CM04-230302Thu

2023년 3월 1일 수요일 20:00

CM04-230302Thu

Announcements

1. Handouts, annotated handouts, and lecture videos are available at http://cisl.postech.ac.kr/class/eece574/schedule.htm . Passwords are …
2. HW#02 due at 10:59am on Tue., Mar. 7th. HW#03 will be posted soon.
3. Clickable time stamps to the lecture videos, please.

Corrections; Clarifications; Common mistakes,misunderstandings, misconceptions, myth; Frequently Asked Questions; Conventional wisdom

1.

Overview (Chapter numbers from the main textbook by Papulis)

9. General Concepts
7. Sequence of Random Variables
10. Random Walks and Other Applications
11. Spectral Representation
12. Spectrum Estimation
13. Mean-Square Estimation

□ average of X_t
- ensemble/statistical
- time

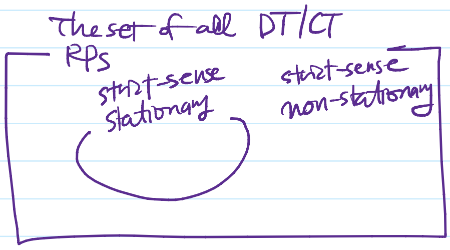
Review

1. expectation of a RP = a mean path/function = the ensemble average of the RP
2. Only the 1st-order PDF/CDF is involved.
3. It may have a drastically different property from sample paths
4. infinity, parallel universe, …

Preview

Venn diagram

1. classification of RPs
   a. CT continuous/discrete, DT continuous/discrete
   b. stationary vs. non-stationary
2. Stationarity in the strict sense vs. in a wide sense vs. 1st-order , 2nd-order, …



6.2 STATIONARITY AND INDEPENDENCE

- a full characterization of a RP
  - Q. Recall how to fully characterize a random vector $\underline{X}$, which is a collection of finite number of jointly distributed random variables?
  - A.
    - Method 1. Provide a probability space $(S, \mathcal{A}, P)$ and a mapping rule $\underline{X}(s)$ or, equivalently, the induced probability space $(\mathbb{R}^n, \mathcal{B}^n, P_{\underline{X}})$.
    - Method 2. Provide a joint CDF/PDF/CF of the finite number of random variables.
      - $X \to$ Provide the CDF $F_X(x)$.
      - $\underline{X} = \begin{bmatrix} X_1 \\ Y_2 \end{bmatrix} \to$ Provide the joint CDF $F_{X_1,X_2}(x_1, x_2)$.
      - $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \to$ Provide the joint CDF $F_{X_1,X_2,X_3}(x_1, x_2, x_3)$.
      - $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{10^6} \end{bmatrix} \to$ Provide the joint CDF $F_{X_1,X_2,\cdots,X_{10^6}}(x_1, x_2, \cdots, x_{10^6})$.

① $\boxed{(S,\mathcal{A},P)} \xrightarrow{s} \boxed{\underline{X}(s)}$

①′ $\boxed{(\mathbb{R}^n, \mathcal{B}^n, P_{\underline{X}})} \to \underline{X}$

② $\boxed{F_{\underline{X}}(\underline{x}), f_{\underline{X}}(\underline{x}), p_{\underline{X}}(\underline{x})} \to \underline{X}$

②′ $\Phi_{\underline{X}}(\underline{\omega})$

- Q. Now, how can we fully characterize a CT random process $X(t)$, a collection of *infinite* number of jointly distributed random variables?
- A.

①

– Q. Now, how can we fully characterize a CT random process $X(t)$, a collection of *infinite* number of jointly distributed random variables?

– A.

∗ Method 1. Provide a probability space $(S, \mathcal{A}, P)$ and a mapping rule $X(t, s)$.

∗ Method 2. Provide all possible joint CDF/PDF/CF of any finite number of random variables.

· Theorem. (Kolmogorov's Extension/Consistency Theorem) A CT random process is fully characterized by providing a joint CDF $F_{X(t_1),X(t_2),\cdots,X(t_N)}(x_1, x_2, \cdots, x_N)$ for every N-tuple $(t_1, t_2, \cdots, t_N) \in \mathbb{R}^N$ for every $N \in \mathbb{N}$.

(1) $F_{X(t)}(x)$ for every $t \in \mathbb{R}$,

(2) $F_{X(t_1),X(t_2)}(x_1, x_2)$ for every pair $(t_1, t_2) \in \mathbb{R}^2$,

(3) $F_{X(t_1),X(t_2),X(t_3)}(x_1, x_2, x_3)$ for every triplet $(t_1, t_2, t_3) \in \mathbb{R}^3$,

(⋮)

and they are all consistent after any marginalization.

· Proof. Omitted.

· Note that this method works even though there are *uncountably infinite* number of jointly distributed random variables but only the joint CDFs of *any finite* number of random variables are provided.

· Similarly, a DT RP can be fully characterized.

① $(S,\mathcal{A},P) \xrightarrow{s} X_t(s) \to X_t$

② $F_{X_t}(x),\ F_{X_{t_1},X_{t_2}}(x_1,x_2), \cdots \to X_t$

Distribution and Density Functions

- CT case

1. Given a particular time t_1, the distribution function associated with the random variable $X_1 = X(t_1)$ is defined as

$$F_X(x_1; t_1) \triangleq P(X(t_1) \le x_1) = F_{X(t_1)}(x_1)$$

for any real number x_1.

2. For two RVs $X_1 = X(t_1)$ and $X_2 = X(t_2)$, the second-order joint distribution function is

$$F_X(x_1, x_2; t_1, t_2) \triangleq P(X(t_1) \le x_1, X(t_2) \le x_2) = F_{X(t_1),X(t_2)}(x_1, x_2)$$

3. For N RVs $X_i = X(t_i), i = 1, 2, \ldots, N$, the Nth-order joint distribution function is

$$F_X(x_1, ..., x_N; t_1, ..., t_N) \triangleq P(X(t_1) \le x_1, ..., X(t_N) \le x_N) = F_{X(t_1),...,X(t_N)}(x_1, ..., x_N) \qquad (6.2\text{-}1)$$

4. joint density functions

$$f_X(x_1; t_1) = dF_X(x_1; t_1)/dx_1 \qquad (6.2\text{-}2)$$

$$f_X(x_1, x_2; t_1, t_2) = \partial^2 F_X(x_1, x_2; t_1, t_2)/(\partial x_1 \partial x_2) \qquad (6.2\text{-}3)$$

$$f_X(x_1, ..., x_N; t_1, ..., t_N) = \partial^N F_X(x_1, ..., x_N; t_1, ..., t_N)/(\partial x_1 \cdots \partial x_N) \qquad (6.2\text{-}4)$$

- Similar for DT cases. Repeat above 1-4 yourself.
- Def. Given a RP $X(t)$, the deterministic signal

$$\mu_X(t) \triangleq \mathbf{E}[X(t)] = \int_{-\infty}^{\infty} x f_{X(t)}(x)\,dx$$

as a function of t is called the mean path, the mean function, or simply the mean of $X(t)$. the ensemble average of $X(t)$.

– can also be defined for DT RPs.

cf. a sample path, a sample function – the time average of $X(t)$

Recall ① $E[X] = \int_{s\in S} X(s)\,P(ds)$

①′ $E[X] = \int_{x\in\mathbb{R}} x\,P_X(dx)$

② $E[X] = \int_{-\infty}^{\infty} x\,dF_X(x)$

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

$= \sum_{x\in\mathcal{X}} x\,p_X(x)$

②′ $\Phi_X(\omega) = E[e^{j\omega X}]$

$\frac{d\Phi_X(\omega)}{d\omega} = E[jX e^{j\omega X}]$

$\therefore E[X] = \frac{1}{j}\frac{d\Phi_X(\omega)}{d\omega}\Big|_{\omega=0}$

Examples

- Fully characterize the following RPs. $\to S = \{0, 1\}$

1. DT discrete: a Bernoulli random process with p

cf) multinoulli $S = \{1, 2, \cdots, M\}$

– $X[n]$ is a Bernoulli RV with p.

– $X[n]$'s are i.i.d.

(i) $p_{X[n]}(x) = (1-p)\delta[x] + p\,\delta[x-1]$

(ii) $p_{X[n_1],X[n_2]}(x_1, x_2) = p_{X[n_1]}(x_1)\,p_{X[n_2]}(x_2)$

(iii) ‥

– Q. Let $N = X[1] + X[2] + \cdots X[n]$ and $p = \lambda/n$. Find $P(N = 0)$, $P(N = 1)$, and $P(N \ge 2)$. Also find the limits of $P(N = 0)$, $P(N = 1)$, and $P(N \ge 2)$ as n tends to ∞.

2. DT continuous: a DT white Gaussian RP with power σ^2

– $X[n] \boxed{\sim} \mathcal{N}(0, \sigma^2)$. $\sigma^2 > 0$

$f_{X[n]}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$,

$f_{X[n_1],X[n_2]}(x_1, x_2) = f_{X[n_1]}(x_1)\,f_{X[n_2]}(x_2)$

⋮

$f_{X[n_1],X[n_2],\cdots X[n_M]}(x_1, x_2, \cdots, x_M)$

CN

$\sum_{n=1}^{100}$

(iii) - -

CN $\quad \sum_{n=1}^{100}$

$f_{X[n_1],X[n_2]}(x_1,x_2) = f_{X[n_1]}(x_1) f_{X[n_2]}(x_2)$

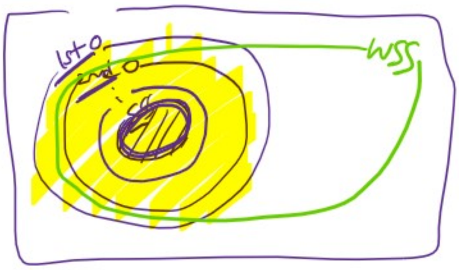
$\vdots$

$f_{X[n_1],X[n_2],\dots,X[n_M]}(x_1,x_2,\dots,x_M) = \prod_{m=1}^{M} f_{X[n_m]}(x_m)$ prod

– $X[n]$'s are i.i.d. $\boxed{(S,\mathcal{F},P)} \to \boxed{X_t(s)} \to$

- In general, it is extremely complicated to fully characterize a random process.
- There are some classes of random processes which have very simple ways to fully characterize.
- A random process is said to be stationary if all its statistical properties do not change with time shift, i.e., with your choice of $t = 0$.

  Otherwise, it is said to be non-stationary.
- There are several "levels" of stationarity, all of which depend on the density functions of the random variables of the process.



Stationary Process to Order One

- A random process is said to be stationary if its statistical properties do not change with time (not change with the choice of $t = 0$, not change with a shift in the time origin).
  - There are several "levels" of stationarity,
  - all of which depend on the distribution functions of the random variables of the process.
- Other processes are called nonstationary.
- Def. A random process $X(t)$ ($X_t(s)$, X_t) is called stationary to order one (or first-order stationary) if its first-order density function does not change with a shift in time origin, that is,

$$f_X(x_1; t_1) = f_X(x_1; t_1 + \Delta)$$

$\forall t, \left(f_{X(t)}(x) = f_{X(0)}(x), \forall x \right)$

for any t_1 and any real number Δ, i.e., $F_{X(t_1)}(x) = F_{X(t_1+\Delta)}(x), \forall x, \forall t_1, \forall \Delta$.

- First-order stationarity
  - $f_X(x_1; t_1)$ is independent of t_1.
  - The process mean value $\mathbf{E}[X(t)]$ is a constant:

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

$$= \int_{-\infty}^{\infty} x f_{X(t)}(x)\,dx = \mu_X, \forall t.$$

Def → existence, uniqueness

↳ consequences: lemma, proposition, theorem, corollary, ⋮

Proof. For $X_1 = X(t_1)$,

$$\mathbf{E}[X_1] = \mathbf{E}[X(t_1)] = \int_{-\infty}^{\infty} x_1 f_X(x_1; t_1) dx_1.$$

For $X_2 = X(t_2)$,

$$\mathbf{E}[X_2] = \mathbf{E}[X(t_2)] = \int_{-\infty}^{\infty} x_1 f_X(x_1; t_2) dx_1$$

By letting $t_2 = t_1 + \Delta$, we get

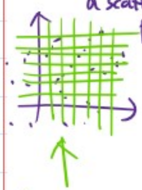
$$\mathbf{E}[X(t_1 + \Delta)] = \mathbf{E}[X(t_1)]$$

X_t

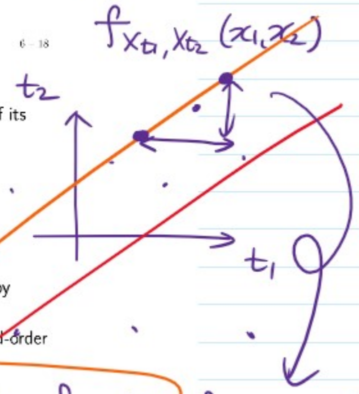
$f_{X_{t_1}, X_{t_2}}(x_1, x_2)$

N.

(x_1, x_2) a scatter plot 산점도



Stationarity to Order Two and Wide-Sense Stationarity



- Def. A process is called stationary to order two (or second-order stationary) if its second-order density function satisfies

$$f_X(x_1, x_2; t_1, t_2) = f_X(x_1, x_2; t_1 + \Delta, t_2 + \Delta)$$

for all t_1, t_2, and Δ, i.e.,
$F_{X(t_1),X(t_2)}(x_1, x_2) = F_{X(t_1+\Delta),X(t_2+\Delta)}(x_1, x_2), \forall (x_1, x_2), \forall (t_1, t_2), \forall \Delta.$

- Second-order stationarity
  - The joint density is a function of time differences but not absolute time (by letting $\Delta = -t_1$).
  - A second-order stationary RP is a first-order stationary because the second-order density function determines the lower, first-order, density.

$f_{X_1, X_2}(x_1, x_2)$

- Def. The autocorrelation function of the random process $X(t)$

$R_{XX}(t_2, t_1) =$
$$R_{XX}(t_1, t_2) = \mathbf{E}[X(t_1)X(t_2)] \triangleq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1 dx_2$$

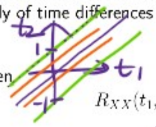
$f_{X_t}(x) = f_{X_0}(x), \forall t$

  - The autocorrelation function of a second-order stationary process is a function only of time differences and not absolute time, that is, if

$f_{X_t}(x) = f_{X_{t+\Delta}}(x), \forall \Delta$

If S to 2 then $R_{XX}(t_1+\Delta, t_2+\Delta)$

$$\boxed{\tau = t_2 - t_1}$$



then

$$R_{XX}(t_1, t_1 + \tau) = \mathbf{E}[X(t_1)X(t_1 + \tau)] = R_{XX}(\tau)$$

$R_{X(t_1),X(t_2)} \triangleq E[X(t_1)X(t_2)]$

$f_{X_{t_1}, X_{t_2}}(x_1, x_2) = f_{X_{t_1+\Delta}, X_{t_2+\Delta}}(x_1, x_2)$ $\forall \Delta.$

$f_{X(t_1),X(t_2),X(t_3)}(x_1, x_2, x_3)$
$= f_{X(t_1+\Delta),X(t_2+\Delta),X(t_3+\Delta)}(x_1, x_2, x_3), \forall \Delta.$

- Def. A process is called wide-sense stationary (WSS) if

$$\mathbf{E}[X(t)] = \bar{X} = \text{constant} \qquad (6.2\text{-}5a)$$
$$\mathbf{E}[X(t)X(t+\tau)] = R_{XX}(\tau) \qquad (6.2\text{-}5b)$$

- Wide-sense stationarity
  - Second-order stationarity is often more restrictive than necessary, and a more relaxed form of stationarity is desirable.
  - A process stationary to order 2 is clearly wide-sense stationary.
    However, the converse is not necessarily true.

Example: Show that the random process

$$X(t) = A\cos(\omega_0 t + \Theta)$$

is WSS, if it is assumed that A and ω_0 are constants and Θ is a uniformly distributed RV on the interval $(0, 2\pi)$.

Proof. The mean value is

$$\mathbf{E}[X(t)] = \int_0^{2\pi} A\cos(\omega_0 t + \theta)\frac{1}{2\pi}d\theta = 0.$$

The autocorrelation function becomes

$$\begin{aligned} R_{XX}(t, t+\tau) &= \mathbf{E}[A\cos(\omega_0 t + \Theta)\, A\cos(\omega_0 t + \omega_0\tau + \Theta)] \\ &= \frac{A^2}{2}\cos(\omega_0 \tau) \end{aligned}$$

$$= \frac{[illegible]}{2}\cos(\omega_0\tau)$$

Therefore, $X(t)$ is WSS.

N-Order and Strict-Sense Stationarity

- Def. A random process is stationary to order N if its Nth-order density function is invariant to a time origin shift; that is, if

$$f_X(x_1, ..., x_N; t_1, ..., t_N) = f_X(x_1, ..., x_N; t_1 + \Delta, ..., t_N + \Delta)$$

  for all $t_1, ..., t_N$ and Δ, i.e.,

$$F_{X(t_1),X(t_2),\cdots,X(t_N)}(x_1, x_2, \cdots, x_N) = F_{X(t_1+\Delta),X(t_2+\Delta),\cdots,X(t_N+\Delta)}(x_1, x_2, \cdots, x_N),$$

  $\forall (t_1, t_2, \cdots, t_N) \in \mathbb{R}^N, \forall \Delta.$

- Stationarity of order N implies stationarity to all orders $k \leq N$.
- Def. A process stationary to all orders $N = 1, 2, ...,$ is called strict-sense stationary (SSS).