I225E Statistical Signal Processing

3. Basics of Stochastic Process

MAWALIM and UNOKI

candylim@jaist.ac.jp and unoki@jaist.ac.jp

School of Information Science

JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY, HOKURIKU 1990

Stochastic processes

- Definitions
  Covariance, (auto- and cross-) correlations, correlation coefficients
- Stationarity (strict sense and wide sense)
  Normal processes
- Random walk
  (De Moivre-Laplace theorem, Stirling’s formula)
- Wiener process
- Ergodicity

1. Introduction

- **Random variable: $\boldsymbol{X}$**

  Trial S: Throw dice/coin toss

  Outcome ω: Throw dice/coin toss

  Random Variable: $\boldsymbol{X}(\omega) = \{1, 2, 3, 4, 5, 6\}$;

  $\boldsymbol{X}(\omega) = \{0, 1\} \rightarrow \boldsymbol{X}(\omega)$ corresponds to ω

- **Stochastic Process: $\boldsymbol{X}(t, \omega)$**

  Outcome ω: Results of all trials

  Time t; $(-\infty, \infty)$

  $\rightarrow$ Function of time $\boldsymbol{X}(t, \omega)$ corresponds to ω.

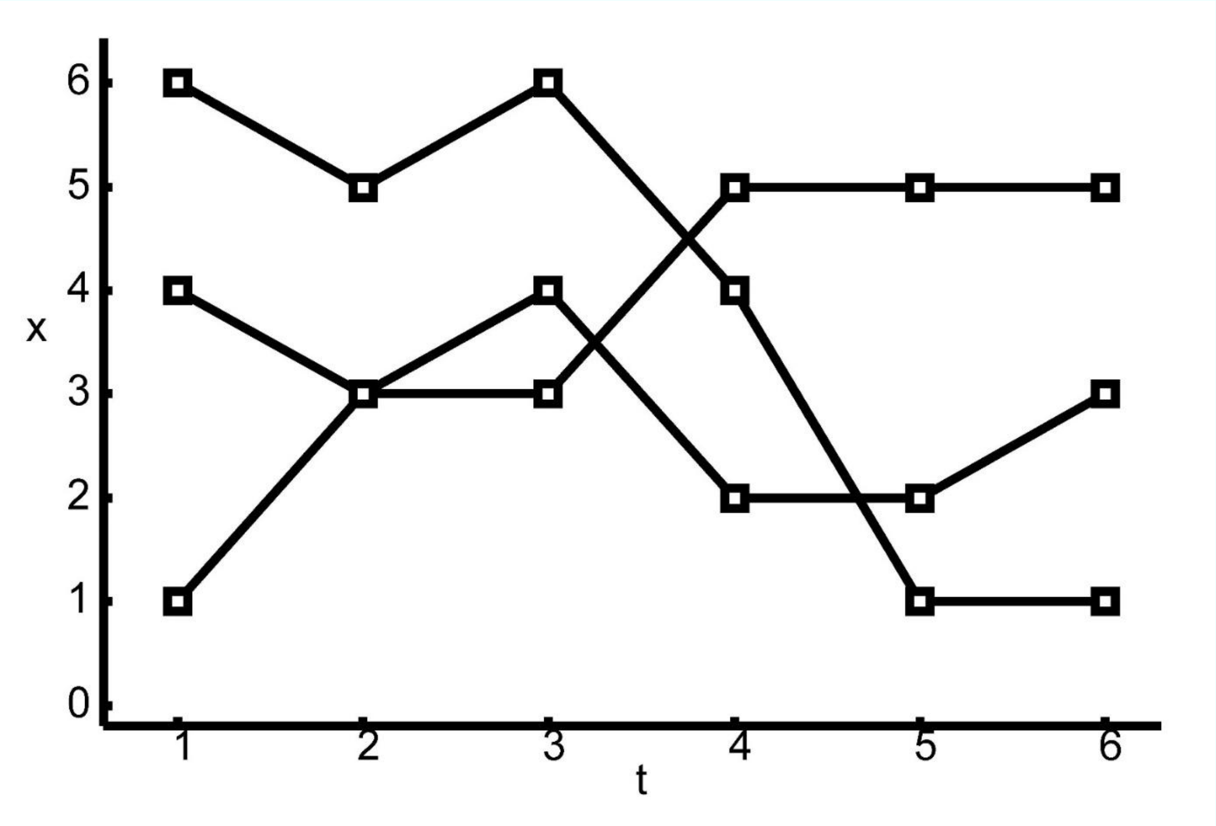
Stochastic process $\boldsymbol{X}(t, \omega)$ represents an ω-parameter family of functions of time.

- For a fixed value t

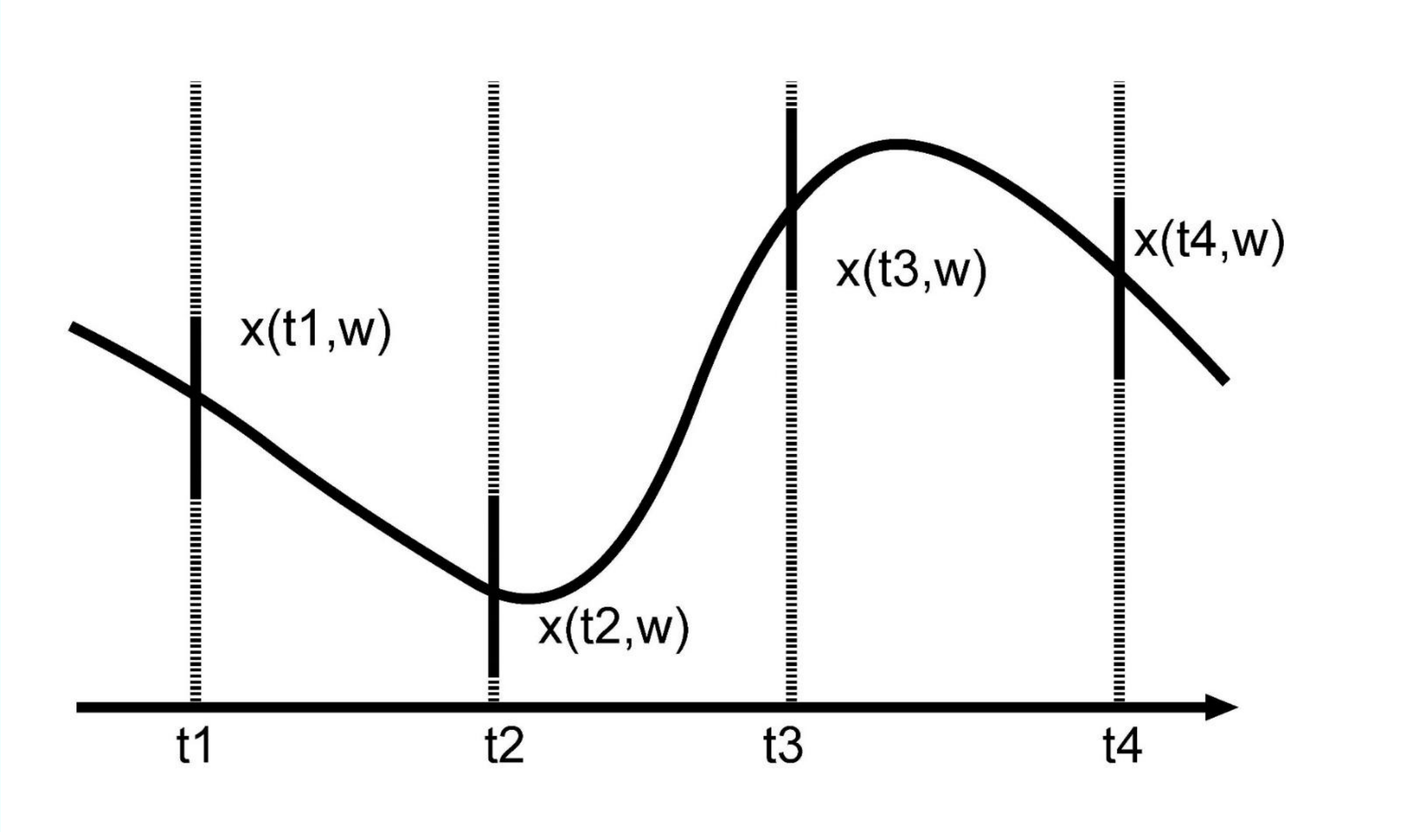
  $\boldsymbol{X}(t)$ is a random variable that corresponds to ω

- Discrete-type vs. Continuous-type
  - Discrete time: $t \in N$ (Integer number)
  - Continuous time: $t \in R$ (Real number)
  - Discrete state: $\boldsymbol{X} \in$ Countable number of state
  - Continuous state: $\boldsymbol{X} \in$ Uncountable number of state

- Example of discrete-time discrete-state process: Series of numbers obtained by throwing a dice for six times.



- Example of continuous-time continuous-state process: If t is fixed, $\boldsymbol{X}(t)$ represents stochastic variable.



2. Definition

Statistical quantities of stochastic process

Stochastic process is a set of uncountable number of random variables. For each t, $\boldsymbol{X}(t)$ represents a random variable.

For a fixed t,

Probability distribution of $\boldsymbol{X}(t)$:

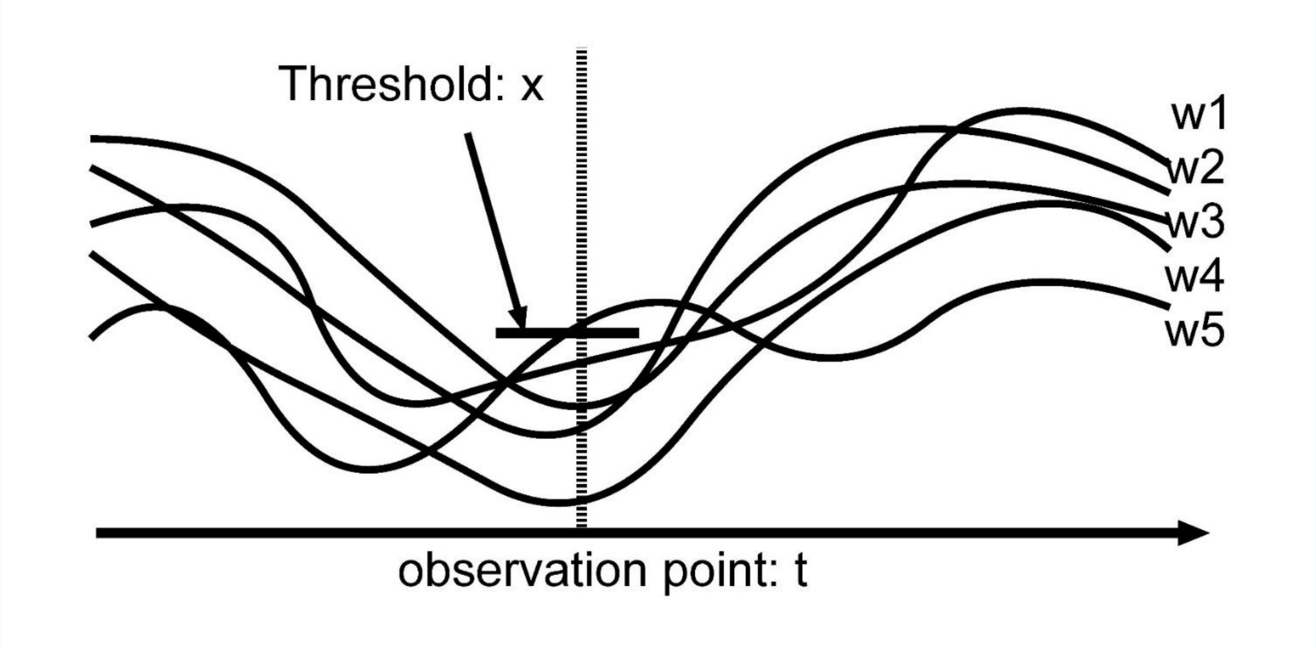
$$F_{\boldsymbol{X}}(x,t) = P\{\boldsymbol{X}(t) \le x\}$$

Probability density of $\boldsymbol{X}(t)$:

$$f_{\boldsymbol{X}}(x,t) = \frac{\partial F_{\boldsymbol{X}}(x,t)}{\partial x}$$

Frequency

For n samples, n functions $\boldsymbol{X}(t, \omega_i)$ $(i = 1, 2, \cdots, n)$ are observed. Denote the number of samples that does not exceed a threshold value x by $n_t(x)$ $(\boldsymbol{X}(t, \omega_i) \leq x)$, $F_{\boldsymbol{X}}(x, t) \approx \dfrac{n_t(x)}{n}$



*n*th-order distribution and *n*th-order probability density

- **Joint distribution of random variable**

$X(t_i)$ $(i = 1, 2, \cdots, n)$

$$F_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = P\{\boldsymbol{X}(t_1) \le x_1, \boldsymbol{X}(t_2) \le x_2, \cdots, \boldsymbol{X}(t_n) \le x_n\}$$

$$f_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = \frac{\partial F_X(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}$$

- **Marginal distribution:**

$$F_{\boldsymbol{X}}(x_1, x_2, \cdots, x_{n-1}; t_1, t_2, \cdots, t_{n-1}) = F_X(x_1, x_2, \cdots, x_{n-1}, \infty; t_1, t_2, \cdots, t_n)$$

$$f_{\boldsymbol{X}}(x_1, x_2, \cdots, x_{n-1}; t_1, t_2, \cdots, t_{n-1}) = \int_{-\infty}^{\infty} f_X(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) dx_n$$

Mean value of random variable $\boldsymbol{X}$ at $\boldsymbol{t}$

$$\eta_{\boldsymbol{X}}(t) = E\{\boldsymbol{X}(t)\} = \int_{-\infty}^{\infty} x f_{\boldsymbol{X}}(x,t)dx$$

where **sample mean** is $\bar{\boldsymbol{X}} = \frac{1}{n}\sum_{t=1}^{n} \boldsymbol{X}(t)$.

Autocorrelation of $\boldsymbol{X}(t)$

$$R_{\boldsymbol{XX}}(t_1,t_2) = E\{\boldsymbol{X}(t_1)\boldsymbol{X}(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{\boldsymbol{X}}(x_1,x_2;t_1,t_2)dx_1 dx_2$$

where **sample autocorrelation** is

$$\bar{R}_{\boldsymbol{XX}}(t_1,t_2) = \frac{1}{n}\sum_{t=1}^{n} \boldsymbol{X}(t+t_1)\boldsymbol{X}(t+t_2)$$

Covariance

$$C_{\boldsymbol{XX}}(t_1,t_2) = R_{\boldsymbol{XX}}(t_1,t_2) - \eta_{\boldsymbol{X}}(t_1)\eta_{\boldsymbol{X}}(t_2)$$

In the case of $t_1 = t_2 = t$, $C_{\boldsymbol{XX}}(t_1,t_2)$ is equal to variance of

$$\boldsymbol{X}(t) \rightarrow C_{\boldsymbol{XX}}(t,t) = E\{\boldsymbol{X}(t)\boldsymbol{X}(t)\} - \eta_{\boldsymbol{X}}^2(t) = Var\big(\boldsymbol{X}(t)\big)$$

Complex process

$\boldsymbol{X}(t) = \boldsymbol{Y}(t) + j\boldsymbol{Z}(t)$: complex variable $\boldsymbol{X}(t)$ is composed of real part $\boldsymbol{Y}(t)$ and imaginary part $\boldsymbol{Z}(t)$.

$$R_{\boldsymbol{XX}}(t_1,t_2) = E\{\boldsymbol{X}(t_1)\boldsymbol{X}^*(t_2)\}$$

$$R_{\boldsymbol{XX}}(t_2,t_1) = R_{\boldsymbol{XX}}^*(t_1,t_2)$$

$$R_{\boldsymbol{XX}}(t,t) = E\{|\boldsymbol{X}(t)|^2\} \geq 0$$

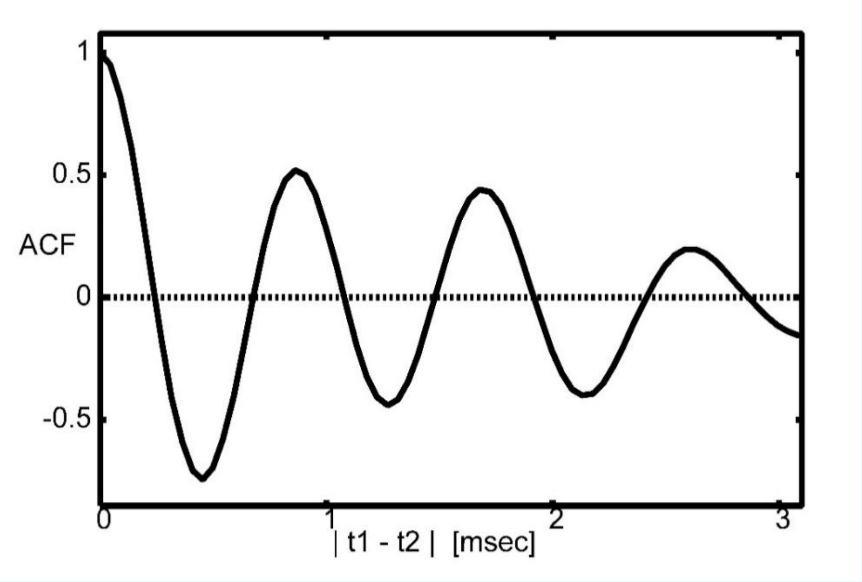
$$C_{\boldsymbol{XX}}(t_1,t_2) = R_{\boldsymbol{XX}}(t_1,t_2) - \eta_{\boldsymbol{X}}(t_1)\eta_{\boldsymbol{X}}^*(t_2)$$

Correlation coefficient

$$r(t_1, t_2) = \frac{C_{XX}(t_1, t_2)}{\sqrt{C_{XX}(t_1, t_1) C_{XX}(t_2, t_2)}}$$

Example

Correlation coefficient $\bar{R}(|t_1 - t_2|) = \bar{R}(t_1, t_2)$ computed for vowel /a/.



- **Cross-correlation** of 2 stochastic processes $\boldsymbol{X}(t), \boldsymbol{Y}(t)$

$$R_{\boldsymbol{XY}}(t_1, t_2) = E\{\boldsymbol{X}(t_1)\boldsymbol{Y}^*(t_2)\} = R^*_{\boldsymbol{YX}}(t_2, t_1)$$

- **Cross-covariance** of 2 stochastic processes $\boldsymbol{X}(t), \boldsymbol{Y}(t)$

$$C_{\boldsymbol{XY}}(t_1, t_2) = R_{\boldsymbol{XY}}(t_1, t_2) - \eta_{\boldsymbol{X}}(t_1)\eta^*_{\boldsymbol{Y}}(t_2)$$

- 2 stochastic processes $\boldsymbol{X}(t), \boldsymbol{Y}(t)$ are (mutually) **orthogonal**.

  For any t_1, t_2, $R_{\boldsymbol{XY}}(t_1, t_2) = 0$

- 2 stochastic processes $\boldsymbol{X}(t), \boldsymbol{Y}(t)$ are **uncorrelated**.

  For any t_1, t_2, $C_{\boldsymbol{XY}}(t_1, t_2) = 0$

- ***a*-dependent**

  $C_{\boldsymbol{XY}}(t_1, t_2) = 0$ for $|t_2 - t_1| > a$

- **White noise $\boldsymbol{W}(t)$**

  For $t_1 \neq t_2, C_{\boldsymbol{WW}}(t_1, t_2) = 0$.

  In other words, $C_{\boldsymbol{WW}}(t_1, t_2) = q(t_1)\delta(t_1 - t_2)$

- **Uncorrelated increments**

  For $t_1 < t_2 \leq t_3 < t_4, \boldsymbol{X}(t_2) - \boldsymbol{X}(t_1)$ and $X(t_4) - X(t_3)$ are not correlated.

Example: Integral of white noise, Brownian motion

Independent increments

For $t_1 < t_2 \leq t_3 < t_4$,
$\boldsymbol{X}(t_2) - \boldsymbol{X}(t_1)$ and $\boldsymbol{X}(t_4) - \boldsymbol{X}(t_3)$ are independent.

Example: Random walk, Wiener process, Poisson process

Independent process

For 2 process $\boldsymbol{X}(t), \boldsymbol{Y}(t)$, random variables $\boldsymbol{X}(t_i), \boldsymbol{Y}(t_j)$ are independent from each other.

Namely, for any t_1, t_2,

$$E\{\boldsymbol{X}(t_i)\boldsymbol{Y}(t_j)\} = E\{\boldsymbol{X}(t_i)\}E\{\boldsymbol{Y}(t_j)\}$$

Normal process

For any $n, t_1, t_2, \cdots, t_n$, joint distribution of random variables $\boldsymbol{X}(t_i)$ $(i = 1, 2, \cdots, n)$ becomes *n*th-order normal distribution.

- In case of $n = 1$,

  setting $\eta_{\boldsymbol{X}}(t) = E\{\boldsymbol{X}(t)\}$, $\sigma_{\boldsymbol{X}}^2(t) = C_{\boldsymbol{XX}}(t, t)$

$$f_{\boldsymbol{X}}(x; t) = \frac{1}{\sqrt{2\pi}\sigma_X(t)} \exp\left[-\frac{1}{2}\left(\frac{x - \eta_X(t)}{\sigma_X(t)}\right)^2\right]$$

- In case of $n = 2$, setting $\eta_{\boldsymbol{X}}(t_i) = E\{\boldsymbol{X}(t_i)\}$,

$\sigma_X^2(t_i) = C_{\boldsymbol{XX}}(t_i t_i)$, $\rho = \frac{C_{\boldsymbol{XX}}(t_1,t_2)}{\sigma_{\boldsymbol{X}}(t_1)\sigma_{\boldsymbol{X}}(t_2)}$,

$$f_{\boldsymbol{X}}(x_1, x_2\,; t_1, t_2) = \frac{1}{2\pi\sigma_{\boldsymbol{X}}(t_1)\sigma_{\boldsymbol{X}}(t_2)\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2} Q(x_1, x_2\,; t_1, t_2)\right]$$

where

$$Q(x_1, x_2\,; t_1, t_2) = \frac{1}{1-\rho^2}\left\{\left(\frac{x_1 - \eta_{\boldsymbol{X}}(t_1)}{\sigma_{\boldsymbol{X}}(t_1)}\right)^2 - 2\rho\left(\frac{x_1 - \eta_{\boldsymbol{X}}(t_1)}{\sigma_{\boldsymbol{X}}(t_1)}\right)\left(\frac{x_2 - \eta_{\boldsymbol{X}}(t_2)}{\sigma_{\boldsymbol{X}}(t_2)}\right) + \left(\frac{x_2 - \eta_{\boldsymbol{X}}(t_2)}{\sigma_{\boldsymbol{X}}(t_2)}\right)^2\right\}$$

3. Stationary process

Strict sense stationary (SSS) process

Statistical property is invariant under time shift. Namely, for any constant c,

$$F_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = F_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n; t_1 + c, t_2 + c, \cdots, t_n + c)$$

$$f_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = f_{\boldsymbol{X}}(x_1, x_2, \cdots, x_n; t_1 + c, t_2 + c, \cdots, t_n + c)$$

Hence

$f_{\boldsymbol{X}}(x; t) = f_X(x) \rightarrow$ 1st-order density is independent of t.

$f_{\boldsymbol{X}}(x_1, x_2; t_1, t_2) = f_{\boldsymbol{X}}(x_1, x_2; \tau) \rightarrow$ 2nd-order density is a function of time lag τ

- **Wide sense stationary (WSS) process**
- Statistical quantities up to 2nd-order are independent of time. Namely,

$E\{\boldsymbol{X}(t)\} = \eta x \rightarrow$ Mean is independent of t.

$E\{\boldsymbol{X}(t+\tau)\boldsymbol{X}^*(t)\} = R_{\boldsymbol{XX}}(\tau) \rightarrow$ Autocorrelation is a function of time lag τ.

- Hence

(a) $R(0) = E\{\boldsymbol{X}(t)\boldsymbol{X}^*(t)\} \rightarrow$ Mean square is independent of t.

(b) Variance $C_{\boldsymbol{XX}}(\tau) = R_{\boldsymbol{XX}}(\tau) - |\eta_{\boldsymbol{X}}|^2$

(c) Correlation coefficient $r(\tau) = C_{\boldsymbol{XX}}(\tau)/C_{\boldsymbol{XX}}(0)$

(d) Joint wide sense stationary

Each of two processes $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ is wide sense stationary, and their cross-correlation depends only on $\tau = t_1 - t_2$.

$$R_{\boldsymbol{XY}}(\tau) = E\{(\boldsymbol{X}(t+\tau)\boldsymbol{Y}^*(t)\}$$

$$C_{\boldsymbol{XY}}(\tau) = R_{\boldsymbol{XY}}(\tau) - \eta_{\boldsymbol{X}}\eta_{\boldsymbol{Y}}^*$$

(e) If white noise $\boldsymbol{W}(t)$ is weakly stationary,

$$E\{\boldsymbol{W}(t)\} = \eta_{\boldsymbol{W}},\ C_{\boldsymbol{WW}} = q\delta(\tau)$$

(where $\eta_{\boldsymbol{W}}$ and q are constants)

In this lecture, we suppose $\eta_{\boldsymbol{W}} = 0$.

(f) If $\boldsymbol{X}(t)$ is an a-dependent process,

$C(\tau) = 0$ for $|\tau| > a$

a is called ***correlation time***.

(g) If $\boldsymbol{X}(t)$ is static sense stationary, then it is wide sense stationary. However, the inverse is not necessarily true.

(h) Since normal process can be described in terms of 2^{nd}-order statistics, inverse of (g) also holds. Namely, if normal process is weakly stationary, it is also strongly stationary.

Sampling

If we set $\boldsymbol{X}[n] = \boldsymbol{X}(n\Delta t)$, statistical quantity of $\boldsymbol{X}[n]$ can be determined by statistical quantity of $\boldsymbol{X}(t)$. Namely,

$$\eta_{\boldsymbol{X}}[n] = \eta_{\boldsymbol{X}}(n\Delta t),$$
$$R_{\boldsymbol{XX}}[n_1, n_2] = R_{\boldsymbol{XX}}(n_1\Delta t, n_2\Delta t).$$

Furthermore, if $\boldsymbol{X}(t)$ is stationary, $\boldsymbol{X}[n]$ is also stationary. Opposite is not necessarily true.

4. Example of stochastic process

Random walk

[Problem]

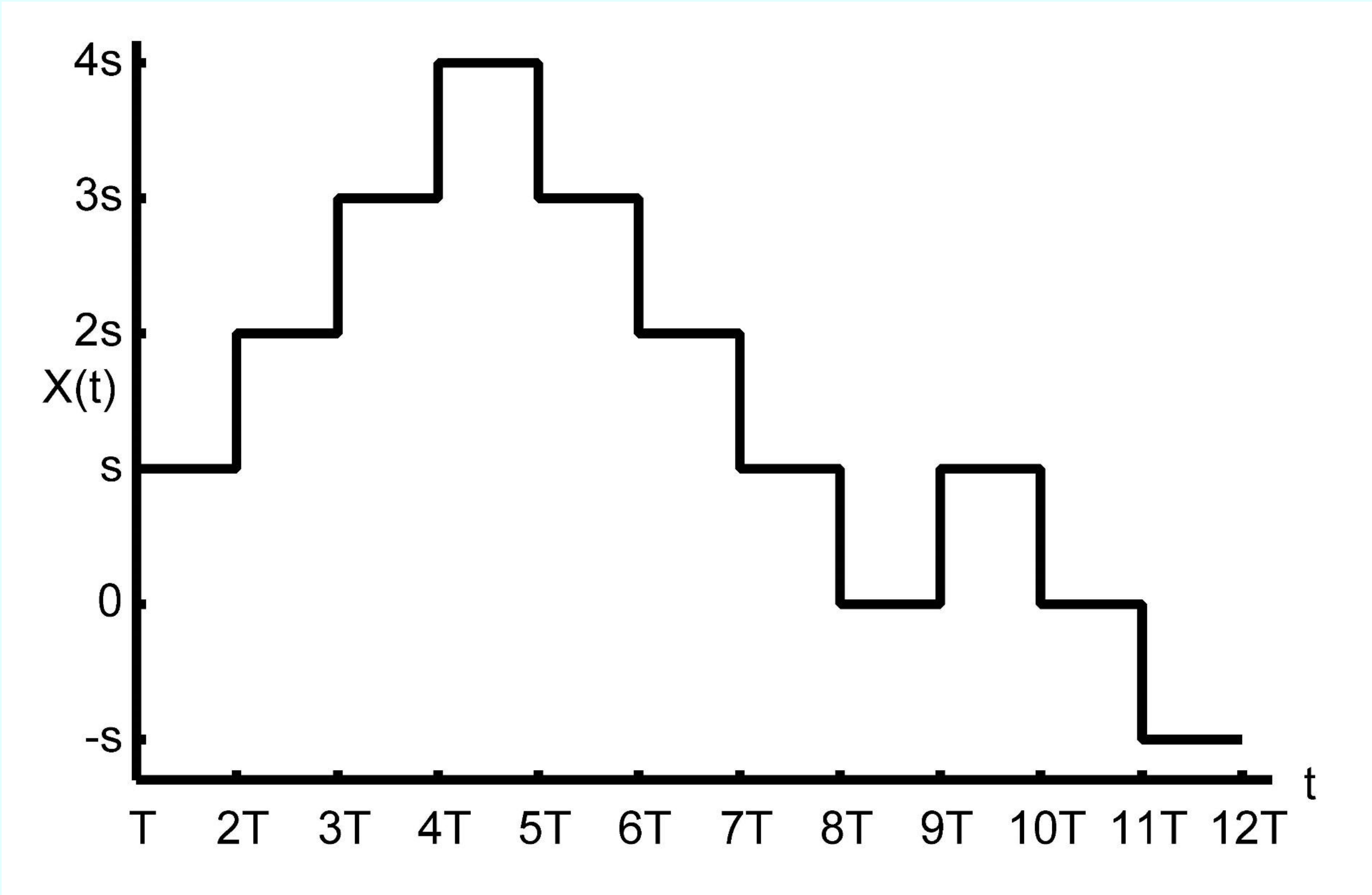
(a) Start at $t = 0$. Every time of T, throw a coin.

(b) If front face is up, proceed to right with s-step.

(c) If back face is up, proceed to left with s-step.

(d) Position at $t = nT$: $\boldsymbol{X}(\mathrm{t})$

Study the statistical quantities (mean, variance, and distribution function) of random variable $\boldsymbol{X}(t)$.


4s
3s
2s
X(t)
s
0
-s
t
T
2T
3T
4T
5T
6T
7T
8T
9T
10T
11T
12T

- If we suppose that, for the first n steps, front face was up for k times, and back face was up for $n-k$ times,

$$\boldsymbol{X}(nT) = ks - (n-k)s = ms$$

where $m = 2k - n, m = -n, n-2, \cdots, n$

- Probability of obtaining front for k times among n trials is

$$P\{\boldsymbol{X}(nT) = ms\} = \binom{n}{k}\frac{1}{2^n} \quad \text{where } k = \frac{m+n}{2}$$

- Denoting the *i*th step by $\boldsymbol{X}(nT)$ can be described as $\boldsymbol{X}(nT) = \Delta\boldsymbol{X}_1 + \Delta\boldsymbol{X}_2 + \cdots + \Delta\boldsymbol{X}_n$. $\Delta\boldsymbol{X}_i$ $(= \pm s)$ is an independent random variable with $E\{\Delta\boldsymbol{X}_i\} = 0$ and $E\{\Delta\boldsymbol{X}_i^2\} = s^2$

$$E\{\boldsymbol{X}(nT)\} = nE\{\Delta\boldsymbol{X}_i\} = 0$$

$$E\{\boldsymbol{X}^2(nT)\} = nE\{\Delta\boldsymbol{X}_i^2\} = ns^2$$

- According to De Moivre-Laplace theorem,

  "If $npq \gg 1$, in $\sqrt{npq}$ neighborhood of $k = np$,

  $$\binom{n}{k} p^k q^{n-k} \cong \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}},$$

  Hence substituting $p = q = 0.5$, $m = 2k - n$,

  $$P\{\boldsymbol{X}(nT) = ms\} \cong \frac{1}{\sqrt{n\pi/2}} e^{-\frac{m^2}{2n}}$$ holds for $|m| \sim \sqrt{n}$.

  Therefore, $P\{\boldsymbol{X}(nT) \le ms\} = \boldsymbol{\Phi}\left(\frac{m}{\sqrt{n}}\right)$ for $nT - T < t \le T$

  where $\Phi(\cdot)$ represents distribution function of standard normal distribution $N(0,1)$. In addition, if $n_1 < n_2 \le n_3 < n_4$, increments $\boldsymbol{X}(n_4T) - \boldsymbol{X}(n_3T)$ and $\boldsymbol{X}(n_2T) - \boldsymbol{X}(n_1T)$ are independent.

De Moivre-Laplace theorem: a derivation

$$\boxed{\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}, \quad p+q=1}$$

[Outline of derivation] Using Stirling's formula for factorial, $n! \approx \sqrt{2\pi n}\left(n/e\right)^n$,

$$\binom{n}{k} p^k q^{n-k} \approx \sqrt{\frac{n}{2\pi k(n-k)}}\left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}.$$

Setting $k = npq + x\sqrt{npq}$ and expanding using a Taylor series $\ln(1+x) = x - x^2/2 + \cdots$,

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{x^2}{2}}.$$

Wiener process

Consider a random walk in the limit of $n \to \infty$. We consider a limit $T \to 0$ under the condition of $s^2 = \alpha T$. Then $\boldsymbol{X}(t)$ becomes continuous-time continuous-state stochastic process

$$\boldsymbol{Y}(t) = \lim_{T \to 0} \boldsymbol{X}(t).$$

$\boldsymbol{Y}(t)$ is called ***Wiener process***.

Mean and Variance

According to the results of random walk,

$$E\{\boldsymbol{Y}(t)\} = 0$$

$$E\left\{\boldsymbol{Y}^{\mathbf{2}}(t)\right\} = ns^2 = \frac{ts^2}{T} = \alpha t.$$

- **Distribution function**: Substituting y$= ms$, $t = nT$ into distribution function of random walk,

$$P\{\boldsymbol{Y}(t) \le y\} = \boldsymbol{\Phi}\left(\frac{m}{\sqrt{n}}\right) = \boldsymbol{\Phi}\left(\frac{y/s}{\sqrt{t/T}}\right) = \boldsymbol{\Phi}\left(\frac{y}{\sqrt{\alpha t}}\right)$$

Hence, probability density of $\boldsymbol{Y}(t)$ is distributed normally as $N(0, \alpha t)$.

$$f_{\boldsymbol{Y}}(y,t) = \frac{1}{\sqrt{2\pi\alpha t}} e^{-\frac{y^2}{2\alpha t}}$$

- **Autocorrelation**:

$$R_{\boldsymbol{YY}}(t_1, t_2) = \alpha \min(t_1, t_2)$$

- **Increment**: if $t_1 < t_2 \le t_3 < t_4$, increments $\boldsymbol{Y}(t_4) - \boldsymbol{Y}(t_3)$ and $\boldsymbol{Y}(t_2) - \boldsymbol{Y}(t_1)$ are intendant.

Generalized random walk

- In random walk, we suppose that probability of obtaining front face is p, whereas probability of obtaining back is $q = 1 - p$. Then,

$$\boldsymbol{X}(t) = \sum_{k=1}^{n} c_k U(t - kT) \quad \text{for} \quad (n-1)T < t \leq T$$

where c_k is a random number, which takes value of s with probability p and takes a value of $-s$ with probability q and

$$U(t) = 0\ (t < 0) \text{ and } U(t) = 1\ (t \geq 0).$$

$\boldsymbol{X}(t)$ is called ***generalized random walk***.

Generalized random walk

Using the following properties of binominal distribution:

$$E\{c_k\} = (p-q)s$$

$$E\{c_k^2\} = s^2, \qquad Var(c_k^2) = 4pqs^2$$

- **Mean and Variance:**

$$E\{\boldsymbol{X}(t)\} = n(p-q)s$$

$$Var(\boldsymbol{X}(t)) = 4npqs^2$$

- **Distribution function:**

  For large *n*, $X(t)$ is normally distributed with

$$E\{\boldsymbol{X}(t)\} \cong \frac{t}{T}(p-q)s$$

$$Var(\boldsymbol{X}(t)) \cong \frac{4t}{T}4pqs^2$$

5. Ergodic property

- **Problem:**

  Consider an estimation of statistical quantity of $\boldsymbol{X}(t)$ such as its mean.

  $\eta(t) = E\{\boldsymbol{X}(t)\}$, from real data.

- **Method:**

  Given n samples $\boldsymbol{X}(t, \omega_i)$ $(i = 1, 2, \cdots, n)$, average is obtained as follows.

$$\hat{\eta}(t) = \frac{1}{n}\sum_{i=1}^{n} X(t, \omega_i)$$

- **Practical Problem:**

  It is rare to have some many samples. In most cases, only a single time series $\boldsymbol{X}(t)$ is given.

- **Non-stationary:**

  If $\boldsymbol{X}(t)$ is non-stationary and mean $E\{\boldsymbol{X}(t)\}$ is a function of t, estimation is impossible.

  However, if $\boldsymbol{X}(t)$ is stationary, time-average, computed as

$$\eta_T = \frac{1}{2T}\int_{-T}^{T} \boldsymbol{X}(t)dt$$

  becomes

$$\eta_T \to E\{\boldsymbol{X}\} \text{ as } T \to \infty$$

- Ergodic property implies **time-average equals to ensemble average.**

Mean-ergodic process

Problem:

Given a stationary real process $\boldsymbol{X}(t)$, compute its average $\eta = E\{\boldsymbol{X}(t)\}$. Define a time average over a duration of $2T$

$$\eta_T = \frac{1}{2T}\int_{-T}^{T}\boldsymbol{X}(t)dt$$

as a new random variable, average of η_T is

$$E\{\eta_T\} = \frac{1}{2T}\int_{-T}^{T}E\{\boldsymbol{X}(t)\}dt = \eta$$

If the variance has a property of $\sigma_T^2 \to 0$ in the limit of $T \to \infty$ time average converges to the true average. Namely,

$$P(\eta_T = \eta) \to 1$$

$\boldsymbol{X}(t)$ is called ***Mean-ergodic process***.

Slutsky theorem:

- If $\frac{1}{T}\int_0^T C(\tau)d\tau \to 0$ as $T \to \infty$, $\boldsymbol{X}(t)$ is a mean-ergodic process.
- Sufficient condition (a): $\int_0^\infty C(\tau)d\tau < \infty$
- Sufficient condition (b): For t $\to \infty$, $C(\tau) \to 0$

$$\mathrm{E}\left[\left(\eta_T-\eta\right)^2\right]=\frac{1}{\left(2T\right)^2}\int_{-T}^{T}dt\int_{-T}^{T}dt'\mathrm{E}\left[\left(x\left(t\right)-\eta\right)\left(x\left(t'\right)-\eta\right)\right]$$

$$=\frac{2}{\left(2T\right)^2}\int_{-2T}^{2T}du\int_{-2T}^{2T}d\tau C\left(\tau\right)$$

$$=\frac{2}{T}\int_{-2T}^{2T}d\tau C\left(\tau\right)$$