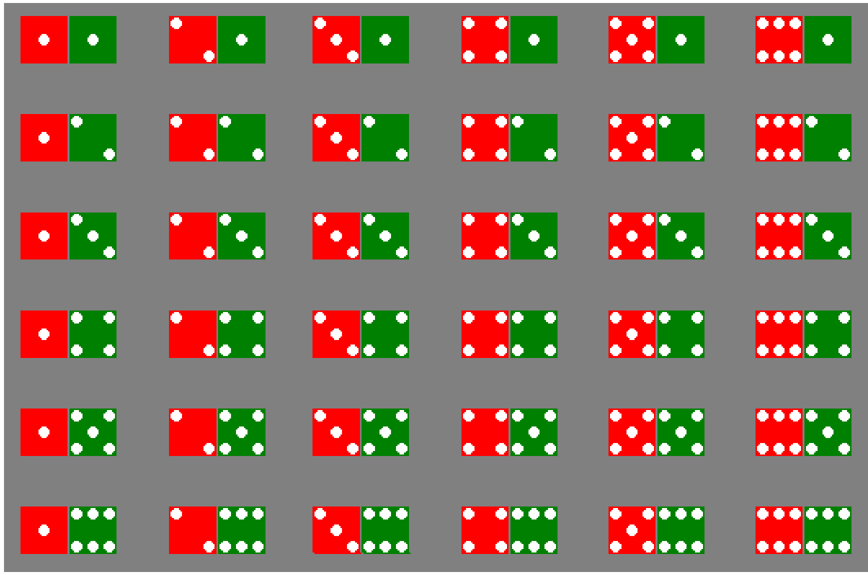
P(A) does change, depending on whether B happens or not…

A and B are dependent!

Example-3: Two Dice

Toss a pair of fair dice.
Define

- A: red die show 1
- B: green die show 1



P(A|B) = P(A and B)/P(B)

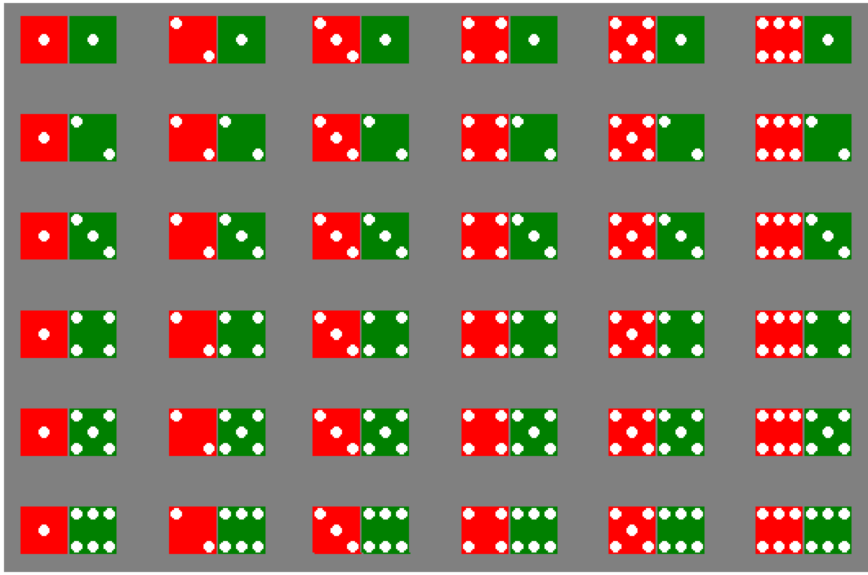
=1/36/1/6=1/6=P(A)

P(A) does not change, whether B happens or not…

A and B are independent!

Example-3: Two Dice (Contd..)

- Toss a pair of fair dice. Define
  - A: add to 3
  - B: add to 6



P(A|B) = P(A and B)/P(B)

=0/36/5/6=0

P(A) does change when B happens

A and B are dependent! In fact, when B happens, A can't

Independence

- **We can redefine independence in terms of conditional probabilities:**

Two events A and B are **independent** if and only if

$$P(A|B) = P(A) \quad \text{or} \quad P(B|A) = P(B)$$

Otherwise, they are **dependent**.

- Once you've decided whether or not two events are independent, you can use the following rule to calculate their intersection.

The Multiplicative Rule for Intersections

- **For any two events, A and B, the probability that both A and B occur is**

$$P(A \cap B) = P(A)\,P(B \text{ given that A occurred}) = P(A)P(B|A)$$

- If the events **A** and **B** are independent, then the probability that both **A** and **B** occur is

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

Example-1

- In a certain population, 10% of the people can be classified as being high risk for a heart attack. Three people are randomly selected from this population. What is the probability that exactly one of the three are high risk?

Define H: high risk   N: not high risk

P(exactly one high risk) = P(HNN) + P(NHN) + P(NNH)

= P(H)P(N)P(N) + P(N)P(H)P(N) + P(N)P(N)P(H)

$= (.1)(.9)(.9) + (.9)(.1)(.9) + (.9)(.9)(.1) = 3(.1)(.9)^2 = .243$

Example-2

- Suppose we have additional information in the previous example. We know that only 49% of the population are female. Also, of the female patients, 8% are high risk. A single person is selected at random. What is the probability that it is a high risk female?

Define H: high risk F: female

From the example, $P(F) = .49$ and $P(H|F) = .08$. Use the Multiplicative Rule:

$$P(\text{high risk female}) = P(H \cap F)$$

$$= P(F)P(H|F) = .49(.08) = .0392$$

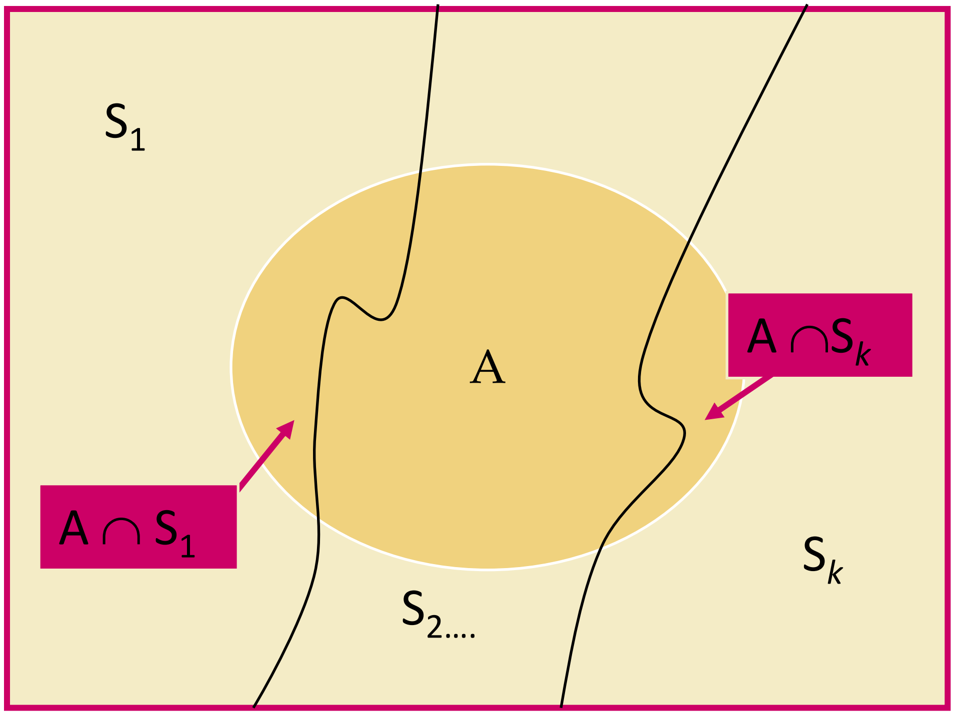
The Law of Total Probability

- Let S_1 , S_2 , S_3 ,..., S_k be mutually exclusive and exhaustive events (that is, one and only one must happen). Then the probability of any event A can be written as

$$P(A) = P(A \cap S_1) + P(A \cap S_2) + \ldots + P(A \cap S_k)$$

$$= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \ldots + P(S_k)P(A|S_k)$$

The Law of Total Probability (Contd..)



$$P(A) = P(A \cap S_1) + P(A \cap S_2) + \ldots + P(A \cap S_k)$$

$$= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \ldots + P(S_k)P(A|S_k)$$

Bayes' Rule

Let S_1 , S_2 , S_3 ,..., S_k be mutually exclusive and exhaustive events with prior probabilities $P(S_1)$, $P(S_2)$,...,$P(S_k)$. If an event A occurs, the posterior probability of S_i, given that A occurred is

$$P(S_i \mid A) = \frac{P(S_i)P(A \mid S_i)}{\sum P(S_i)P(A \mid S_i)} \text{ for } i = 1, 2, \ldots k$$

Proof

$$P(A \mid S_i) = \frac{P(AS_i)}{P(S_i)} \longrightarrow P(AS_i) = P(S_i)P(A \mid S_i)$$

$$P(S_i \mid A) = \frac{P(AS_i)}{P(A)} = \frac{P(S_i)P(A \mid S_i)}{\sum P(S_i)P(A \mid S_i)}$$

Example-1

From a previous example, we know that 49% of the population are female. Of the female patients, 8% are high risk for heart attack, while 12% of the male patients are high risk. A single person is selected at random and found to be high risk. What is the probability that it is a male?

Define H: high risk F: female M: male

We know:

P(F) = .49

P(M) = .51

P(H|F) = .08

P(H|M) = .12

$$P(M|H) = \frac{P(M)P(H|M)}{P(M)P(H|M) + P(F)P(H|F)}$$

$$= \frac{.51(.12)}{.51(.12) + .49(.08)} = .61$$

Example-2

- Suppose a rare disease infects one out of every 1000 people in a population. And suppose that there is a good, but not perfect, test for this disease: if a person has the disease, the test comes back positive 99% of the time. On the other hand, the test also produces some false positives: 2% of uninfected people are also test positive. And someone just tested positive. What are his chances of having this disease?

Example-2 (contd..)

Define A: has the disease B: test positive

We know:

$P(A) = .001$ $P(A^c) = .999$

$P(B|A) = .99$ $P(B|A^c) = .02$

We want to know $P(A|B) = ?$

$$P(A \mid B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^C)P(B|A^C)}$$

$$= \frac{.001 \times .99}{.001 \times .99 + .999 \times .02} = .0472$$

Example-3

A survey of job satisfaction of teachers was taken, giving the following results

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	74	43	117
	High School	224	171	395
	Elementary	126	140	266
	Total	424	354	778

Example-3 (contd..)

If all the cells are divided by the total number surveyed, 778, the resulting table is a table of empirically derived probabilities.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

Example-3 (contd..)

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

For convenience, let C stand for the event that the teacher teaches college, S stand for the teacher being satisfied and so on. Let's look at some probabilities and what they mean.

$P(C) = 0.150$ is the proportion of teachers who are college teachers.

$P(S) = 0.545$ is the proportion of teachers who are satisfied with their job.

$P(C \cap S) = 0.095$ is the proportion of teachers who are college teachers and who are satisfied with their job.

Example-3 (contd..)

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

$$P(C \mid S) = \frac{P(C \cap S)}{P(S)}$$

$$= \frac{0.095}{0.545} = 0.175$$

is the proportion of teachers who are college teachers given they are satisfied. Restated: This is the proportion of satisfied that are college teachers.

$$P(S \mid C) = \frac{P(S \cap C)}{P(C)}$$

$$= \frac{P(C \cap S)}{P(C)} = \frac{0.095}{0.150}$$

$$= 0.632$$

is the proportion of teachers who are satisfied given they are college teachers. Restated: This is the proportion of college teachers that are satisfied.

Example-3 (contd..)

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.342
	Total	0.545	0.455	1.000

Are C and S independent events?

$$P(C) = 0.150 \text{ and } P(C \mid S) = \frac{P(C \cap S)}{P(S)} = \frac{0.095}{0.545} = 0.175$$

$P(C|S) \neq P(C)$ so C and S are dependent events.

Example-3 (contd..)

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.658
	Total	0.545	0.455	1.000

$P(C \cup S)$?

$P(C) = 0.150$, $P(S) = 0.545$ and

$P(C \cap S) = 0.095$, so

$$P(C \cup S) = P(C)+P(S) - P(C \cap S)$$

$$= 0.150 + 0.545 - 0.095$$

$$= 0.600$$

Example-4

- Tom and Dick are going to take a driver's test at the nearest DMV office. Tom estimates that his chances to pass the test are 70% and Dick estimates his as 80%. Tom and Dick take their tests independently.
- Define D = {Dick passes the driving test}
- Define T = {Tom passes the driving test}

T and D are independent.

P (T) = 0.7, P (D) = 0.8

Example-4 (contd..)

What is the probability that at most one of the two friends will pass the test?

P(At most one person pass)

$= P(D^c \cap T^c) + P(D^c \cap T) + P(D \cap T^c)$

$= (1 - 0.8)\ (1 - 0.7) + (0.7)\ (1 - 0.8) + (0.8)\ (1 - 0.7)$

$= .44$

P(At most one person pass)

$= 1 - P(\text{both pass}) = 1 - 0.8 \times 0.7 = .44$

Example-4 (contd..)

What is the probability that at least one of the two friends will pass the test?

P(At least one person pass)

$= P(D \cup T)$

$= 0.8 + 0.7 - 0.8 \times 0.7$

$= .94$

P(At least one person pass)

$= 1 - P(\text{neither passes}) = 1 - (1-0.8) \times (1-0.7) = .94$

Example-4 (contd..)

Suppose we know that only one of the two friends passed the test. What is the probability that it was Dick?

P(D | exactly one person passed)

= P(D $\cap$ exactly one person passed) / P(exactly one person passed)

= $P(D \cap T^c) / (P(D \cap T^c) + P(D^c \cap T))$

= 0.8 x (1-0.7)/(0.8 x (1-0.7)+(1-.8) x 0.7)

= .63

Index

- Random Variables- Definition
- Conditions for a Function to be a Random Variable, Discrete,
- Continuous and Mixed Random Variable,
- Distribution and Density functions, Properties,
- Binomial, Poisson, Uniform, Gaussian, Exponential, Rayleigh,
- Methods of defining Conditioning Event,
- Conditional Distribution, Conditional Density and their Properties.

Random Variables

- **A quantitative variable *x* is a random variable if the value that it assumes, corresponding to the outcome of an experiment is a chance or random event.**

- **Random variables can be discrete or continuous.**

- **Examples:**
  - ✓ *x* = Exam score for a randomly selected student
  - ✓ *x* = number of people in a room at a randomly selected time of day
  - ✓ *x* = number on the upper face of a randomly tossed die

Probability Distributions for Discrete Random Variables

- **The probability distribution for a discrete random variable *x* resembles the relative frequency distributions. It is a graph, table or formula that gives the possible values of *x* and the probability *p(x)* associated with each value.**

We must have

$$0 \le p(x) \le 1 \text{ and } \sum p(x) = 1$$

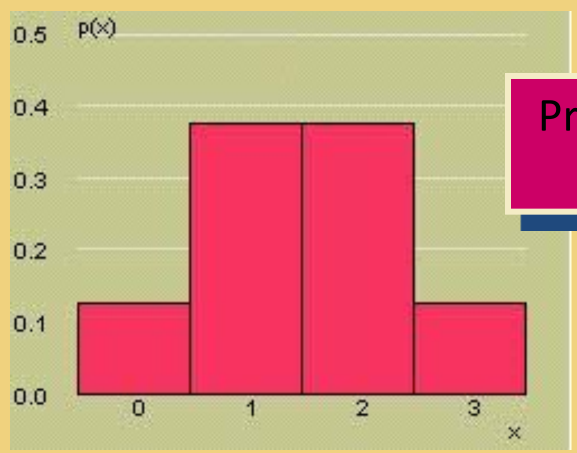
Example-1

- Toss a fair coin three times and define
- *x* = number of heads.

		x
HHH	1/8	3
HHT	1/8	2
HTH	1/8	2
THH	1/8	2
HTT	1/8	1
THT	1/8	1
TTH	1/8	1
TTT	1/8	0

$P(x = 0) = 1/8$
$P(x = 1) = 3/8$
$P(x = 2) = 3/8$
$P(x = 3) = 1/8$

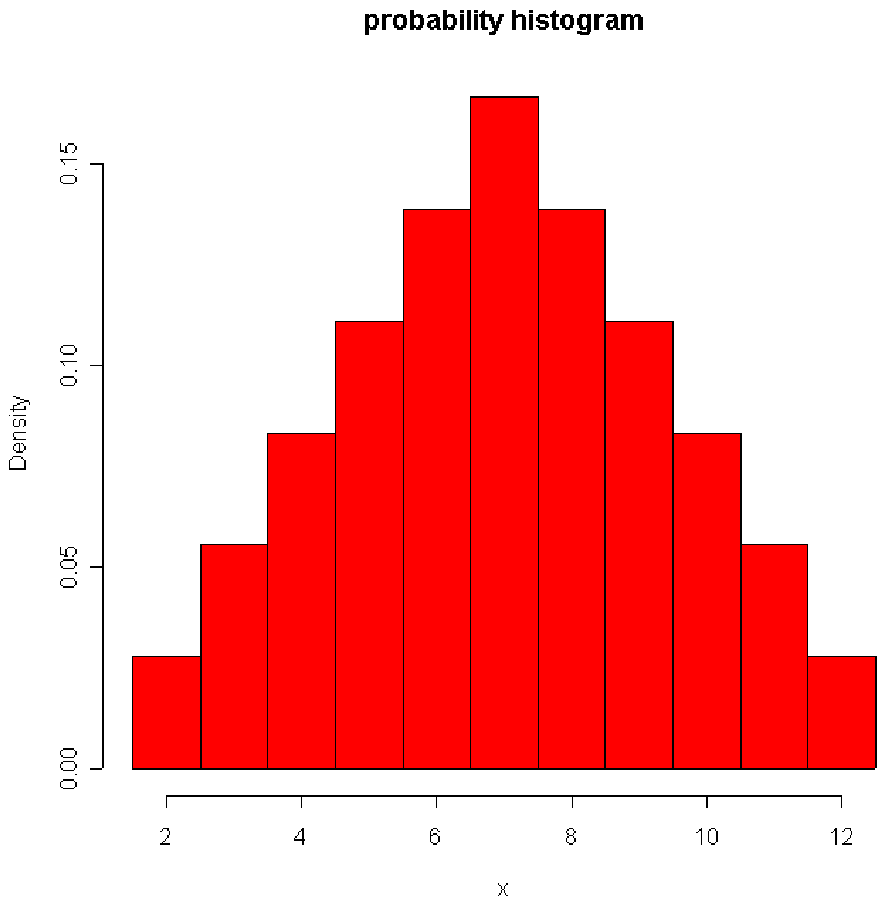
x	$p(x)$
0	1/8
1	3/8
2	3/8
3	1/8



Probability Histogram for *x*

Example-2

- Toss two dice and define
- ***x* = sum of two dice.**



x	*p(x)*
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Probability Distributions

- Probability distributions can be used to describe the population.
  - **Shape:** Symmetric, skewed, mound-shaped…
  - **Outliers:** unusual or unlikely measurements
  - **Center and spread:** mean and standard deviation. A population mean is called μ and a population standard deviation is called σ.

The Mean and Standard Deviation

Let *x* be a discrete random variable with probability distribution *p(x)*. Then the mean, variance and standard deviation of *x* are given as

$$\text{Mean}: \mu = \sum xp(x)$$

$$\text{Variance}: \sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\text{Standard deviation}: \sigma = \sqrt{\sigma^2}$$

Example

Toss a fair coin 3 times and record *x* the number of heads.

x	$p(x)$	$xp(x)$	$(x-\mu)^2 p(x)$
0	1/8	0	$(-1.5)^2(1/8)$
1	3/8	3/8	$(-0.5)^2(3/8)$
2	3/8	6/8	$(0.5)^2(3/8)$
3	1/8	3/8	$(1.5)^2(1/8)$

$$\mu = \sum xp(x) = \frac{12}{8} = 1.5$$

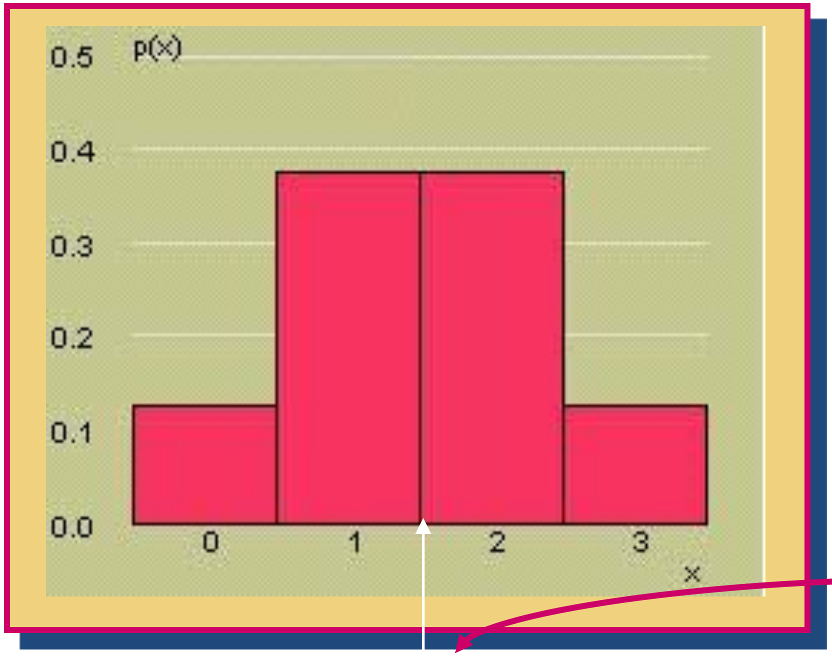
$$\sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\sigma^2 = .28125 + .09375 + .09375 + .28125 = .75$$

$$\sigma = \sqrt{.75} = .688$$

Example

The probability distribution for *x* the number of heads in tossing 3 fair coins.



μ

- Shape?
- Outliers?
- Center?
- Spread?

Symmetric; mound-shaped

None

$\mu = 1.5$

$\sigma = .688$

Probability Distribution Function

The probability $P(X \leq x)$ is the probability of the event $\{X \leq x\}$. i,e

$$F_x(x) = P\{X \leq x\}, \quad -\infty \leq x \leq \infty$$

Properties of Distribution

The properties of a distribution function:

- $F_x(-\infty) = 0$
- $F_x(\infty) = 1$
- $0 \leq F_x(x) \leq 1$
- $F_x(x_1) \leq F_x(x_2)$, if $x_1 < x_2$ (Non-decreasing function)
- $P\{x_1 < X < x_2\} = F_x(x_2) - F_x(x_1)$
- $F_x(x^+) = F_x(x)$ (Continuous from the right)

Properties of Distribution

Proof for $F_x(x_2) - F_x(x_1)$

- The events $\{X \leq x_1\}$ and $\{x_1 < X < x_2\}$ are mutually exclusive, i.e. $\{X < x_2\} = \{X \leq x_1\} \cup \{x_1 < X < x_2\}$
- $P\{X < x_2\} = P\{X \leq x_1\} + P\{x_1 < X < x_2\}$
- $P\{x_1 < X < x_2\} = P\{X < x_2\} - P\{X \leq x_1\}$

$$= F_x(x_2) - F_x(x_1)$$

Properties of Distribution

If X is a discrete random variable taking values $x_i, \; i = 1,2,\ldots,N$, then $F_x(x)$ must have a staircase function given by

$$F_x(x) = \sum_{i=1}^{N} P\{X = x_i\}\, u(x - x_i)$$

$$= \sum_{i=1}^{N} P(x_i)\, u(x - x_i)$$

where $u(.)$ is the unit-step function defined by:

$$u(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

If N is infinite, then

$$P(x_i) = P\{X = x_i\}$$

Probability Density Function

The probability density function of the random variable X is defined as the derivative of the distribution function:

$$f_x(x) = \frac{dF_x(x)}{dx}$$

Probability Density

1. If the derivative of $F_x(x)$ exists then $f_x(x)$ exists
2. There may be places where $\frac{dF_x(x)}{dx}$ is not defined at points of abrupt change, then we shall assume that the number of points where $F_x(x)$ is not differentiable is countable.
3. For discrete random variables having a stair step form of distribution function.

$$f_x(x) = \sum_{i=1}^{N} P(x_1)\delta(x - x_1)$$

Properties of Density Functions.

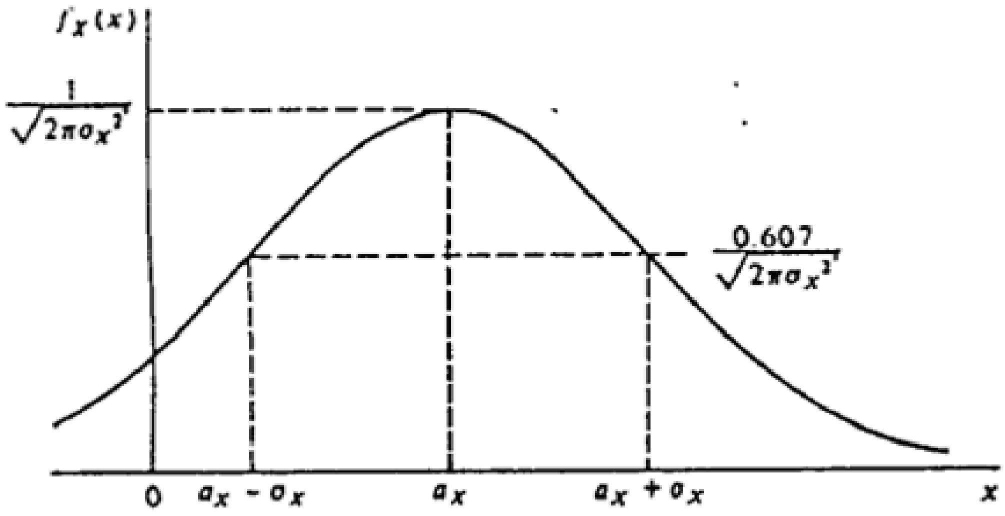
- $0 \leq f_X(x)$ all x
- $\int_{-\infty}^{\infty} f_X(x)dx = 1$
- $F_X(x) = \int_{-\infty}^{x} f_X(x)dx = 1$
- $P\{x_1 < X < x_2\} = \int_{x_1}^{x_2} f_X(x)dx$

Gaussian Density Function

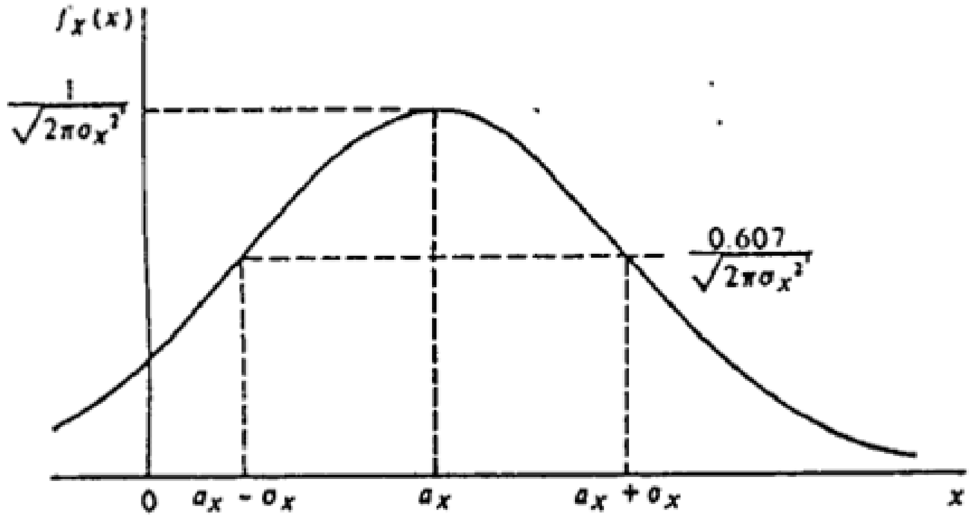
A random variable X is called Gaussian if its density function has the form

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(x-a_x)^2/2\sigma x^2}$$

Where $\sigma_x > 0$ and $-\infty < a_x < \infty$ are real constants.



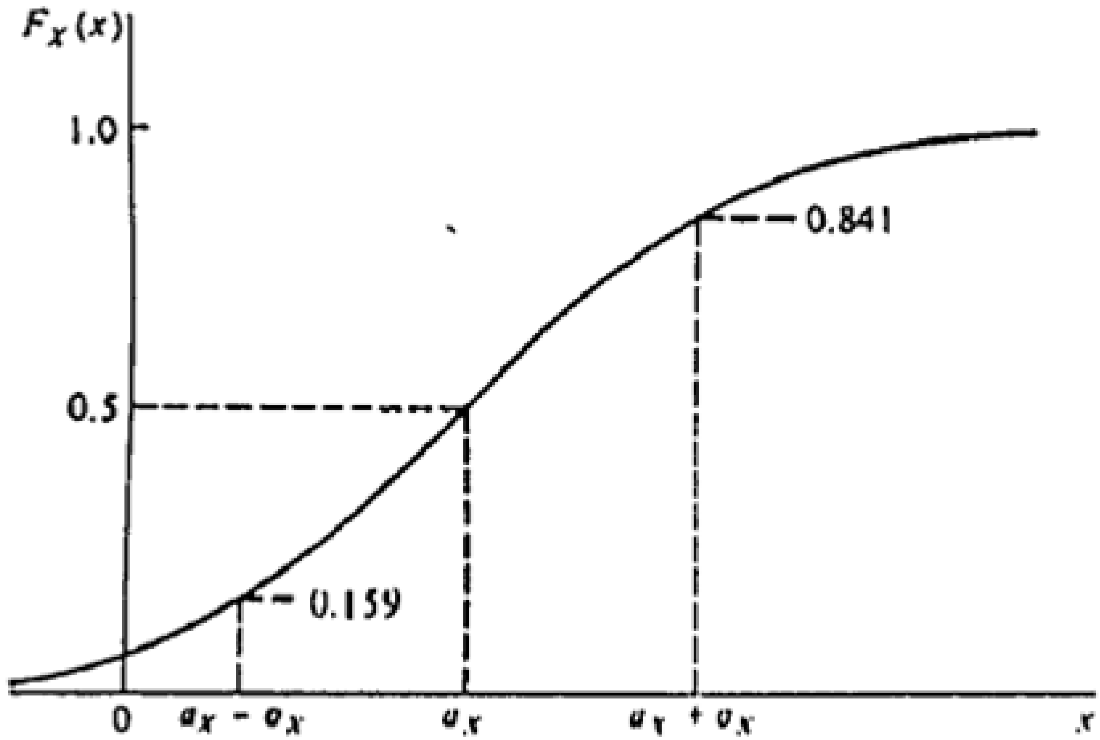
Gaussian Density



1. Its maximum value $(2\pi\sigma_2^x)^{-\frac{1}{2}}$ occurs at $x = a_x$.
2. Its "spread" about the point $x = a_x$ is related to $\sigma_{x.}$
3. The function decreases to 0.607 times its maximum at $x = a_x + \sigma_x$ and $x = a_x - \sigma_x$.
4. The Gaussian density is the most important of all densities. It enters into nearly all areas of engineering

Gaussian Distribution

$$F_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{-\infty}^{x} e^{-(\xi - ex)^2/2\sigma x^2} d\xi$$



Gaussian Distribution

- This integral has no known closed-form solution and must be evaluated by numerical methods.
- We could develop a set of tables of $F_x(x)$ for various x and α_x and σ_x as parameters (infinite number of tables).
- Only one table of $F_x(x)$ for the normalized (specific) values $\alpha_x = 0$ and $\sigma_{x=1}$ given by

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\xi^2/2} d\xi$$

which is a function of x only & tabulated for $x \geq 0$.

- For negative values of x we have

$$F(-x) = 1 - F(x)$$

Gaussian Distribution

- Making the variable change $u = \frac{\xi - a_x}{\sigma_x}$, we get

$$F_x(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-a_x)/\sigma_x} e^{-ut/2}\, du = F\left(\frac{x - a_x}{\sigma_x}\right)$$

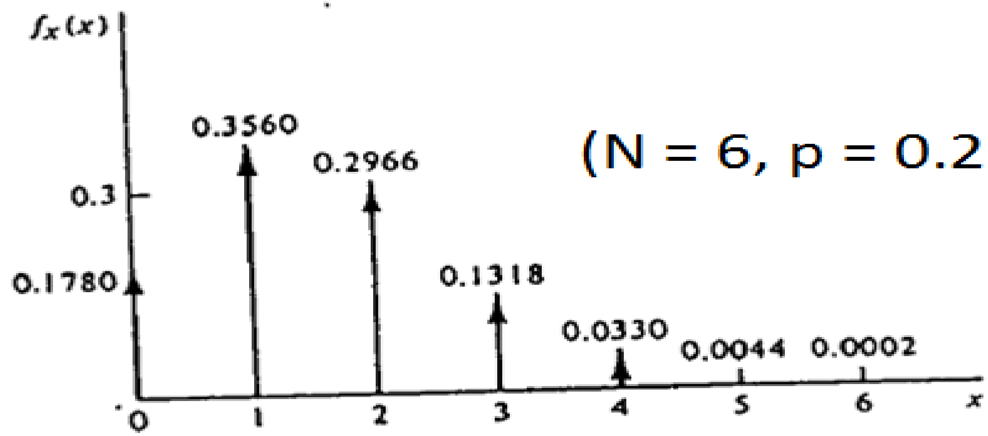
Binomial Density Function

$$f_X(x) = \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \delta(x-k)$$

where $\binom{N}{k}$ is the binomial coefficient defined as

$$\binom{N}{k} = \frac{N!}{k!\,(N-k)!}$$

and $0 < p < 1,\ N = 1,2 \ldots \ldots \ldots.$



(N = 6, p = 0.25)

Binomial Density

1. The binomial density is applied to Bernoulli trail experiment, having only two possible outcomes on any given trial.
2. It applies to many games of chance, detection problems in radar and sonar, and many experiments

Binomial Distribution

By integration, the binomial distribution function is found:

$$F_X(x) = \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \delta(x-k)$$

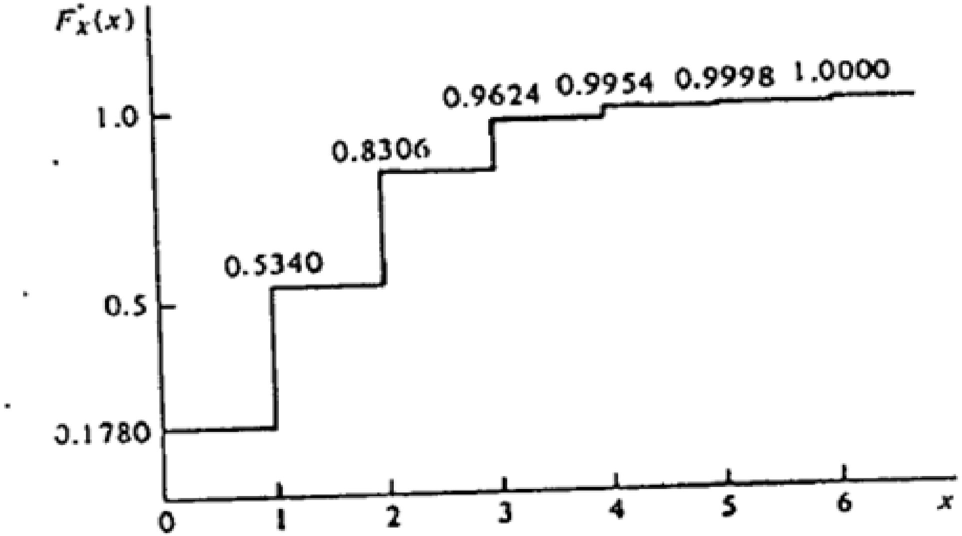


Figure: Binomial distribution function (N = 6, p = 0.25)

Poisson Density Function

$$f_X(x) = e^{-b} \sum_{k=0}^{\infty} \frac{b^k}{k!} \delta(x-k)$$

$$F_X(x) = e^{-b} \sum_{k=0}^{\infty} \frac{b^k}{k!} u(x-k)$$

where $b > 0$ is a real constant.

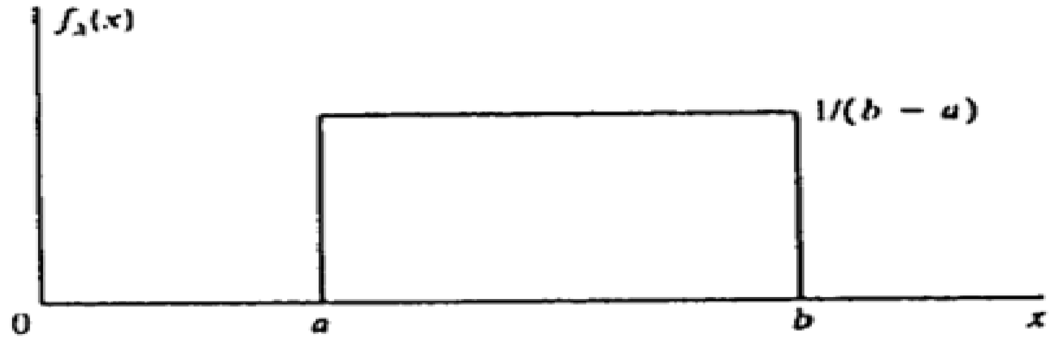
- These functions appear quite similar to binomial
- If $N \to \infty$ and $p \to 0$ for the binomial case in such a way that $N_p = b$, a constant, the Poisson case results.
- The Poisson random variable applies to a wide variety of counting-type applications.

Poisson Density

- It describes
  - the number of defective units in a sample taken from a production line,
  - the number of telephone calls during a period of time,
  - the number of electrons emitted from a small section of a cathode in a given time interval, etc.
  - If the time interval of interest has duration T, and the events being counted are known to occur at an average rate λ and have a Poisson distribution, then $b = \lambda T$

Uniform Density Function

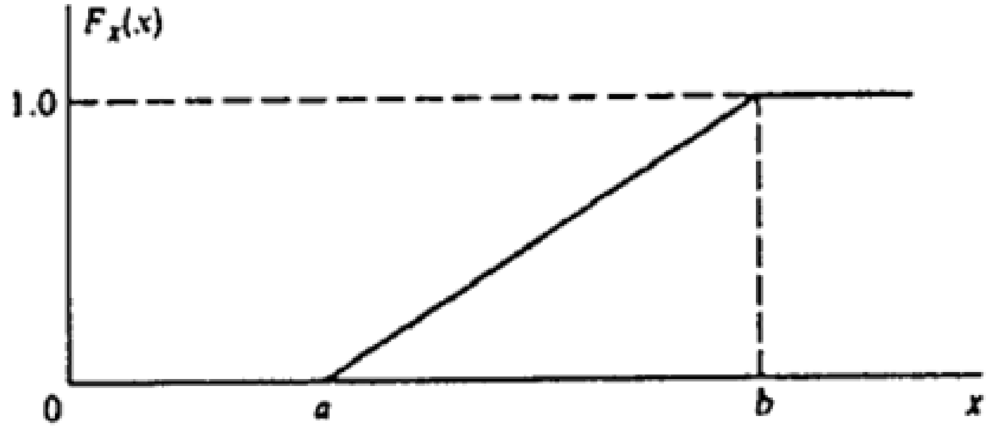
$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & elsewhere \end{cases}$$



for real constants $-\infty < a < \infty$ and $b > a$.

Uniform Distribution

$$F_X(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \end{cases}$$

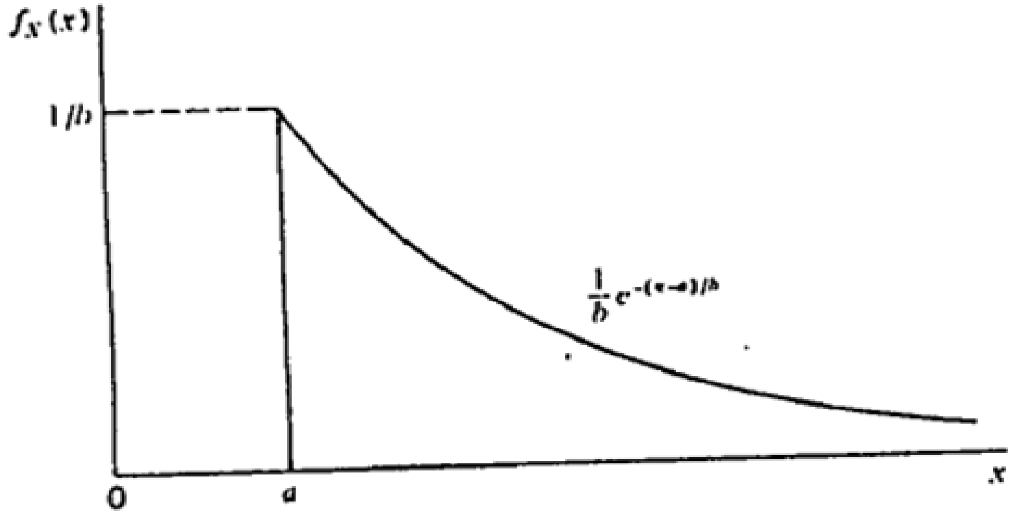


Uniform Distribution

- The error of quantization of signal samples prior to encoding in digital communication systems.
- Quantization amounts to “rounding off” the actual sample to the nearest of discrete quantum level.
- The quantization error introduced in the round-off process are uniformly distributed.

Exponential Density Function

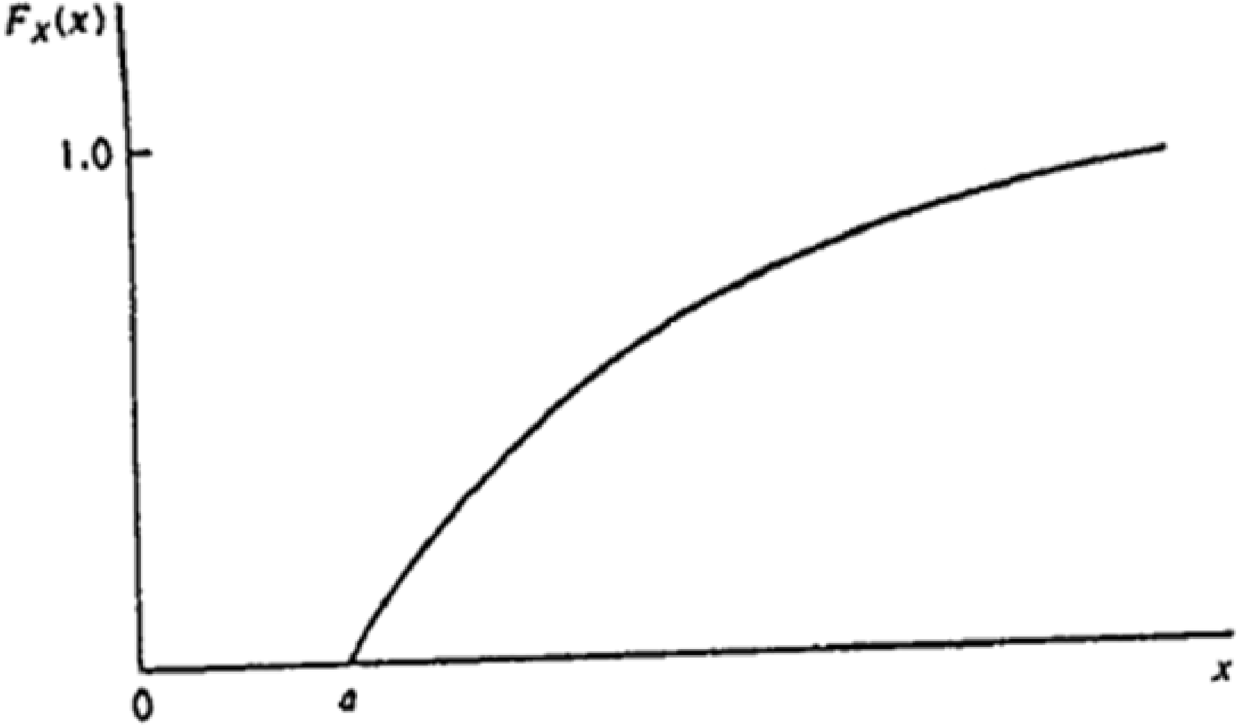
$$f_X(x) = \begin{cases} \frac{1}{b} e^{-(x-a)/b} & x > a \\ 0, & x < a \end{cases}$$



for real numbers $-\infty < a < \infty$ and $b > 0$

Exponential Distribution

$$F_X(x) = \begin{cases} 1 - e^{-(x-a)/b} & x > a \\ 0, & x < a \end{cases}$$

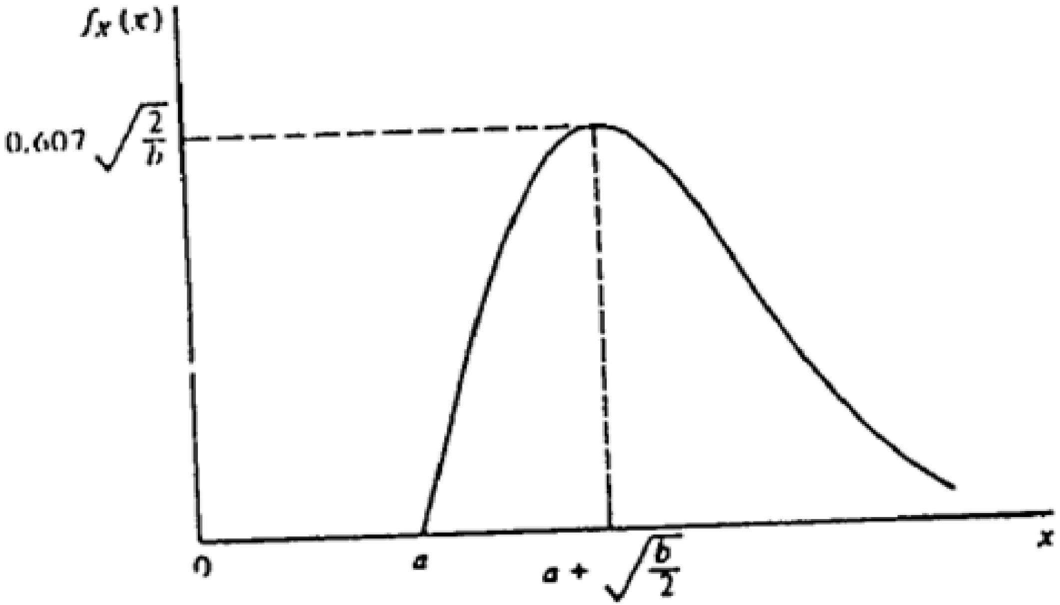


Exponential Density

- The exponential density is useful in describing raindrop sizes when a large number of rainstorm measurements are made.
- It is also known to approximately describe the fluctuations in signal strength received by radar from certain types of aircraft.

Rayleigh Density Function

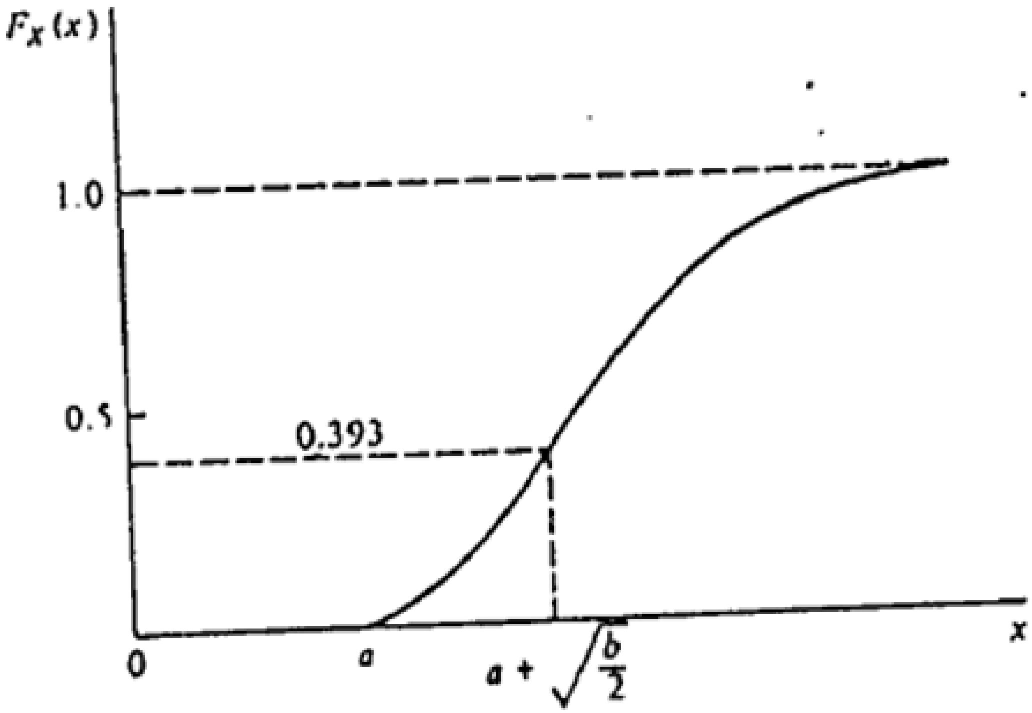
$$f_X(x) = \begin{cases} \frac{2}{b}(x-a)e^{-(x-a)2/b}, & x \geq a \\ 0, & x < a \end{cases}$$



for the real constants $-\infty < a < \infty$ and $b > 0$

Rayleigh Distribution

$$F_X(x) = \begin{cases} 1 - e^{-(x-a)2/b} & x \geq a \\ 0, & x < a \end{cases}$$



Rayleigh Density

- The Rayleigh density describes the envelope of white gaussian noise when passed through a band-pass filter.
- It is also is important in analysis of errors in various measurement systems.

Conditional Distribution

Conditional Distribution Function

- Let A and B be the two events & $P(B) \neq 0$, then
$$P(A|B) = \frac{P\{A \cap B\}}{P(B)}$$
- Let A be defined as the event $\{X \leq x\}$ for the random variable X.
- The resulting probability $P\{X \leq x|B\}$ is defined as the conditional distribution function of X , which is denote d by
$$F_X(x|B) = P\{X \leq x|B\} = \frac{P\{X \leq x \cap B\}}{P(B)}$$

where $\{X \leq x \cap B\}$ is the joint event $\{X \leq x\} \cap B$. This joint event consists of all outcomes ***s*** such that $X(s) \leq x$ and ***s*** ϵ B

Properties of Conditional Distribution Function

- $F_X(-\infty|B) = 0$
- $F_X(\infty|B) = 1$
- $0 \leq F_X(\infty|B) \leq 1$
- $F_X(x_1|B) \leq F_X(x_2|B) \quad if\ x_1 < x_2$
- $P\{x_1 < X \leq x_2|B\} = F_X(x_2|B) - F_X(x_1|B)$
- $F_X(x^+|B) = F_X(x|B)$

Conditional Density Function

The conditional density function of the random variable X is defined as the derivative of the conditional distribution function, and is given by

$$f_X(x|B) = \frac{dF_X(x|B)}{dx}$$

If $F_X(x|B)$ contains step discontinuities (when X is a discrete or mixed random variable), we assume that impulse functions are present in $f_X(x|B)$ to account for the derivatives at the discontinuities.

Properties of Conditional Density Function

- $f_X(x|B) \geq 0$
- $\int_{-\infty}^{\infty} f_X(x|B)dx = 1$
- $F_X(x|B) = \int_{-\infty}^{x} f_X(\xi|B)\, d\xi$
- $P\{x_1 < X < x_1|B\} = \int_{x_1}^{x_2} f_X(x|B)dx$

Methods of Conditioning

Methods of Defining Conditioning Event

If event B is defined in terms of the random variable X as $B = \{X \leq b\}$, where b is some real number $-\infty < b < \infty$ & $P\{X \leq b\} \neq 0$, then we have

$$F_X(x|B) = P\{X \leq x|B\}$$

$$= P\{X \leq x|X \leq b\}$$

$$= \frac{P\{X \leq x \cap X \leq b\}}{P\{X \leq b\}}$$

Methods of Conditioning

Case (i):

If $b \leq x$, then the event $\{X \leq b\}$ is an subset of the event $\{X \leq x\}$, so $\{X \leq x\} \cap \{X \leq b\} = \{X \leq b\}$. Then we have

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

$$= \frac{P\{X \leq b\}}{P\{X \leq b\}} = 1 \qquad x \geq b$$

Methods of Conditioning

Case (ii):

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

$$= \frac{P\{X \leq x\}}{P\{X \leq b\}} = \frac{F_X(x)}{F_X(b)} \qquad x < b$$

By combining the last two expressions, we have

$$F_X(x|X \leq b) = \begin{cases} \dfrac{F_X(x)}{F_X(b)} & x < b \\ 1, & x \geq b \end{cases}$$

From our assumption that the conditioning event has nonzero probability, $0 < F_X(b) \leq 1$, so the conditional distribution function is never smaller than the ordinary distribution function $F_X(x|X \leq b) \geq F_X(x)$

Methods of Conditioning

Similarly the conditional density function is

$$f_X(x|X \leq b) = \begin{cases} \dfrac{f_X(x)}{F_X(x)} = \dfrac{f_X(x)}{\int_{-\infty}^{b} f_X(x)dx} & x < b \\ 0, & x \geq b \end{cases}$$

From our assumption $0 < f_X(x) \leq 1$, so the conditional density function is never smaller than the ordinary density function

$$f_X(x|X \leq b) \geq f_X(x) \qquad x < b$$

The result can be extended to more general event $B = \{a < X \leq b\}$

Expected Value of a Random variable

In general, the expected value of any random variable X is defined by

$$E[X] = \bar{X} = \int_{-\infty}^{\infty} x f_X(x)\,dx$$

Expected value

If X is discrete with N possible values x_i having probabilities $P(x_i)$ of occurrence, then

$$f_x(x) = \sum_{i=1}^{N} P(x_i)\delta(x - x_i)$$

Then we have

$$E[x] = \sum_{i=1}^{N} x_i P(x_i)$$

Conditional Expected Value

If $f_X(x|B)$ is the conditional density where B is any event defined on the sample space of X, then the *conditional expected value* of X, is given by

$$E[X|B] = \int_{-\infty}^{\infty} x f_X(x|B)\,dx$$

Conditional Expected value

If the event $B = \{X \leq b\}$, $\infty < b < \infty$

$$f_X(x|X \leq b) = \begin{cases} \dfrac{f_X(x)}{\int_{-\infty}^{b} f_X(x)dx} & x < b \\ 0 & x \geq b \end{cases}$$

Then, the conditional expected value is given by

$$E[x|X \leq b] = \frac{\int_{-\infty}^{b} x f_X(x)dx}{\int_{-\infty}^{b} f_X(x)dx}$$

which is the mean value of X when X is constrained to the set $\{X \leq b\}$.

Moments About the Origin

The expected value of X^n, $n = 0,1,2, \ldots \ldots \ldots$ is given by

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\,dx$$

gives the moments about the origin of the random variable X. These are also called standard moments and are denoted as m_n

Moments about origin

For $n = 0$,

$$m_0 = E[X^0] = \int_{-\infty}^{\infty} x^0 f_X(x)dx = \int_{-\infty}^{\infty} f_X(x)dx$$

is the area of under the function $f_X(x)$.

For $n = 1$,

$$m_1 = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \bar{X}$$

is the expected value of X.

Moments About the Mean

The expected value of $(X-\bar{X})^n,\ n=0,1,2,\dots\dots\dots$ is given by

$$E[(X-\bar{X})^n]=\int_{-\infty}^{\infty}(x-\bar{X})^n f_X(x)dx$$

gives the moments about the mean of the random variable X. These are also called central moments and are denoted as μ_n

Moments about mean

For $n = 0$,

$$\mu_n = E[(X - \bar{X})^n] = \int_{-\infty}^{\infty} (x - \bar{X})^n f_x(x) dx$$

$$\mu_0 = E[(X - \bar{X})^0] = \int_{-\infty}^{\infty} f_x(x) dx$$

is the area of under the function $f_x(x)$.

For $n = 1$,

$$\mu_1 = E[(X - \bar{X})] = E[X] - \bar{X} = 0$$

Variance

The second central moment μ_2 is given by

$$\mu_2 = E[(X - \bar{X})^2] = \int_{-\infty}^{\infty} (X - \bar{X})^2 f_x(x) dx$$

1. It is popularly known as the *variance* σ_x^2 of the random variable X.
2. The positive square root σ_x of variance is called the standard deviation of X.
3. It is a measure of the spread in the function $f_x(x)$ about the mean.

Variance

The second central moment is given by

$$\mu_2 = E[(X - \bar{X})^2]$$

By expanding we get

$$\mu_2 = E[X^2 - 2\bar{X}X + \bar{X}^2]$$
$$= E[X^2] - 2\bar{X}E[X] + \bar{X}^2$$
$$= E[X^2] - \bar{X}^2 = m_2 - \mathrm{m}_1{}^2$$

Skew

The third central moment is given by

$$\mu_3 = E[(X - \bar{X})^3]$$

$$\mu_3 = E[X^3 - 3X^2\bar{X} + 3X\bar{X}^2 - \bar{X}^3]$$

$$= E[X^3] - 3E[X^2]\bar{X} + 3\bar{X}^2E[X] - \bar{X}^3$$

$$= m_3 - 3m_2 m_1 + 3m_1{}^3 - m_1{}^3$$

$$= m_3 - 3m_2 m_1 + 2m_1{}^3$$

Skew

- μ_3 is a measure of asymmetry of $f_x(x)$ about the mean.
- It will be called the *skew* of the density function.
- If a density is symmetric about $x = \bar{X}$, it has zero skew. For this case, $\mu_n = 0$ for all odd values of n.
- The normalized third central moment $\mu_3/{\sigma_x}^3$ is known as the coefficient of *skewness*.

Module-II

SINGLE RANDOM VARIABLE TRANSFORMATIONS-MULTIPLE RANDOM VARIABLES

Functions That Give Moments

Two functions can be defined that allow moments to be calculated for a random variable

- The characteristic function
- The moment generating function

Characteristic Function

The *characteristic function* of a random variable X is defined by

$$\Phi_x(\omega) = E\left[e^{j\omega x}\right] = \int_{-\infty}^{\infty} e^{j\omega x} f_x(x)dx$$

where $= \sqrt{-1}$. It is a function of the real number $-\infty < \omega < \infty$.

$\Phi_x(\omega)$ is seen as the *Fourier transform* (with the sign of ω reversed) of $f_x(x)$

Characteristic function

- $f_x(x)$ can be found from the *Inverse Fourier transform* (with sign of x reversed) i.e.

$$f_x(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_x(\omega)e^{j\omega x}dx$$

- By differentiating $\Phi_x(\omega)$ *n* times with respect to ω and setting $\omega = 0$, we can show that the nth moment of X is given by

$$m_n = (-j)^n \frac{d^n\Phi_x(\omega)}{d\omega^n}\bigg|_{\omega=0}$$

- A major advantage is that $\Phi_x(\omega)$ always exist, so the moments can always be found if $\Phi_x(\omega)$ is known, provided the derivatives of $\Phi_x(\omega)$exist.
- It can be shown that the maximum magnitude of a characteristic function is unity and occurs at $\omega = 0$ i.e., $|\Phi_x(\omega)| \leq \Phi_x(0) = 1$

Moment Generating Function

The *moment generating function* of a random variable X is defined by

$$M_x(v) = E[e^{vx}] = \int_{-\infty}^{\infty} f_x(x) e^{vx} dx$$

Where v is a real number $-\infty < v < \infty$.

Moment generating function

- Moments are related to $M_x(v)$ by the expression

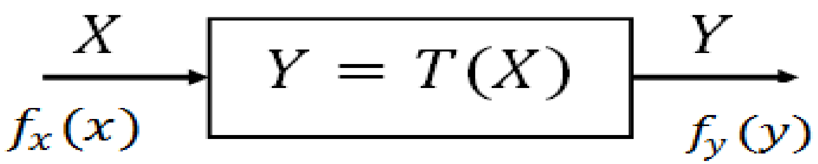
$$m_n = \left.\frac{d^n M_x(v)}{dv^n}\right|_{v=0}$$

- The main disadvantage of the moment generating function is that it may not exist for all random variables.
- In fact, $M_x(v)$ exists only if all the moments exist

Transformations of a Random Variable

Transformations of A Random Variable

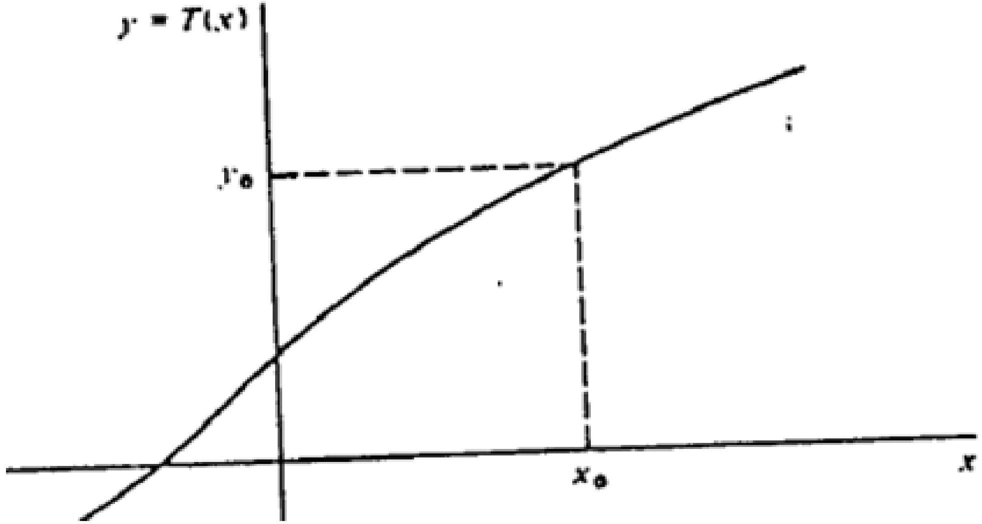
- Quite often one may wish to transform one random variable X into a new random variable Y by means of a transformation



- Typically, the density function $f_x(x)$ or distribution function $F_x(x)$ of X is known , and the problem is to determine either the density function $f_y(y)$ or distribution function $F_y(y)$ of Y.
- The transformation T can be linear, nonlinear, segmented ,staircase, etc

Monotonic Transformation of a Continuous Random variable

- A transformation T is called *monotonically increasing* if $T(x_1) < T(x_2)$ for any $x_1 < x_2$.

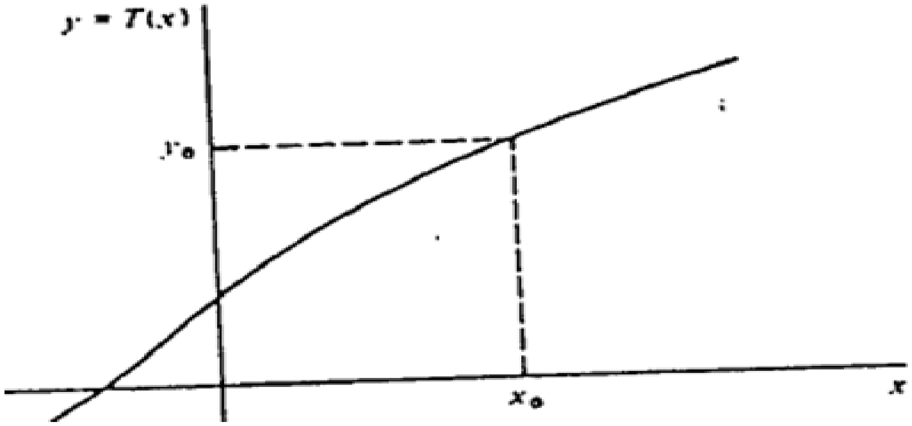


Monotonic Transformations of a Random Variable

- Consider first the increasing transformation as shown in figure. we assume that T is continuous and differentiable at all values of x for which $f_x(x) \neq 0$.
- Let Y have a particular value y_0 corresponding to the particular value of x_0 of X. The two numbers are related by

$$y_0 = T(x_0) \qquad \text{or} \qquad x_0 = T^{-1}(y_0)$$

where T^{-1} represents the inverse transformation.



Monotonic Transformations of a Random Variable

- Now the probability of the event $\{Y \leq y_0\}$ must equal the probability of the event $\{X \leq x_0\}$ because of the one-to-one correspondence between X and Y i.e.

$$F_y(y_0) = P(Y \leq y_0) = P(X \leq x_0) = F_X(x_0)$$

or

$$\int_{-\infty}^{y_0} f_y(y)dy = \int_{-\infty}^{x_0=T^{-1}(y_0)} f_x(x)dx$$

- Differentiating both sides of the above equation with respect to y_0 & using Leibniz's rule, we get

$$f_y(y_0) = f_x[T^{-1}(y_0)]\frac{dT^{-1}(y_0)}{dy_0}$$

Monotonic Transformations of a Random Variable

- Since this result applies for any y, then we write

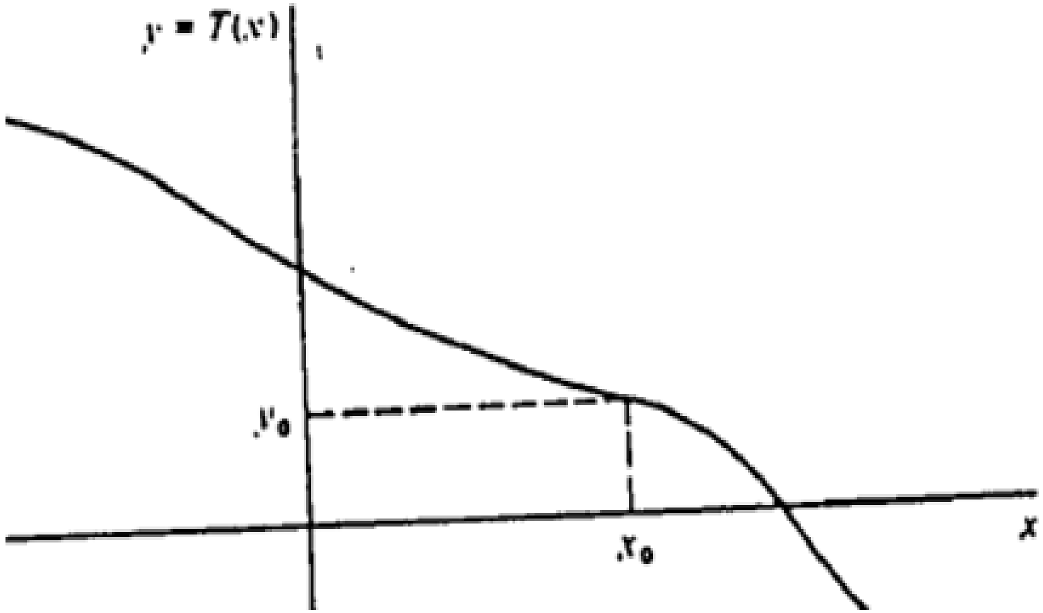
$$f_y(y) = f_x[T^{-1}(y)]\frac{dT^{-1}(y)}{dy}$$

or, more compactly,

$$f_y(y) = f_x(x)\frac{dx}{dy}$$

Monotonic Transformations of a Random Variable

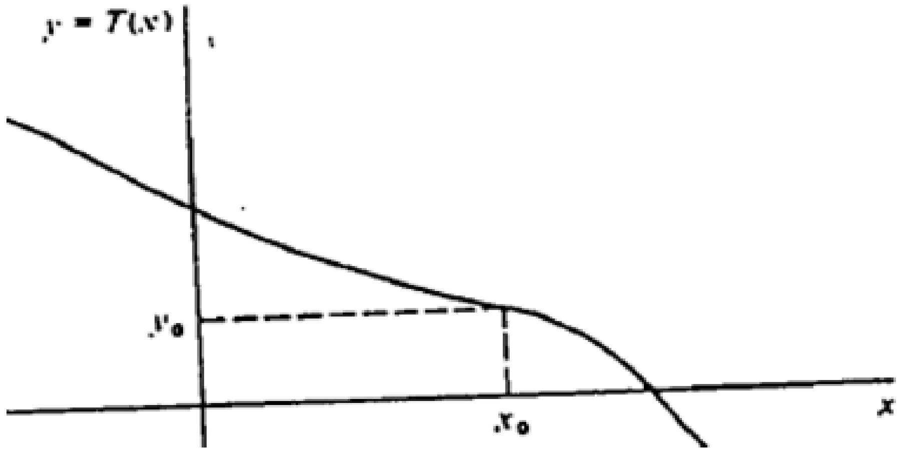
- It is *monotonically decreasing* if $T(x_1) > T(x_2)$ for any $x_1 < x_2$.



Monotonic Transformations of a Random Variable

Similarly for a decreasing transformation, we get

$$F_y(y_0) = p\{Y \leq y_0\} = p\{X \leq x_0\} = 1 - F_x(x)$$



Repeating the above steps, we get

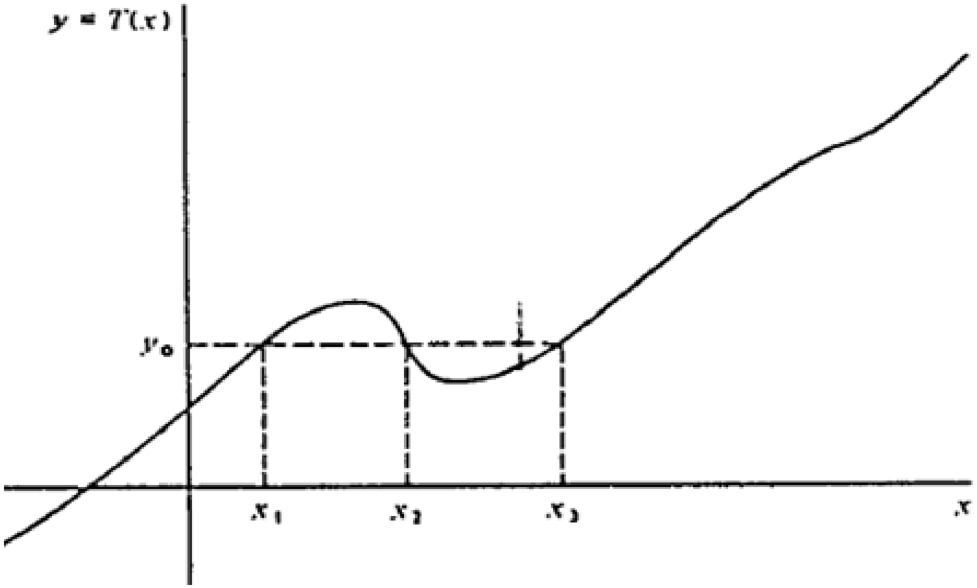
$$-f_y(y_0) = f_x[T^{-1}(y_0)] \frac{dT^{-1}(y_0)}{dy_0}$$

Since the slope of $T^{-1}(y)$is also negative, we again obtain

$$f_y(y) = f_x[T^{-1}(y)] \left|\frac{dT^{-1}(y)}{dy}\right| \text{or} f_y(y) = f_x(x) \left|\frac{dx}{dy}\right||$$

(same for both increasing and decreasing transformations)

Nonmonotonic transformations of a continuous random variable



- In this case, there may be more than one interval of values of X that correspond to the event $\{Y \le y_0\}$ corresponds to the event $\{X \le x_1 \; and \; x_2 \le X \le x_3\}$

Non-Monotonic Transformations of a Random Variable

- Thus, the probability of the event $\{Y \le y_0\}$ now equals the probability of the event $\{x \text{ values yielding } Y \le y_0\}$, which we shall write as $\{x|Y \le y_0\}$ i.e.,

$$F_y(y_0) = p\{Y \le y_0\} = p\{x|Y \le y_0\} = \int_{\{x|Y \le y_0\}} f_x(x)dx$$

- Differentiating we get the density function of Y as

$$f_y(y_0) = \frac{d}{dy_0}\int_{\{x|Y \le y_0\}} f_x(x)\,dx$$

Non-Monotonic Transformations of a Random Variable

- The density function is also given by

$$f_y(y) = \sum_n \frac{f_x(x_n)}{\left|\left.\frac{dT(x)}{dx}\right|_{x=x_0}\right|}$$

where the sum is taken so as to include all the roots $x_n, n = 1, 2, \ldots \ldots$,which are the real solutions of the equation

$$y = T(x)$$

Transformation of a Discrete Random Variable

- If X is a discrete random variable

$$f_X(x) = \sum_n p(x_n)\delta(x - x_n)$$

$$F_X(x) = \sum_n p(x_n)u(x - x_n)$$

where the sum is taken to include all the possible values x_n , $n = 1, 2, \ldots\ldots,$ of X.

- If the transformation $Y = T(X)$ is continuous and monotonic, there is a one-to-one correspondence between X and Y so that a set $\{x_n\}$, through the equation $y_n = T\{x_n\}$ so that $P\{y_n\} = P\{x_n\}$.

Transformations of a Discrete Random Variable

- If the transformation $Y = T(X)$ is continuous and monotonic, there is a one-to-one correspondence between X and Y so that a set $\{x_n\}$, through the equation $y_n = T\{x_n\}$ so that $P\{y_n\} = P\{x_n\}$.
- Thus, we have

$$f_y(y) = \sum_n p(y_n)\delta(y - y_n)$$

$$F_y(y) = \sum_n p(y_n)u(y - y_n) \text{ where } y_n = T(x_n)$$

- If T is not monotonic , the above procedure remains same, but $P(y_n)$ will equal the sum of the probabilities of the various x_n for which $y_n = T(x_n)$

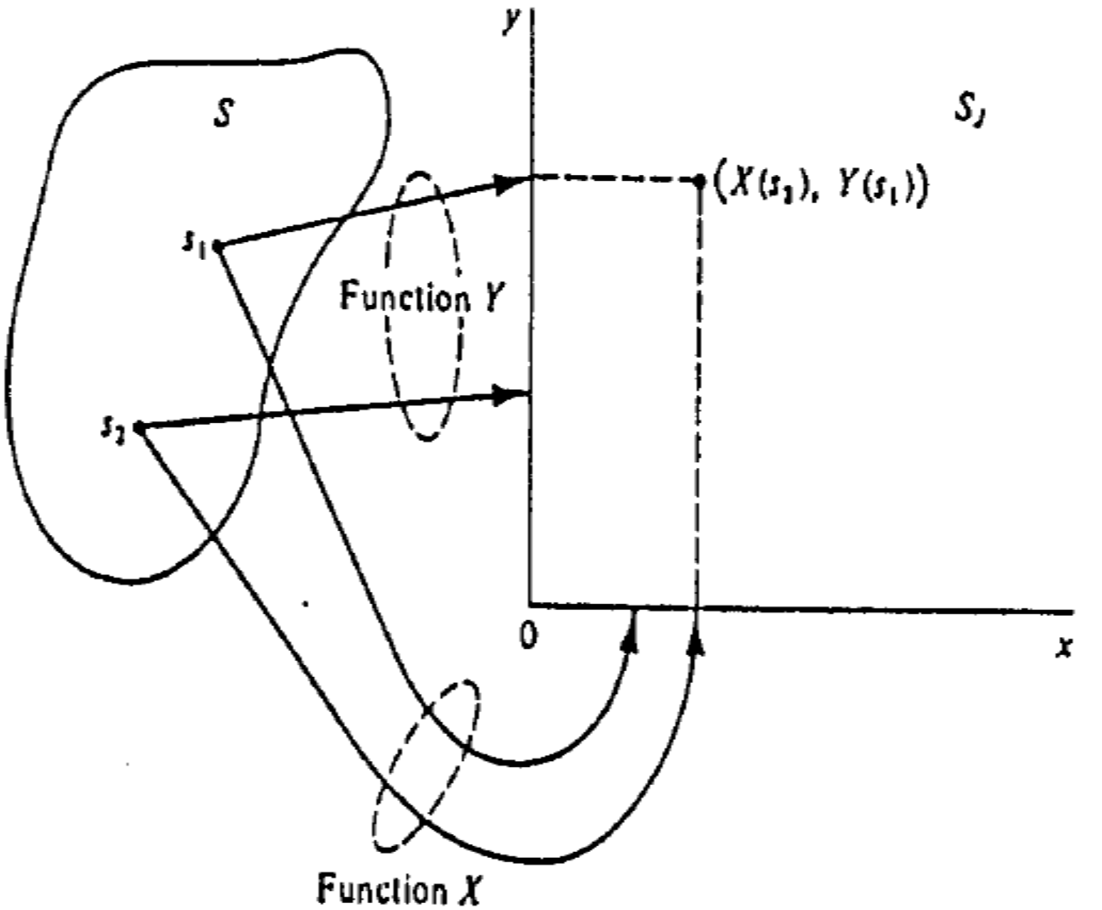
Vector random variables

- **There are many cases where the outcome is a vector of numbers. We have already seen one such experiment, in, where a dart is thrown at random on a dartboard of radius *r. The outcome is a pair (X, Y) of* random variables that are such that $X^2 + Y^2 \leq r^2$.**

- **we measure voltage and current in an electric circuit with known resistance. Owing to random fluctuations and measurement error, we can view this as an outcome (*V, I*)of a pair of random variables.**

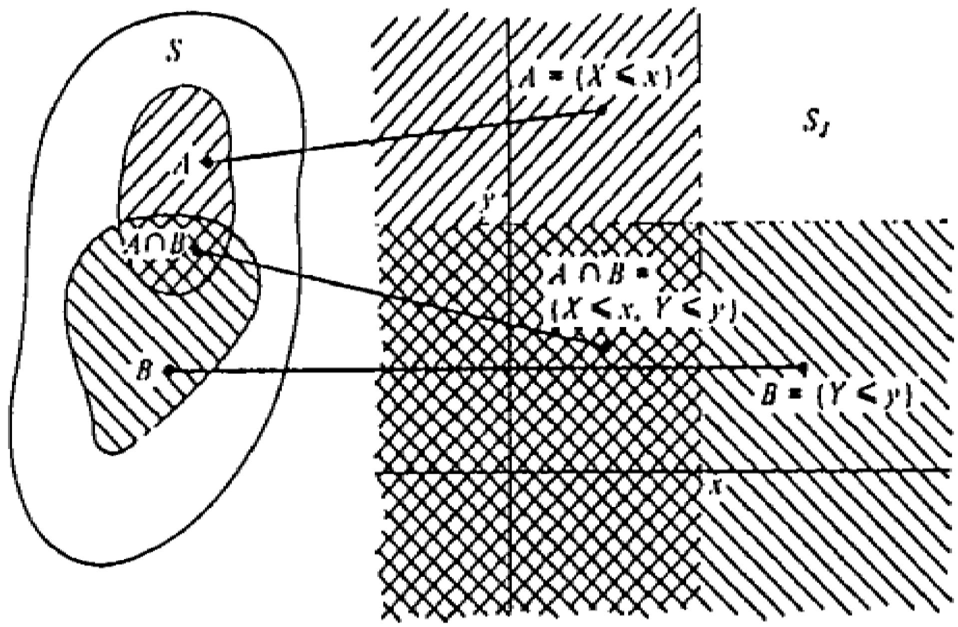
Vector random variables

Mapping the sample space to joint sample space



Vector random variables

Comparison of sample space s with sj



Joint distribution function

- **Let *X and Y be random variables. The pair (X, Y) is then called a* (two-dimensional) *random vector.*t**

- **The *joint distribution function (joint cdf) of (X, Y) is defined as* $F(x, y) = P(X \leq x, Y \leq y)$**

  for $x, y \in R$.

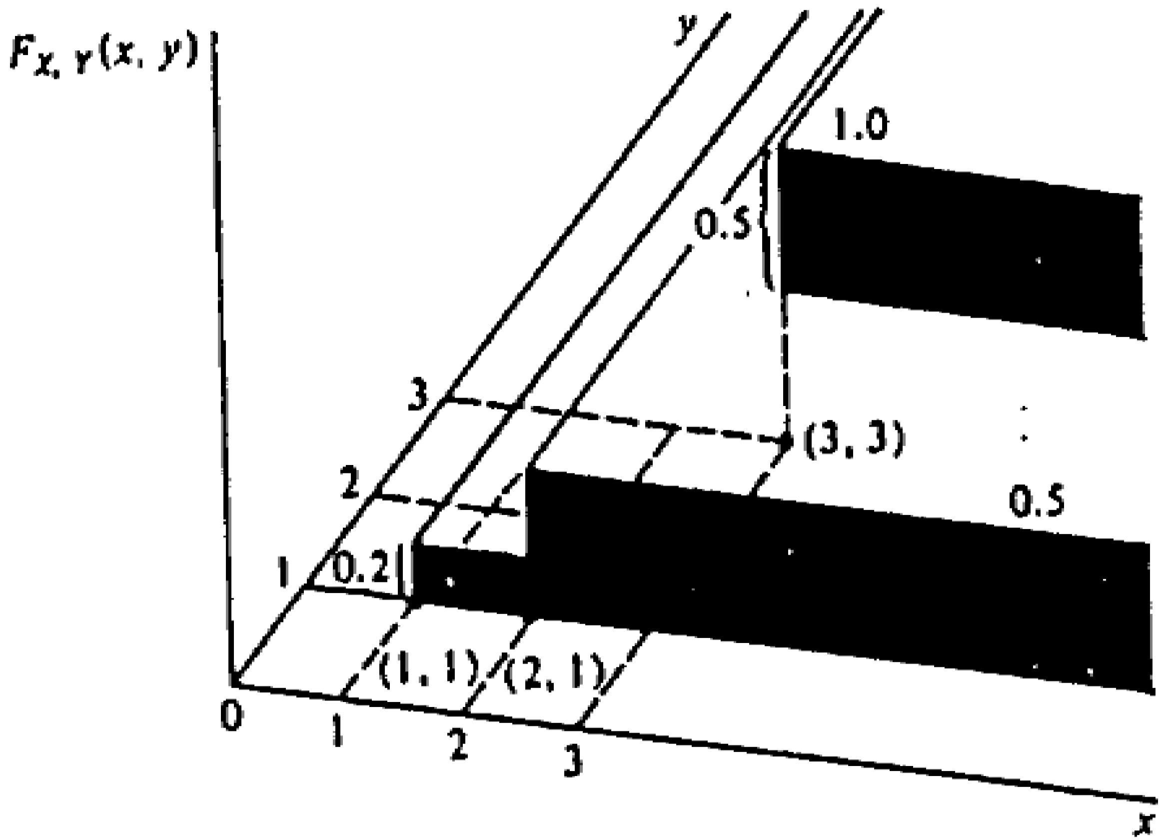
Joint distribution function

- **Assume the joint sample space S_J has only three possible elements (1,1),(2,1),(3,3).The probabilities of the elements are to be P(1,1)=0.2,P(2,1)=0.3 ,P(3,3)=0.5.We find $F_{X,Y}(X,Y)$**

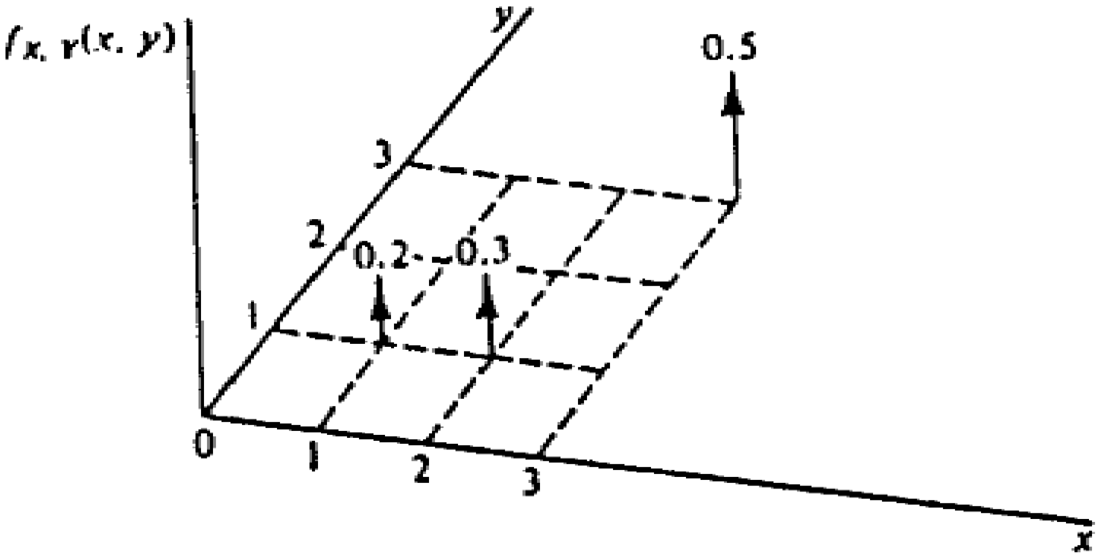
Joint distribution function

- **In constructing joint distribution function we observe that has no elements for $x<1, y<1$.only at the point (1,1)does the function assume a step value. So long as $x\geq1, y\geq1$ this probability is maintained.For larger x and y the point(2,1) produces a second stair step of 0.3 which holds the region $x\geq2, y\geq1$.The second step is added to the first.Finally third step of 0.5 is added to the two for $x\geq3, y\geq3$**

Joint distribution function



Joint density function



Properties of Joint Distribution

- **properties:**

1)

$$F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$$

$$\{X \leq -\infty, Y \leq y\} \subseteq \{X \leq -\infty\}$$

2) $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$ if $x_1 \leq x_2$ and $y_1 \leq y_2$

If $x_1 < x_2$ and $y_1 < y_2$,

$\{X \leq x_1, Y \leq y_1\} \subseteq \{X \leq x_2, Y \leq y_2\}$

$\therefore P\{X \leq x_1, Y \leq y_1\} \leq P\{X \leq x_2, Y \leq y_2\}$

$\therefore F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$

Properties of joint distribution

3)

$$F_{X,Y}(\infty,\infty) = 1$$

4) $F_{X,Y}(x,y)$

$$F_X(x) = F_{XY}(x,+\infty)$$

5) If $x_1 < x_2$ and $y_1 < y_2$

$$P(\{x_1 < X \leq x_2,\ y_1 < Y \leq y_2\}) = F_{X,Y}(x_2,y_2) - F_{X,Y}(x_1,y_2) - F_{X,Y}(x_2,y_1) + F_{X,Y}(x_1,y_1)$$

$$F_{X,Y}(x,y),\ -\infty < x < \infty, -\infty < y < \infty$$

Properties of joint distribution

6)

$$F_X(x) = F_{XY}(x, +\infty)$$

$$\{X \le x\} = \{X \le x\} \cap \{Y \le +\infty\}$$

$$\therefore F_X(x) = P(\{X \le x\}) = P(\{X \le x, Y \le \infty\}) = F_{X,Y}(x, +\infty)$$

$$F_Y(y) = F_{XY}(\infty, y)$$

$$F_{X,Y}(x,y), \; -\infty < x < \infty, -\infty < y < \infty$$

$F_X(x)$ and $F_Y(y)$ Called marginal cumulative distribution function

Marginal distribution functions

- **The distribution of one random variable can be obtained by setting the other value to infinity in $F_{X,Y}(x,y)$.The functions obtained in this manner $F_X(x)$,$F_Y(y)$ are called marginal distribution functions.**

Marginal distribution functions

- **Example:**

$F_{X,Y}(x,y)$=P(1,1)u(x-1)u(y-1)+P(2,1)u(x-2)u(y-1)+ P(3,3)u(x-3)u(y-3)

P(1,1)=0.2, P(2,1)=0.3, P(3,3)=0.5 if we set y=∞ then

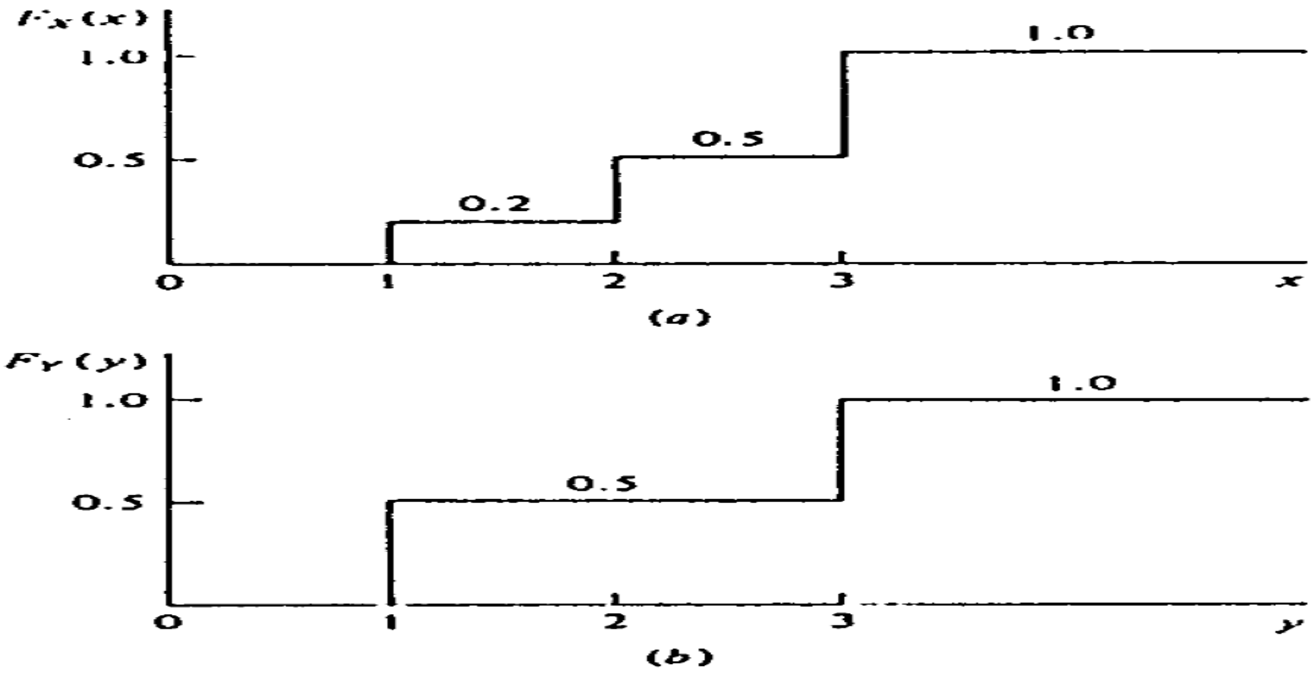
$F_X(x)$= 0.2u(x-1)+0.3u(x-2)+ 0.5u(x-3)

similarly

$F_Y(y)$= 0.2u(y-1)+0.3u(y-1)+ 0.5u(y-3)

- **=0.5u(y-1)+0.5u(y-3)**

Marginal distribution functions



Marginal distribution functions

- **Consider two jointly distributed random variables and with the joint CDF**

$$F_{X,Y}(x,y)=\begin{cases}(1-e^{-2x})(1-e^{-y}) & x\geq 0, y\geq 0\\ 0 & \text{otherwise}\end{cases}$$

1)Find the marginal CDFs

2) Find the probability P(1<x≤2, 1<y≤2)

Marginal distribution functions

a)

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y) = \begin{cases} 1-e^{-2x} & x \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y) = \begin{cases} 1-e^{-y} & y \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$P\{1 < X \leq 2,\ \ 1 < Y \leq 2\} = F_{X,Y}(2,2) + F_{X,Y}(1,1) - F_{X,Y}(1,2) - F_{X,Y}(2,1)$$

$$= (1-e^{-4})(1-e^{-2}) + (1-e^{-2})(1-e^{-1}) - (1-e^{-2})(1-e^{-2}) - (1-e^{-4})(1-e^{-1})$$

$$= 0.0272$$

Joint Probability Density Function

If and are two continuous random variables and their joint distribution function is continuous in both and then we can define *joint probability density function* by

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y),$$

provided it exists.

Clearly $$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v) dv du$$

Properties of Joint Probability Density Function

1 $f_{X,Y}(x,y)$ **is always a non-negative quantity. That is,**

$$f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$$

The probability of any Borel can be obtained by

2)

3)

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

$$P(B) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\,dx\,dy$$

Marginal density functions

- **The marginal density functions and of two joint RVs are given by the derivatives of the corresponding marginal distribution functions. Thus**

$$
\begin{aligned}
f_X(x) &= \frac{d}{dx} F_X(x) \\
&= \frac{d}{dx} F_X(x, \infty) \\
&= \frac{d}{dx} \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{X,Y}(u, y)dy \right) du \\
&= \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy
\end{aligned}
$$

and similarly $\quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx$

Marginal density functions

- The marginal CDF and pdf are same as the CDF and pdf of the concerned single random variable. The *marginal* term simply refers that it is derived from the corresponding joint distribution or density function of two or more jointly random variables.
- With the help of the two-dimensional Dirac Delta function, we can define the joint pdf of two discrete jointly random variables. Thus for discrete jointly random variables and

$$f_{X,Y}(x,y) = \sum_{(x_i,y_j)\in R_X \times R_Y.} \sum p_{X,Y}(x,y)\delta(x-x_i, y-y_j)$$

Marginal density functions

- **The joint density function**

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-2x})(1-e^{-y}) & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

$$= \frac{\partial^2}{\partial x \partial y}[(1-e^{-2x})(1-e^{-y})] \quad x \geq 0, y \geq 0$$

$$= 2e^{-2x}e^{-y} \quad x \geq 0, y \geq 0$$

Marginal density functions

- **The joint pdf of two random variables X and Y are given by**

$$f_{X,Y}(x,y) = cxy \quad 0 \le x \le 2,\ 0 \le y \le 2$$
$$= 0 \ \text{otherwise}$$

1)Find c

2)Find $\mathrm{f_{XY}(x, y)}$

3)Find $f_X(x)$ **and** $f_Y(y)$

4)What is the probability P(0<x≤1, 0<y≤1)

Marginal density functions

- **Solution:**

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = c\int_0^2\int_0^2 xydydx$$

$$\therefore c = \frac{1}{4}$$

$$F_{X,Y}(x,y) = \frac{1}{4}\int_0^y\int_0^x uvdudv$$

$$= \frac{x^2y^2}{16}$$

$$f_X(x) = \int_0^2 \frac{xy}{4}dy \quad 0 \le y \le 2$$

$$= \frac{x}{2} \qquad 0 \le y \le 2$$

Similarly

$$f_Y(y) = \frac{y}{2} \qquad 0 \le y \le 2$$

Marginal density functions

$$P(0 < X \le 1, 0 < Y \le 1)$$

$$= F_{X,Y}(1,1) + F_{X,Y}(0,0) - F_{X,Y}(0,1) - F_{X,Y}(1,0)$$

$$= \frac{1}{16} + 0 - 0 - 0$$

$$= \frac{1}{16}$$

CONDITIONAL DISTRIBUTIONS

We discussed the conditional CDF and conditional PDF of a random variable conditioned on some events defined in terms of the same random variable. We observed that

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

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

Conditional Probability Density Functions

Suppose and are two discrete jointly random variable with the joint PMF fxy(x,y) . The conditional PMF of y given x=x is denoted by and defined as

$$f_{y/x}(y/x)$$

$$p_{Y/X}(y/x) = P(\{Y = y\}/\{X = x\})$$

$$= \frac{P(\{X = x\} \cap \{Y = y\})}{P\{X = x\}}$$

$$= \frac{p_{X,Y}(x,y)}{p_X(x)} \quad \text{provided} \quad p_X(x) \neq 0$$

Thus,

$$p_{Y/X}(y/x) = \frac{p_{X,Y}(x,y)}{p_X(x)} \quad \text{provided} \quad p_X(x) \neq 0$$

Conditional Probability Distribution Function

- **Consider two continuous jointly random variables and with the joint probability distribution function We are interested to find the conditional distribution function of one of the random variables on the condition of a particular value of the other random variable.**
- **We *cannot* define the conditional distribution function of the random variable on the condition of the event by the relation**

$$F_{Y/X}(y/x) = P(Y \le y / X = x)$$

$$= \frac{P(Y \le y, X = x)}{P(X = x)}$$

Conditional Probability Distribution Function

$$F_{Y/X}(y/x) = \lim_{\Delta x \to 0} P(Y \le y / x < X \le x + \Delta x)$$

$$= \lim_{\Delta x \to 0} \frac{P(Y \le y, x < X \le x + \Delta x)}{P(x < X \le x + \Delta x)}$$

$$= \lim_{\Delta x \to 0} \frac{\int_{\infty}^{y} f_{X,Y}(x,u)\Delta x\, du}{f_X(x)\Delta x}$$

$$= \frac{\int_{\infty}^{y} f_{X,Y}(x,u)\, du}{f_X(x)}$$

$$\therefore F_{Y/X}(y/x) = \frac{\int_{\infty}^{y} f_{X,Y}(x,u)\, du}{f_X(x)}$$

CONDITIONAL PROBABILITY DENSITY FUNCTION

- $f_{Y/X}(y/X=x)=f_{Y/X}(y/x)$ is called the *conditional probability density function*

define the conditional distribution function .

$$\lim_{\Delta y\to 0,\Delta x\to 0}\left(F_{Y/X}(y+\Delta y/x<X<x+\Delta x)-F_{Y/X}(y/x<X<x+\Delta x)\right)/\Delta y$$

$$=\lim_{\Delta y\to 0,\Delta x\to 0}\left(P(y<Y\le y+\Delta y/x<X\le x+\Delta x)\right)/\Delta y$$

$$=\lim_{\Delta y\to 0,\Delta x\to 0}\left(P(y<Y\le y+\Delta y,x<X\le x+\Delta x)\right)/P(x<X\le x+\Delta x)\Delta y$$

$$=\lim_{\Delta y\to 0,\Delta x\to 0}f_{X,Y}(x,y)\Delta x\Delta y/f_X(x)\Delta x\Delta y$$

$$=f_{X,Y}(x,y)/f_X(x)$$

Point conditioning

First consider the case when X and Y are both discrete. Then the marginal pdf's

$f_Y(y)=P(Y=y)$ $f_X(x)=P(X=x)$

The joint pdf is, similarly

$f_{X,Y}(x,y)=P(X\leq x,Y\leq y)$

Conditional density function is given by

$f_X(x/B)=$ $\frac{dFX(x|B)}{dx}$

Point conditioning

$$F_{Y/X}(y/x) = P(Y \le y / X = x)$$

$$= \frac{P(Y \le y, X = x)}{P(X = x)}$$

Distribution function of one random variable X conditioned by that second variable Y has some specific values of y. This is called point conditioning

B={y-Δy<Y≤y+Δy}

Where Δy is a small quantity that we eventually let approach 0.

Point conditioning

Fx(x/ y-Δy<Y≤y+Δy)=

$$\frac{\int_{y-\Delta y}^{y+\Delta y}\int_{-\infty}^{x} f_{X,Y}(\xi_1,\xi_2)\,d\xi_1\,d\xi_2}{\int_{y-\Delta y}^{y+\Delta y} f_Y(\xi)\,d\xi}$$

$$F_{X,Y}(x,y)=\sum_{i=1}^{N}\sum_{j=1}^{M}P(x_{i,}y_j)\delta(x-x_i)\delta(y-y_j)$$

$$\mathrm{Fx}(x/Y = yk)=\sum_{i=1}^{N}\frac{P(x_i,y_k)}{P(y_k)}u(x-x_i)$$

$$\mathrm{fx}(x/Y = yk)=\sum_{i=1}^{N}\frac{P(x_i,y_k)}{P(y_k)}\delta(x-x_i)$$

Point conditioning

- **As Δy->0 denominator becomes zero. For smaller Δy values conditional density may exist.**

Fx(x/ y-Δy<Y≤y+Δy)=

$$\frac{\int_{-\infty}^{x} f_{X,Y}(\xi_1, y)d\xi_1 2\Delta y}{f_Y(y)2\Delta y}$$

$$\text{Fx}(x/\ Y = y) = \frac{\int_{-\infty}^{x} f_{X,Y}(\xi, y)d\xi}{f_Y(y)}$$

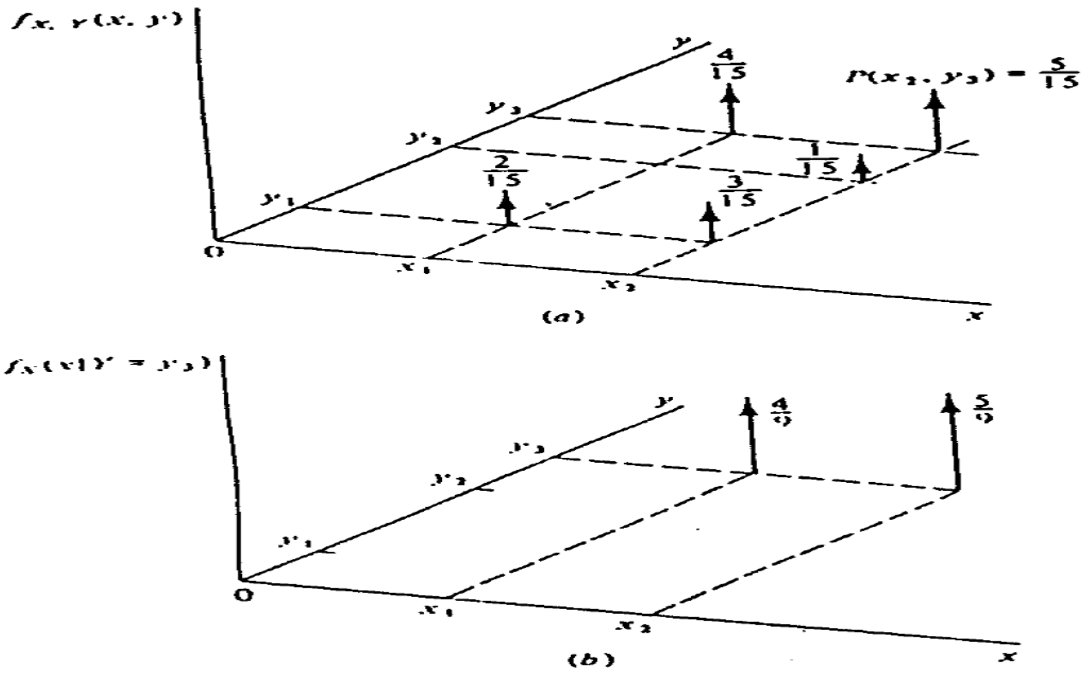
Interval Conditioning

- **Distribution function of one random variable X conditioned by that second variable Y has some specific values of y. This is called point conditioning**

$B=\{y_a<Y\leq y_b\}$

Example

$P(x_1,y_1)=2/15, P(x_2,y_1)=3/15$.etc.since $P(y_3)=4/15+5/15=9/15$ find $f_x(x/y=y_3)$



Statistical independence

$$F_{X,Y}(x,y) = P\{X \le x, Y \le y\}$$

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \, \partial y} F_{X,Y}(x,y)$$

Statistical independence

$$F_{X,Y}(x,y) = P\{X \le x, Y \le y\}$$

$$= P\{X \le x\} P\{Y \le y\}$$

$$= F_X(x) F_Y(y)$$

$$\therefore f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

$$= \frac{dF_X(x)}{dx} \frac{dF_Y(y)}{dy}$$

$$= f_X(x) f_Y(y)$$

$$\therefore f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

Statistical independence

- **Statistical independence**

$$f_{Y/X}(y) = f_Y(y)$$

$$F_X(x \mid Y \le y) = \frac{P\{X \le x, Y \le y\}}{P\{Y \le y\}} = \frac{F_{X,Y}(x, y)}{F_Y(y)}$$

$$F_X(x \mid Y \le y) = F_X(x)$$

In other words the conditional distribution ceases to be conditional and simply equals the marginal distribution for independent random variables. It can also be shown that

$$f_X(x \mid Y \le y) = f_X(x)$$
$$f_Y(y \mid X \le x) = f_Y(y)$$

EXAMPLE

- For discrete variables independence means the probability in a cell must be the product of the marginal probabilities of its row and column. In the first table below this is true: every marginal probability is 1/6 and every cell contains 1/36, i.e. the product of the marginal's. Therefore X and Y are independent. In the second table below most of the cell probabilities are not the product of the marginal probabilities. For example, none of marginal probabilities are 0, so none of the cells with 0 probability can be the product of the marginal's.

EXAMPLE

$X\backslash Y$	1	2	3	4	5	6	$p(x_i)$
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
$p(y_j)$	1/6	1/6	1/6	1/6	1/6	1/6	1

$X\backslash T$	2	3	4	5	6	7	8	9	10	11	12	$p(x_i)$
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0	1/6
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	1/6
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	1/6
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	1/6
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	1/6
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	1/6
$p(y_j)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

EXAMPLE

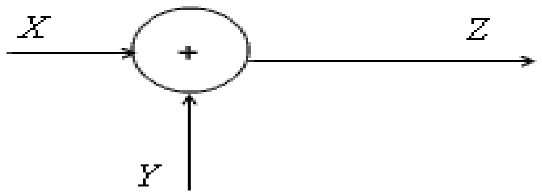
- For continuous variables independence means you can factor the joint pdf or cdf as the product of a function of x and a function of y. (i) Suppose X has range [0, 1/2], Y has range [0, 1] and $f(x, y) = 96x^2 y^3$ then X and Y are independent. The marginal densities are $f_X(x) = 24x^2$ and $f_Y(y) = 4y^3$. (ii) If $f(x, y) = 1.5(x^2+y^2)$ over the unit square then X and Y are not independent because there is no way to factor f(x, y) into a product $f_X(x)f_Y(y)$. (iii) If $F(x, y) = 1/2\,(x^3y + xy^3)$ over the unit square then X and Y are not independent .because the cdf does not factor into a product $F_X(x)F_Y(y)$.

Sum of two random variables

We are often interested in finding out the probability density function of a function of two or more RVs

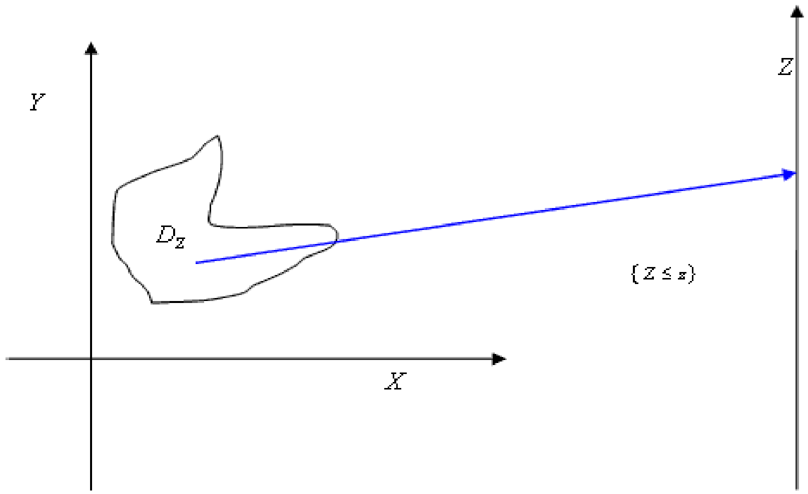
The received signal by a communication receiver is given by

$$Z = X + Y$$



Sum of two random variables

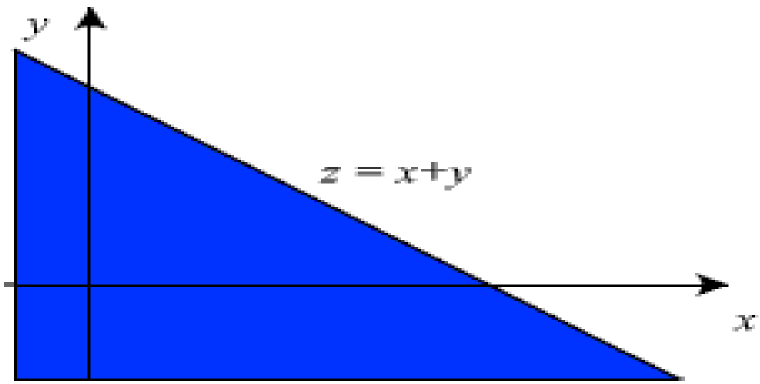
corresponding to each z. $\{Z \le z\}$ We can find a variable subset $D_Z = \{(x,y) \mid g(x,y) \le z\}$



$$\therefore F_Z(z) = P(\{Z \le z\})$$
$$= P\{(x,y) \mid (x,y) \in D_z\}$$
$$= \iint_{(x,y)\in D_z} f_{X,Y}(x,y)\,dy\,dx$$

Probability density function Sum of two random variables

Probability density function of *Z* = *X* + *Y* .



$Z \leq z$

$\Rightarrow X + Y \leq z$

$$F_Z(z) \stackrel{D_Z}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{z-x'} f_{X,Y}(x', y') dy' dx'.$$

Probability density function Sum of two random variables

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{z-x} f_{X,Y}(x,y)dy\right]dx$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{z} f_{X,Y}(x,u-x)du\right]dx \quad \text{substituting } y = u - x$$

$$= \int_{-\infty}^{z}\left[\int_{-\infty}^{\infty} f_{X,Y}(x,u-x)dx\right]du \quad \text{interchanging the order of integration}$$

$$\therefore f_Z(z) = \frac{d}{dz}\int_{-\infty}^{z}\left[\int_{-\infty}^{\infty} f_{X,Y}(x,u-x)dx\right]du$$

$$= \int_{-\infty}^{\infty} f_{X,Y}(x,u-x)dx$$

$$\therefore f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x,u-x)dx$$

Probability density function Sum of two random variables

$$f_{X,Y}(x, z-x) = f_X(x) f_Y(z-x)$$

$$\therefore f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$$

$$= f_X(z) * f_Y(z)$$

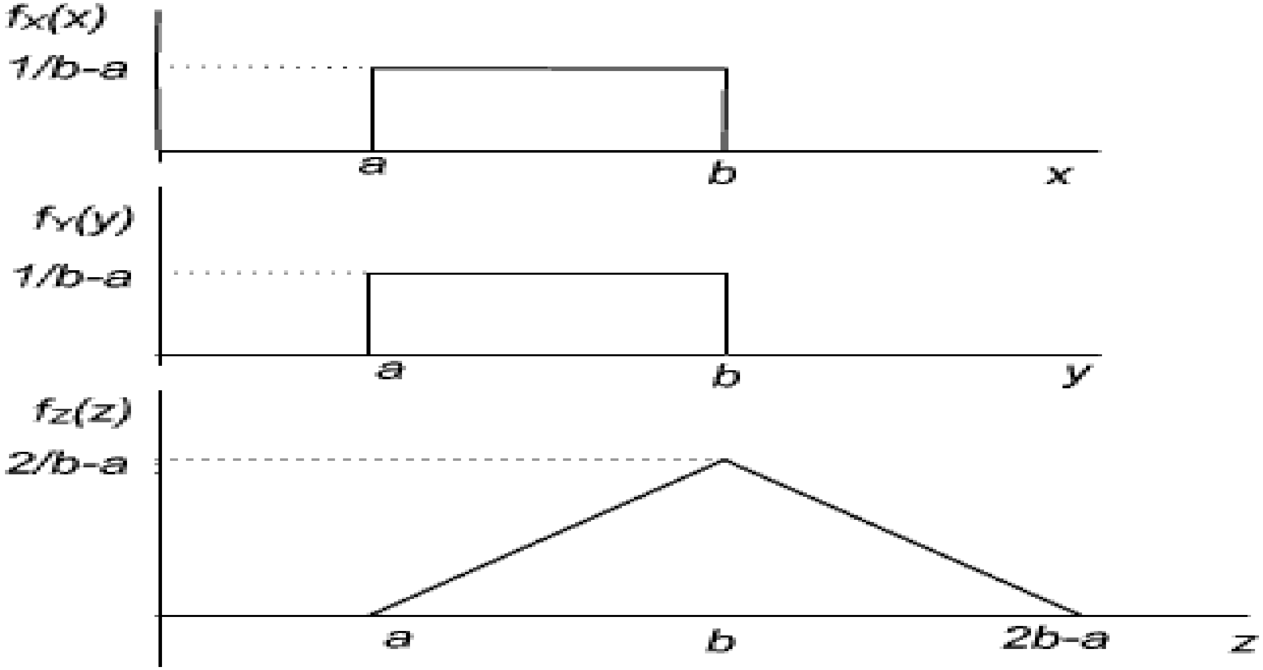
Thus the pdf for the sum of two random variables is given by a superposition integral. If *X* and *Y* are independent random variables, then the pdf is given by the convolution integral of the margial pdf's of *X* and *Y* :

Example

- **Suppose *X* and *Y* are independent random variables and each uniformly distributed over (*a*, *b*). And are as shown in the figure below.**

$f_X(x)$ $f_Y(y)$

Example



Central Limit Theorem

Consider n **independent** random variables x_1, x_2, x_3......xn ,The mean and variance of each of the random variables are assumed to be known. Suppose $E[x]=\mu_x$ $var(x)=\sigma_x^2$ and . Form a random variable

$$Y_N=X_1+X_2+.......X_N$$

The mean and variance of Y_N are given by

$$E[yn]= \mu_{x\,1} + \mu_{x\,2} + \mu_{x\,3}......... + \mu_{x\,n}$$

Central Limit Theorem

$$\operatorname{var}(Y_N) = \sigma^2_{Y_N} = E\left\{ \sum_{i=1}^{N} (X_i - \mu_{X_i}) \right\}^2$$

$$= \sum_{i=1}^{N} E(X_i - \mu_{X_i})^2 + \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} E(X_i - \mu_i)(X_j - \mu_j)$$

$$= \sigma^2_{X_1} + \sigma^2_{X_2} + \ldots + \sigma^2_{X_n}$$

$\because X_i$ and X_j are independent for $i \neq j$.

Central Limit Theorem

The CLT states that under very general conditions $\left\{ Y_n = \sum_{i=1}^{n} X_i \right\}$ converges *in distribution* to $Y \sim N(\mu_Y, \sigma_Y^2)$ as $n \rightarrow \infty$

1. The random variables are independent and identically distributed.
2. The random variables are independent with same mean and variance, but not identically distributed.
3. The random variables are independent with different means and same variance and not identically distributed.
4. The random variables are independent with different means and each variance being neither too small nor too large.

D_Z

Proof of Central Limit Theorem

$$\mu_{Y_n} = 0,$$
$$\sigma^2_{Y_n} = \sigma^2_X,$$
$$E(Y_n^3) = E(X^3)/\sqrt{n} \text{ and so on.}$$

$$\phi_{Y_n}(\omega) = E\left(e^{j\omega Y_n}\right) = E\left(e^{\left(j\omega \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i\right)}\right)$$

$$e^{j\omega y_n}$$

$$e^{j\omega Y_n} = 1 + j\omega Y_n + \frac{(j\omega)^2}{2!} Y_n^2 + \frac{(j\omega)^3}{3!} Y_n^3 + \ldots$$

Proof of Central Limit Theorem

$$\phi_{Y_n}(\omega) = E\left(e^{j\omega Y_n}\right) = 1 + j\omega\mu_{Y_n} + \frac{(j\omega)^2}{2!}E(Y_n^2) + \frac{(j\omega)^2}{3!}E(Y_n^3) + \ldots$$

$\mu_{Y_n} = 0$ and $E(Y_n^2) = \sigma_{Y_n}^2 = \sigma_X^2$, we get

$$\phi_{Y_n}(\omega) = 1 - (\omega^2 / 2!)\sigma_X^2 + R(\omega, n)$$

$$\lim_{n \to \infty} R(\omega, n) = 0$$

Proof of Central Limit Theorem

$$\therefore \lim_{n\to\infty} \phi_{Y_n}(\omega) \simeq 1 - \frac{\omega^2}{2!}\sigma^2{}_X \simeq e^{-\frac{\omega^2\sigma_X^2}{2}}$$

which is the characteristic function of a Gaussian random variable with 0 mean and variance ${\sigma_x}^2$

$$Y_n \xrightarrow{d} N(0, \sigma_X^2)$$

Module-III

OPERATIONS ON MULTIPLE RANDOM VARIABLES – EXPECTATIONS

Expected Values of Random Variables

If g(x,y) is a function of a continuous random variables X and Y then then the expected value of is given by

$$\bar{g} = E\left[g(X,Y)\right] = \begin{cases} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dxdy & \text{Continuous} \\ \sum_{i}\sum_{k} g(x_i,y_k)P_{X,Y}(x_i,y_k) & \text{Discrete} \end{cases}$$

Example

The joint pdf of two random variables is given by

$$f_{XY}(x,y)=\frac{1}{4}xy \quad 0\leq x\leq 2 \quad 0\leq y\leq 2$$

Find the joint expectation of $g(X,Y)=x^2y$

$E(g(x,y)= E(x^2y)$

Example

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{XY}(x,y)\,dxdy$$

$$= \int_{0}^{2}\int_{0}^{2} x^2 y \frac{1}{4} xydxdy$$

$$= \frac{1}{4}\int_{0}^{2} x^3dx\int_{0}^{2} y^2dy$$

$$= \frac{1}{4}\times\frac{2^4}{4}\times\frac{2^3}{3}$$

$$= \frac{8}{3}$$

Example

Consider the discrete random variables x and y. The joint probability mass function of the random variables are tabulated in Table . Find the joint expectation of g(x,y)=xy.

$Y \backslash X$	0	1	2	$p_Y(y)$
0	0.25	0.1	0.15	0.5
1	0.14	0.35	0.01	0.5
$p_X(x)$	0.39	0.45	0.16	

Example

$$E[XY] = \sum_{x}\sum_{y} g(x, y)\, p_{XY}(x, y)$$

$$= 1 \times 1 \times 0.35 + 1 \times 2 \times 0.01$$

$$= 0.37$$

Properties

- Expectation is a linear operator. We can generally write

$E[a_1g_1(x,y)+a_2g_2(x,y)=a_1E(g_1(x,y)+a_2E(g_2(x,y))$

$E[xy+5\log_e xy]=E[xy]+5E[\log_e xy]$

Properties

If x and y are independent random variables and g(x,y)=g1(x,y)×g2(x,y) then E[g(x,y)]=E[g1(x,y)]×E[g2(x,y]

$$
\begin{aligned}
Eg(X,Y) &= Eg_1(X)g_2(Y) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(X)g_2(Y)f_{X,Y}(x,y)dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(X)g_2(Y)f_X(x)f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} g_1(X)f_X(x)dx \int_{-\infty}^{\infty} g_2(Y)f_Y(y)dy \\
&= Eg_1(X)Eg_2(Y)
\end{aligned}
$$

Joint moments about the origin

For two continuous random variables X and Y, *the joint moment of order m+n* is defined as

$$E(X^m Y^n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^m y^n f_{XY}(x,y)dxdy$$

And the joint central moment of order m+n is defined as

$$E(X-\mu_x)^m E(Y-\mu_y)^n = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)^m (y-\mu_y)^n f_{XY}(x,y)dxdy$$

$$\mu_x = E[x]$$

$$\mu_y = E[y]$$

Joint moments about the origin

For two discrete random variables X and Y, *the joint moment of order m+n* is defined as

$$E(X^m Y^n) = \sum_x \sum_y x^m y^n f_{XY}(x,y)dxdy$$

$$E(X-\mu_x)^m E(Y-\mu_y)^n = \sum_x \sum_y (x-\mu_x)^m (y-\mu_y)^n f_{XY}(x,y)$$

$$\mu_x = E[x]$$

$$\mu_y = E[y]$$

Covariance of two random variables

The *covariance* of two random variables X and Y is defined as

$$Cov(X,Y)=E(X-\mu_x)E(Y-\mu_y)$$

Cov(X, Y) is also denoted as σ_{XY}.

$$\begin{aligned} Cov(X,Y) &= E(X-\mu_x)^m E(Y-\mu_y)^n \\ &= E(XY-\mu_y X-\mu_x Y+\mu_x\mu_y) \\ &= E(XY)-\mu_Y E(X)-\mu_X E(y)+\mu_X\mu_Y \\ &= E(XY)-\mu_x\mu_y \end{aligned}$$

Covariance

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)].$$

By expanding and simplifying the right side of (10-10), we also get

$$Cov(X,Y) = E(XY) - \mu_X \mu_Y = E(XY) - E(X)E(Y)$$

$$= \overline{XY} - \overline{X}\,\overline{Y}.$$

It is easy to see that

$$|Cov(X,Y)| \le \sqrt{Var(X)Var(Y)}.$$

To see (10-12), let $U = aX + Y$, so that

$$Var(U) = E\left[\{a(X-\mu_X) + (Y-\mu_Y)\}^2\right]$$

$$= a^2 Var(X) + 2a\,Cov(X,Y) + Var(Y) \ge 0.$$

Covariance of two random variables

The ratio $\rho(X,Y)=\frac{Cov(X,Y)}{\sigma_x\sigma_y}$ is called the **correlation coefficient**.

If $\rho_{X,Y}>0$ then are called positively correlated.

If $\rho_{X,Y}<0$ then are called negatively correlated

If $\rho_{X,Y}=0$ then are uncorrelated.

We will also show that $|\rho(X,Y)|\le 1$

Uncorrelated random variables

Two random variables are called *uncorrelated* if

Cov(X,Y)=0

Which also means E(XY)=$\mu_x \mu_y$

If are independent random variables, then

$$f_{XY}(x,y) = f_X(x) f_Y(y)$$

Thus two independent random variables are always uncorrelated.

Uncorrelated random variables

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)\,f_Y(y)dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X\,Y}(x,y)dxdy$$

$$= \int_{-\infty}^{\infty} xf_X(x)dx\int_{-\infty}^{\infty} yf_Y(y)dy$$

$$= E[x]E[y]$$

joint characteristic function

The *joint characteristic function* of two random variables X and Y is defined by

$$\phi_{XY}(\omega_1, \omega_2) = E[e^{j\omega_1 x + j\varpi_2 y}]$$

If and are jointly continuous random variables, then

$$\phi_{X,Y}(\omega_1, \omega_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y) e^{j\omega_1 x + j\varpi_2 y} dxdy$$

Joint moments about the origin

For two discrete random variables X and Y, *the joint moment of order m+n* is defined as

$$E(X^mY^n) = \sum_x \sum_y x^m y^n f_{XY}(x, y)dxdy$$

And the joint central moment of order m+n is defined as

$$E(X-\mu_x)^m E(Y-\mu_y)^n = \sum_x \sum_y (x-\mu_x)^m (y-\mu_y)^n f_{XY}(x, y)$$

$$\mu_x = E[x]$$

$$\mu_y = E[y]$$

Covariance of two random variables

The *covariance* of two random variables X and Y is defined as

$$Cov(X,Y)=E(X-\mu_x)E(Y-\mu_y)$$

Cov(X, Y) is also denoted as σ_{XY}.

$$Cov(X,Y) = E(X-\mu_x)^m E(Y-\mu_y)^n$$
$$= E(XY - \mu_y X - \mu_x Y + \mu_x \mu_y)$$
$$= E(XY) - \mu_Y E(X) - \mu_X E(y) + \mu_X \mu_Y$$
$$= E(XY) - \mu_x \mu_y$$

Transformations of multiple random variables

The joint density function of new random variable $Y_i=T(X_1,X_2,\ldots\ldots X_N)$ i=1,2,3….n

The random variable Xj can be obtained from inverse transformation

$$X_j=T_j^{-1}(Y_1,Y_2,\ldots\ldots Y_N)$$

$$\left.\begin{aligned} x_1 &= g_1^{-1}(y_1, y_2, \cdots, y_k) \\ x_2 &= g_2^{-1}(y_1, y_2, \cdots, y_k) \\ &\vdots \\ x_n &= g_n^{-1}(y_1, y_2, \cdots, y_{k=n}) \end{aligned}\right\} **$$

Transformations of multiple random variables

- **Assuming that the partial derivatives $\partial g_i^{-1} / \partial y_i$ exist at every point $(y_1, y_2,\ldots,y_{k=n})$. Under these assumptions, we have the following determinant *J***

Transformations of multiple random variables

$$J = det\begin{bmatrix} \frac{\partial g_1^{-1}}{\partial y_1} & \cdots & \frac{\partial g_1^{-1}}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n^{-1}}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}}{\partial y_n} \end{bmatrix}$$

called as the Jacobian of the transformation specified by (**).

Then, the joint pdf of $Y_1, Y_2, \ldots, Y_k$ can be obtained by using the change of variable technique of multiple variables.

Transformations of multiple random variables

- **As a result, the new p.d.f. is defined as follows:**

$$g(y_1, y_2, \cdots, y_n) = \begin{cases} f_{X_1,\cdots,X_n}\left(g_1^{-1}, g_2^{-1}, \cdots, g_n^{-1}\right)|J|, \text{for } (y_1, y_2, \cdots, y_n) \in \Psi \\ 0, otherwise \end{cases}$$

Linearly transformation of Gaussian random variables

- Linearly transforming set of Gaussian random variables $X_1, X_2, \ldots..X_N$ for which the joint density function exists.The new variables $Y_1, Y_2, \ldots..Y_N$ are
- $Y_1 = a_{11}X_1 + a_{12}X_2 + \ldots\ldots + a_{1N}X_N$
- $Y_2 = a_{21}X_1 + a_{22}X_2 + \ldots\ldots + a_{2N}X_{N.}$
- $Y_N = a_{N1}X_1 + a_{N2}X_2 + \ldots\ldots + a_{NN}X_N$

Linearly transformation of Gaussian random variables

- **where the coefficients aij I and j=1,2,..N are real numbers. Now we define the following matrices**

$$[T] = \begin{bmatrix} a11 & a12 \ldots. & a1N \\ a21 & a22 \ldots & a2N \\ aN1 & aN2 \ldots & aNN \end{bmatrix}$$

$$[Y] \quad \begin{bmatrix} Y1 \\ \vdots \\ YN \end{bmatrix}$$

$$X_i = Ti^{-1}(Y_1 \ldots . Y_N) = a^{i1}Y_1 + a^{i2}y_2 + \ldots . + a^{iN}Y_N$$

$$X_i - = Ti^{-1}(Y_1 \ldots . Y_N) = a^{i1}Y_1 + a^{i2}y_2 + \ldots . + a^{iN}Y_N$$

N random variable case

$$Cij = E[(X_i - \hat{X}_i)(X_j - \hat{X}_j)] = \begin{cases} \sigma_{Xi}^{2} & \text{For } i=j \\ C_{XiXj} & \text{For } i \neq j \end{cases}$$

For the special case N=2 The covariance matrix becomes

$$[Cx] = \begin{bmatrix} \sigma_{X1}^{2} & \rho\sigma_{X1}\sigma_{X2} \\ \rho\sigma_{X1}\sigma_{X2} & \sigma_{X2}^{2} \end{bmatrix}$$

N random variable case

$$[Cx]^{-1} = \frac{1}{(1-\rho^2)}\begin{bmatrix} \dfrac{1}{\sigma_{X1}{}^2} & -\rho/\sigma_{X1}\sigma_{X2} \\ -\rho/\sigma_{X1}\sigma_{X2} & 1/\sigma_{X2}{}^2 \end{bmatrix}$$

$$\left|[C_X{}^{-1}\right| = 1/\sigma_{X1}{}^2\sigma_{X2}{}^2(1-\rho^2)$$

Example :Suppose *X* and *Y* are two jointly-Gaussian 0-mean random variables with variances of 1 and 4 respectively and a covariance of 1. Find the joint PDF

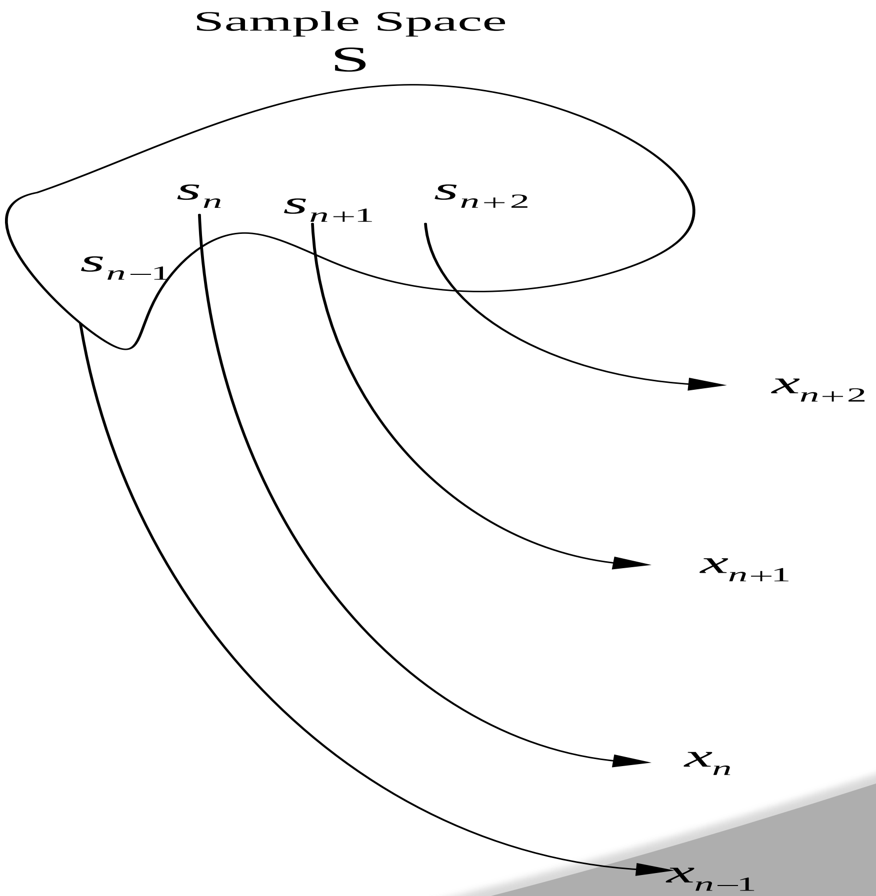
Module-IV

RANDOM PROCESSES – TEMPORAL CHARACTERISTICS

Random Process

- The concept of random variable was defined previously as mapping from the **Sample Space S** to the real line as shown below

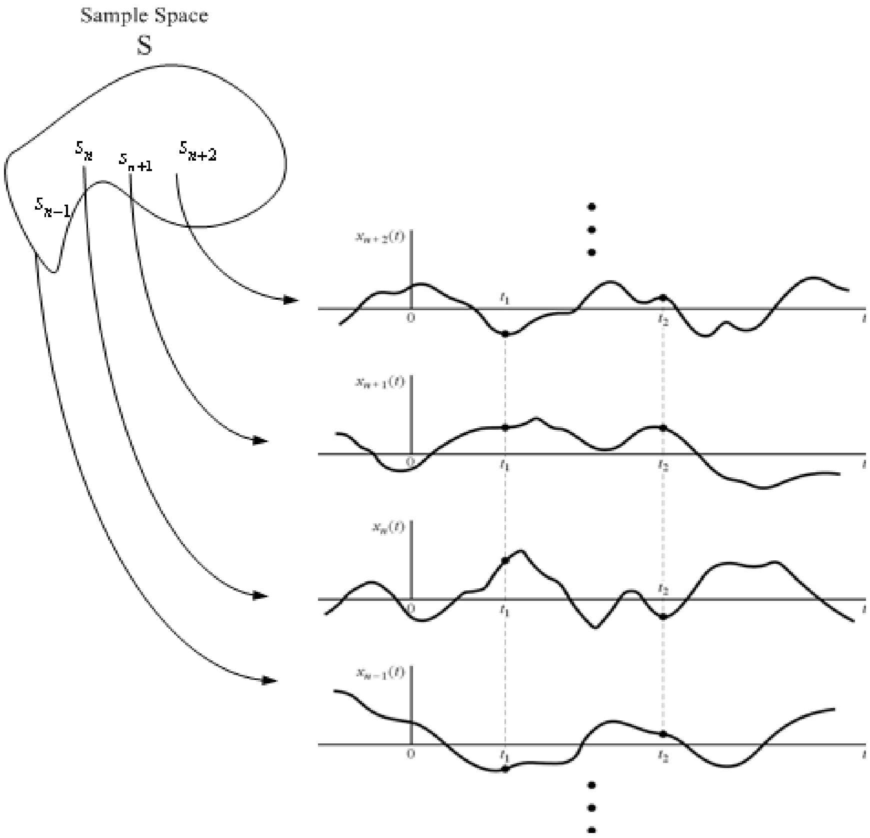
- A random process is a process (i.e., variation in time or one dimensional space) whose behavior is not completely predictable and can be characterized by statistical laws.



- Examples of random processes
  - Daily stream flow
  - Hourly rainfall of storm events
  - Stock index

Random Process (Contd..)

- The concept of random process can be extended to include time and the outcome will be random functions of time as shown beside $x(t,s)$
- Where s is the outcome of an experiment



- The functions

  $\cdots x_{n+2}(t), x_{n+1}(t), x_n(t), x_{n-1}(t), \cdots$

  are one realizations of many of the random process X(t)

- A random process also represents a random variable when time is fixed

  $X(t_1)$ is a random variable

Classification of Random Process

- Classification of random process
  - Continuous random process
  - Discrete random process
  - Continuous random sequence
  - Discrete random sequence

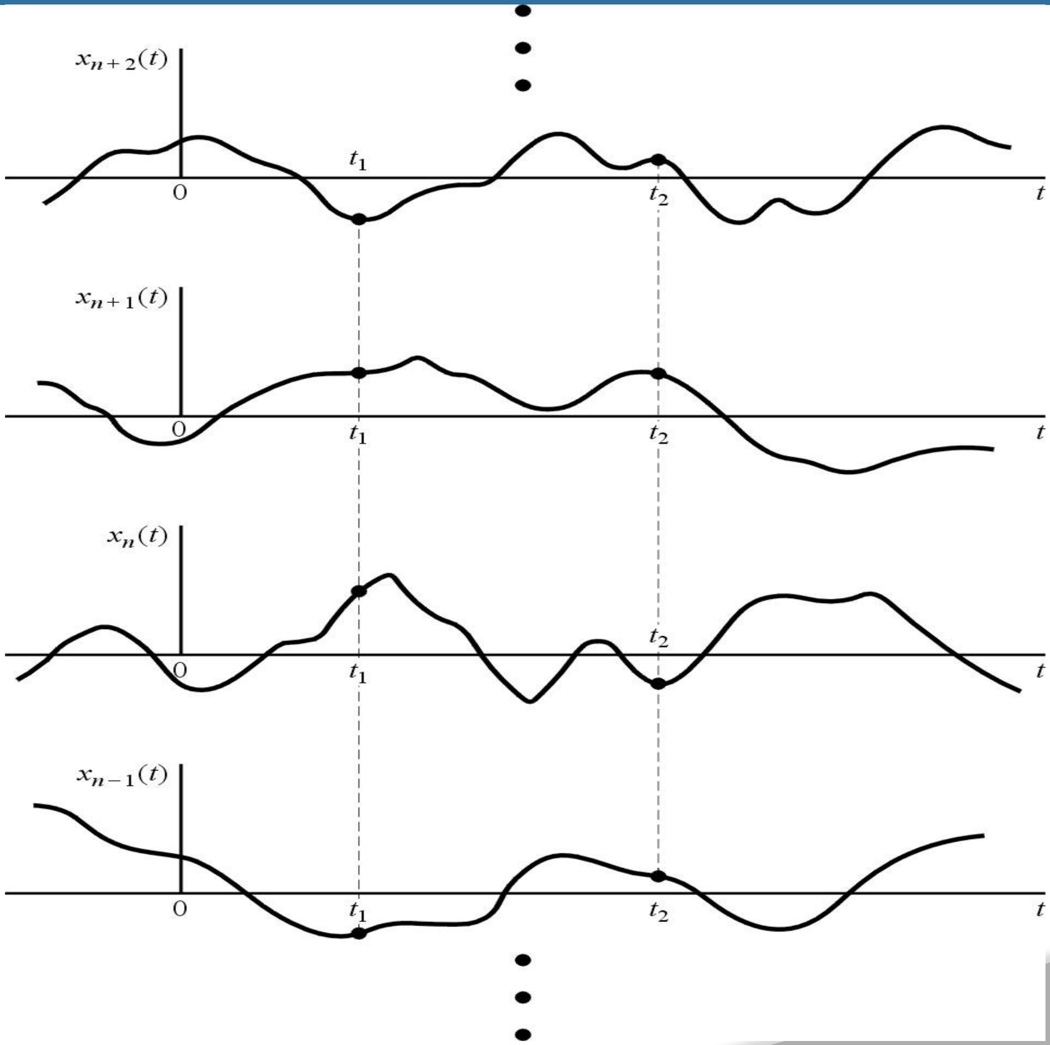
Continuous time t => x(t) = Random process
Discrete time n => x[n] = Random sequence

Continuous Random Process

- ❑ Continuous random process

  Continuous time t

  x(t) = Continuous Random process

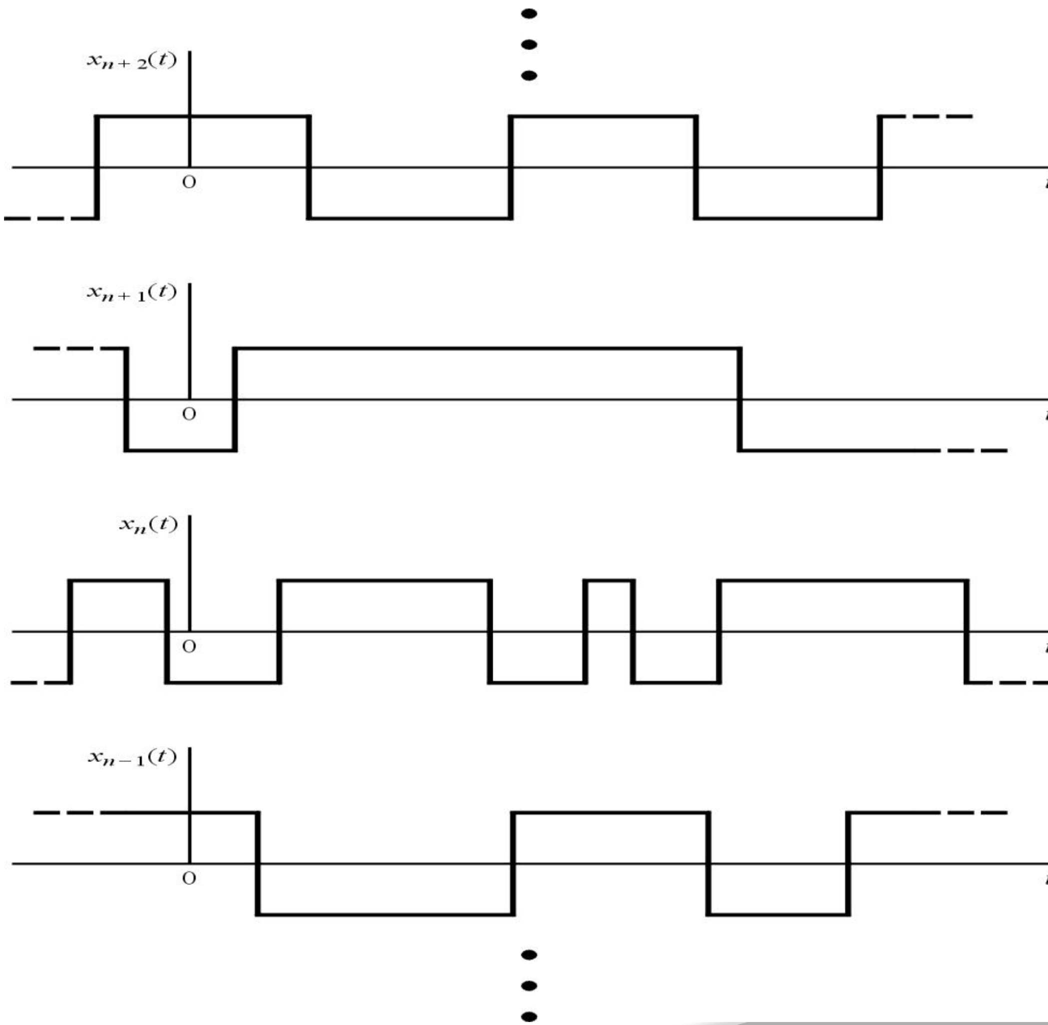


Discrete Random Process

- Discrete random process

Continuous time t

x(t) = Discrete Random process

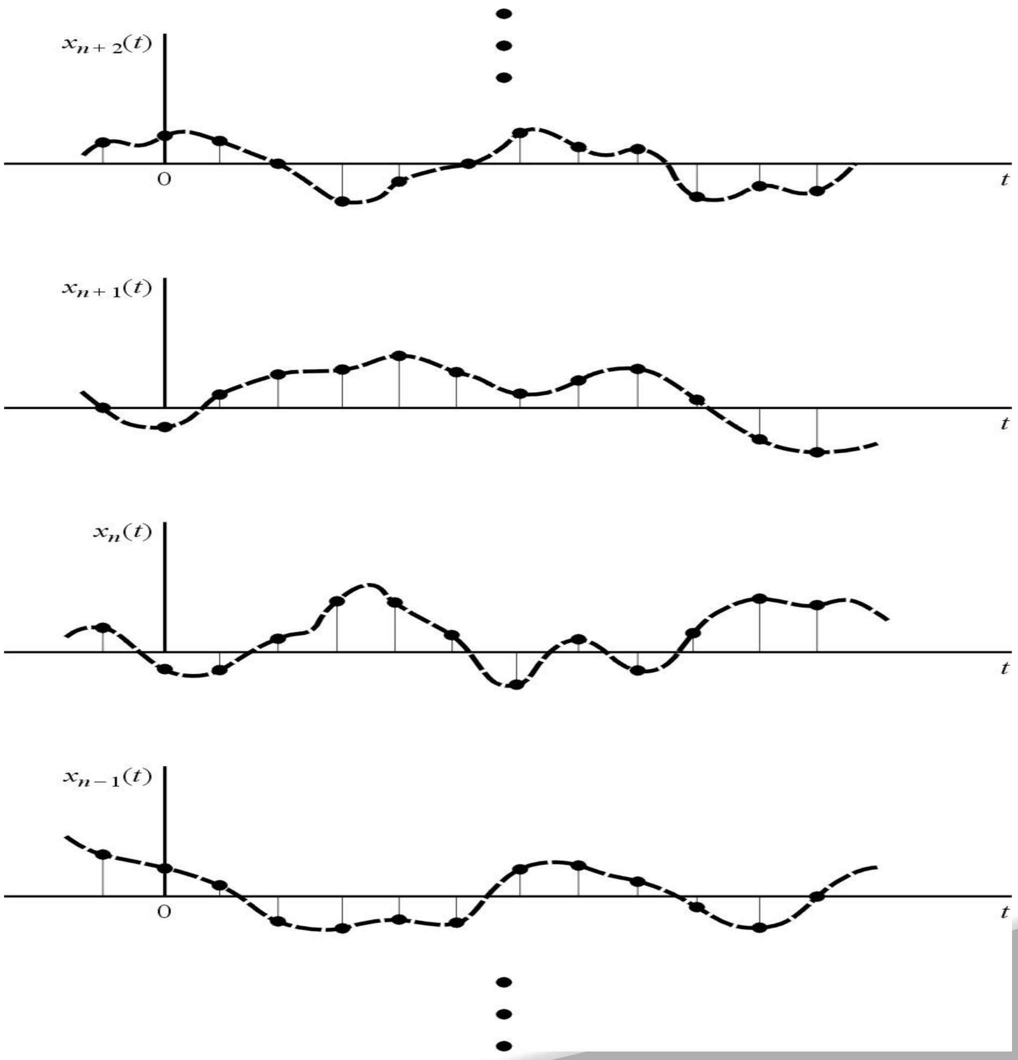


Continuous Random Sequence

- Continuous random sequence

  discrete time n

  x(n) = Continuous Random sequence

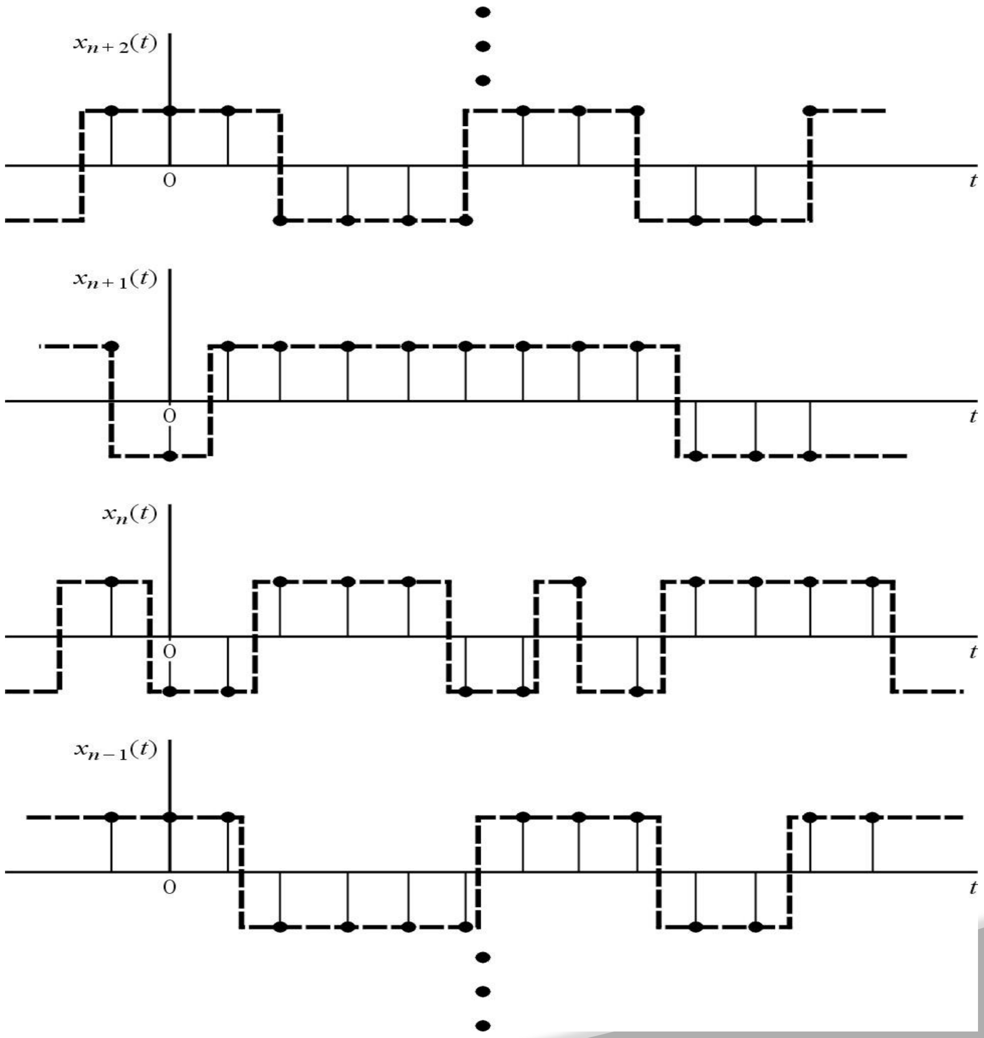


Discrete Random Sequence

- Discrete random sequence

  discrete time n

  x(n) = discrete Random sequence



Random Process Concept

- Deterministic random process
  - Future values of any sample function can be predicted exactly from the past values

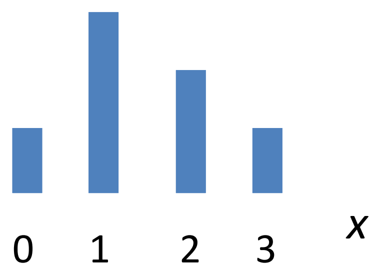
$$X(t) = A\cos(\omega_0 t + \theta), \qquad A, \omega_0, \theta: \quad \text{r.v.'s}$$

- Non deterministic random process
  - Future values of any sample function can not be predicted exactly from the past values

What is a distribution and density?

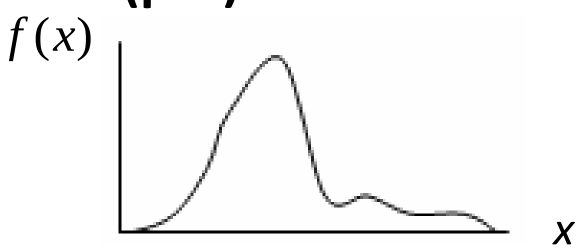
- A distribution characterises the probability (mass) associated with each possible outcome of a stochastic process
- Distributions of discrete data characterised by probability mass functions

$$P(X_i = x)$$



$$\sum_i P(X = x_i) = 1$$

- Distributions of continuous data are characterised by probability density functions (pdf)

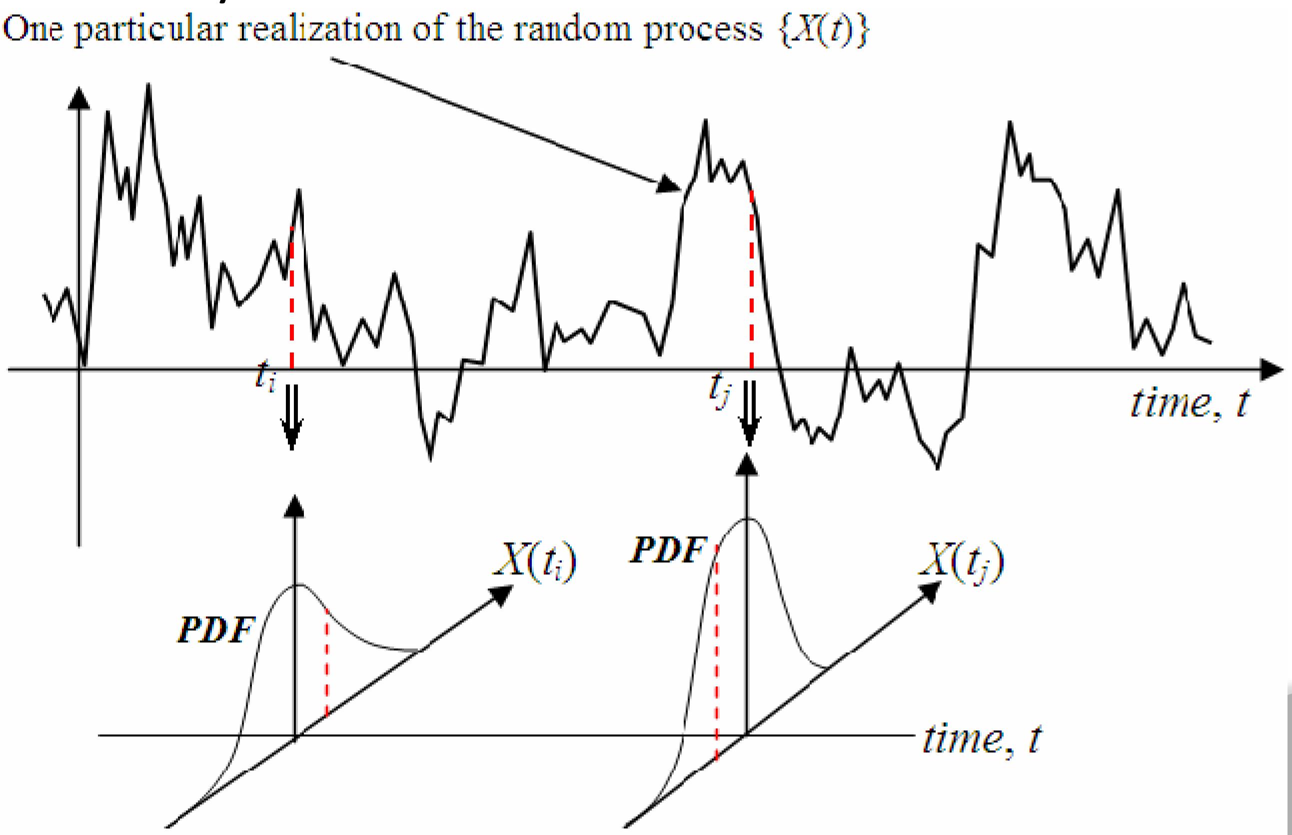


$$\int_{-\infty}^{\infty} f(x)dx = 1$$

- For RVs that map to the integers or the real numbers, the cumulative density function (cdf) is a useful alternative representation

Stationary and Independence

- Stationary Random Process
  - all its statistical properties do not change with time
- Non Stationary Random Process
  - not stationary



Stationary and Independence (Contd..)

- **First-order densities of a random process**
  - **A stochastic process is defined to be completely or totally characterized if the joint densities for the random variables $X(t_1), X(t_2), \cdots X(t_n)$ are known for all times $t_1, t_2, \cdots, t_n$ and all n.**

- For a specific t, X(t) is a random variable with distribution

$$F(x,t) = p[X(t) \leq x]$$

- The function F(x,t) is defined as the first-order distribution of the random variable X(t). Its derivative with respect to x

$$f(x,t) = \frac{\partial F(x,t)}{\partial x}$$

is the first-order density of X(t).

Stationary and Independence (Contd..)

- **If the first-order densities defined for all time t, i.e. f(x,t), are all the same, then f(x,t) does not depend on t and we call the resulting density the first-order density of the random process {x(t)} ; otherwise, we have a family of first-order densities.**

- **The first-order densities (or distributions) are only a partial characterization of the random process as they do not contain information that specifies the joint densities of the random variables defined at two or more different times.**

Stationary and Independence (Contd..)

- For $t = t_1$ and $t = t_2$, $X(t)$ represents two different random variables $X_1 = X(t_1)$ and $X_2 = X(t_2)$ respectively. Their joint distribution is given by

$$F_x(x_1, x_2, t_1, t_2) = P\{X(t_1) \le x_1, X(t_2) \le x_2\}$$

and

$$f_x(x_1, x_2, t_1, t_2) = \frac{\partial^2 F_x(x_1, x_2, t_1, t_2)}{\partial x_1 \; \partial x_2}$$

represents the second-order density function of the process $X(t)$.

- Similarly $f_x(x_1, x_2, \cdots x_n, \; t_1, t_2 \cdots, t_n)$ represents the n^{th} order density function of the process $X(t)$.

Mean and variance of a random process

- **The first-order density of a random process, *f*(*x*,*t*), gives the probability density of the random variables *X*(*t*) defined for all time *t*. The mean of a random process, $m_X(t)$, is thus a function of time specified by**

$$m_X(t) = E[X(t)] = E[X_t] = \int_{-\infty}^{+\infty} x_t f(x_t, t) dx_t$$

- For the case where the mean of *X*(*t*) does not depend on *t*, we have

$$m_X(t) = E[X(t)] = m_X \quad \text{(a constant)}$$

- The variance of a random process, also a function of time, is defined by

$$\sigma_X^2(t) = E\left\{[X(t) - m_X(t)]^2\right\} = E[X_t^2] - [m_X(t)]^2$$

Stationary and Independence

- The random process X(t) can be classified as follows:
- **First-order stationary**
- A random process is classified as **first-order stationary** if its first-order probability density function remains equal regardless of any shift in time to its time origin.
- If we X_{t1}let represent a given value at time t1then we define a first-order stationary as one that satisfies the following equation:

$$f_X(x_{t1}) = f_X(x_{t1} + \tau)$$

- The physical significance of this equation is that our density function,

$f_X(x_{t1})$ is completely independent of t1 and thus any time shift t

- For first-order stationary the mean is a constant, independent of any time shift

Stationary and Independence (Contd..)

❑Second-order stationary

❑A random process is classified as **second-order stationary** if its second-order probability density function does not vary over any time shift applied to both values.

❑ In other words, for values X_{t1} and X_{t2} then we will have the following be equal for an arbitrary time shift t

$$f_X(x_{t1}, x_{t2}) = f_X(x_{t1+\tau}, x_{t2+\tau})$$

❑From this equation we see that the absolute time does not affect our functions, rather it only really depends on the time difference between the two variables.

Stationary and Independence (Contd..)

- For a second-order stationary process, we need to look at the **autocorrelation function** (will be presented later) to see its most important property.

- Since we have already stated that a second-order stationary process depends only on the time difference, then all of these types of processes have the following property:

$$R_{XX}(t,t+\tau) = E[X(t)X(t+\tau)]$$
$$= R_{XX}(\tau)$$

Wide-Sense Stationary (WSS)

- A process that satisfies the following:
- The mean is a constant and the autocorrelation function depends only on the difference between the time indices

$$E[X(t)] = \bar{X} = \text{constant}$$

$$E[X(t)X(t+\tau)] = R_{XX}(\tau)$$

is a Wide-Sense Stationary (WSS)

Second-order stationary → Wide-Sense Stationary

The converse is not true in general

Wide-Sense Stationary (Example)

$$X(t) = a\cos(\omega_0 t + \varphi), \qquad \varphi \sim U(0,2\pi).$$

This gives

$$\mu_X(t) = E\{X(t)\} = aE\{\cos(\omega_0 t + \varphi)\}$$
$$= a\cos\omega_0 t\, E\{\cos\varphi\} - a\sin\omega_0 t\, E\{\sin\varphi\} = 0,$$ Constant

since $E\{\cos\varphi\} = \frac{1}{2\pi}\int_0^{2\pi} \cos\varphi\, d\varphi = 0 = E\{\sin\varphi\}.$

Similarly

$$R_{XX}(t_1,t_2) = a^2 E\{\cos(\omega_0 t_1 + \varphi)\cos(\omega_0 t_2 + \varphi)\}$$
$$= \frac{a^2}{2} E\{\cos\omega_0(t_1 - t_2) + \cos(\omega_0(t_1 + t_2) + 2\varphi)\}$$
$$= \frac{a^2}{2}\cos\omega_0(t_1 - t_2).$$

So given X(t) is WSS

Nth order and Strict-Sense Stationary

- In strict terms, the statistical properties are governed by the joint probability density function. Hence a process is n^{th}-order **Strict-Sense Stationary (S.S.S)** if

$$f_X(x_1, x_2, \cdots x_n,\ t_1, t_2 \cdots, t_n) \equiv f_X(x_1, x_2, \cdots x_n,\ t_1 + c, t_2 + c \cdots, t_n + c) \rightarrow (1)$$

- For *any c,* where the left side represents the joint density function of the random variables $X_1 = X(t_1),\ X_2 = X(t_2),\ \cdots,\ X_n = X(t_n)$ and the right side corresponds to the joint density function of the random variables $X_1' = X(t_1 + c),\ X_2' = X(t_2 + c),\ \cdots,\ X_n' = X(t_n + c).$

- A process *X*(*t*) is said to be **strict-sense stationary** if equation (1) true for all $t_i,\ \ i = 1, 2, \cdots, n,\ \ n = 1, 2, \cdots$ and *any* $c.$

Ergodic Process

A stationary random process for which time averages equal ensemble averages is called an ergodic process:

$$\langle x[n] \rangle = m_x$$

$$\langle x[n+m]x[n]^* \rangle = \phi_{xx}[m]$$

Ergodic Process (Contd..)

It is common to assume that a given sequence is a sample sequence of an ergodic random process, so that averages can be computed from a single sequence.

In practice, we cannot compute with the limits, but instead the quantities.

Similar quantities are often computed as estimates of the mean, variance, and autocorrelation.

$$\hat{m}_x = \frac{1}{L}\sum_{n=0}^{L-1} x[n]$$

$$\sigma_x^2 = \frac{1}{L}\sum_{n=0}^{L-1}\left(x[n]-\hat{m}_x\right)^2$$

$$\left\langle x[n+m]x^*[n]\right\rangle_L = \frac{1}{L}\sum_{n=0}^{L-1} x[n+m]x^*[n]$$

Time Average and Ergodicity

- The time average of a quantity is defined as

$$A[\bullet] = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} [\bullet] dt$$

Here A is used to denote time average in a manner analogous to E for the statistical average.

- The time average is taken over all time because, as applied to random processes, sample functions of processes are presumed to exist for all time.

Time Average and Ergodicity (Contd..)

- Let x(t) be a sample of the random process X(t) were the lower case letter imply a sample function.

- We define the mean value $\overline{x} = A\left[x(t)\right]$

  (a lowercase letter is used to imply a sample function) and the time autocorrelation function $\Re_{XX}(\tau)$ as follows:

$$\overline{x} = A\left[x(t)\right] = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, dt$$

$$\Re_{XX}(\tau) = A\left[x(t)x(t+\tau)\right] = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)x(t+\tau)\, dt$$

- For any one sample function (i.e., x(t)) of the random process X(t), the last two integrals simply produce two numbers.

- A number for the average $\overline{x}$ and a number for $\Re_{XX}(\tau)$ for a specific value of τ

Time Average and Ergodicity (Contd..)

- Since the sample function x(t) is one out of other samples functions of the random process X(t),
- The average $\overline{x}$ and the autocorrelation $\Re_{XX}(\tau)$ are random variables
- By taking the expected value for $\overline{X}$ and $\Re_{XX}(\tau)$,we obtain

$$E[\overline{x}] = E[A[x(t)]] = E\left[\lim_{T \to \infty} \frac{1}{2T}\int_{-T}^{T} x(t)\, dt\right] = \lim_{T \to \infty} \frac{1}{2T}\int_{-T}^{T} E[x(t)]\, dt$$

$$= \lim_{T \to \infty} \frac{1}{2T}\int_{-T}^{T} \overline{X}\, dt = \lim_{T \to \infty} \overline{X}(1) = \overline{X}$$

$$E[\Re_{XX}(\tau)] = E\,[A[x(t)x(t+\tau)]\,] = E\left[\lim_{T \to \infty} \frac{1}{2T}\int_{-T}^{T} x(t)x(t+\tau)\, dt\right]$$

$$= \lim_{T \to \infty} \frac{1}{2T}\int_{-T}^{T} E[x(t)x(t+\tau)]\, dt = \lim_{T \to \infty} \frac{1}{2T}\int_{-T}^{T} R_{XX}(\tau)\, dt = R_{XX}(\tau)$$

Time Average and Ergodicity (Contd..)

- Time cross correlation

$$\Re_{xy}(\tau) = A[x(t)y(t+\tau)] = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)y(t+\tau)dt$$

- Ergodic =>

$$\bar{x} = \bar{X}$$

$$\Re_{xx}(\tau) = R_{XX}(\tau)$$

- Jointly Ergodic => Ergodic X(t) and Y(t)

$$\Re_{xy}(\tau) = R_{XY}(\tau)$$

Autocorrelation

Introduction

- Autocorrelation occurs in time-series studies when the errors associated with a given time period carry over into future time periods.

- For example, if we are predicting the growth of stock dividends, an overestimate in one year is likely to lead to overestimates in succeeding years.

- Times series data follow a natural ordering over time.

- It is likely that such data exhibit intercorrelation, especially if the time interval between successive observations is short, such as weeks or days.

Introduction (contd..)

- We expect stock market prices to move or move down for several days in succession.
- We experience autocorrelation when

$$E(u_i u_j) \neq 0$$

- Tintner defines autocorrelation as ‘lag correlation of a given series within itself, lagged by a number of times units’ whereas serial correlation is the ‘lag correlation between two different series’.

Autocorrelation and its Properties

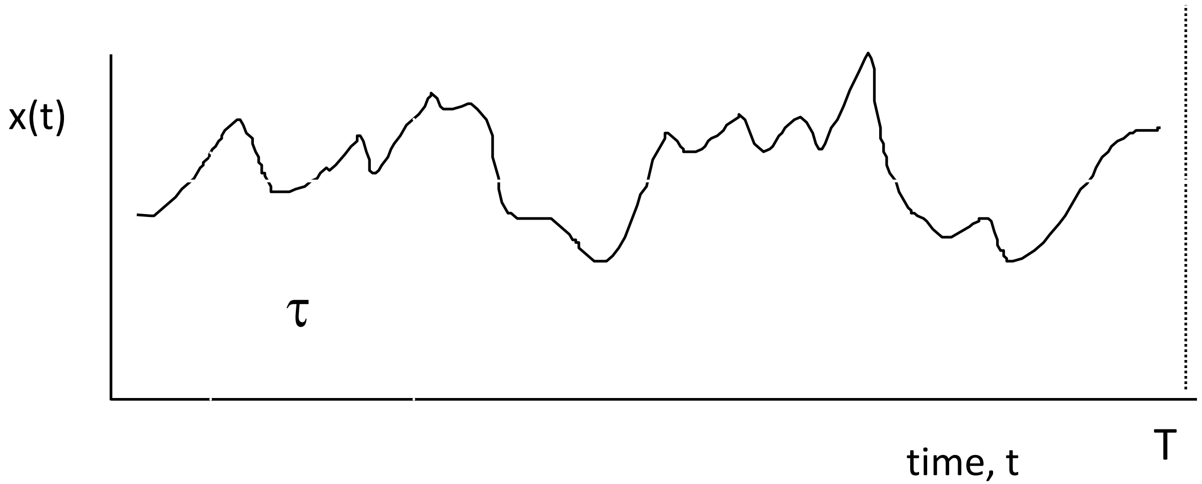
- The autocorrelation function of a random process X(t) is the correlation $E\left[X_1X_2\right]$ of two random variables $X_1 = X(t_1)$ and $X_2 = X(t_2)$ by the process at times t1 and t2

$$R_{XX}(t_1,t_2) = E\left[X(t_1)X(t_2)\right]$$

- Assuming a second-order stationary process

$$R_{XX}(t, t+\tau) = E\left[X(t)X(t+\tau)\right] \qquad R_{XX}(\tau) = E\left[X(t)X(t+\tau)\right]$$

- **Autocorrelation :**

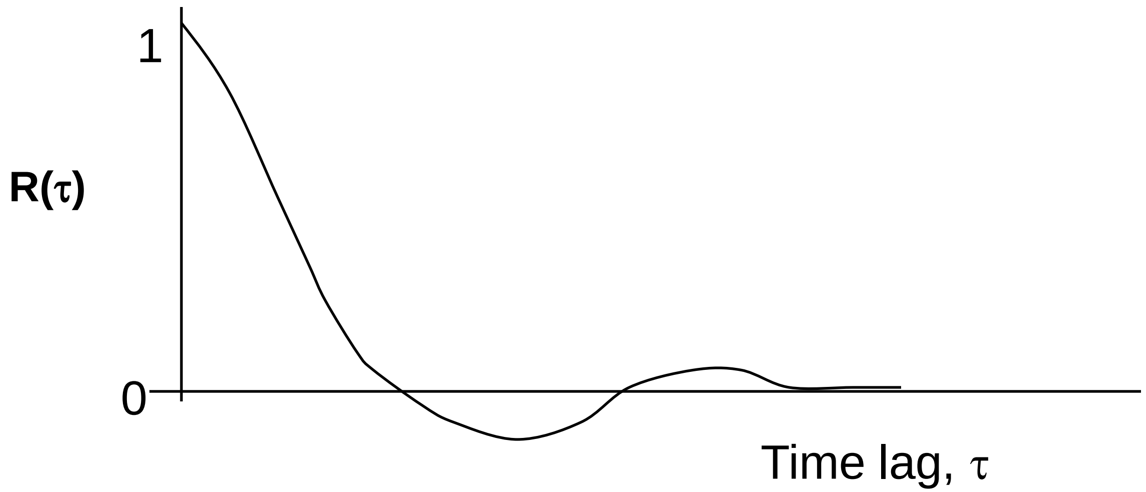


- The autocorrelation, or auto covariance, describes the general dependency of x(t) with its value at a short time later, x(t+τ)

$$\rho_x(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \left[x(t) - \bar{x}\right]\left[x(t+\tau) - \bar{x}\right]dt$$

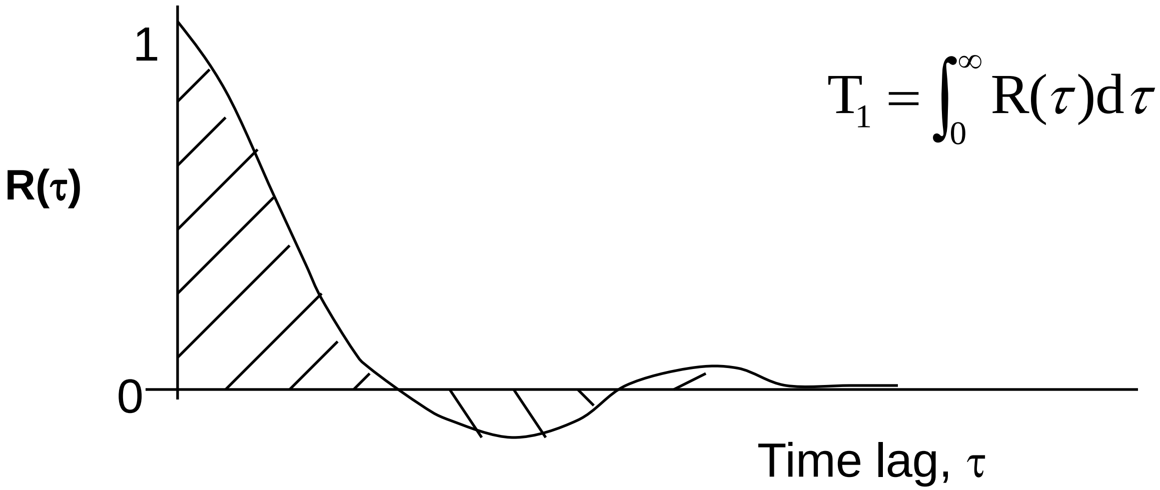
- The value of $\rho_x(\tau)$ at τ equal to 0 is the variance, σ_x^2
- Normalized auto-correlation : $R(\tau)=\rho_x(\tau)/\sigma_x^2$ $R(0)= 1$

Autocorrelation and its Properties (Contd..)



- The autocorrelation for a random process eventually decays to zero at large τ
- The autocorrelation for a sinusoidal process (deterministic) is a cosine function which does not decay to zero

Autocorrelation and its Properties (Contd..)



- The area under the normalized autocorrelation function for the fluctuating wind velocity measured at a point is a measure of the average time scale of the eddies being carried passed the measurement point, say T_1
- If we assume that the eddies are being swept passed at the mean velocity, $\bar{U}.T_1$ is a measure of the average length scale of the eddies. This is known as the 'integral length scale', denoted by l_u

Autocorrelation and its Properties (Contd..)

- **Properties of Autocorrelation function**

$$R_{XX}(t,t+\tau)=E[X(t)X(t+\tau)]=R_{XX}(\tau)$$

(1) $\left|R_{XX}(\tau)\right|\le R_{XX}(0)$

(2) $R_{XX}(-\tau)=R_{XX}(\tau)$

(3) $R_{XX}(0)=E[X(t)^2]$

(4) stationary & ergodic $X(t)$ with no periodic components

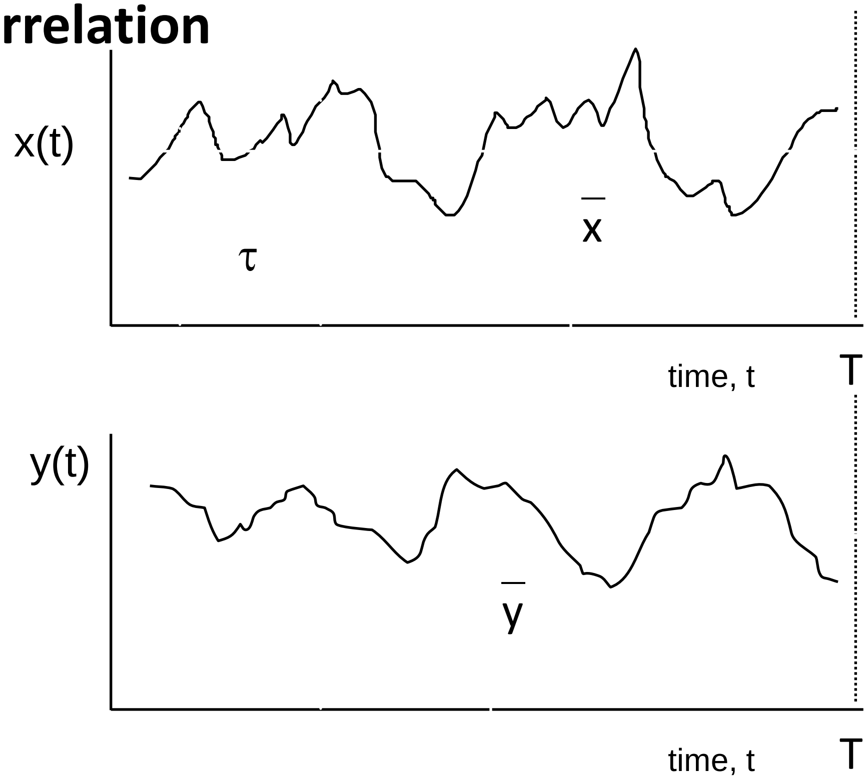
$$\Rightarrow \lim_{|\tau|\to\infty} R_{XX}(\tau)=\overline{X}^2$$

(5) stationary $X(t)$ has a periodic component

$\Rightarrow$ $R_{XX}(\tau)$ has a periodic component with the same period.

Cross-correlation

- **Cross-correlation**



- The cross-correlation function describes the general dependency of x(t) with another random process y(t+τ), delayed by a time delay, τ

$$c_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \left[x(t) - \bar{x}\right]\left[y(t+\tau) - \bar{y}\right] dt$$

Correlation coefficient

- Correlation coefficient
  - The correlation coefficient, ρ, is the covariance normalized by the standard deviations of x and y

$$\rho = \frac{\overline{x'(t).y'(t)}}{\sigma_x.\sigma_y}$$

When x and y are identical to each other, the value of ρ is +1 (full correlation)

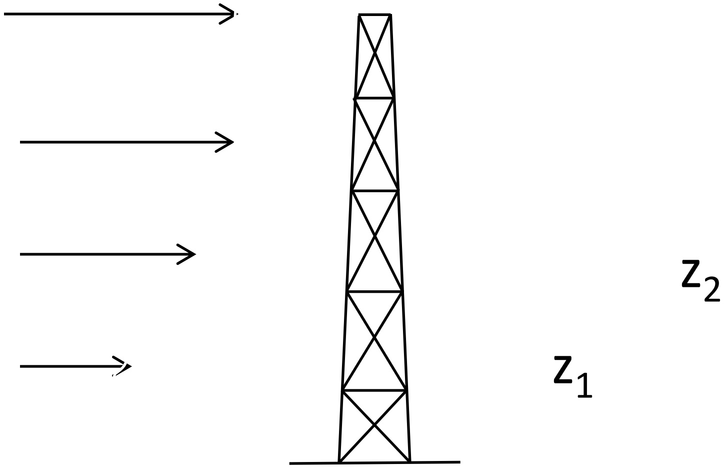
When y(t)=–x(t), the value of ρ is – 1

In general, $-1 < \rho < +1$

Application of correlation

- **Correlation - application :**

  - The fluctuating wind loading of a tower depends on the correlation coefficient between wind velocities and hence wind loads, at various heights



For heights, z_1, and z_2 :

$$\rho(z_1, z_2) = \frac{\overline{u'(z_1).u'(z_2)}}{\sigma_u(z_1).\sigma_u(z_2)}$$

Properties of Cross Correlation

Properties of cross-correlation function of jointly w.s.s. r.p.'s:

$$R_{XY}(\tau) = E[X(t)Y(t+\tau)]$$

(1) $R_{XY}(-\tau) = R_{YX}(\tau)$

(2) $|R_{XY}(\tau)| \le \sqrt{R_{XX}(0)R_{YY}(0)}$

(3) $|R_{XY}(\tau)| \le \frac{1}{2}[R_{XX}(0) + R_{YY}(0)]$

$$E[\{Y(t+\tau) + \alpha X(t)\}^2] \ge 0, \quad \forall \alpha$$

$$\sqrt{R_{XX}(0)R_{YY}(0)} \le \frac{1}{2}[R_{XX}(0) + R_{YY}(0)]$$

Example of Cross Correlation

$$A, B : \text{r.v.'s} \qquad \omega_0 = \text{const}$$

$$E[A] = E[B] = 0, \quad E[AB] = 0, \quad E[A^2] = E[B^2] = \sigma^2$$

$$X(t) = A\cos(\omega_0 t) + B\sin(\omega_0 t), \quad Y(t) = B\cos(\omega_0 t) - A\sin(\omega_0 t)$$

$$E[X(t)] = E[A\cos(\omega_0 t) + B\sin(\omega_0 t)] = E[A]\cos(\omega_0 t) + E[B]\sin(\omega_0 t) = 0$$

$$R_{XX}(t, t+\tau) = E[X(t)X(t+\tau)]$$

$$= E[A^2\cos(\omega_0 t)\cos(\omega_0 t + \omega_0 \tau) + AB\cos(\omega_0 t)\sin(\omega_0 t + \omega_0 \tau)$$

$$+ AB\sin(\omega_0 t)\cos(\omega_0 t + \omega_0 \tau) + B^2\sin(\omega_0 t)\sin(\omega_0 t + \omega_0 \tau)]$$

$$= \sigma^2\{\cos(\omega_0 t)\cos(\omega_0 t + \omega_0 \tau) + \sin(\omega_0 t)\sin(\omega_0 t + \omega_0 \tau)\} = \sigma^2\cos(\omega_0 \tau)$$

$$\Rightarrow \quad X(t) : \text{w.s.s.}$$

Example of Cross Correlation

$Y(t)$: w.s.s.

$$
\begin{aligned}
R_{XY}(\tau) &= E[X(t)Y(t+\tau)] \\
&= E\{[A\cos(\omega_0 t) + B\sin(\omega_0 t)][B\cos(\omega_0(t+\tau)) - A\sin(\omega_0(t+\tau))]\} \\
&= E[AB\cos(\omega_0 t)\cos(\omega_0 t + \omega_0 \tau) + B^2 \sin(\omega_0 t)\cos(\omega_0 t + \omega_0 \tau) \\
&\qquad - A^2 \cos(\omega_0 t)\sin(\omega_0 t + \omega_0 \tau) - AB\sin(\omega_0 t)\sin(\omega_0 t + \omega_0 \tau)] \\
&= \sigma^2[\sin(\omega_0 t)\cos(\omega_0 t + \omega_0 \tau) - \cos(\omega_0 t)\sin(\omega_0 t + \omega_0 \tau)] \\
&= -\sigma^2 \sin(\omega_0 \tau)
\end{aligned}
$$

$\Rightarrow$ $X(t)$ & $Y(t)$: jointly w.s.s.

Covariance

- **Covariance**

  - The covariance is the cross correlation function with the time delay, τ, set to zero

$$c_{xy}(0) = \overline{x'(t).y'(t)} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \left[x(t) - \overline{x}\right]\left[y(t) - \overline{y}\right]dt$$

  - Note that here x'(t) and y'(t) are used to denote the fluctuating parts of x(t) and y(t) (mean parts subtracted)

Auto Covariance

- The auto covariance Cx(t1,t2) of a random process X(t) is defined as the covariance of X(t1) and X(t2)

  Cx(t1,t2)=E[{X(t1)-mx(t1)}{X(t2)-mx(t2)}]

  Cx(t1,t2) = Rx(t1,t2)-mx(t1)mx(t2)

- The variance of X(t) can be obtained from Cx(t1,t2)

  VAR[X(t)] = E[(X(t)-mx(t))2] = Cx(t,t)

- The correlation coefficient of X(t) is given by

$$\rho_x(t_1,t_2) = \frac{C_X(t_1,t_2)}{\sqrt{C_X(t_1,t_1)}\sqrt{C_X(t_2,t_2)}}$$

$$|\rho_x(t_1,t_2)| \leq 1$$

Auto Covariance Example#1

Example:

Let X(t) = Acos2πt, where A is some random variable

The mean of X(t) is given by

$$m_X(t) = E[A\cos 2\pi t] = E[A]\cos 2\pi t$$

The autocorrelation is

$$R_X(t_1,t_2) = E[A\cos(2\pi t_1)A\cos(2\pi t_2)]$$

$$R_X(t_1,t_2) = E[A^2]\cos(2\pi t_1)\cos(2\pi t_2)$$

And the autocovariance

$$C_X(t_1,t_2) = R_X(t_1,t_2) - m_X(t_1)m_X(t_2)$$

$$C_X(t_1,t_2) = \left\{E[A^2] - E[A]^2\right\}\cos(2\pi t_1)\cos(2\pi t_2)$$

$$C_X(t_1,t_2) = VAR[A]\cos(2\pi t_1)\cos(2\pi t_2)$$

Auto Covariance Example#2

Example:

Let X(t) = cos(ωt+θ), where θ is uniformly distributed in the interval (-π, π). The mean of X(t) is given by

$$m_X(t) = E[\cos(\omega t + \theta)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(\omega t + \theta) = 0$$

The autocorrelation and autocovariance are then

$$C_X(t_1, t_2) = R_X(t_1, t_2) = E[\cos(\omega t_1 + \theta)\cos(\omega t_2 + \theta)]$$

$$C_X(t_1, t_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}\{\cos(\omega(t_1 - t_2)) + \cos(\omega(t_1 + t_2) + 2\theta)\}d\theta$$

$$C_X(t_1, t_2) = \frac{1}{2}\cos(\omega(t_1 - t_2))$$

Cross Covariance

- The cross covariance Cx,y(t1,t2) of a random process X(t) and Y(t) is defined as

$$C_{x,y}(t1,t2)=E[\{X(t1)-m_x(t1)\}\{Y(t2)-m_y(t2)\}]$$

$$C_x(t1,t2) = R_{x,y}(t1,t2)-m_x(t1)m_y(t2)$$

- The process X(t) and Y(t) are said to be uncorrelated if

$$C_{x,y}(t1,t2) = 0 \text{ for all } t1, t2$$

Random sequence

Random Sequence (=Discrete-time R.P)

$$X(nT_s) = X[n]$$

Mean $= E(X[n])$

$$R_{XX}(n, n+k) = E(X[n]X[n+k])$$

$$C_{XX}(n, n+k) = E\{(X[n] - \bar{X}[n])(X[n+k] - \bar{X}[n+k])\}$$
$$= R_{XX}(n, n+k) - \bar{X}[n]\bar{X}[n+k]$$

$$R_{XY}(n, n+k) = E(X[n]Y[n+k])$$

$$C_{XY}(n, n+k) = E\{(X[n] - \bar{X}[n])(Y[n+k] - \bar{Y}[n+k])\}$$
$$= R_{XY}(n, n+k) - \bar{X}[n]\bar{Y}[n+k]$$

Gaussian Random Process

- Let X(t) be a random process and let X(t1), X(t2), ….X(tn) be the random variables obtained from X(t) at t=t1,t2……..tn sec respectively

- Let all these random variables be expressed in the form of a matrix

$$X = \begin{bmatrix} X(t_1) \\ X(t_2) \\ \vdots \\ X(t_n) \end{bmatrix}$$

- Then, X(t) is referred to as normal or Gaussian process if all the elements of X are jointly Gaussian

Gaussian Random Process

- continuous r.p. $X(t), \quad -\infty < t < \infty$

$$f_X(x_1,\cdots,x_N;t_1,\cdots,t_N) = \frac{1}{\sqrt{(2\pi)^N |C_X|}} \exp\{-\frac{1}{2}[x-\bar{X}]^t C_X^{-1}[x-\bar{X}]\}$$

$$\bar{X}_i = E[X(t_i)] \qquad C_{ik} = C_{XX}(t_i,t_k)$$

stationary $\Rightarrow$ $E[X(t)] = \bar{X}$ (const) & $R_{XX}(t_i,t_k) = R_{XX}(t_k - t_i)$

$$C_{XX}(t_i,t_k) = C_{XX}(t_k - t_i)$$

w.s.s. Gaussian $\Rightarrow$ strictly stationary w.s.s. gaussian r.p. $X(t)$

$$\bar{X} = 4 \qquad R_{XX}(\tau) = 25e^{-3|\tau|} \qquad t_i = t_0 + \frac{i-1}{2}, \quad i = 1,2,3.$$

$$C_{ik} = C_{XX}(t_i, t_k) = R_{XX}(t_i, t_k) - \bar{X}^2 = 25e^{-3\frac{|k-i|}{2}} - 16$$

$$C_X = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} = \begin{bmatrix} 25-16 & 25e^{-\frac{3}{2}}-16 & 25e^{-3}-16 \\ 25e^{-\frac{3}{2}}-16 & 25-16 & 25e^{-\frac{3}{2}}-16 \\ 25e^{-3}-16 & 25e^{-\frac{3}{2}}-16 & 25-16 \end{bmatrix}$$

Properties of Gaussian Process

- If a gaussian process $X(t)$ is applied to a stable linear filter, then the random process $Y(t)$ developed at the output of the filter is also gaussian.
- Considering the set of random variables or samples $X(t1)$, $X(t2)$,.....$X(tn)$ obtained by observation of a random process $X(t)$ at instants t1,t2,.......tn, if the process $X(t)$ is gaussian, then this set of random variables are jointly gaussian for any n, with their n-fold joint p.d.f. being completely determined by the set of means.

$$mx(ti) = E[X(ti)] \text{ for } i=1,2,....n$$

and the set of auto covariance function

$$Cxx(t1,t2) = E[\{X(t1)-E[X(t1)]\}\{X(t2)-E[X(t2)]\}]$$

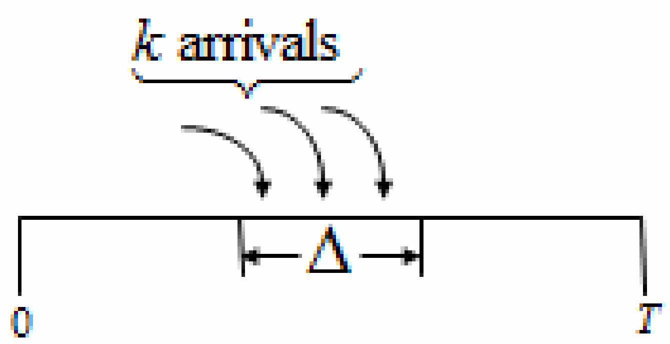
- If a gaussian process is wide sense stationary, then the process is also stationary in the strict sense
- If the set of random variables $X(t1)$,$X(t2)$...$X(tn)$ are uncorrelated then they are statistically independent

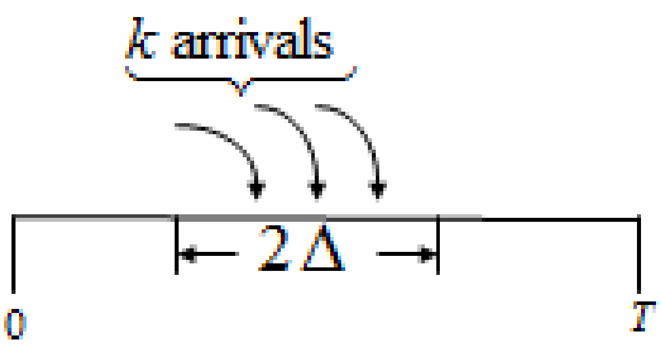
Poisson Random Process

- we introduced Poisson arrivals as the limiting behavior of Binomial random variables

where
$$P\left\{\begin{array}{l}\text{"}k \text{ arrivals occur in an} \\ \text{interval of duration } \Delta\text{"}\end{array}\right\} = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \cdots$$

$$\lambda = np = \mu T \cdot \frac{\Delta}{T} = \mu\Delta$$





Poisson Random Process

- It follows that

$$P\left\{\begin{array}{l}\text{"}k \text{ arrivals occur in an} \\ \text{ interval of duration } 2\Delta\text{"}\end{array}\right\} = e^{-2\lambda}\frac{(2\lambda)^k}{k!}, \qquad k = 0,\ 1,\ 2, \cdots,$$

since in that case

$$np_1 = \mu T \cdot \frac{2\Delta}{T} = 2\mu\Delta = 2\lambda.$$

- From the above equations, Poisson arrivals over an interval form a Poisson random variable whose parameter depends on the duration of that interval.

- The Bernoulli nature of the underlying basic random arrivals, events over non overlapping intervals are independent. We shall use these two key observations to define a Poisson process formally.

Poisson Random Process

- **Definition:** $X(t) = n(0, t)$ represents a Poisson process if

(i) the number of arrivals $n(t_1, t_2)$ in an interval (t_1, t_2) of length $t = t_2 - t_1$ is a Poisson random variable with parameter λt.

Thus

$$P\{n(t_1, t_2) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \cdots, \quad t = t_2 - t_1$$

and

(ii) If the intervals (t_1, t_2) and (t_3, t_4) are non overlapping, then the random variables $n(t_1, t_2)$ and $n(t_3, t_4)$ are independent.

Since $n(0, t) \sim P(\lambda t)$,we have

$$E[X(t)] = E[n(0, t)] = \lambda t$$

and

$$E[X^2(t)] = E[n^2(0, t)] = \lambda t + \lambda^2 t^2.$$

Poisson Random Process

- To determine the autocorrelation function $R_{XX}(t_1, t_2)$, let $t_2 > t_1$, then from (ii) above $n(0, t_1)$ and $n(t_1, t_2)$ are independent Poisson random variables with parameters λt_1 and $\lambda(t_2 - t_1)$ respectively. Thus

$$E[n(0,t_1)n(t_1,t_2)] = E[n(0,t_1)]E[n(t_1,t_2)] = \lambda^2 t_1(t_2 - t_1).$$

But

$$n(t_1,t_2) = n(0,t_2) - n(0,t_1) = X(t_2) - X(t_1)$$

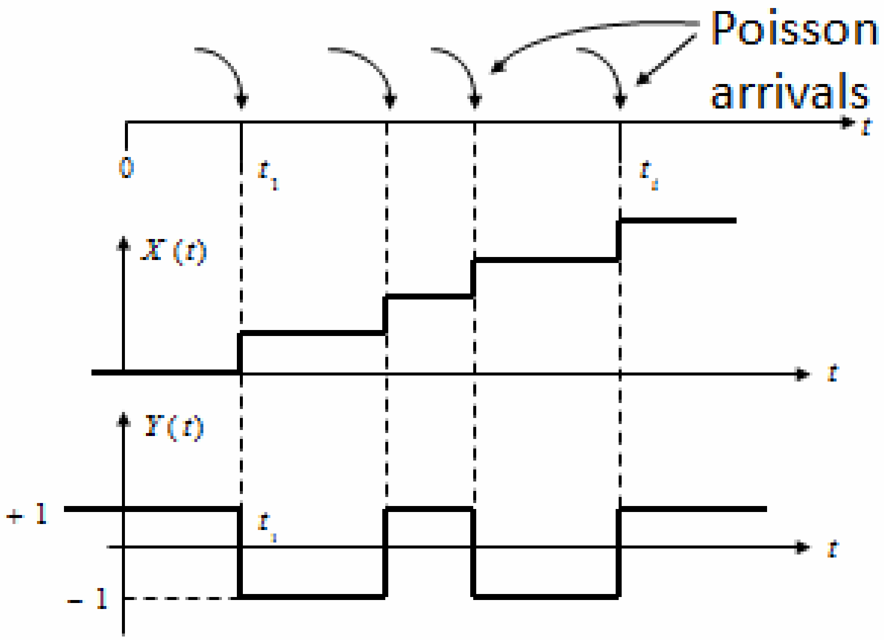
and hence the left side of above equation can be rewritten as

$$E[X(t_1)\{X(t_2) - X(t_1)\}] = R_{XX}(t_1,t_2) - E[X^2(t_1)].$$

$$R_{XX}(t_1,t_2) = \lambda^2 t_1(t_2 - t_1) + E[X^2(t_1)]$$
$$= \lambda t_1 + \lambda^2 t_1\, t_2, \qquad t_2 \ge t_1 .$$

Similarly $$R_{XX}(t_1,t_2) = \lambda t_2 + \lambda^2 t_1\, t_2, \qquad t_2 < t_1 .$$

Thus $$R_{XX}(t_1,t_2) = \lambda^2 t_1\, t_2 + \lambda \min(t_1, t_2).$$



- Notice that the Poisson process X(t) does not represent a wide sense stationary process.

- Define a binary level process

$$Y(t) = (-1)^{X(t)}$$

that represents a telegraph signal Notice that the transition instants $\{t_i\}$ are random Although $X(t)$ does not represent a wide sense stationary process, its derivative $X'(t)$ *does* represent a wide sense stationary process.

$$X(t) \qquad \frac{d(\cdot)}{dt} \qquad X'(t)$$

(Derivative as a LTI system)

From there

$$\mu_{X'}(t) = \frac{d\mu_X(t)}{dt} = \frac{d\lambda t}{dt} = \lambda, \quad a\ constant$$

and

$$R_{XX'}(t_1, t_2) = \frac{\partial R_{XX}(t_1, t_2)}{\partial t_2} = \begin{cases} \lambda^2 t_1 & t_1 \le t_2 \\ \lambda^2 t_1 + \lambda & t_1 > t_2 \end{cases}$$

$$= \lambda^2 t_1 + \lambda\, U(t_1 - t_2)$$

and

$$R_{X'X'}(t_1, t_2) = \frac{\partial R_{XX'}(t_1, t_2)}{\partial t_1} = \lambda^2 + \lambda\, \delta(t_1 - t_2).$$

Poisson Random Process

Define the processes

$$Y(t) = \sum_{i=1}^{X(t)} N_i \quad ; \quad Z(t) = \sum_{i=1}^{X(t)} (1 - N_i) = X(t) - Y(t)$$

we claim that both *Y*(*t*) and *Z*(*t*) are independent Poisson processes with parameters λpt and λqt respectively.

Proof:

$$Y(t) = \sum_{n=k}^{\infty} P\{Y(t) = k \mid X(t) = n\} P\{X(t) = n)\}.$$

But given *X*(*t*) = *n*, we have $Y(t) = \sum_{i=1}^{n} N_i \sim B(n, p)$ so that

$$P\{Y(t) = k \mid X(t) = n\} = \binom{n}{k} p^k q^{n-k}, \qquad 0 \le k \le n,$$

and

$$P\{X(t) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Poisson Random Process

$$P\{Y(t)=k\} = e^{-\lambda t}\sum_{n=k}^{\infty}\frac{n!}{(n-k)!k!}p^k q^{n-k}\frac{(\lambda t)^n}{n!} = \frac{p^k e^{-\lambda t}}{k!}(\lambda t)^k \underbrace{\sum_{n=k}^{\infty}\frac{(q\lambda t)^{n-k}}{(n-k)!}}_{e^{q\lambda t}}$$

$$= (\lambda pt)^k \frac{e^{-(1-q)\lambda t}}{k!} = e^{-\lambda pt}\frac{(\lambda pt)^k}{k!}, \qquad k = 0, 1, 2, \cdots$$

$$\sim \quad P(\lambda pt).$$

More generally,

$$P\{Y(t)=k, Z(t)=m\} = P\{Y(t)=k, X(t)-Y(t)=m\}$$

$$= P\{Y(t)=k, X(t)=k+m\}$$

$$= P\{Y(t)=k \mid X(t)=k+m\}P\{X(t)=k+m\}$$

$$= \binom{k+m}{k}p^k q^m \cdot e^{-\lambda t}\frac{(\lambda t)^{k+m}}{(k+m)!} = \underbrace{e^{-\lambda pt}\frac{(\lambda pt)^n}{k!}}_{P(Y(t)=k)} \underbrace{e^{-\lambda qt}\frac{(\lambda qt)^n}{m!}}_{P(Z(t)=m)}$$

$= P\{Y(t)=k\}P\{Z(t)=m\},$ which completes the proof.

Poisson Random Process

-- integer-valued discrete r.p. $X(t), \quad -\infty < t < \infty$

$$X(0) = 0 \qquad\qquad t_b < t_a \quad \Rightarrow \quad X(t_b) \le X(t_a)$$

$$P[X(t_a) - X(t_b) = k] = \frac{[\lambda(t_a - t_b)]^k}{k!} e^{-\lambda(t_a - t_b)}, \quad k = 0,1,2,\cdots$$

$t_d < t_c \le t_b < t_a \quad \Rightarrow \quad X(t_a) - X(t_b)$ & $X(t_c) - X(t_d)$ are indep.

$$\bar{X}(t) = E[X(t)] = \lambda t \qquad\qquad R_{XX}(t,t) = E[X(t)^2] = \lambda t + (\lambda t)^2$$

$$C_{XX}(t,t) = \lambda t$$

Poisson Random Process

$$0 < t_1 < t_2 \quad \Rightarrow$$

$$P[X(t_1) = k_1, X(t_2) = k_2] = P[X(t_1) = k_1, X(t_2) - X(t_1) = k_2 - k_1]$$

$$= \begin{cases} \dfrac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \dfrac{[\lambda(t_2 - t_1)]^{(k_2 - k_1)}}{(k_2 - k_1)!} e^{-\lambda(t_2 - t_1)}, & k_2 \geq k_1 \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{(\lambda t_1)^{k_1} [\lambda(t_2 - t_1)]^{(k_2 - k_1)}}{k_1!(k_2 - k_1)!} e^{-\lambda t_2}, & k_2 \geq k_1 \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Poisson Random Process

$$0 < t_1 < t_2 \quad \Rightarrow$$

$$P[X(t_2) = k_2 \,|\, X(t_1) = k_1] = P[X(t_2) - X(t_1) = k_2 - k_1 \,|\, X(t_1) = k_1]$$

$$= P[X(t_2) - X(t_1) = k_2 - k_1]$$

$$= \begin{cases} \dfrac{[\lambda(t_2 - t_1)]^{(k_2 - k_1)}}{(k_2 - k_1)!} e^{-\lambda(t_2 - t_1)}, & k_2 \geq k_1 \\ 0, & \text{otherwise} \end{cases}$$

Example

$$X(t) = \text{Poisson r.p.}$$

$$0 < t_1 < t_2 < t_3$$

$$0 \le k_1 \le k_2 \le k_3 \quad \Rightarrow$$

$$P[X(t_1) = k_1, X(t_2) = k_2, X(t_3) = k_3]$$

$$= P[X(t_1) = k_1, X(t_2) - X(t_1) = k_2 - k_1, X(t_3) - X(t_2) = k_3 - k_2]$$

$$= P[X(t_1) = k_1]P[X(t_2) - X(t_1) = k_2 - k_1]P[X(t_3) - X(t_2) = k_3 - k_2]$$

$$= \frac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \frac{[\lambda(t_2 - t_1)]^{(k_2 - k_1)}}{(k_2 - k_1)!} e^{-\lambda(t_2 - t_1)} \frac{[\lambda(t_3 - t_2)]^{(k_3 - k_2)}}{(k_3 - k_2)!} e^{-\lambda(t_3 - t_2)}$$

$$= \frac{(\lambda t_1)^{k_1}[\lambda(t_2 - t_1)]^{(k_2 - k_1)}[\lambda(t_3 - t_2)]^{(k_3 - k_2)}}{k_1!(k_2 - k_1)!(k_3 - k_2)!} e^{-\lambda t_3}$$

Tutorial#1

- Problem-1:

A discrete random process is defined as d(n)=x(n)-x(n-1), where x(n) is a stationary process with zero mean. If var[d(n)]=1/10 var[x(n)], find Rxx(1)/var[x(n)].

- Solution:

$$\mathrm{var}[d(n)] = E[d^2(n)] - \{E[d(n)]\}^2$$

$$E[d(n)] = E[x(n) - x(n-1)] = 0$$

- var[d(n)]= 2*var[x(n)]-2*Rxx(-1)
- But Rxx(-1)=Rxx(1)
- Since var[d(n)]=1/10 var[x(n)]
- Then Rxx(1)/var[x(n)]=1.9/2=0.95

Tutorial#1

- Problem-2:

For two random variables X and Y

$$f_{XY}(x,y)=0.5\delta(x+1)\delta(y)+0.1\delta(x)\delta(y)+0.1\delta(x)\delta(y-2)+0.4\delta(x-1)\delta(y+2)+0.2\delta(x-1)\delta(y-1)+0.5\delta(x-1)\delta(y-3)$$

find

a) Correlation

b) Covariance

c) Correlation coefficient of X and Y

- Solution:
- Correlation: $R_{XY}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x,y)dxdy=0.9$
- Covariance: Cxy=Rxy-E[X]E[Y]=0.9-(0.6)(1.1)=0.24
- Correlation coefficient of X and Y: $\rho_{XY}=\frac{C_{XY}}{\sqrt{\sigma_X{}^2}\sqrt{\sigma_Y{}^2}}=\frac{0.24}{\sqrt{(1.24)(5.49)}}=0.092$

Tutorial#1

- Problem-3:

Mean of X=6 and Rxx(t,t+τ)=36+25$e^{-|\tau|}$ For a random process X(t). Indicate with of the following statements are true based on what is known with certainty X(t)

a) is first order stationary
b) has total average power of 61W
c) is ergodic
d) is wide sense stationary

- Solution:
- a) A random process X(t) is said to be first order stationary if

  fx(x1,t) = fx(xi,t+Δ) (i.e., no change in time shift)

  here mean is constant, hence X(t) is first order stationary (true)

- b) The average power of random process with autocorrelation function Rxx(τ) is Pavg=Rxx(τ)at τ=0, hence Pavg=61W (true)

Tutorial#1

- c) if E[X(t)]= mean(X(t)) is not equal to zero, with no periodic components then X(t) is ergodic.

- d) a random variable X(t) is said to be wide sense stationary if it satisfies two conditions

  i) mean is constant

  ii) autocorrelation function is a function of τ

  here both the conditions are true, hence X(t) is wide sense stationary

Tutorial#2

- Problem-1:

Given the autocorrelation function for a stationary ergodic process with no periodic components is Rxx(τ)= $25+\frac{4}{1+6\tau^2}$. Find the mean and variance of X(t).

- Solution:
- if X(t) is ergodic with no periodic components then

$$\lim_{\tau\to\infty} R_{XX}(\tau) = \overline{X}^2$$
$$\overline{X} = \pm 5 = \mu_1$$

- where μ_1 is the 1st mean or expectation of X(t)
- autocorrelation at origin is the mean square value Rxx(0)= $\overline{X^2}$ (2nd moment or μ_2)
- variance $\sigma_X{}^2 = \mu_2 - \mu_1{}^2 = 29 - 25 = 4$

Tutorial#2

- Problem-2:

Consider a random process X(t)=Acoswt, where w is a constant and A is a random variable uniformly distributed over (0,1). Find the auto correlation and covariance of X(t).

- Solution:
- f(A)=1 in (0,1)
- Rxx(t1,t2)=E[X(t1).X(t2)]= 1/3[coswt1.coswt2]
- covariance Cxx(t1,t2)=Rxx(t1,t2)-E[X(t1)].E[X(t2)]=1/12 coswt1.coswt2

Tutorial#2

- Problem-3:

X(t) is a gaussian process with mean 2 and autocorrelation function of

Rxx(τ)=5 $e^{-0.2|\tau|}$

Find P[X(4)≤1]

- Solution:
- for a gaussian random variable X

$$P(X \le k) = 1 - Q\left(\frac{k-m}{\sigma}\right)$$

$$P(X(4) \le 1) = 1 - Q\left(\frac{1-m}{\sigma}\right)$$

- finally $P(X(4) \le 1) = Q(1)$

Tutorial#3

- Problem-1:

The relationship between the input X(t) and Y(t) of a system is Y(t)=$X^2(t)$
X(t) is a zero mean stationary gaussian random process with auto correlation function Rxx(τ)=$e^{-\alpha|\tau|} \; for \, \alpha > 0$
Find E[Y(t)]and Ryy(τ)

- Solution:

$$E[Y(t)] = E[X^2(t)] = R_{XX}(0) = 1$$

$$R_{YY}(\tau) = E[Y(t_1)Y(t_2)] = E[X^2(t_1)X^2(t_2)]$$

$$R_{YY}(\tau) = 1 + 2e^{-2\alpha|\tau|}$$

Tutorial#3

- Problem-2:

Aircraft arrive at an airport according to a poisson process at a rate of 12 per hour. All aircrafts are handled by one air traffic controller. If the controller takes a 2minute coffee break, what is the probability that he will miss one or more arriving aircrafts.

- Solution:
- λ=12; t=2
- probability that he miss one or more arriving aircrafts

  =1-(probability that does not miss any aircraft)

  =1-(k=0)
- According to poisson process,

$$P[X(t)=k]=\frac{(\lambda t)^k e^{-\lambda t}}{k!} \, where k=0,1,2...$$

- at k=0 P[X(t)=0]= e^{-24}
- probability that he miss one or more arriving aircrafts is 1- e^{-24}

Linear Systems with Random Inputs

Random Signal Response of Linear Systems

- Linear system fundamentals
- Consider: linear, stable, time-invariant system.
- $y(t) = \int_{-\infty}^{+\infty} h(\lambda)x(t-\lambda)d\lambda = \int_{-\infty}^{+\infty} x(\lambda)h(t-\lambda)d\lambda$
- $Y(t) = \int_{-\infty}^{+\infty} h(\lambda)X(t-\lambda)d\lambda$

Mean and Mean Squared Value of System Response

- Assume $X(t)$ is w.s.s.
- Assume integration and expectations are exchangeable
- **Mean:**
  - $E[Y(t)] = E\left[\int_{-\infty}^{+\infty} h(\lambda)X(t-\lambda)d\lambda\right] = \int_{-\infty}^{+\infty} h(\lambda)E[X(t-\lambda)]d\lambda = \bar{X}\int_{-\infty}^{+\infty} h(\lambda)d\lambda = \bar{Y}$ constant
- **Mean squared**
  - $E[Y^2(t)] = E\left[\int_{-\infty}^{+\infty} h(\lambda_1)X(t-\lambda_1)d\lambda_1 \int_{-\infty}^{+\infty} h(\lambda_2)X(t-\lambda_2)d\lambda_2\right]$
  - $= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} E[X(t-\lambda_1)X(t-\lambda_2)]h(\lambda_1)h(\lambda_2)d\lambda_1 d\lambda_2$
  - $\overline{Y^2} = E[Y^2(t)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} R_{XX}(\lambda_1-\lambda_2)h(\lambda_1)h(\lambda_2)d\lambda_1 d\lambda_2$
  - Not always easy to integrate!

Example

Find $\overline{Y^2}$ if input is white noise

$$R_{XX}(\lambda_1 - \lambda_2) = \left(\frac{N_0}{2}\right)\delta(\lambda_1 - \lambda_2)$$

N_0 is a positive real constant.

$$\overline{Y^2} = E[Y^2(t)]$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\frac{N_0}{2}\right)\delta(\lambda_1 - \lambda_2)h(\lambda_1)h(\lambda_2)d\lambda_1 d\lambda_2$$

$$= \frac{N_0}{2}\int_{-\infty}^{+\infty} h^2(\lambda_2)d\lambda_2$$

Output power is proportional to the area under the square of $h(t)$ in this case.

Autocorrelation Function of Response

Assume w.s.s.

$$R_{YY}(\tau) = E[Y(t)Y(t+\tau)]$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} R_{XX}(\tau + \lambda_1 - \lambda_2)h(\lambda_1)h(\lambda_2)d\lambda_1 d\lambda_2$$

$$R_{YY}(\tau) = R_{XX}(\tau) * h(-\tau) * h(\tau)$$

Two fold convolution

Cross Correlation Function of input and output

- $R_{XY}(\tau) = R_{XX}(\tau) * h(\tau)$
- $R_{YX}(\tau) = R_{XX}(\tau) * h(-\tau)$
- $R_{YY}(\tau) = R_{XY}(\tau) * h(-\tau) = R_{YX}(\tau) * h(\tau)$
- $X(t)$ & $Y(t)$ are jointly w.s.s. if $X(t)$ is w.s.s. because $Y(t)$ will be w.s.s.
- Example: For the same white noise example, find $R_{XY}(\tau)$ & $R_{YX}(\tau)$
- $R_{XY}(\tau) = \int_{-\infty}^{+\infty} \left(\frac{N_0}{2}\right) \delta(\tau - \lambda) h(\lambda) d\lambda = \frac{N_0}{2} h(\tau)$
- $R_{YX}(\tau) =$
  $\int_{-\infty}^{+\infty} \left(\frac{N_0}{2}\right) \delta(\tau - \lambda) h(-\lambda) d\lambda = \frac{N_0}{2} h(-\tau) = R_{XY}(-\tau)$

Module-V
Random Process-Spectral Characteristics

Power density spectrum

Introduction

- Fourier integral

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}[\int_{-\infty}^{\infty}x(\tau)e^{-j\omega\tau}d\tau]e^{j\omega t}d\omega$$

- Fourier transform

$$X(\omega)=\int_{-\infty}^{\infty}x(t)e^{-j\omega t}dt$$

- Inverse Fourier transform

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)e^{j\omega t}d\omega$$

Introduction (Contd..)

$$x_T(t) = \begin{cases} x(t), & -T < t < T \\ 0, & o/w \end{cases}$$

Assume $\int_{-T}^{T} |x_T(t)|\, dt < \infty,$ for all finite T.

$$X_T(\omega) = \int_{-\infty}^{\infty} x_T(t) e^{-j\omega t} dt = \int_{-T}^{T} x(t) e^{-j\omega t} dt$$

- Energy contained in x(t) in the interval (-T,T)

$$E(T) = \int_{-\infty}^{\infty} x_T(t)^2 dt = \int_{-T}^{T} x(t)^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X_T(\omega)|^2 d\omega$$

Introduction (Contd..)

- Average power in x(t) in the interval (-T,T)

$$P(T) = \frac{1}{2T}\int_{-T}^{T} x(t)^2\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{|X_T(\omega)|^2}{2T}\,d\omega$$

$$x(t) \rightarrow X(t), \quad \text{take expectation,} \quad \text{let } T \rightarrow \infty.$$

- Average power in random process x(t)

$$P_{XX} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} E[X(t)^2]dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{T\to\infty}\frac{E[|X_T(\omega)|^2]}{2T}\,d\omega$$

$$P_{XX} = A\{E[X(t)^2]\} \qquad P_{XX} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega)d\omega$$

$$S_{XX} = \lim_{T\to\infty}\frac{E[|X_T(\omega)|^2]}{2T} \quad \text{power density spectrum}$$

Example-1

$$P_{XX} = A\{E[X(t)^2]\}$$

$$\text{w.s.s.} \Rightarrow \quad P_{XX} = R_{XX}(0)$$

- Example-1 $X(t) = A_0 \cos(\omega_0 t + \Theta)$ Θ -- uniformly distributed on $(0, \frac{\pi}{2})$

$$E[X(t)^2] = E[A_0^2 \cos^2(\omega_0 t + \Theta)] = E[\frac{A_0^2}{2} + \frac{A_0^2}{2}\cos(2\omega_0 t + 2\Theta)]$$

$$= \frac{A_0^2}{2} + \frac{A_0^2}{2}\int_0^{\frac{\pi}{2}} \frac{2}{\pi}\cos(2\omega_0 t + 2\theta)d\theta = \frac{A_0^2}{2} + \frac{A_0^2}{2\pi}\sin(2\omega_0 t + 2\theta)\Big|_{\theta=0}^{\frac{\pi}{2}}$$

$$= \frac{A_0^2}{2} - \frac{A_0^2}{\pi}\sin(2\omega_0 t)$$

$$P_{XX} = A\{E[X(t)^2]\} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}[\frac{A_0^2}{2} - \frac{A_0^2}{\pi}\sin(2\omega_0 t)]dt = \frac{A_0^2}{2}$$

Example-2

- Example-2 $X(t) = A_0 \cos(\omega_0 t + \Theta)$

$$X_T(\omega) = \int_{-T}^{T} A_0 \cos(\omega_0 t + \Theta) e^{-j\omega t} dt = \int_{-T}^{T} A_0 \frac{1}{2} [e^{j\Theta} e^{j\omega_0 t} + e^{-j\Theta} e^{-j\omega_0 t}] e^{-j\omega t} dt$$

$$= \frac{A_0}{2} e^{j\Theta} \int_{-T}^{T} e^{j(\omega_0 - \omega)t} dt + \frac{A_0}{2} e^{-j\Theta} \int_{-T}^{T} e^{-j(\omega_0 + \omega)t} dt$$

$$= A_0 T e^{j\Theta} \frac{\sin[(\omega - \omega_0)T]}{(\omega - \omega_0)T} + A_0 T e^{-j\Theta} \frac{\sin[(\omega + \omega_0)T]}{(\omega + \omega_0)T}$$

$$\int_{-T}^{T} e^{j\beta t} dt = \frac{1}{j\beta} e^{j\beta t} \Big|_{t=-T}^{T} = \frac{e^{j\beta T} - e^{-j\beta T}}{j\beta} = 2T \frac{\sin(\beta T)}{\beta T}$$

$$S_{XX} = \lim_{T \to \infty} \frac{E[|X_T(\omega)|^2]}{2T} \quad \text{power density spectrum}$$

$$X_T(\omega) = A_0 T e^{j\Theta} \frac{\sin[(\omega-\omega_0)T]}{(\omega-\omega_0)T} + A_0 T e^{-j\Theta} \frac{\sin[(\omega+\omega_0)T]}{(\omega+\omega_0)T}$$

$$X_T(\omega)^* = A_0 T e^{-j\Theta} \frac{\sin[(\omega-\omega_0)T]}{(\omega-\omega_0)T} + A_0 T e^{j\Theta} \frac{\sin[(\omega+\omega_0)T]}{(\omega+\omega_0)T}$$

$$\left|X_T(\omega)\right|^2 = X_T(\omega) X_T(\omega)^* = A_0^2 [T^2 \frac{\sin^2[(\omega-\omega_0)T]}{(\omega-\omega_0)^2 T^2} + T^2 \frac{\sin^2[(\omega+\omega_0)T]}{(\omega+\omega_0)^2 T^2}]$$

$$+ A_0^2 T^2 (e^{j2\Theta} + e^{-j2\Theta}) \frac{\sin[(\omega-\omega_0)T]}{(\omega-\omega_0)T} \frac{\sin[(\omega+\omega_0)T]}{(\omega+\omega_0)T}$$

$$E[e^{j2\Theta} + e^{-j2\Theta}] = E[2\cos 2\Theta] = \int_0^{\frac{\pi}{2}} \frac{2}{\pi} 2\cos 2\theta \, d\theta = \frac{2}{\pi} \sin 2\theta \Big|_0^{\pi/2} = 0$$

$$\frac{E[\left|X_T(\omega)\right|^2]}{2T} = \frac{A_0^2 \pi}{2} [\frac{T}{\pi} \frac{\sin^2[(\omega-\omega_0)T]}{(\omega-\omega_0)^2 T^2} + \frac{T}{\pi} \frac{\sin^2[(\omega+\omega_0)T]}{(\omega+\omega_0)^2 T^2}]$$

Example-2 (Contd..)

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \pi$$

$$\int_{-\infty}^{\infty} \frac{T}{\pi} \frac{\sin^2(\alpha T)}{(\alpha T)^2} d\alpha = \int_{-\infty}^{\infty} \frac{T}{\pi} \frac{\sin^2 x}{x^2} \frac{1}{T} dx = 1 \qquad \text{(a)}$$

$$\lim_{T\to\infty} \frac{T}{\pi} \frac{\sin^2(\alpha T)}{(\alpha T)^2} = \begin{cases} \infty, & \text{if } \alpha = 0 \\ 0, & \text{if } \alpha \neq 0 \end{cases} \qquad \text{(b)}$$

$$\text{(a) \& (b)} \Rightarrow \lim_{T\to\infty} \frac{T}{\pi} \frac{\sin^2(\alpha T)}{(\alpha T)^2} = \delta(\alpha)$$

$$S_{XX}(\omega) = \lim_{T\to\infty} \frac{E[|X_T(\omega)|^2]}{2T} = \frac{A_0^2 \pi}{2}[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]$$

$$P_{XX} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega) d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{A_0^2\pi}{2}[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)] d\omega = \frac{A_0^2}{2}$$

Properties Power density spectrum

Properties of the power density spectrum:

(1) $S_{XX}(\omega) \geq 0$

(2) $X(t)$ real $\Rightarrow$ $S_{XX}(-\omega) = S_{XX}(\omega)$

(3) $S_{XX}(\omega)$ is real

(4) $\frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega)d\omega = A\{E[X(t)^2]\}$

$$S_{XX}(\omega) = \lim_{T\to\infty} \frac{E[|X_T(\omega)|^2]}{2T}$$

PF of (2): $X_T(\omega) = \int_{-T}^{T} X(t)e^{-j\omega t}dt$

$$X_T(\omega)^* = \int_{-T}^{T} X(t)^* e^{j\omega t}dt = \int_{-T}^{T} X(t)e^{j\omega t}dt = X_T(-\omega)$$

$$S_{XX}(-\omega) = \lim_{T\to\infty} \frac{E[X_T(-\omega)X_T(-\omega)^*]}{2T} = \lim_{T\to\infty} \frac{E[X_T(\omega)^* X_T(\omega)]}{2T} = S_{XX}(\omega)$$

Properties Power density spectrum

Properties of the power density spectrum

(5) $S_{\dot{X}\dot{X}}(\omega)=\omega^2 S_{XX}(\omega)$ $\qquad \frac{d}{dt}X(t)=\lim_{\varepsilon\to 0}\frac{X(t+\varepsilon)-X(t)}{\varepsilon}$

PF of (5):

$$\dot{X}_T(t)=\begin{cases}\lim_{\varepsilon\to 0}\frac{X(t+\varepsilon)-X(t)}{\varepsilon}, & -T<t<T\\ 0, & \text{o/w}\end{cases}$$

$$f(t-a)\xleftrightarrow{FT} F(\omega)e^{-j\omega a}$$

$$\dot{X}_T(t)\xleftrightarrow{\text{FT}}\lim_{\varepsilon\to 0}\frac{X_T(\omega)e^{j\omega\varepsilon}-X_T(\omega)}{\varepsilon}=j\omega X_T(\omega)$$

$$S_{\dot{X}\dot{X}}(\omega)=\lim_{T\to\infty}\frac{E[|\dot{X}_T(\omega)|^2]}{2T}=\lim_{T\to\infty}\frac{E[|j\omega X_T(\omega)|^2]}{2T}=\omega^2\lim_{T\to\infty}\frac{E[|X_T(\omega)|^2]}{2T}=\omega^2 S_{XX}(\omega)$$

Properties Power density spectrum

Bandwidth of the power density spectrum

$$X(t) \text{ real} \;\Rightarrow\; S_{XX}(\omega) \text{ even}$$

$S_{XX}(\omega)$ lowpass form $\Rightarrow$

$$W_{\text{rms}}^2 = \frac{\int_{-\infty}^{\infty} \omega^2 S_{XX}(\omega)d\omega}{\int_{-\infty}^{\infty} S_{XX}(\omega)d\omega}$$

root mean square Bandwidth

$S_{XX}(\omega)$ bandpass form $\Rightarrow$ $\bar{\omega}_0 = \dfrac{\int_0^{\infty} \omega S_{XX}(\omega)d\omega}{\int_0^{\infty} S_{XX}(\omega)d\omega}$ mean frequency

$$W_{\text{rms}}^2 = \frac{4\int_0^{\infty} (\omega - \bar{\omega}_0)^2 S_{XX}(\omega)d\omega}{\int_0^{\infty} S_{XX}(\omega)d\omega}$$

rms BW

Example

$$S_{XX}(\omega) = \frac{10}{[1+(\omega/10)^2]^2} \qquad S_{XX}(\omega) \text{ lowpass form}$$

$$\int_{-\infty}^{\infty} S_{XX}(\omega) d\omega = \int_{-\infty}^{\infty} \frac{10}{[1+(\omega/10)^2]^2} d\omega = \int_{-\pi/2}^{\pi/2} \frac{10}{[1+\tan^2\theta]^2} 10\sec^2\theta\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} \frac{100}{\sec^2\theta} d\theta = \int_{-\pi/2}^{\pi/2} 100\cos^2\theta\, d\theta = \int_{-\pi/2}^{\pi/2} 100\frac{1+\cos 2\theta}{2} d\theta = 50\pi$$

$$\omega = 10\tan\theta \quad \Rightarrow \quad d\omega = 10\sec^2\theta\, d\theta$$

$$\int_{-\infty}^{\infty} \omega^2 S_{XX}(\omega) d\omega = \int_{-\infty}^{\infty} \frac{10\omega^2}{[1+(\omega/10)^2]^2} d\omega = \int_{-\pi/2}^{\pi/2} \frac{10^3\tan^2\theta}{[1+\tan^2\theta]^2} 10\sec^2\theta\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} \frac{10^4\tan^2\theta}{\sec^2\theta} d\theta = \int_{-\pi/2}^{\pi/2} 10^4\sin^2\theta\, d\theta = \int_{-\pi/2}^{\pi/2} 10^4\frac{1-\cos 2\theta}{2} d\theta = 5000\pi$$

Example

$$W_{\text{rms}}^2 = \frac{\int_{-\infty}^{\infty} \omega^2 S_{XX}(\omega) d\omega}{\int_{-\infty}^{\infty} S_{XX}(\omega) d\omega} = 100$$

rms BW $\quad W_{\text{rms}} = 10 \text{ rad/sec}$

$$S_{XX}(\omega) = \frac{10}{[1+(\omega/10)^2]^2}$$

Relationship between PSD and autocorrelation

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega)e^{j\omega\tau}d\omega = A[R_{XX}(t,t+\tau)]$$

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} A[R_{XX}(t,t+\tau)]e^{-j\omega\tau}d\tau$$

$$S_{XX}(\omega) = \lim_{T\to\infty}\frac{E[X_T(\omega)^* X_T(\omega)]}{2T} = \lim_{T\to\infty}\frac{1}{2T}E[\int_{-T}^{T} X(t_1)e^{j\omega t_1}dt_1 \int_{-T}^{T} X(t_2)e^{-j\omega t_2}dt_2]$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} E[X(t_1)X(t_2)]e^{j\omega(t_1-t_2)}dt_2 dt_1$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_{XX}(t_1,t_2)e^{j\omega(t_1-t_2)}dt_2 dt_1$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega)e^{j\omega\tau}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_{XX}(t_1,t_2)e^{j\omega(t_1-t_2)}dt_2 dt_1\, e^{j\omega\tau}d\omega$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_{XX}(t_1,t_2)\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega(\tau+t_1-t_2)}d\omega dt_2 dt_1$$

Relationship between PSD and autocorrelation

$$\delta(t) \xleftrightarrow{FT} 1$$

$$\int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt = 1$$

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(\omega) e^{j\omega\tau} d\omega = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \int_{-T}^{T} R_{XX}(t_1, t_2) \delta(\tau + t_1 - t_2) dt_2 dt_1$$

$$= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} R_{XX}(t_1, t_1 + \tau) dt_1 = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} R_{XX}(t, t + \tau) dt$$

$$= A[R_{XX}(t, t+\tau)]$$

$$A[R_{XX}(t, t+\tau)] \xleftrightarrow{FT} S_{XX}(\omega)$$

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} A[R_{XX}(t, t+\tau)] e^{-j\omega\tau} d\tau$$

Relationship between PSD and autocorrelation

- $X(t)$ w.s.s. $\Rightarrow A[R_{XX}(t,t+\tau)] = R_{XX}(\tau)$

$$R_{XX}(\tau) \xleftrightarrow{FT} S_{XX}(\omega)$$

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-j\omega\tau} d\tau$$

$$R_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(\omega) e^{j\omega\tau} d\omega$$

Example#1

- $X(t) = A\cos(\omega_0 t + \Theta)$

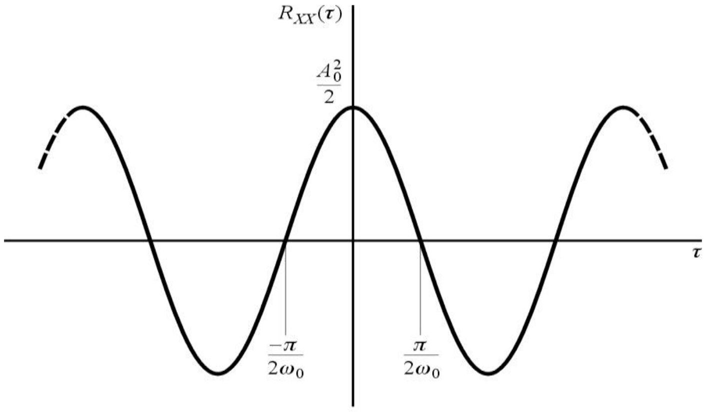
$$R_{XX}(\tau) = \frac{A_0^2}{2}\cos(\omega_0 \tau)$$

$$R_{XX}(\tau) = \frac{A_0^2}{4}(e^{j\omega_0 \tau} + e^{-j\omega_0 \tau})$$

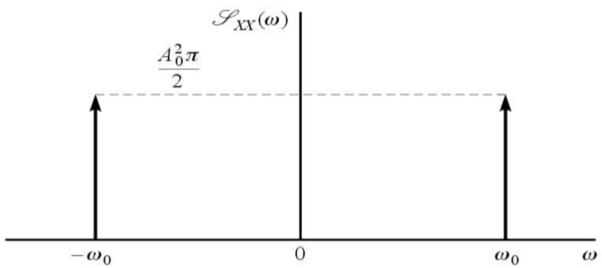
$$x(t)e^{j\alpha t} \overset{FT}{\longleftrightarrow} X(\omega - \alpha)$$

$$1 \overset{FT}{\longleftrightarrow} 2\pi\delta(\omega)$$

$$S_{XX}(\omega) = \frac{A_0^2\pi}{2}[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$$



(a)



(b)

Example#2

$$X(t) \text{ -- w.s.s.}$$

$$R_{XX}(\tau) = \begin{cases} A_0[1 - \frac{|\tau|}{T}], & -T < \tau < T \\ 0 \quad , & \text{elsewhere} \end{cases}$$

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-j\omega\tau} d\tau = A_0 \int_{-T}^{0} (1 + \frac{\tau}{T}) e^{-j\omega\tau} d\tau + A_0 \int_{0}^{T} (1 - \frac{\tau}{T}) e^{-j\omega\tau} d\tau$$

$$\int_{-T}^{0} (1 + \frac{\tau}{T}) e^{-j\omega\tau} d\tau = \int_{T}^{0} (1 - \frac{\alpha}{T}) e^{j\omega\alpha} (-d\alpha) = \int_{0}^{T} (1 - \frac{\tau}{T}) e^{j\omega\tau} d\tau$$

$$S_{XX}(\omega) = A_0 \int_0^T (1 - \frac{\tau}{T})(e^{j\omega\tau} + e^{-j\omega\tau}) d\tau = 2A_0 \int_0^T (1 - \frac{\tau}{T}) \cos(\omega\tau) d\tau$$

$$v = 1 - \frac{\tau}{T} \quad \Rightarrow \quad v' = \frac{-1}{T} \qquad u' = \cos(\omega\tau) \quad \Rightarrow \quad u = \frac{1}{\omega} \sin(\omega\tau)$$

Example#2

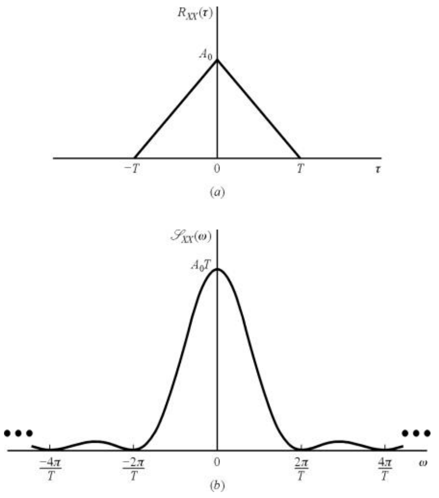
$$S_{XX}(\omega) = 2A_0\left(1-\frac{\tau}{T}\right)\frac{\sin(\omega\tau)}{\omega}\bigg|_0^T$$

$$+\frac{2A_0}{T}\int_0^T \frac{\sin(\omega\tau)}{\omega}\,d\tau$$

$$=\frac{2A_0}{T}\,\frac{-\cos(\omega\tau)}{\omega^2}\bigg|_0^T = \frac{2A_0}{\omega^2 T}[1-\cos(\omega T)]$$

$$=4A_0\frac{\sin^2(\omega T/2)}{\omega^2 T} = A_0 T\,\frac{\sin^2(\omega T/2)}{(\omega T/2)^2}$$

$$=A_0 T\,\mathrm{Sa}^2(\omega T/2)$$



Cross-power density spectrum

$$W(t) = X(t) + Y(t)$$

$$R_{WW}(t, t+\tau) = E[W(t)W(t+\tau)] = E\{[X(t)+Y(t)][X(t+\tau)+Y(t+\tau)]\}$$
$$= R_{XX}(t, t+\tau) + R_{YY}(t, t+\tau) + R_{XY}(t, t+\tau) + R_{YX}(t, t+\tau)$$

$$S_{WW}(\omega) = S_{XX}(\omega) + S_{YY}(\omega) + F\{A[R_{XY}(t, t+\tau)]\} + F\{A[R_{YX}(t, t+\tau)]\}$$

Cross-power density spectrum

$$x_T(t) = \begin{cases} x(t), & -T < t < T \\ 0, & o/w \end{cases} \qquad y_T(t) = \begin{cases} y(t), & -T < t < T \\ 0, & o/w \end{cases}$$

Assume $\int_{-T}^{T} |x_T(t)|\,dt < \infty$ & $\int_{-T}^{T} |y_T(t)|\,dt < \infty$, for all finite T.

$$x_T(t) \xleftrightarrow{\text{FT}} X_T(\omega) \qquad y_T(t) \xleftrightarrow{\text{FT}} Y_T(\omega)$$

Cross Power contained in $x(t), y(t)$ in the interval $(-T, T)$

$$P_{XY}(T) = \frac{1}{2T}\int_{-\infty}^{\infty} x_T(t)\,y_T(t)\,dt = \frac{1}{2T}\int_{-T}^{T} x(t)\,y(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{X_T(\omega)^* Y_T(\omega)}{2T}\,d\omega$$

Parseval's theorem

Cross-power density spectrum

average Cross Power contained in $X(t), Y(t)$ in the interval $(-T, T)$

$$\bar{P}_{XY}(T) = \frac{1}{2T}\int_{-T}^{T} R_{XY}(t,t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{E[X_T(\omega)^* Y_T(\omega)]}{2T} d\omega$$

total average Cross Power contained in $X(t), Y(t)$

$$P_{XY} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R_{XY}(t,t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{E[X_T(\omega)^* Y_T(\omega)]}{2T} d\omega$$

cross-power density spectrum $$S_{XY}(\omega) = \lim_{T\to\infty} \frac{E[X_T(\omega)^* Y_T(\omega)]}{2T}$$

Cross-power density spectrum

$$P_{XY} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XY}(\omega)d\omega$$

$$S_{YX}(\omega) = \lim_{T\to\infty}\frac{E[Y_T(\omega)^* X_T(\omega)]}{2T}$$

$$P_{YX} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{YX}(\omega)d\omega = P_{XY}^*$$

$$\text{Total cross power } = P_{XY} + P_{YX}$$

$$X(t), Y(t) \text{ orthogonal } \Rightarrow P_{XY} = P_{YX} = 0$$

Properties of cross-power density spectrum

$X(t), Y(t)$ real

Properties of the cross-power density spectrum:

(1) $S_{XY}(\omega) = S_{YX}(-\omega) = S_{YX}(\omega)^*$

PF of (1): $X_T(\omega) = \int_{-T}^{T} X(t) e^{-j\omega t} dt$

$$X_T(\omega)^* = \int_{-T}^{T} X(t)^* e^{j\omega t} dt = \int_{-T}^{T} X(t) e^{j\omega t} dt = X_T(-\omega)$$

$$S_{YX}(-\omega) = \lim_{T\to\infty} \frac{E[Y_T(-\omega)^* X_T(-\omega)]}{2T} = \lim_{T\to\infty} \frac{E[Y_T(\omega) X_T(\omega)^*]}{2T} = S_{XY}(\omega)$$

$$S_{YX}(-\omega) = \lim_{T\to\infty} \frac{E[Y_T(-\omega)^* X_T(-\omega)]}{2T} = \lim_{T\to\infty} \frac{E[Y_T(\omega) X_T(\omega)^*]}{2T} = S_{YX}(\omega)^*$$

Properties of cross-power density spectrum

(2) $\mathrm{Re}[S_{XY}(\omega)]$ & $\mathrm{Re}[S_{YX}(\omega)]$ -- even

(3) $\mathrm{Im}[S_{XY}(\omega)]$ & $\mathrm{Im}[S_{YX}(\omega)]$ -- odd

$$A\left[R_{XY}(t,t+\tau)\right] \xleftrightarrow{FT} S_{XY}(\omega)$$

$$A\left[R_{YX}(t,t+\tau)\right] \xleftrightarrow{FT} S_{YX}(\omega)$$

(4) $X(t)$ & $Y(t)$ orthogonal $\Rightarrow S_{XY}(\omega) = S_{YX}(\omega) = 0$

$X(t)$ & $Y(t)$ orthogonal $\Rightarrow R_{XY}(t,t+\tau) = 0 \Rightarrow \mathrm{A}[R_{XY}(t,t+\tau)] = 0$

(5) $X(t)$ & $Y(t)$ uncorrelated & have constant mean $\bar{X}, \bar{Y}$

$$\Rightarrow S_{XY}(\omega) = S_{YX}(\omega) = 2\pi \bar{X}\bar{Y}\delta(\omega)$$

Properties of cross-power density spectrum

PF of (5): $R_{XY}(t,t+\tau)=\overline{X}\,\overline{Y} \;\Rightarrow\; A[R_{XY}(t,t+\tau)]=\overline{X}\,\overline{Y}$

$$\Rightarrow\; S_{XY}(\omega)=2\pi\overline{X}\,\overline{Y}\,\delta(\omega)=S_{YX}(\omega)^*$$

$X(t),Y(t)$ -- jointly w.s.s. $\Rightarrow$ $R_{XY}(\tau)\xleftrightarrow{\text{FT}}S_{XY}(\omega)$

$$R_{YX}(\tau)\xleftrightarrow{\text{FT}}S_{YX}(\omega)$$

Example#1

$$X(t), Y(t) \text{ -- jointly w.s.s.}$$

$$S_{XY}(\omega) = \begin{cases} a + jb\omega/W, & -W < \omega < W \\ 0, & \text{o/w} \end{cases}$$

$$R_{XY}(\tau) = \frac{1}{2\pi}\int_{-W}^{W}(a + j\frac{b\omega}{W})e^{j\omega\tau}d\omega = \frac{1}{2\pi}(a + j\frac{b\omega}{W})\frac{e^{j\omega\tau}}{j\tau}\Bigg|_{-W}^{W} - \frac{1}{2\pi}\int_{-W}^{W}\frac{b}{W\tau}e^{j\omega\tau}d\omega$$

$$= \frac{1}{2\pi}\left[\frac{(a + jb)e^{jW\tau}}{j\tau} - \frac{(a - jb)e^{-jW\tau}}{j\tau}\right] - \frac{1}{2\pi}\frac{be^{j\omega\tau}}{jW\tau^2}\Bigg|_{-W}^{W}$$

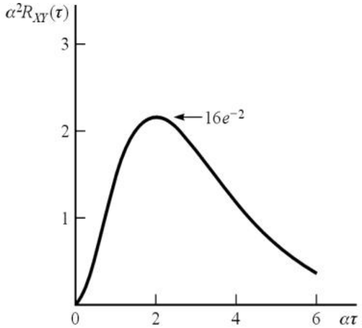
$$= \frac{1}{2\pi}\left[\frac{a}{\tau}\frac{e^{jW\tau} - e^{-jW\tau}}{j} + \frac{b}{\tau}(e^{jW\tau} + e^{-jW\tau})\right] - \frac{1}{2\pi}\frac{b}{W\tau^2}\frac{e^{j\omega\tau} - e^{-j\omega\tau}}{j}$$

$$= \frac{a}{\pi\tau}\sin(W\tau) + \frac{b}{\pi\tau}\cos(W\tau) - \frac{b}{\pi W\tau^2}\sin(W\tau)$$

Example#2

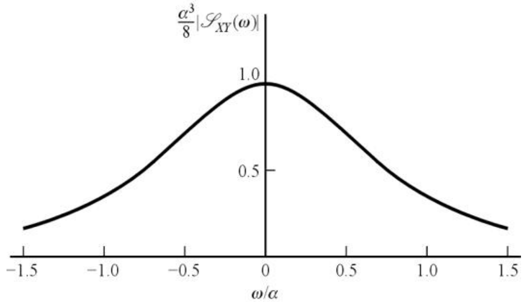
$$X(t), Y(t) \text{ -- jointly w.s.s.}$$

$$S_{XY}(\omega) = \frac{8}{(\alpha + j\omega)^3} \qquad \alpha > 0$$



(a)

$$u(\tau)\tau^2 e^{-\alpha\tau} \xleftrightarrow{FT} \frac{2}{(\alpha + j\omega)^3}$$



(b)

$$R_{XY}(\tau) = 4u(\tau)\tau^2 e^{-\alpha\tau}$$

Relationship between C-PSD and cross-correlation

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XY}(\omega)e^{j\omega\tau}d\omega = A[R_{XY}(t,t+\tau)]$$

$$S_{XY}(\omega) = \int_{-\infty}^{\infty} A[R_{XY}(t,t+\tau)]e^{-j\omega\tau}d\tau$$

$$S_{XY}(\omega) = \lim_{T\to\infty}\frac{E[X_T(\omega)^* Y_T(\omega)]}{2T} = \lim_{T\to\infty}\frac{1}{2T}E[\int_{-T}^{T} X(t_1)e^{j\omega t_1}dt_1 \int_{-T}^{T} Y(t_2)e^{-j\omega t_2}dt_2]$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} E[X(t_1)Y(t_2)]e^{j\omega(t_1-t_2)}dt_2 dt_1$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_{XY}(t_1,t_2)e^{j\omega(t_1-t_2)}dt_2 dt_1$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XY}(\omega)e^{j\omega\tau}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_{XY}(t_1,t_2)e^{j\omega(t_1-t_2)}dt_2 dt_1\, e^{j\omega\tau}d\omega$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_{XY}(t_1,t_2)\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega(\tau+t_1-t_2)}d\omega dt_2 dt_1$$

Relationship between C-PSD and cross-correlation

$$\int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt = 1$$

$$\delta(t) \xleftrightarrow{FT} 1$$

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XY}(\omega) e^{j\omega\tau} d\omega = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \int_{-T}^{T} R_{XY}(t_1, t_2) \delta(\tau + t_1 - t_2) dt_2 dt_1$$

$$= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} R_{XY}(t_1, t_1 + \tau) dt_1 = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} R_{XY}(t, t+\tau) dt$$

$$= A[R_{XY}(t, t+\tau)]$$

$$A[R_{XY}(t, t+\tau)] \xleftrightarrow{FT} S_{XY}(\omega)$$

$$S_{XY}(\omega) = \int_{-\infty}^{\infty} A[R_{XY}(t, t+\tau)] e^{-j\omega\tau} d\tau$$

Example:

$$R_{XY}(t,t+\tau)=\frac{AB}{2}\{\sin(\omega_0\tau)+\cos[\omega_0(2t+\tau)]\}$$

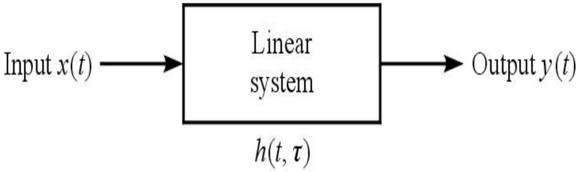
$$A[R_{XY}(t,t+\tau)]=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}R_{XY}(t,t+\tau)dt$$

$$=\frac{AB}{2}\sin(\omega_0\tau)+\frac{AB}{2}\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\cos[\omega_0(2t+\tau)]dt$$

$$=\frac{AB}{2}\sin(\omega_0\tau)=\frac{AB}{4j}[e^{j\omega_0\tau}-e^{-j\omega_0\tau}]$$

$$S_{XY}(\omega)=\frac{AB}{4j}[2\pi\delta(\omega-\omega_0)-2\pi\delta(\omega+\omega_0)]=\frac{-j\pi AB}{2}[\delta(\omega-\omega_0)-\delta(\omega+\omega_0)]$$

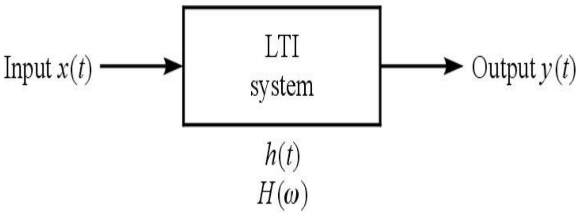
Linear System $$y(t)=\int_{-\infty}^{\infty} x(\xi)h(t,\xi)d\xi$$



(a)

$$\delta(t-\xi) \quad \rightarrow \quad h(t,\xi)$$ impulse response

Linear Time-Invariant System (LTI system)



(b)

$$y(t)=\int_{-\infty}^{\infty} x(\xi)h(t-\xi)d\xi=\int_{-\infty}^{\infty} h(\xi)x(t-\xi)d\xi$$

$$y(t)=x(t)*h(t)=h(t)*x(t)$$ convolution integral

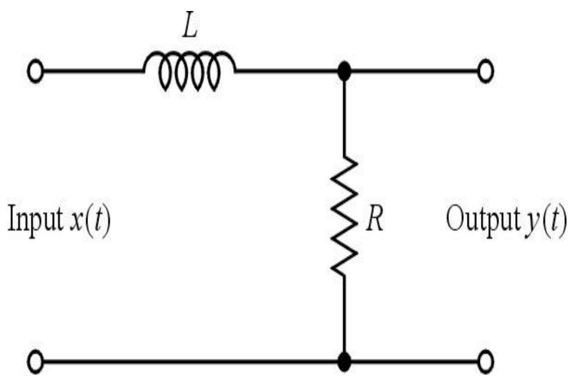
$$Y(\omega)=X(\omega)H(\omega)$$

$$x(t)=e^{j\omega t} \quad \Rightarrow \quad \frac{y(t)}{x(t)}=\frac{\int_{-\infty}^{\infty} h(\xi)e^{j\omega(t-\xi)}d\xi}{e^{j\omega t}}=\int_{-\infty}^{\infty} h(\xi)e^{-j\omega\xi}d\xi=H(\omega)$$

Linear system fundamentals

Example-1: $$H(s) = \frac{R}{sL + R}$$

$$H(\omega) = \frac{R}{j\omega L + R}$$



LTI causal $\Leftrightarrow$ $h(t) = 0 \quad \text{for } t < 0$

LTI stable $\Leftrightarrow$ $\int_{-\infty}^{\infty} |h(t)| \, dt < \infty$

Ideal lowpass filter

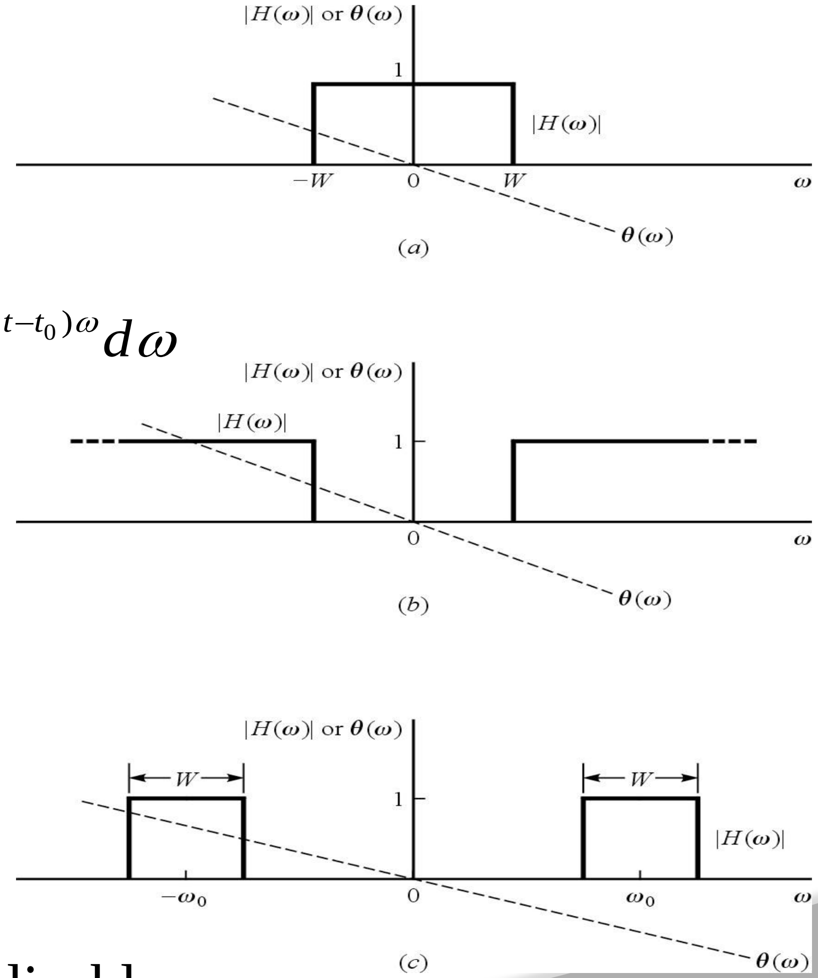
$$H(\omega)=\begin{cases} e^{-jt_0\omega}, & |\omega|<W \\ 0, & \text{o/w} \end{cases}$$

$$h(t)=\frac{1}{2\pi}\int_{-W}^{W} e^{-jt_0\omega}e^{jt\omega}d\omega=\frac{1}{2\pi}\int_{-W}^{W} e^{j(t-t_0)\omega}d\omega$$

$$=\frac{1}{2\pi}\frac{1}{j(t-t_0)}e^{j(t-t_0)\omega}\Big|_{-W}^{W}$$

$$=\frac{1}{2\pi}\frac{e^{j(t-t_0)W}-e^{-j(t-t_0)W}}{j(t-t_0)}$$

$$=\frac{W}{\pi}\frac{\sin[(t-t_0)W]}{(t-t_0)W}$$



Not causal $\Rightarrow$ Not physically realizable

Random signal response of linear systems

$X(t)$ -- w.s.s. random input $\qquad Y(t)=\int_{-\infty}^{\infty} h(\xi)X(t-\xi)d\xi$

$$E[Y(t)] = E[\int_{-\infty}^{\infty} h(\xi)X(t-\xi)d\xi] = \int_{-\infty}^{\infty} h(\xi)E[X(t-\xi)]d\xi$$

$$= \bar{X}\int_{-\infty}^{\infty} h(\xi)d\xi = \bar{Y}$$

$$R_{YY}(t,t+\tau) = E[Y(t)Y(t+\tau)]$$

$$= E[\int_{-\infty}^{\infty} h(\xi_1)X(t-\xi_1)d\xi_1 \int_{-\infty}^{\infty} h(\xi_2)X(t+\tau-\xi_2)d\xi_2]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E[X(t-\xi_1)X(t+\tau-\xi_2)]h(\xi_1)h(\xi_2)d\xi_1 d\xi_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{XX}(\tau+\xi_1-\xi_2)h(\xi_1)h(\xi_2)d\xi_1 d\xi_2$$

$X(t)$ w.s.s. $\Rightarrow$ $Y(t)$ w.s.s.

Random signal response of linear systems

$$R_{YY}(\tau) = \int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} R_{XX}(\tau + \xi_1 - \xi_2) h(\xi_1) d\xi_1] h(\xi_2) d\xi_2$$

$$= \int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} R_{XX}(\tau - \xi_2 - \xi_1) h(-\xi_1) d\xi_1] h(\xi_2) d\xi_2$$

$$= \int_{-\infty}^{\infty} R_{XX}(\xi_1) * h(-\xi_1) \Big|_{\xi_1 = \tau - \xi_2} h(\xi_2) d\xi_2$$

$$= R_{XX}(\tau) * h(-\tau) * h(\tau)$$

$$E[Y(t)^2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{XX}(\xi_1 - \xi_2) h(\xi_1) h(\xi_2) d\xi_1 d\xi_2$$

Example-1: white noise $X(t)$ $\qquad R_{XX}(\tau) = (N_0 / 2)\delta(\tau)$

$$E[Y(t)^2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (N_0 / 2)\delta(\xi_1 - \xi_2) h(\xi_1) h(\xi_2) d\xi_1 d\xi_2$$

$$= (N_0 / 2) \int_{-\infty}^{\infty} h(\xi_2)^2 d\xi_2$$

Random signal response of linear systems

$$R_{XY}(t,t+\tau)=E[X(t)Y(t+\tau)]=E[X(t)\int_{-\infty}^{\infty}h(\xi)X(t+\tau-\xi)d\xi]$$

$$=\int_{-\infty}^{\infty}E[X(t)X(t+\tau-\xi)]h(\xi)d\xi$$

$$=\int_{-\infty}^{\infty}R_{XX}(\tau-\xi)h(\xi)d\xi$$

$$=R_{XX}(\tau)*h(\tau)=R_{XY}(\tau)$$

$$R_{YX}(\tau)=R_{XY}(-\tau)=R_{XX}(-\tau)*h(-\tau)=R_{XX}(\tau)*h(-\tau)$$

$$=\int_{-\infty}^{\infty}R_{XX}(\tau-\xi)h(-\xi)d\xi$$

$X(t)$ w.s.s. $\Rightarrow$ $X(t)$ & $Y(t)$ jointly w.s.s.

$$R_{YY}(\tau)=R_{XY}(\tau)*h(-\tau)=R_{YX}(\tau)*h(\tau)$$

Random signal response of linear systems

Example-2: white noise $X(t)$ $\quad R_{XX}(\tau) = (N_0/2)\delta(\tau)$

$$R_{XY}(\tau) = R_{XX}(\tau) * h(\tau) = \int_{-\infty}^{\infty} R_{XX}(\tau - \xi) h(\xi) d\xi$$

$$= \int_{-\infty}^{\infty} (N_0/2)\delta(\tau - \xi) h(\xi) d\xi = (N_0/2) h(\tau)$$

$$R_{YX}(\tau) = R_{XY}(-\tau) = (N_0/2) h(-\tau)$$

Spectral characteristics of system response

$$R_{XY}(\tau) = R_{XX}(\tau) * h(\tau) \qquad S_{XY}(\omega) = S_{XX}(\omega)H(\omega)$$

$$R_{YX}(\tau) = R_{XX}(\tau) * h(-\tau) \qquad S_{YX}(\omega) = S_{XX}(\omega)H(-\omega) = S_{XX}(\omega)H(\omega)^*$$

$$R_{YY}(\tau) = R_{XY}(\tau) * h(-\tau) = R_{XX}(\tau) * h(\tau) * h(-\tau)$$

$$S_{YY}(\omega) = S_{XY}(\omega)H(\omega)^* = S_{XX}(\omega)H(\omega)H(\omega)^* = S_{XX}(\omega)\left|H(\omega)\right|^2$$

$$h(\tau) \quad \xleftrightarrow{FT} \quad H(\omega)$$

$$h(\tau) \ \text{real} \ \Rightarrow \quad h(-\tau) \quad \xleftrightarrow{FT} \quad H(-\omega) = H(\omega)^*$$

Spectral characteristics of system response

average power $$P_{YY} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{YY}(\omega)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega)\left|H(\omega)\right|^2 d\omega$$

Example-1:

$$S_{XX}(\omega) = \frac{N_0}{2} \qquad H(\omega) = \frac{1}{1+(j\omega L/R)}$$

$$S_{YY}(\omega) = S_{XX}(\omega)\left|H(\omega)\right|^2 = \frac{N_0/2}{1+(\omega L/R)^2}$$

$$P_{YY} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{YY}(\omega)d\omega = \frac{N_0}{4\pi}\int_{-\infty}^{\infty}\frac{1}{1+(\omega L/R)^2}d\omega$$

$$= \frac{N_0}{4\pi}\int_{-\pi/2}^{\pi/2}\frac{1}{1+\tan^2\theta}\frac{R}{L}\sec^2\theta d\theta = \frac{N_0 R}{4\pi L}\int_{-\pi/2}^{\pi/2} d\theta = \frac{N_0 R}{4L}$$

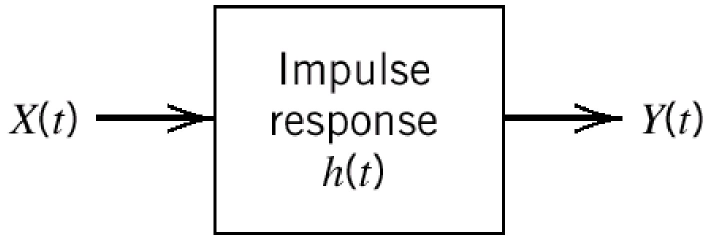
Spectral characteristics of system response

$$h(t) = (R/L)u(t)e^{-Rt/L} \quad \xleftrightarrow{FT} \quad H(\omega) = \frac{1}{1+(j\omega L/R)}$$

By Example-1,

$$P_{YY} = \frac{N_0}{2}\int_{-\infty}^{\infty} h(t)^2\,dt = \frac{N_0}{2}\int_{0}^{\infty} (R/L)^2 e^{-2Rt/L}dt = \frac{N_0 R}{4L} e^{-2Rt/L}\Big|_0^{\infty} = \frac{N_0 R}{4L}$$

Random process through a LTI System



$$Y(t) = \int_{-\infty}^{\infty} h(\tau_1) X(t-\tau_1)\, d\tau_1$$

where h(t) is the impulse response of the system

$$\mu_Y(t) = E[Y(t)]$$

If *E*[*X*(*t*)] is finite and system is stable

$$= E\left[\int_{-\infty}^{\infty} h(\tau_1) X(t-\tau_1)\, d\tau_1\right]$$

$$= \int_{-\infty}^{\infty} h(\tau_1) E[x(t-\tau_1)]\, d\tau_1$$

If *X*(*t*) is stationary, *H*(0) :System DC response.

$$= \int_{-\infty}^{\infty} h(\tau_1) \mu_X(t-\tau_1)\, d\tau_1$$

$$\mu_Y = \mu_X \int_{-\infty}^{\infty} h(\tau_1)\, d\tau_1 = \mu_X H(0),$$

Random process through a LTI System

Consider autocorrelation function of *Y*(*t*):

$$R_Y(t,\mu) = E\left[Y(t)Y(\mu)\right]$$

$$= E\left[\int_{-\infty}^{\infty} h(\tau_1)X(t-\tau_1)\, d\tau_1 \int_{-\infty}^{\infty} h(\tau_2)X(\mu-\tau_2)\, d\tau_2\right]$$

If $\mathrm{E}[X^2(t)]$ **is finite and the system is stable,**

$$R_Y(t,\mu) = \int_{-\infty}^{\infty} d\tau_1\, h(\tau_1) \int_{-\infty}^{\infty} d\tau_2\, h(\tau_2) R_X(t-\tau_1, \mu-\tau_2)$$

If $R_X(t-\tau_1, \mu-\tau_2) = R_X(t-\mu-\tau_1+\tau_2)$ **(stationary)**

$$R_Y(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau-\tau_1+\tau_2)\, d\tau_1\, d\tau_2$$

Stationary input, Stationary output

$$R_Y(0) = E\left[Y^2(t)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau_2-\tau_1)\, d\tau_1\, d\tau_2$$

Power Spectral Density (PSD)

Consider the Fourier transform of *g*(t),

$$G(f) = \int_{-\infty}^{\infty} g(t)\exp(-j2\pi ft)\,dt$$

$$g(t) = \int_{-\infty}^{\infty} G(f)\exp(j2\pi ft)\,df$$

Let *H*(*f*) denote the frequency response,

$$\tau = \tau_2 - \tau_1$$

$$h(\tau_1) = \int_{-\infty}^{\infty} H(f)\exp(j2\pi f\tau_1)\,df$$

$$E\left[Y^2(t)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} H(f)\exp(j2\pi f\tau_1)\,df\right]h(\tau_2)R_X(\tau_2-\tau_1)\,d\tau_1\,d\tau_2$$

$$= \int_{-\infty}^{\infty} df\; H(f)\int_{-\infty}^{\infty} d\tau_2 h(\tau_2)\int_{-\infty}^{\infty} R_X(\tau_2-\tau_1)\exp(j2\pi f\tau_1)\,d\tau_1$$

$$= \int_{-\infty}^{\infty} df\; H(f)\underline{\int_{-\infty}^{\infty} d\tau_2 h(\tau_2)\exp(j2\pi f\tau_2)}\int_{-\infty}^{\infty} R_X(\tau)\exp(-j2\pi f\tau)\,d\tau$$

$H^{*}(f)$ (complex conjugate response of the filter)


Thank you