Digital Communication

Youlong Wu

ShanghaiTech University

wuyl1@shanghaitech.edu.cn

Probability and Stochastic Process

Probability and Stochastic Process

Objectives:

- Know some useful probability distributions
- Can calculate $E(g(X))$ and $\mathsf{Var}(g(X))$

Probability

Given events A, B

- $P(A), P(\bar{A})$,
- $P(A \cap B)$ or $P(A, B)$, $P(A \cup B)$
- $P(A|B) = P(A \cap B)/P(B) = P(A, B)/P(B)$

Two important properties

- $P(A \cup B) = P(A) + P(B)$, **only if** $A \cap B = \emptyset$
- $P(A|B) = P(A)$, **only if** A and B are *independent*

Questions:

- $A \cap B = \emptyset$ v.s. $A \perp B$
- For $P(A|B)$ and $P(A)$, which one is larger?
  (hint: check when $A \cap B = \emptyset$ and $A \cap B = B$)

Discrete Random variable

Given a discrete random variable X:

- Probability :

$$p(x), x \in \mathcal{X}$$

- Cumulative distribution function (CDF):

$$F(x) = P(X \leq x) = \sum\nolimits_{X \leq x} p(x)$$

- Expectation:

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i), \quad E(g(X)) = \sum_{i=1}^{\infty} g(x_i) p(x_i)$$

- Variance

$$\mathsf{Var}(X) = E(X^2) - E^2(X), \quad \mathsf{Var}(g(X)) = E(g(X)^2) - E^2(g(X))$$

Continuous Random variable

Given a continuous random variable X:

- Cumulative distribution function (CDF):

$$F(x) = P(X \leq x), -\infty \leq x \leq +\infty$$

- Probability density function (PDF):

$$f(x) = \frac{dF(x)}{dx} \Longrightarrow F(X) = \int_{-\infty}^{x} f(u)du$$

- Expectation (e.g, $g(X) = X$):

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

- Variance

$$\mathsf{Var}(X) = E(X^2) - E^2(X)$$

Function of Random Variable

Example 1: Find PDF of

$$Y = aX + b$$

where $a > 0$ and $b > 0$ are constant.

Solution:

$$F_Y(y) = P(aX + b \leq y) = P(x \leq \frac{y-b}{a})$$
$$= \int_{-\infty}^{\frac{y-b}{a}} f_X(x)dx$$

By differentiation $\frac{\partial F_Y(y)}{\partial y}$, we obtain

$$\boxed{f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)}$$

Function of Random Variable

Example 2: Find PDF of

$$Y = aX^2 + b$$

where $a > 0$ and $b > 0$ are constant.

Solution:

Function of Random Variable

Example 2: Find PDF of

$$Y = aX^2 + b$$

where $a > 0$ and $b > 0$ are constant.

Solution:

$$F_Y(y) = P(aX^2 + b \leq y) = P\left(|x| \leq \sqrt{\frac{y-b}{a}}\right)$$

$$= \int_{-\sqrt{\frac{y-b}{a}}}^{\sqrt{\frac{y-b}{a}}} f_X(x)dx$$

By differentiation $\frac{\partial F_Y(y)}{\partial y}$, we obtain

$$\boxed{f_Y(y) = \frac{f_X(\sqrt{y-b/a})}{2a\sqrt{y-b/a}} + \frac{f_X(-\sqrt{y-b/a})}{2a\sqrt{y-b/a}}}$$

Conditional Expectation

- $E(X|Y=y)$

$$E(X|Y=y) = \int_x xf(x|y)dx \tag{1}$$

- $E(X|Y)$ a function of RV Y, i.e., $g(Y) = E(X|Y)$
- $E(E(X|Y)) = E(X)$

$$\begin{aligned} E(E(X|Y)) &= \int_x f(y)E(X|Y=y)dy \\ &= \int_x f(y) \int_y \frac{f(x,y)}{f(y)} xdxdy \\ &= \int_x x \int_y f(x,y)dydx \\ &= \int_x f(x)xdx \end{aligned}$$

Useful Probability Distributions

Binomial distribution: (e.g. tossing coin)

Given X, where $P(X = 1) = p$ and $P(X = 0) = 1 - p$. Now

$$Y = \sum_{i=1}^{n} X_i$$

where $X_i, i = 1, 2, \ldots, n$ are statistically independent and identically distributed (i.i.d), with pdf same as X. Then,

$$\boxed{P(Y = k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{(n-k)}}$$

where $\left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!}$.

$$\boxed{E(Y) = np, \quad E(Y^2) = np(1-p) + n^2p^2}$$

Uniform Distribution

Given a RV X, with $a \leq X \leq b$. Its PDF follows

$$\boxed{f_X(x) = \frac{1}{b-a}}$$

And,

$$\boxed{E(X) = \frac{a+b}{2}, \quad \mathsf{Var}(X) = \frac{(a-b)^2}{12}}$$

Gaussian (Normal) Distribution

Given a RV $X \sim \mathcal{N}(u, \sigma^2)$, its PDF follows

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

And,

$$\boxed{E(X) = \mu, \quad \text{Var}(X) = \sigma^2}$$

Note:
The sum of n statistically independent Gaussian RVs is also Gaussian RV.

Question: $Y = 3X_1 + 2X_2 + X_3$, with X_i i.i.d $\sim \mathcal{N}(0, 1)$, Y's distribution is?

Q function: (not a distribution!)

$$\boxed{Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt}$$

Properties of $Q(x)$:

$$Q(0) = \frac{1}{2}, \quad Q(-x) = 1 - Q(x),$$
$$Q(\infty) = 0, \quad Q(-\infty) = 1$$

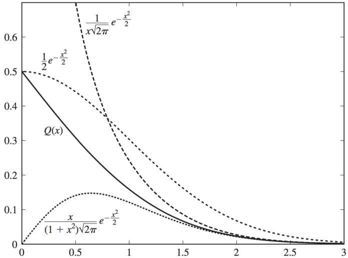
Useful bounds for $Q(x)$: if $x > 0$

$$Q(x) < \frac{1}{2} e^{-\frac{x^2}{2}}$$

$$Q(X) < \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$Q(X) > \frac{x}{(1+x^2)\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$ when x is large;



Q function

$$\boxed{Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt}$$

CDF of Gaussian distribution: $F_X(x) = P(X \le x)$

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = 1 - \int_x^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \\ &= 1 - \int_{\frac{x-\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{\mathrm{u}^2}{2}} d\mathrm{u} = 1 - Q(\frac{x-\mu}{\sigma}) \end{aligned}$$

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$\boxed{P(X > a) = Q(\frac{a-\mu}{\sigma}), \quad P(X < a) = Q(\frac{\mu - a}{\sigma})}$$

Usefulness: Given μ and σ, we obtain $P(X < a)$ directly by checking table of Q function values.

Complementary Error Function (erfc)

x	$Q(x)$	x	$Q(x)$	x	$Q(x)$	x	$Q(x)$
0	0.500000	1.8	0.035930	3.6	0.000159	5.4	3.3320×10^{-8}
0.1	0.460170	1.9	0.028717	3.7	0.000108	5.5	1.8990×10^{-8}
0.2	0.420740	2	0.022750	3.8	7.2348×10^{-5}	5.6	1.0718×10^{-8}
0.3	0.382090	2.1	0.017864	3.9	4.8096×10^{-5}	5.7	5.9904×10^{-9}
0.4	0.344580	2.2	0.013903	4	3.1671×10^{-5}	5.8	3.3157×10^{-9}
0.5	0.308540	2.3	0.010724	4.1	2.0658×10^{-5}	5.9	1.8175×10^{-9}
0.6	0.274250	2.4	0.008198	4.2	1.3346×10^{-5}	6	9.8659×10^{-10}
0.7	0.241960	2.5	0.006210	4.3	8.5399×10^{-6}	6.1	5.3034×10^{-10}
0.8	0.211860	2.6	0.004661	4.4	5.4125×10^{-6}	6.2	2.8232×10^{-10}
0.9	0.184060	2.7	0.003467	4.5	3.3977×10^{-6}	6.3	1.4882×10^{-10}
1	0.158660	2.8	0.002555	4.6	2.1125×10^{-6}	6.4	7.7689×10^{-11}
1.1	0.135670	2.9	0.001866	4.7	1.3008×10^{-6}	6.5	4.0160×10^{-11}
1.2	0.115070	3	0.001350	4.8	7.9333×10^{-7}	6.6	2.0558×10^{-11}
1.3	0.096800	3.1	0.000968	4.9	4.7918×10^{-7}	6.7	1.0421×10^{-11}
1.4	0.080757	3.2	0.000687	5	2.8665×10^{-7}	6.8	5.2309×10^{-12}
1.5	0.066807	3.3	0.000483	5.1	1.6983×10^{-7}	6.9	2.6001×10^{-12}
1.6	0.054799	3.4	0.000337	5.2	9.9644×10^{-8}	7	1.2799×10^{-12}
1.7	0.044565	3.5	0.000233	5.3	5.7901×10^{-8}	7.1	6.2378×10^{-13}

$$\boxed{\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt}$$

Relation between $\mathrm{erfc}(x)$ and $Q(x)$

$$Q(x) = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right), \quad \mathrm{erfc}(x) = 2Q(\sqrt{2}x)$$

χ^2 Distribution

Let $X \sim \mathcal{N}(\mu, \sigma^2)$, and

$$Y = X^2$$

Then, the Y follows χ^2 distribution:

$$\boxed{f_Y(y) = \frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{y^2}{2\sigma^2}}, \quad y \geq 0}$$

Let $X_1, X_2, \ldots, X_n$ are i.i.d.$\mathcal{N}(\mu, \sigma^2)$, and

$$Y = \sum_{i=1}^{n} X_i^2$$

Then, the Y is a χ^2 RV with n degree of freedom. see [Proakis'45]

Rayleigh Distribution

Let X_1 and X_2 be i.i.d $\mathcal{N}(0, \sigma^2)$, and

$$Y = X_1^2 + X_2^2$$

Then, the PDF of

$$R = \sqrt{X_1^2 + X_2^2}$$

follows Rayleigh distribution:

$$\boxed{f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, \quad r \geq 0}$$

Rice distribution

Let $X_1 \sim \mathcal{N}(\mu_1, \sigma^2)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma^2)$, $X_1 \perp X_2$ and

$$Y = X_1^2 + X_2^2$$

Then, the PDF of

$$R = \sqrt{X_1^2 + X_2^2}$$

follows Rayleigh distribution:

$$\boxed{f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2+s^2}{2\sigma^2}} I_0(rs/\sigma^2), \quad r \geq 0}$$

where $s^2 = \mu_1^2 + \mu_2^2$ and $I_\alpha(\cdot)$ is αth-order modified Bessel function [Proakis'B].

Jointly Gaussian Distribution

- Given an $n \times 1$ column vector $\boldsymbol{X}$, its pdf follows:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}(\text{det}\boldsymbol{C})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T \boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{m})}$$

  where

$$\begin{aligned} \boldsymbol{m} &= E[\boldsymbol{X}] \\ \boldsymbol{C} &= E[(\boldsymbol{X}-\boldsymbol{m})(\boldsymbol{X}-\boldsymbol{m})^t] \end{aligned}$$

  with $C_{i,j} = \text{cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)]$.

- $\text{cov}(\boldsymbol{X}, \boldsymbol{Y}) = \text{cov}(\boldsymbol{Y}, \boldsymbol{X})^T$
- $\text{cov}(A\boldsymbol{X} + \boldsymbol{b}, \boldsymbol{Y}) = A\text{cov}(\boldsymbol{X}, \boldsymbol{Y})$
- $\text{cov}(\boldsymbol{X} + \boldsymbol{Y}, \boldsymbol{Z}) = \text{cov}(\boldsymbol{X}, \boldsymbol{Y}) + \text{cov}(\boldsymbol{X}, \boldsymbol{Z})$

Jointly Gaussian Distribution: $n = 2$

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}(\text{det}\boldsymbol{C})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T \boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{m})}$$

- Special case: $n = 2$

$$\boldsymbol{m} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

where $\rho = \frac{\text{cov}[X_1, X_2]}{\sigma_1\sigma_2}$. Thus:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{(\frac{x_1-\mu_1}{\sigma_1})^2 + (\frac{x_2-\mu_2}{\sigma_2})^2 - 2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2})}{2(1-\rho^2)}}$$

- When $\rho = 0$

$$f(x_1, x_2) = \mathcal{N}(\mu_1, \sigma_1^2) \times \mathcal{N}(\mu_2, \sigma_2^2)$$

For jointly Gaussian RVs, uncorrelated = independent (In general?)

Jointly Gaussian Distribution: $n = 2$

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}(\text{det}\boldsymbol{C})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T \boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{m})}$$

- Special case: $n = 2$

$$\boldsymbol{m} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

where $\rho = \frac{\text{cov}[X_1, X_2]}{\sigma_1\sigma_2}$. Thus:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{(\frac{x_1-\mu_1}{\sigma_1})^2 + (\frac{x_2-\mu_2}{\sigma_2})^2 - 2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2})}{2(1-\rho^2)}}$$

- When $\rho = 0$

$$f(x_1, x_2) = \mathcal{N}(\mu_1, \sigma_1^2) \times \mathcal{N}(\mu_2, \sigma_2^2)$$

For jointly Gaussian RVs, uncorrelated $=$ independent (In general?)

In general: independent $\Rightarrow$uncorrelated; uncorrelated $\nRightarrow$independent

Example: Independent v.s. Uncorrelated

Let $f_{XY}(x,y) = \frac{1}{\pi}$, if $x^2 + y^2 \leq 1$; $f_{XY}(x,y) = 0$, otherwise. Find

- $f_X(x)$, $f_Y(y)$, $E(X)$, $E(X|y)$, $E(X,Y)$,
- Prove X and Y are uncorrelated, but not independent

Solution:

Example: Independent v.s. Uncorrelated

Let $f_{XY}(x,y) = \frac{1}{\pi}$, if $x^2 + y^2 \leq 1$; $f_{XY}(x,y) = 0$, otherwise. Find

- $f_X(x)$, $f_Y(y)$, $E(X)$, $E(X|y)$, $E(X,Y)$,
- Prove X and Y are uncorrelated, but not independent

Solution:

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} f_{XY}(x,y)dy = \frac{2}{\pi}\sqrt{1-x^2}$$

$$E[X] = \frac{2}{\pi}\int_{-1}^{1} x\sqrt{1-x^2}dy = 0$$

$$E[X|y] = \int_{-1}^{1} xf(x|y)dx = 0, E[X|Y] = E[X,Y] = 0$$

By symmetry: $E[Y|X] = E[Y] = 0$, Thus:
$E[XY] = E[X]E[Y] = 0$, X and Y are uncorrelated, but,

$$f_X(x)f_Y(y) = \frac{2}{\pi}\sqrt{1-x^2}\frac{2}{\pi}\sqrt{1-y^2} \neq f_{XY}(x,y)$$

Homework: Linear Combination of Gaussian RVs

Given $(X_1, X_2) \sim \mathcal{N}(0, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix})$, find the pdf of $Y = \frac{X_1+X_2}{2}$.

Properties of Jointly Gaussian RVs

- Linear combinations of jointly Gaussian RVs are also jointly Gaussian

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}$$

where invertible matrix $\boldsymbol{A}$ represents linear combination.

$$\boldsymbol{m_Y} = \boldsymbol{A}\boldsymbol{m_X}$$
$$\boldsymbol{C_Y} = \boldsymbol{A}\boldsymbol{C_X}\boldsymbol{A}^T$$

- See problem 2.23 in [Proakis'Book]

Properties of Jointly Gaussian RVs

- For jointly Gaussian RVs, uncorrelated = independent
- Linear combinations of jointly Gaussian RVs are also jointly Gaussian
- All subset and all conditional subsets of joint Gaussian RVs are jointly Gaussian

Random Process

- Random variable X v.s. Random Process $X(t)$
- Mean: $\mu_X(t) = E[X(t)]$
- Autocorrelation function: $R_X(t_1, t_2) = E[X(t_1)X^*(t_2)]$
- Cross-correlation function: $R_{XY}(t_1, t_2) = E[X(t_1)Y^*(t_2)]$

$$\begin{aligned} R_X(t_2, t_1) &= R_X^*(t_1, t_2) \\ R_{YX}(t_2, t_1) &= R_{XY}^*(t_1, t_2) \end{aligned}$$

Stationary Process

- Stationary Process: for all time shift τ, time duration k, sample time $t_1, \ldots, t_k$

$$F_X(x_{t_1+\tau,\ldots,x_{t_k+\tau}}) = F_X(x_{t_1,\ldots,x_{t_k}})$$

  CDF or PDF does not change when shift time. (Imply constant mean and variance if they exist)
- Wide-Sense Stationary (WSS) Process:
  1) $\mu_X(t)$ is constant;
  2) $R_X(t_1, t_2) = R(\tau)$, where $\tau = t_1 - t_2$
- Jointly WSS: $R_{XY}(t_1, t_2) = R_{XY}(\tau)$

Power and Energy

Given a signal (or a random process) $x(t)$

- energy (energy signal, $\mathcal{E}_X \nrightarrow \infty$):

$$\mathcal{E}_X = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad \text{(unit } J\text{)}$$

- (average) power (power signal, $\mathcal{P}_X \nrightarrow \infty$)::

$$P_X = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt \quad \text{(unit } W\text{)}$$

- Power spectral density (PSD) $\mathcal{S}_X(f)$: the distribution of power as a function of frequency, unit (W/Hz).

$$S_X(f) = \lim_{T\to\infty} E[|\hat{x}_T(f)|^2]$$

$\hat{x}_T(f) = \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x(t) e^{-i2\pi f t} dt$ is truncated Fourier transform of $x(t)$

Theorems

- Parseval's theorem:

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{x}(f)|^2 df$$

  Energy is constant from time and frequency perspectives.

- Wiener-Khinchin Theorem:

$$\mathcal{S}_X(f) = \mathcal{F}[R_X(\tau)], \quad \mathcal{S}_{XY}(f) = \mathcal{F}[R_{XY}(\tau)]$$

- Power of WSS $X(t)$ (sum of powers at all frequencies):

$$P_X = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt \overset{(a)}{=} E[|X(t)|^2] \overset{(b)}{=} R_X(0) \overset{(c)}{=} \int_{-\infty}^{\infty} \mathcal{S}_X(f) df$$

  (a) gives physical meaning; (b) is from definition of $R_X(\tau)$; (c) is from Wiener-Khinchin, and provides a convenient way to compute P_X

Proofs

- Prove Wiener-Khinchin Theorem: $\mathcal{S}_X(f) = \mathcal{F}[R_X(\tau)]$

$$
\begin{aligned}
S_X(f) &= \lim_{T\to\infty} \mathbf{E}\left[\left|\hat{x}_T(f)\right|^2\right] \\
&= \lim_{T\to\infty} E\left[\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x^*(t)e^{i2\pi ft}\,dt \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t')e^{-i2\pi ft'}\,dt'\right] \\
&= \lim_{T\to\infty} \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{E}\left[x^*(t)x(t')\right] e^{i2\pi f(t-t')}\,dt\,dt' \\
&= \cdots \\
&= \int_{-\infty}^{\infty} R_X(\tau)e^{iw\tau}d\tau \\
&= \mathcal{F}[R_X(\tau)]
\end{aligned}
$$

- If $X(t)$ is real WSS, $S_X(f)$ is real, nonnegative, and even function of f; For complex processes, not necessarily even.

Gaussian Random Process and White Process

Gaussian Random Process

- $X(t)$, for all n and t_i, jointly Gaussian vector

$$(X(t_1), \ldots, X(t_n))$$

- Jointly Gaussian Process: for all n, t_i, t_i', jointly Gaussian vector

$$(X(t_1), \ldots, X(t_n), Y(t_1'), \ldots, Y(t_n'))$$

- Complex Process: $Z(t) = X(t) + jY(t)$ is Gaussian if $X(t)$ and $Y(t)$ are joint Gaussian process

White Process: a process whose PSD is constant

$$S_X(f) = \frac{N_0}{2}$$

(Infinite power, not practical, but still very useful)

Markov Chain

Physical meaning: a random process of current value depends on the entire values only through the most recent values: jth-order Markov Chain

$$\begin{aligned}&P(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots)\\&\qquad = P(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_{n-j} = x_{n-j})\end{aligned}$$

Further look: $X - Y - Z$ forms Markov Chain, then $H(Z|Y, X) = H(Z|Y)$