03 September 2017 21:12

Probability, Random Variables and Random Processes

Probability Space $(\Omega, \mathbb{F}, P)$

$\Omega \rightarrow$ sample space
(set of all outcomes of random experiment)

$\mathbb{F} \rightarrow$ Collection of subsets of Ω
(set of all events)

$P \rightarrow$ Probability measure assigned to every entry in $\mathbb{F}$

Valid Prob. measure satisfies 3 axioms

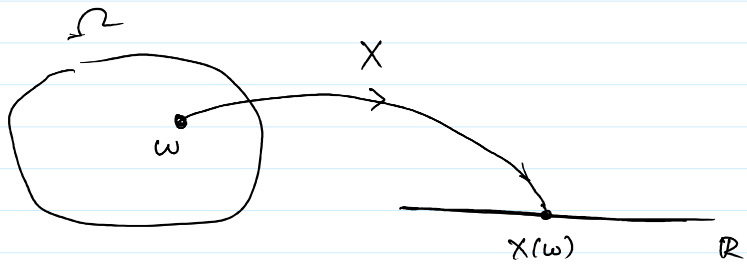
① $P(\Omega) = 1$

② For any $A \in \mathbb{F}$,
$P(A) \geq 0$

③ If $A \cap B = \phi$
(A & B are disjoint events)
then $P(A \cup B) = P(A) + P(B)$

Random Variable is a mapping from sample space Ω

to a real line $\mathbb{R}$



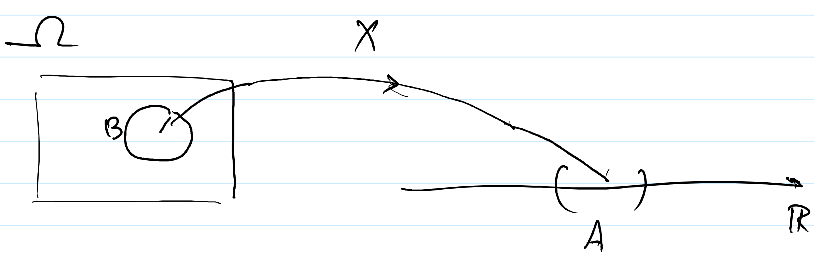
Random Variable defines/induces a probability space in $\mathbb{R}$

$(\mathbb{R}, \mathbb{F}, P_X)$

$\mathbb{R} \rightarrow$ set of all real numbers ($-\infty$ to ∞)

$\mathbb{F} \rightarrow$ collection of subsets of $\mathbb{R}$ (Borel sets)

$P_X \rightarrow$ Prob. measure on $\mathbb{F}$ induced by X



Let $A \subset \mathbb{R}$ & $A \in \mathbb{F}$ (Borel set)

Let B be <u>inverse image</u> of A in Ω

$$B = \{ \omega : X(\omega) \in A \}$$

Now $\quad P_X(A) = P(B)$

Probability induced by random variable X

Prob. in original space

$$X \longrightarrow x$$

A random variable is characterized completely (statistically) by its cumulative distribution function (cdf)

cdf of X denoted by F_X is defined as

$$F_X(x) = P_r\{X \le x\}$$

Probability of <u>any event</u> related to RV X can be obtained using the cdf.

Probability density function (pdf) of random variable X is

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

(if the derivative exists)

Notes :

- $F_X(\infty) = P(X \leq \infty) = 1$
- $F_X(-\infty) = P(X \leq -\infty) = 0$
- $F_X(x)$ is non decreasing function
- $P(a < X \leq b) = F_X(b) - F_X(a)$

$$= \int_a^b f_X(x)\, dx$$

$$X \longrightarrow x$$

Expectation

Mean Value of r.v. X

(denoted by μ_X)

or expected value of X

(denoted by $E(X)$)

is defined as

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

Mean Squared Value is

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx$$

Variance of r.v. X is defined as

$$Var(X) = E(X^2) - \mu_X^2$$

Standard deviation of X is

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

Let $g(\cdot)$ be some function

$$\underset{\Downarrow\ \text{r.v.}}{X} \longrightarrow \boxed{g(\cdot)} \longrightarrow \underset{\Downarrow\ \text{r.v.}}{g(X)}$$

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

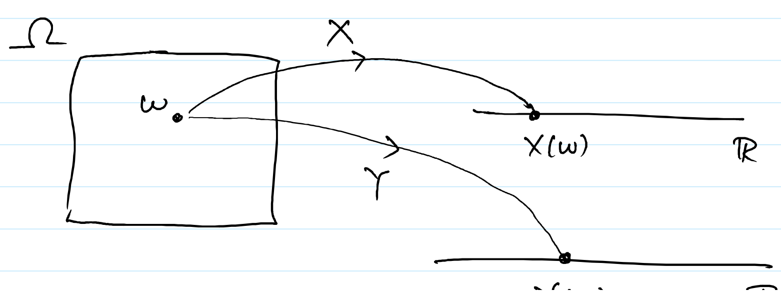
Easy to verify that (a, b are constants)

① $E(aX+b) = a\,E(X) + b$

② $Var(X) = E\{(X-\mu_X)^2\}$

③ $Var(aX+b) = a^2\,Var(X)$

X and Y be two random variables defined on a common probability space



$Y(\omega)$ $\quad$ *

Two r.v. X and Y are completely characterized by joint cdf

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

"," denotes "AND"

Joint pdf is

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

We have

$$P(a < X \le b, c < Y \le d)$$

$$= \int_{x=a}^{b} \int_{y=c}^{d} f_{XY}(x,y)\, dx\, dy$$

From joint cdf, we get marginal cdf as

$$F_X(x) = P(X \le x)$$

$$= P(X \le x, Y \le \infty)$$

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

$$= \int_{t=-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(t,y)\, dt\, dy$$

$$f_X(x) = \frac{d}{}\left(\qquad\qquad \right)$$

$$f_X(x) = \frac{d}{dx}\left(\qquad \right)$$

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

Similarly

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

Covariance of two ~~random variables~~

defined as

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

$$= E(XY) - \mu_X \mu_Y$$

Two random variables are called <u>uncorrelated</u> if their covariance is zero

ie) $E(XY) = E(X)\,E(Y)$

then X and Y are uncorrelated

(X & Y)
Two random variables are called <u>independent</u> if

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

or equivalently

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

Theorem:

If X and Y are independent

then X and Y are uncorrelated

Proof:

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

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f_X(x) f_Y(y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x\, f_X(x)\, dx \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

$$= E(X)\, E(Y)$$

uncorrelated does not imply independence

(except under one special case)

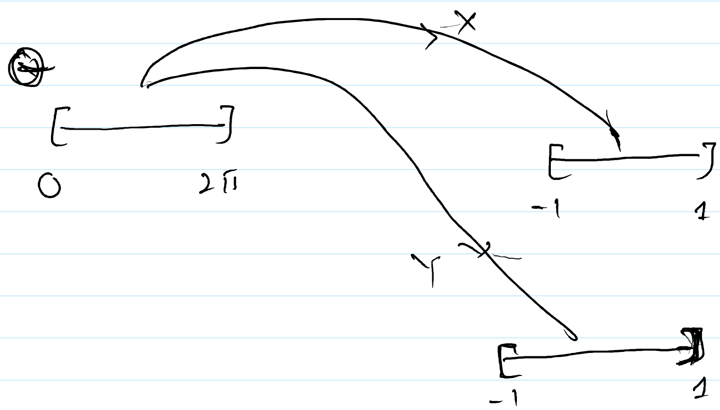
Example:

Let Θ be a r.v. uniformly distributed in $[0, 2\pi]$

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi} & \text{if } \theta \in [0, 2\pi] \\ 0 & \text{else} \end{cases}$$

Let $X = \cos\Theta$

$Y = \sin\Theta$



X and Y are clearly <u>dependent</u>

$$X^2 + Y^2 = 1$$

Now, $E(X) = E(\cos\Theta)$

$$= \int_{-\infty}^{\infty} \cos\theta \, f_\Theta(\theta) \, d\theta$$

$$= \int_0^{2\pi} \cos\theta \, \frac{1}{2\pi} \, d\theta$$

$$= 0$$

Similarly $E(Y) = 0$

Now, $E(XY) = \int_{-\infty}^{\infty} \cos\theta \sin\theta \, f_\theta(\theta) d\theta$

$$= \int_0^{2\pi} \frac{1}{2} \sin 2\theta \, \frac{1}{2\pi} \, d\theta$$

$$= 0$$

$$E(XY) = E(X)\,E(Y)$$

X and Y are uncorrelated

Random vector $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}_{p\times 1}$

$X_1, X_2, \ldots, X_p$ are random variables

$\underline{X}$ completely characterized by cdf

$$F_{\underline{X}}(x_1, x_2, \ldots, x_p) = P\{X_1 \le x_1, \ldots, X_p \le x_p\}$$

$\downarrow$

joint cdf of entries in $\underline{X}$

Similarly pdf of $\underline{X}$ $f_{\underline{X}}$

Similarly pdf of $\underline{X}$ is $f_{\underline{X}}$

is the joint pdf of ~~its~~ entries

Mean vector $\underset{p\times 1}{\underline{\mu}_X} = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$

Covariance Matrix

$$C_X = E\{(\underline{X} - \underline{\mu}_X)(\underline{X} - \underline{\mu}_X)^T\}$$

$p \times p$ matrix

$$C_X = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ \vdots & & & \vdots \\ C_{p1} & C_{p2} & \cdots & C_{pp} \end{bmatrix}$$

$$c_{ij} = E\{(X_i - \mu_i)(X_j - \mu_j)\}$$

$i = j \Rightarrow c_{ii} \rightarrow \mathrm{Var}(X_i)$

$i \neq j \Rightarrow c_{ij} \rightarrow \mathrm{Cov}(X_i, X_j)$

Theorem

Let $A_{M \times p}$ constant matrix

$\underline{b}_{p \times 1}$ constant vector

Let $\underset{(M \times 1)}{\underline{Y}} = A \underline{X} + \underline{b}$

$\underline{X}$ is random vector

with Mean vector $\underline{\mu}_X$

& covariance matrix C_X

Then Mean vector of $\underline{Y} = E(\underline{Y})$

$$\underline{\mu}_Y = A \underline{\mu}_X + \underline{b}$$

$$= A\, E(\underline{X}) + \underline{b}$$

Covariance matrix of $\underline{Y}$

$$C_{\underline{Y}} = E\left((\underline{Y} - \underline{\mu}_Y)(\underline{Y} - \underline{\mu}_Y)^T \right)$$

(MxM) matrix $= A\, C_X\, A^T$

<u>Proof:</u>

$\underline{Y} = A \underline{X} + \underline{b}$

Proof:

$$\underline{Y} = A\underline{X} + \underline{b}$$

$$\underline{\mu}_Y = A\underline{\mu}_X + \underline{b}$$

$$\boxed{(CD)^T = D^T C^T}$$

$$(\underline{Y} - \underline{\mu}_Y) = A(\underline{X} - \underline{\mu}_X)$$

$$(\underline{Y} - \underline{\mu}_Y)(\underline{Y} - \underline{\mu}_Y)^T = A(\underline{X} - \underline{\mu}_X)(\underline{X} - \underline{\mu}_X)^T A^T$$

Taking Expectation

$$E(\quad) = E(\quad)$$

(Since A is constant) $= A \; E\left\{ (\underline{X} - \underline{\mu}_X)(\underline{X} - \underline{\mu}_X)^T \right\} A^T$

$$= A\, C_X\, A^T$$

Random Process

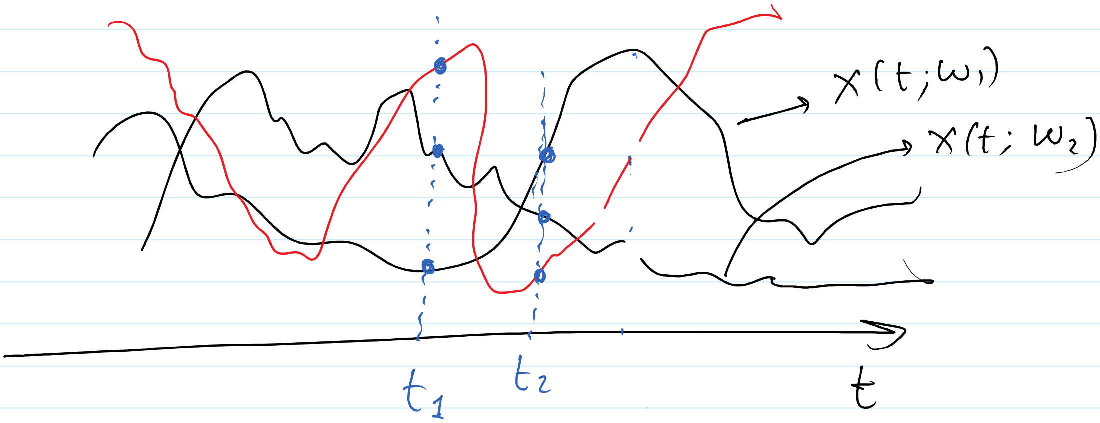
Let $\omega_1, \omega_2, \ldots$ be outcomes

of random experiment

$\omega_i \in \Omega$ (Sample Space)

For each outcome ω_i, we

assign a waveform $x(t;\omega_3)$



The set of waveforms

$\{ x(t;\omega), \omega \in \Omega \}$ is

called a random process if

$x(t;\omega) \Big|_{t=t_i}$ waveforms sampled at $t = t_i$

is a random variable for any t_i

We usually write $\{X(t)\}$ is a random process without explicitly denoting ω

→ $\{X(t)\}$ is an infinite (uncountable) collection of random variables.

→ R.P. $\{X(t)\}$ is completely specified if joint cdf/pdf of all the possible (finite number) of samples of $\{X(t)\}$ is specified

ie) joint cdf of samples of RP

$$\{X(t_1), X(t_2), \dots, X(t_N)\}$$

is specified for all possible values of $t_1, t_2, \dots, t_N$

and all possible integers N

x ———————— x

A R.P. $\{X(t)\}$ is called Stationary

if statistics of $\{X(t)\}$ &

$\{X(t-d)\}$ are identical
$\quad\downarrow$
delay

for any value of delay d.

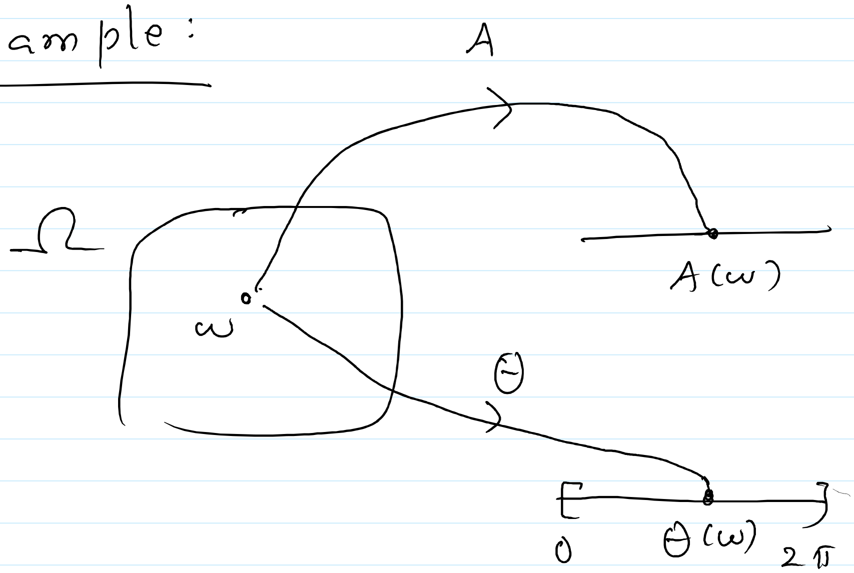
that is, joint cdf of

$\{X(t_1), X(t_2), \dots, X(t_N)\}$ is identical

to that of $\{X(t_1-d), \dots, X(t_N-d)\}$

for all possible t_i, N, d.

Example:



Let

$$X(t,\omega) = A(\omega)\cos(2\pi f_c t + \Theta(\omega))$$

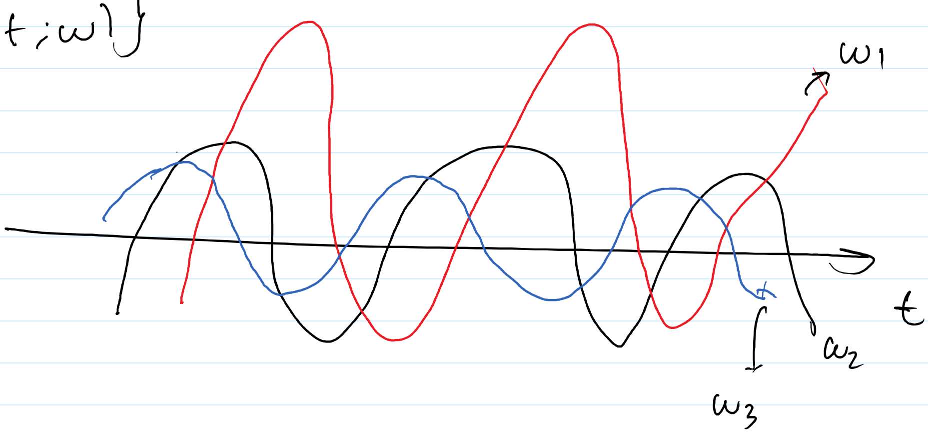
$A \rightarrow$ random variable amplitude

$\Theta \rightarrow$ random variable phase

Let Θ be uniformly distributed in interval $[0 \text{———} 2\pi]$

$$f_{\Theta}(\theta) = \begin{cases} \frac{1}{2\pi}, & 0 \le \theta < 2\pi \\ 0, & \text{else} \end{cases}$$

$\{X(t;\omega)\}$



$$X(t-d) = A\cos(2\pi f_c(t-d) + \Theta)$$

$$= A\cos(2\pi f_c t + \tilde{\Theta})$$

$$\tilde{\Theta} = [\Theta - 2\pi f_c d] \bmod 2\pi$$

Easy to verify that $\tilde{\Theta}$ is uniform

$$f_{\tilde{\Theta}}(a) = \begin{cases} \frac{1}{2\pi} & 0 \le a \le 2\pi \\ 0 & \text{else} \end{cases}$$

$\{X(t)\}$ & $\{X(t-d)\}$ have same statistics.

$\{X(t)\}$ is stationary Random Process

x ——————————— x

Definitions :

Mean function $\mu_X(t)$ of RP

$$\mu_X(t) = E(X(t)) \qquad -\infty < t < \infty$$

↓

as a function of time

Auto Correlation function (ACF)

$$R_x(t;t-\tau) = E\{x(t)\, x(t-\tau)\}$$

$-\infty < t < \infty$

$-\infty < \tau < \infty$

$\tau \rightarrow$ lag (delay) Variable

A process is called wide sense stationary (WSS)

if

① $\mu_x(t) = \mu$ for all t

Mean function is a constant (dc) function

② $R_x(t;t-\tau)$ depends only on τ and not on exact value of t

For WSS, we denote ACF as $R_x(\tau) = E\{x(t)\, x(t-\tau)\}$

Theorem:

A stationary process is also WSS.

But WSS is not stationary (with one exception)

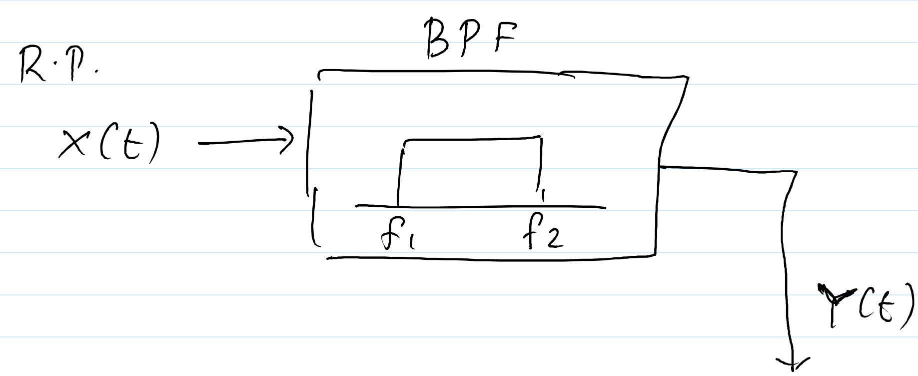
Power Spectral density

PSD of a r.p. $\{X(t)\}$

gives average power in each frequency

$S_X(f) \rightarrow$ PSD of $X(t)$

↳ unit watts / Hz



$\lim \frac{1}{\ } \int^{T_0} |Y(t)|^2 \, dt$ $\boxed{\text{Power}}$

$$\lim_{T_0 \to \infty} \frac{1}{2T_0} \int_{-T_0}^{T_0} |x(t)|^2 \, dt$$

$$\boxed{\text{Power Meter}} \downarrow \boxed{E(\cdot)} \downarrow$$

(Average over all possible outcomes

$$\int_{f_1}^{f_2} S_x(f) \, df$$

Wiener - Khinchine theorem

For a WSS process with ACF $R_x(\tau)$

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} \, d\tau$$

PSD = Fourier Transform of ACF

$x \longrightarrow$