

## Basics of Probability & Random Processes

Ref Material : Stark & Woods book.

Let us briefly understand what a probability space is.  
For this, we need to understand the defns. behind  
fields &  $\sigma$ -fields.

Consider a universal set  $\Omega$  and a collection of subsets of  $\Omega$ .

Let  $E, F, \dots$  denote the subsets in this collection.

This collection of subsets forms a field  $\mathcal{M}$  if

- 1)  $\emptyset \in \mathcal{M}, \Omega \in \mathcal{M}$  (null set & universal set are part of the field)
- 2) If  $E \in \mathcal{M}$  and  $F \in \mathcal{M}$ , then  $E \cup F \in \mathcal{M}, E \cap F \in \mathcal{M}$
- 3) If  $E \in \mathcal{M}$ , then  $E^c \in \mathcal{M}$ .

A  $\sigma$ -field  $\mathcal{F}$  is a field that is closed under any

countable set of unions, intersections & combinations.

i.e., if  $E_1, E_2, \dots, E_n, \dots$  belong to  $\mathcal{F}$  so do

$\bigcup_{i=1}^{\infty} E_i$  and  $\bigcap_{i=1}^{\infty} E_i$  (Set of all elements in at least one  $E_i$ )

NOTE :

A set  $S$  is called countable if there is an injective function  $f$  from  $S \rightarrow \{0, 1, 2, \dots\}$

Ques

- 1) Rationals are countably infinite
- 2) Real nos are uncountably infinite

Consider an experiment  $\mathcal{E}$  with a sample description space  $\Omega$ .  
 If  $\Omega$  has a countable number of elements, then every subset of  $\Omega$  may be assigned a 'probability' consistent with the axioms such that for every event  $E \in \mathcal{F}$

a)  $P(E) \geq 0$

b)  $P(\Omega) = 1$

c)  $P(E \cup F) = P(E) + P(F)$  if  $P(E \cap F) = 0$ .

Now, the class of all subsets make up a  $\sigma$  field and each subset is an event.

When  $\Omega$  is not countable i.e.,  $\Omega = \mathbb{R}$  (real line)  
not every subset of  $\Omega$  can be assigned a probability  
consistent with the above axioms.

Only those subsets for which prob. can be assigned are  
called 'events'.

The collection of those subsets is smaller than the  
collection of all possible subsets that one can define on  $\Omega$ .  
This smaller collection is called a  $\sigma$ -field a.k.a Borel field.

So, the 3 objects  $(\Omega, \mathcal{F}, \mathbb{P})$  form a "probability space"

```
graph TD; A[Sample space] --> B[Omega]; C[Field] --> D[F]; E[Prob. measure] --> F[P]; B --- G["(Omega, F, P)"]; D --- G; F --- G; G --- H["probability space"];
```

Example :

Suppose we do a fair coin toss 'once'

$$\Omega = \{H, T\}$$

$\sigma$  field of events consists of the following sets:

$$\{H\}, \{T\}, \phi, \Omega$$

Prob. measure

$$P(H) = \frac{1}{2}$$

$$P(T) = \frac{1}{2}$$

$$P(\phi) = 0$$

$$P(\Omega) = 1$$

$\Omega$

$\mathcal{F}$

$\mathcal{P}$

## Recap of basics

---

$P(A \cap B)$  : joint prob. of events  $A$  and  $B$ .

Example :  $A$  : Event when it rains

$B$  : Event when it is sunny

$C$  : Event when it rains & is sunny.

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

## In dependence

Two events  $A$  &  $B$  are statistically independent iff

$$P(AB) = P(A)P(B)$$

## Mutually Exclusive

$$P(AB) = 0$$

i.e.,  $P(A \cup B) = P(A) + P(B)$

(occurrence of one  $\Rightarrow$  non-occurrence of the other)



## Baye's Theorem

If  $A_i$ 's  $i = 1, \dots, n$  be a set of disjoint & exhaustive events over a prob. space  $\mathcal{P}$

$$\bigcup_{i=1}^n A_i = \Omega$$

$$A_i \cap A_j = \emptyset \text{ for } i \neq j$$

$i=1$

For any event

$$B \text{ over } \mathcal{P}$$
$$P(B|A_j) P(A_j)$$

$$P(A_j | B) =$$

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

$$P(B)$$

# Probability Distribution Function $F_X(x)$

The PDF contains the information necessary to compute  $P(E)$  for any event  $E$  in the Borel field of events.

$$F_X(x) = P \left[ \left\{ \xi : X(\xi) \leq x \right\} \right] = P_X \left( [-\infty, x] \right)$$

Set of all outcomes  $\xi$  in the underlying sample space such that  $X(\xi)$  assumes values less than  $x$ .

Such that  $\left\{ \xi : X(\xi) \leq x \right\} \subset \Omega$  is a subset of outcomes  
 i.e., that under the mapping  $X(\cdot)$  generates the set  $[-\infty, x]$

## Properties of PDF

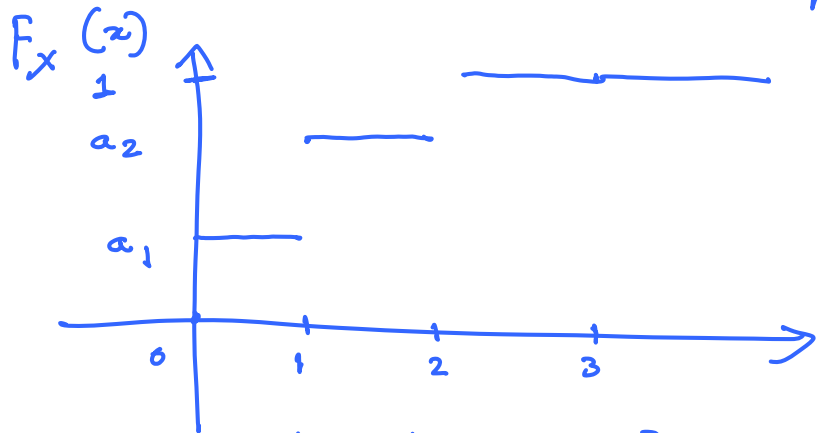
- 1)  $F_X(\infty) = 1$        $F_X(-\infty) = 0$
- 2)  $x_1 \leq x_2 \implies F_X(x_1) \leq F_X(x_2)$  (Non decreasing property)
- 3)  $F_X(x)$  is continuous from the right  
i.e.,  $F_X(x) = \lim_{\epsilon \rightarrow 0} F_X(x + \epsilon)$        $\epsilon > 0$

If  $F_X(x)$  is discontinuous @ a point, say  $x_0$ ,  
then  $F_X(x_0)$  will be taken to mean the value of  
the PDF immediately to the right of  $x_0$ .

Example : Suppose  $x$  is a binomial random variable

$$P(x = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$F_X(x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1-p)^{n-k}$$



Evaluated at  $n=3$

$$\begin{aligned} (a) \text{ Evaluate } P(1.5 < x < 3) \\ &= F_X(3) - P(x=3) - F_X(1.5) \\ &= 1 - P(x=3) - a_2 \end{aligned}$$

## Probability density function

If  $F_X(x)$  is continuous and differentiable,

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

## Properties

$$1) f_X(x) \geq 0$$

$$2) \int_{-\infty}^{\infty} f_X(x) dx = 1$$

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

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

$$\left( F_X(\infty) - F_X(-\infty) = 1 - 0 \right)$$

$$= \Pr(X \leq x)$$

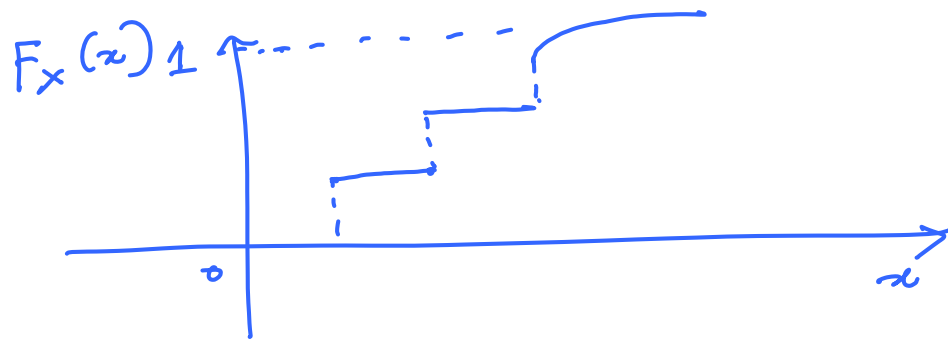
## Continuous, discrete, mixed random variables

---

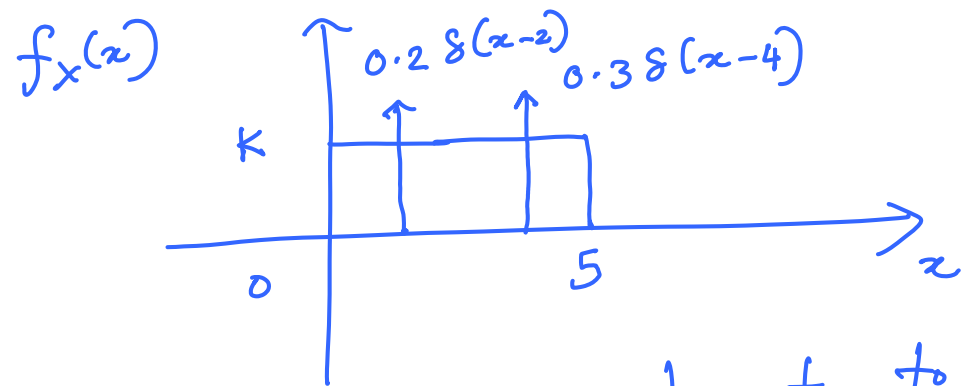
- Examples:
- (a) Gaussian random variable (Continuous)
  - (b) Binomial distribution (Discrete r.v.)
  - (c) Mixture of continuous & discrete random variables

For a mixed random variable

---



Example: Suppose the pdf of a mixed random variable is

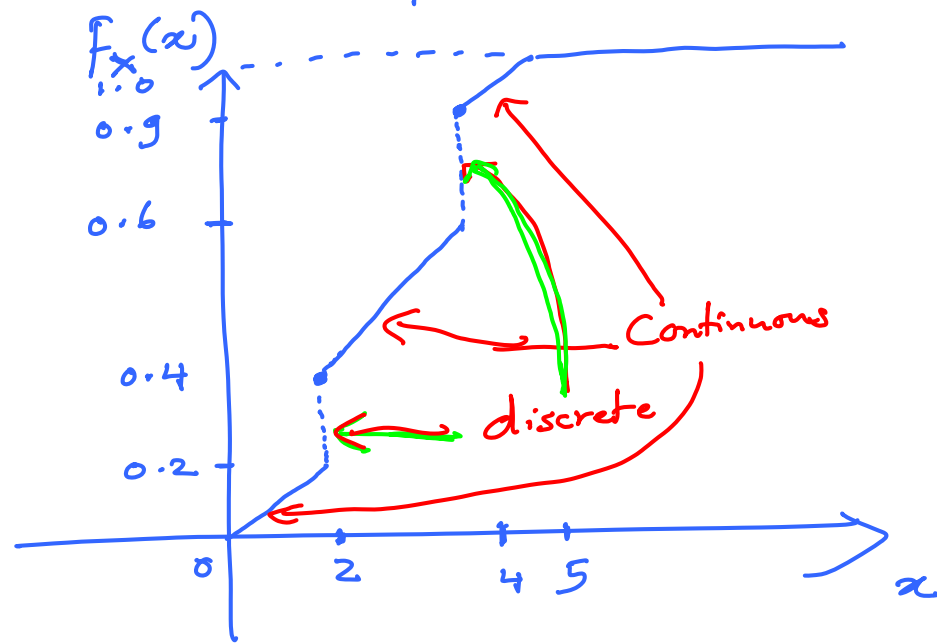


Since density must integrate to '1'

$$5k + 0.2 + 0.3 = 1$$

$$\Rightarrow k = \frac{0.5}{5} = 0.1$$

Sketch of the prob. distribution function



## Mean and Variance of a random variable (Discrete)

---

The mean of a random variable  $X$  with a prob. mass function (pmf)  $P_X$  is  $E(X)$  given by  $u_i$ 's are possible values of  $X$

$$E(X) = \sum_i u_i P_X(u_i)$$

Variance of a random variable  $X$  indicates the "spread" of the pmf of  $X$ .

$$\text{Var} = E\left((X - \mu_X)^2\right) = \sigma_X^2$$

$\sigma_X$  is the std. deviation

$$\sigma_X = \sqrt{\text{Var}(X)}$$



NOTE: For a cont. r.v.

$$E(X) = \int_{\mathcal{R}} x f_x(x) dx$$

if the integral exists

Expectation is a linear operation

$$E(a g(x) + b h(x) + c) =$$

$$a E(g(x)) + b E(h(x)) + c$$

One can compute higher moments  $E(X^k)$  for different values of 'k'

Additional material

- a) characteristic functions
- b) moment generating functions

If  $X$  and  $Y$  are random variables with finite second moments,

$$\text{Correlation : } E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{xy}(x, y) dx dy$$

$$\begin{aligned} \text{Covariance : } \text{cov}(X, Y) &= E((X - \mu_X)(Y - \mu_Y)) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

$$\begin{aligned} \text{Correlation : } \rho_{XY} &= \frac{\text{Cov}(X, Y)}{\sqrt{\text{var}(X) \text{var}(Y)}} \\ \text{Coefft} & \end{aligned}$$

## Few things to note

---

1) If either  $X$  or  $Y$  has zero mean  
 $E(XY) = \text{Cov}(X, Y)$

2) Random variables  $X$  and  $Y$  are uncorrelated if  
 $\text{Cov}(X, Y) = 0 \implies \rho_{X, Y} = 0$

3) If  $X$  and  $Y$  are independent,  
 $E(XY) = E(X)E(Y) \implies \text{Cov}(X, Y) = 0$

The Converse is not true i.e.,  
uncorrelated  $\not\Rightarrow$  Statistical independence  
 $X$  and  $Y$  are uncorrelated

Example : Consider 2 rvs  $X$  and  $Y$  with joint pmf populated below in the table

	$x_1 = -1$	$x_2 = 0$	$x_3 = 1$
$y_1 = 0$	0	$\frac{1}{3}$	0
$y_2 = 1$	$\frac{1}{3}$	0	$\frac{1}{3}$

$P_{XY}(0,1) = \text{Joint Prob} (X=0, Y=1) = 0$

Let us consider compute

$P_X(0)$

$P_Y(1)$

$P_{XY}(0,1) = 0$

$\neq \frac{1}{3}$

$P_X(0) P_Y(1) = \frac{2}{9}$

$= \frac{2}{9} = \sum_{x \in \{-1, 0, 1\}} P_{XY}(x, 1)$

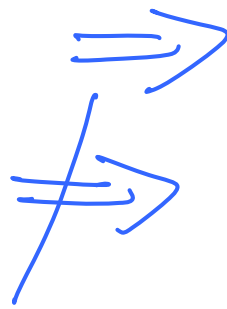
(They are not independent)

$$E(X) = -1 \times \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} = 0$$

$$E(XY) = \text{cov}(X, Y) = -1 \times 1 \cdot \frac{1}{3} + 1 \cdot 1 \cdot \frac{1}{3} + 0 \cdot 0 \cdot \frac{1}{3} = 0$$

$$\text{cov}(X, Y) = 0$$

So, uncorrelated



RVs  $X$  and  $Y$  are uncorrelated  
Statistical independence

## Orthogonal random variables

---

$E(xy) = 0 \implies$  rvs are orthogonal

Now,  $Cov(X, Y) = 0 \implies$  rvs are uncorrelated  
When either  $E(X)$  or  $E(Y)$  are zero and  $(X \text{ and } Y)$  are orthogonal

$$Cov(X, Y) = E(xy) - E(X)E(Y)$$

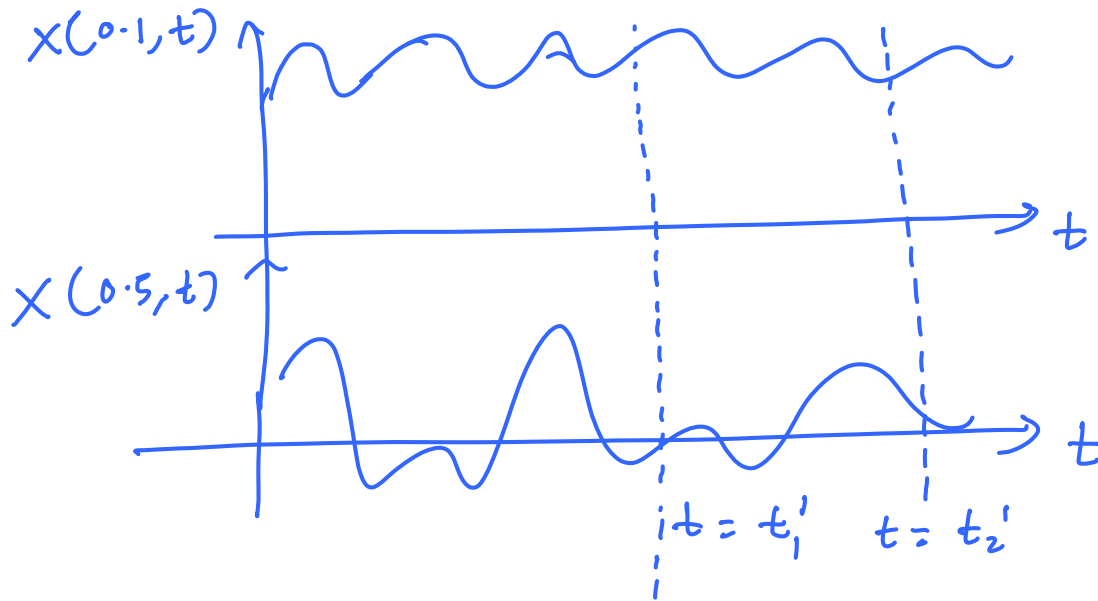
For zero mean rvs  $\overset{0}{=} \implies$  orthogonality  $\implies$  Uncorrelatedness

## Overview of basics of random processes

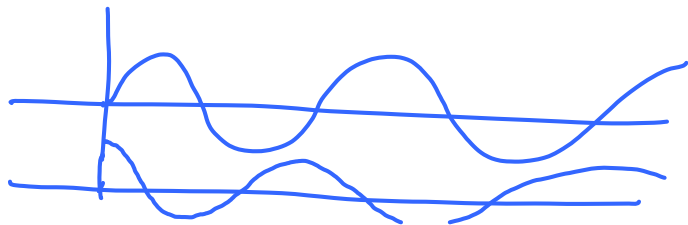
A random process  $x(t)$  (continuous/discrete) is a family of functions real/complex, scalar/vector defined on a probability space. At specified times  $t_1, t_2, \dots$  the samples  $x(t_1), x(t_2), \dots$  are random variables / random vectors.

Given  $(\Omega, \mathcal{F}, P)$   $X(t, \xi) \in \mathcal{F}$  for a fixed 't' on the real line  $-\infty < t < \infty$

When  $\xi$  is fixed,  $X(t, \xi)$  is an ordinary function (time function)  
when  $t$  is fixed,  $X(t, \xi)$  becomes a random variable



$$x(t, \phi) = A \sin(\omega t + \phi)$$



$$\xi = 0.1$$

$$\xi = 0.5$$

$$t = \pi/\omega; \quad A \sin(\pi + \phi)$$

$$0, -A, 0 \quad \underbrace{\quad}_{1/3} \quad 2/3$$

$$\phi \in U [0, 2\pi]$$

$$t = 0$$

$$\phi = 0, \pi/2, \pi$$

$$P_{(0)} = 2/3; \quad P_{(A)} = 1/3$$

$$A \sin(\phi) = \{ 0, A, 0 \}$$

$$\{ 0, A \}$$



## Mean & Correlations

The mean of a R.P. is  $E(x(t))$  and can be a function of the time index.

$$\mu_x(t) \triangleq E[x(t)] \quad -\infty < t < \infty$$

Auto correlation function

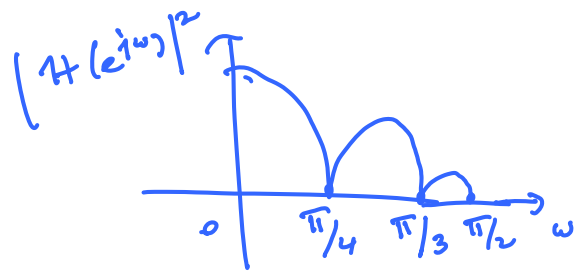
$$R_{xx}(t_1, t_2) = E[x(t_1)x^*(t_2)] \quad -\infty < t_1, t_2 < \infty$$

Eg.

$$x(t) = A e^{j2\pi ft}$$

Covariance function

$$\begin{aligned} \text{Cov}_{xx}(t_1, t_2) &= E\left[(x(t_1) - \mu_x(t_1))(x(t_2) - \mu_x(t_2))^*\right] \\ &= R_{xx}(t_1, t_2) - \mu_x(t_1)\mu_x^*(t_2) \end{aligned}$$



$$\begin{array}{cccccccc}
 1 & 1 & 2 & 1 & \dots & & & \\
 1 & -1 & 1 & -1 & & & & \\
 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \dots \\
 1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 \dots
 \end{array}$$

Example : (Sinusoidal random process)

Suppose  $X(t) = A \sin(\omega_0 t + \theta)$  where  $\theta \sim U[-\pi, \pi]$   
Suppose  $A$  is also a r.v. & let  $A$  &  $\theta$  be statistically independent.

$$\begin{aligned}\mu_x(t) &= E[A \sin(\omega_0 t + \theta)] \\ &= E[A] E[\sin(\omega_0 t + \theta)] \\ &= E[A] \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin(\omega_0 t + \theta) d\theta \\ &= E[A] \cdot 0 = 0\end{aligned}$$

## Auto correlation

$$\begin{aligned} R_{xx}(t_1, t_2) &= E(X(t_1) X^*(t_2)) \\ &= E(A^2 \sin(\omega_0 t_1 + \theta) \sin(\omega_0 t_2 + \theta)) \\ &= E(A^2) E(\sin(\omega_0 t_1 + \theta) \sin(\omega_0 t_2 + \theta)) \\ &= \frac{1}{2} E(A^2) \left[ \underbrace{\cos(\omega_0(t_1 - t_2))}_{\text{constant}} - \underbrace{\cos(2\theta + (t_1 + t_2)\omega_0)}_{\text{depends on } \theta} \right] \\ &= \frac{1}{2} E(A^2) \cdot \cos[\omega_0(t_1 - t_2)] \end{aligned}$$

that gets averaged to 0 over the cycle.

Auto correlation depends on the time lag!

## Statistical Specification of a random sequence

A random sequence  $X(n)$  is said to be statistically 'specified' by knowing the  $N^{\text{th}}$  order prob. dist. fns. for all integers  $N \geq 1$  & times  $n, n+1, \dots, n+N-1$

$$F_X(x_n, x_{n+1}, \dots, x_{n+N-1}; n, n+1, \dots, n+N-1) \\ = \Pr(X[n] \leq x_n, X[n+1] \leq x_{n+1}, \dots, X[n+N-1] \leq x_{n+N-1})$$

The representation we had is an infinite set of PDFs for each  $N$  because for all times  $-\infty < n < \infty$ , we need to know the joint PDF / CDF

$$\begin{aligned}\mu_x[n] = E\{x[n]\} &= \int_{-\infty}^{\infty} x f_x(x; n) dx && \text{Cont. case} \\ &= \sum_{k=-\infty}^{\infty} x_k P[X[n] = x_k]\end{aligned}$$

## Some extensions/classifications of R.P.

Let  $X$  and  $Y$  be R.P. They are

a) uncorrelated:  $R_{xy}(t_1, t_2) = \mu_x(t_1) \mu_y^*(t_2)$

where  $R_{xy}(t_1, t_2) \triangleq E[X(t_1) Y^*(t_2)]$

b) Orthogonal:  $R_{xy}(t_1, t_2) = 0 \quad \forall t_1, t_2$

c) Independent: If for +ve integers  $n$ , the  $n^{\text{th}}$  order PDF of  $X$  and  $Y$  factors i.e.,

$$F_{xy}(x_1, y_1, x_2, y_2, \dots, x_n, y_n; t_1, t_2, \dots, t_n) \\ = F_x(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) F_y(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n)$$

## Stationarity

A R. P.  $X(t)$  is stationary if it has the same  $n^{\text{th}}$  order prob. dist. fns. as  $X(t+T)$

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = F_X(x_1, x_2, \dots, x_n; t_1+T, \dots, t_n+T)$$

If the PDF is differentiable,

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_X(x_1, x_2, \dots, x_n; t_1+T, \dots, t_n+T)$$



## Properties

1) Mean of a stationary process is a 'constant'

$$f(x; t) = f(x; t+T)$$

$$\Rightarrow f(x; t) = f(x; t+T) \Big|_{t=-T} = f(x; 0)$$

$$E[X(t)] = \mu_x(0)$$

2) 2<sup>nd</sup> order density is shift invariant

$$f(x_1, x_2; t_1, t_2) = f(x_1, x_2; t_1+T, t_2+T)$$

Upon choosing  $T = -t_2$

$$f(x_1, x_2; t_1, t_2) = f(x_1, x_2; t_1, -t_2, 0)$$

$$E(X(t_1)X^*(t_2)) = R_{XX}(t_1-t_2, 0)$$

Weak form of stationarity (Wide Sense Stationary)

Defn: A R.P. is WSS

if  $E[X(t)] = \mu_x = \text{a constant}$

$$E[X(t+\tau) X^*(t)] = R_{xx}(\tau) \quad -\infty < \tau < \infty$$

Mean is a const.

Auto corr. depends only on the time lag



Example : Suppose  $x(t) = A e^{j2\pi f t}$  ;  $f$  is known  
(real const.)

$A$  is a real valued r.v. with  $E(A) = 0$   $E(A^2) < \infty$

$$E[X(t)] = E(A e^{j2\pi f t}) = 0 \quad \checkmark \quad (\text{mean is a const.})$$

$$E[X(t+\tau) X^*(t)] = E[A e^{j2\pi f (t+\tau)} A e^{-j2\pi f t}]$$

$$= E(A^2) e^{j2\pi f \tau}$$

$$= R_{xx}(\tau) \quad \checkmark \quad (\text{depends only on lag } \tau)$$

$X(t)$  is WSS

## Exercises

- 1) If  $X(t) = \sum_{k=1}^M A_k e^{j2\pi f_k t}$   
 $E(A_k) = 0$   $\xi$   $A_k$ 's are uncorrelated  
Examine if  $X(t)$  is a W.S.S.

- 2) Prove the following:

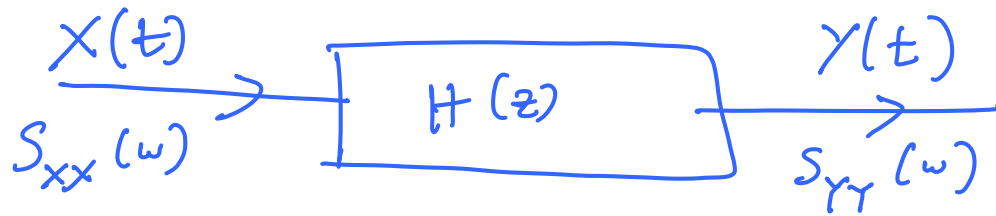
1)  $|R_{xx}(\tau)| \leq R_{xx}(0)$

2)  $|R_{xy}(\tau)| \leq \sqrt{R_{xx}(0) R_{yy}(0)}$

4)  $\forall N > 0 \quad \forall t_1 < t_2 \dots < t_N \quad \xi \text{ complex } a_1, a_2, \dots, a_N$   
 $\sum_{k=1}^N \sum_{l=1}^N a_k a_l^* R_{xx}(t_k - t_l) > 0$

3)  $R_{xx}(\tau) = R_{xx}^*(-\tau)$  complex  
 $R_{xx}(\tau) = R_{xx}(-\tau)$  real

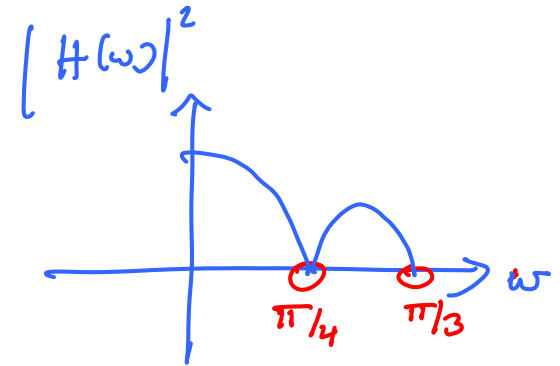
Filtering a WSS process through an LTI system



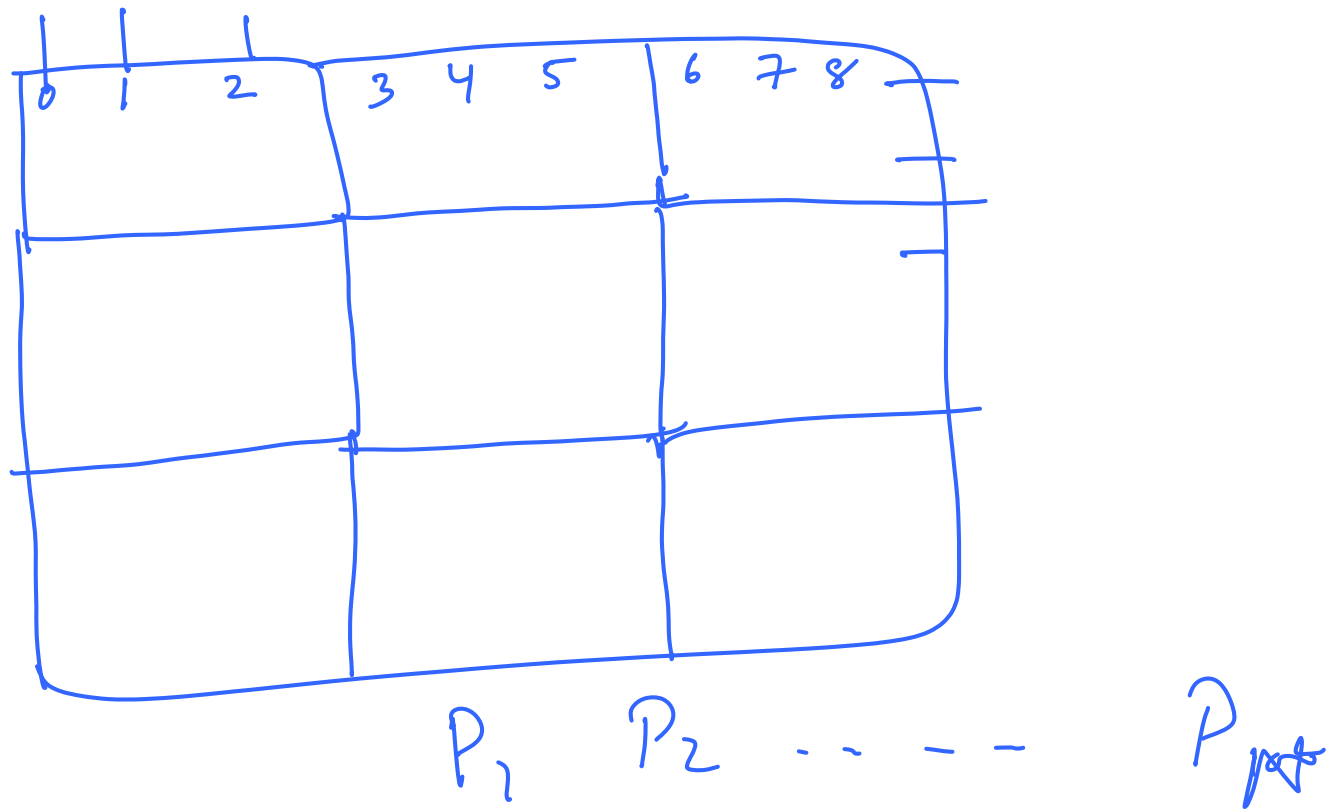
$$S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega)$$

where  $S_{YY}(\omega) \stackrel{\text{IID}}{=} \int_{-\infty}^{\infty} R_{YY}(\tau) e^{-j\omega\tau} d\tau$

$$S_{XX}(\omega) \stackrel{\text{IID}}{=} \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-j\omega\tau} d\tau$$



The role of modulation codes is crucial to examine  $S_{XX}(\omega)$  so that it does not coincide with a spectral null of  $|H(\omega)|^2$



$\{P_i\}$