

1.1 Difference between deterministic and stochastic world

	deterministic world	stochastic world
Single variable: Temp of a sick man	R $T = 39^\circ C$	random variable E, Var, \dots
Variables changing over time: T in first 3 days	$R_+ \rightarrow R$ $T(1) = 39$ $T(2) = 38.5$ $T(3) = 38$ \vdots	stochastic process

1.2 Difference between various fields of stochastics

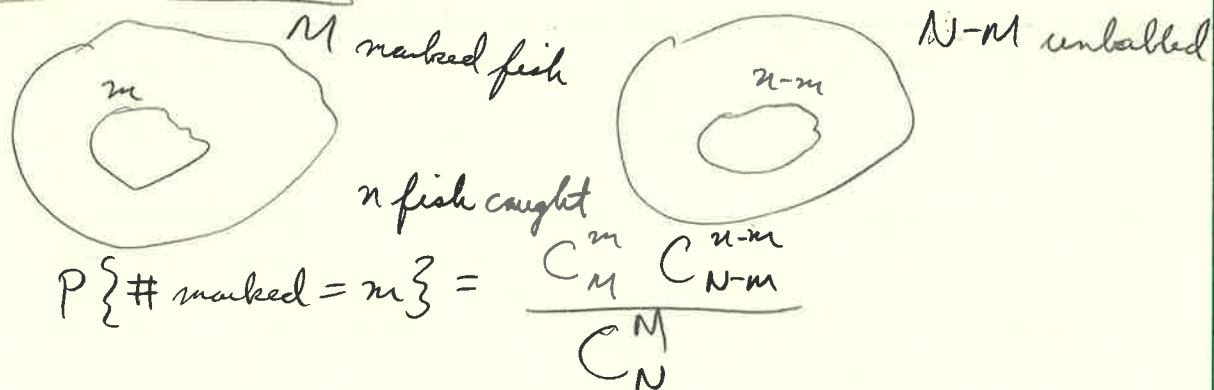
Stochastics

- probability theory
- mathematical statistics
- stochastic processes

Consider a pond that contains fish

Prob theory: # of fish at some given time (N)
 $E, Var, \text{ or limit laws}$

Mathematical Stats:

Repeat m_1, m_2, \dots, m_g

$$(\text{log likelihood}) \sum_{k=1}^g P\{\# \text{ marked} = m_k\} \rightarrow \max_N \quad (\text{MLE})$$

1.3 Probability space (Ω, \mathcal{F}, P)

General theory	Bernoulli Scheme [1, success 0, failure] $(a_1, \dots, a_n), a_i \in \{0, 1\}$	$[0, 1]$ Select point from
Ω -sample space	$\#\Omega = 2^n$, set of all vectors with components $\in \{0, 1\}$	$\Omega = [0, 1]$
\mathcal{F} - σ -algebra 1) $\Omega \in \mathcal{F}$ 2) $A \in \mathcal{F} \Rightarrow \Omega \setminus A \in \mathcal{F}$ 3) $A_1, \dots, A_n, \dots \in \mathcal{F}$ \Downarrow $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$	\mathcal{F} = power set $\#\mathcal{F} = 2^{\#\Omega} = 2^{2^n}$	$P\{x \in [\alpha, \beta]\}$ $\Rightarrow [\alpha, \beta), (\alpha, \beta],$ $(\alpha, \beta), [\alpha, \beta), \{\beta\} \in \mathcal{F}$ Borel σ -algebra
P -probability measure 1) $P(\Omega) = 1$ 2) $A_1, A_2, \dots \in \mathcal{F}$ (disjoint) $\Rightarrow P\{\bigcup_i A_i\} = \sum P(A_i)$ $P: \mathcal{F} \rightarrow [0, 1]$	$P\{1\} = p$ $P\{0\} = 1 - p$	$P\{[\alpha, \beta]\} = \beta - \alpha$

1.4 Definition of a stochastic function. Types of stochastic functions.
 (Ω, \mathcal{F}, P) Random variable - measurable function $\xi: \Omega \rightarrow \mathbb{R}$.

$$\forall B \in \mathcal{B}(\mathbb{R}) : \xi^{-1}(B) \subset \mathcal{F}$$

T - time

 $X: T \times \Omega \rightarrow \mathbb{R}$ - random function, if $\forall t \in T: X(t, \cdot)$ is
a random variable on (Ω, \mathcal{F}, P) , denoted X_t

If $T = \mathbb{R}_+$, this is called a random process or stochastic process

$T = \mathbb{R}_+^n$, random field or stochastic field

$T = \mathbb{N}$, discrete time stochastic process
or \mathbb{Z}

$T = \mathbb{R}_+ \text{ or } \mathbb{R}$, continuous time stochastic process

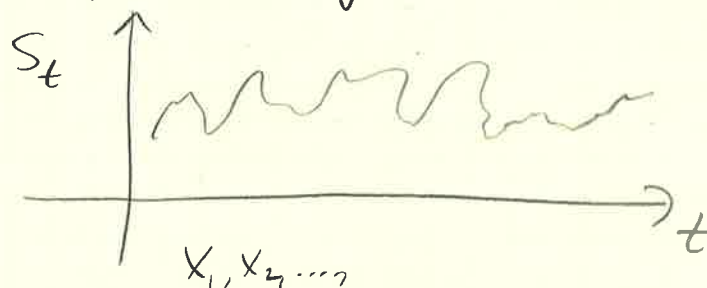
1.5 Trajectories and finite-dimensional distributions

$$X: T \times \Omega \rightarrow \mathbb{R}, \quad T = \mathbb{R}_+$$

$\forall t \in T: X_t = X(t, \cdot)$ is a r.v. on (Ω, \mathcal{F}, P)

Trajectory (= path)

X_t fix ω and get mapping $T \rightarrow \mathbb{R}$

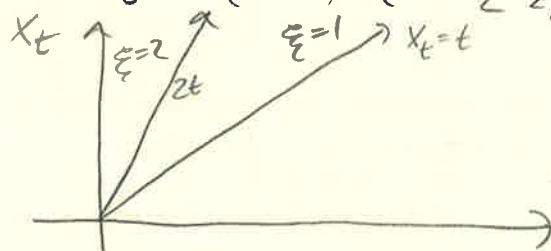


Finite-dimensional distribution $(X_{t_1}, X_{t_2}, \dots, X_{t_n}), t_1, \dots, t_n \in \mathbb{R}$

In mathematics stats, $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ are independent

In stochastic process, $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ are dependent

Ex: $X_t = \xi t$, $\xi = \begin{cases} 1, & \text{w.p. } 1/2 \\ 2, & \text{w.p. } 1/2 \end{cases}$



$$P\{X_{t_1} \leq x_1, X_{t_2} \leq x_2\} = \begin{cases} 0, & \min(\frac{x_1}{t_1}, \frac{x_2}{t_2}) < 1 \\ 1/2, & \text{if } \in [1, 2] \\ 1, & \text{if } \geq 2 \end{cases}$$

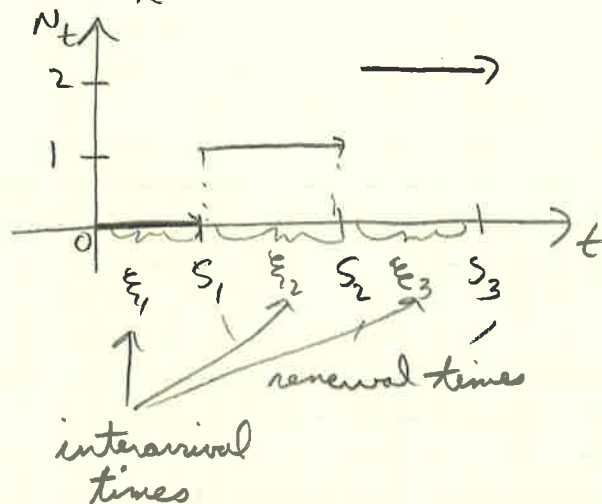
1.6 Renewal process. Counting process.

Renewal processes (discrete time)

$$S_0 = 0, S_n = S_{n-1} + \xi_n, \text{ where } \xi_1, \xi_2, \dots - \text{i.i.d.} > 0 \text{ a.s.}$$

$$P\{\xi_i > 0\} = 1 \Leftrightarrow F(0) = 0$$

$$N_t = \arg\max_k \{S_k \leq t\} \quad (\text{Counting process})$$



$$\{S_n > t\} = \{N_t < n\}$$

$$F \rightarrow \mathbb{E} N_t$$

$$S_n = \xi_1 + \dots + \xi_n$$

1.7. Convolution

Convolution $X \perp\!\!\!\perp Y$

$$X \sim F_X, Y \sim F_Y$$

$$F_{X+Y}(x) = \int_{\mathbb{R}} F_X(x-y) dF_Y(y) =: F_X * F_Y$$

conv in terms of distribution functions

$$X \sim p_X, Y \sim p_Y$$

(If Y, X have densities)

$$p_{X+Y}(x) = \int_{\mathbb{R}} p_X(x-y) p_Y(y) dy =: p_X * p_Y$$

conv in terms of densities

$$S_n = \xi_1 + \dots + \xi_n$$

$$\text{let } F^{n*} := \underbrace{F * \dots * F}_n$$

$$1) F^{n*}(x) \leq F^n(x) \text{ if } F(0)=0$$

$$\xi_1, \dots, \xi_n \stackrel{i.i.d.}{\sim} F$$

$$\{\xi_1 + \dots + \xi_n \leq x\} \subset \{\xi_1 \leq x, \dots, \xi_n \leq x\} \quad \text{Since } \xi_i \geq 0 \text{ a.s.}$$

$$P\{\xi_1 + \dots + \xi_n \leq x\} \leq \prod_{k=1}^n P\{\xi_k \leq x\}$$

$$\stackrel{||}{F^{n*}(x)} \qquad \qquad \qquad F(x)$$

$$2) F^{n*}(x) \geq F^{(n+1)*}(x)$$

$$\{\xi_1 + \dots + \xi_n \leq x\} \supset \{\xi_1 + \dots + \xi_{n+1} \leq x\}$$

Theorem: $S_n = S_{n-1} + \xi_n$ where $\xi_1, \xi_2, \dots \stackrel{i.i.d.}{\sim} F, F(0)=0$

$$(1) \boxed{U(t) = \sum_{n=1}^{\infty} F^{n*}(t) < \infty}$$

$$(2) \boxed{\mathbb{E}N_t = U(t)}$$

proof for (2)

$$\begin{aligned} \mathbb{E}N_t &= \mathbb{E}[\#\{n: S_n \leq t\}] \\ &= \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{S_n \leq t\}}\right] = \sum_{n=1}^{\infty} P\{S_n \leq t\} \\ &= \sum_{n=1}^{\infty} F^{n*}(t) \end{aligned}$$

1.8 Laplace transform. Calculation of an expectation of a counting process (1)

Laplace transform

$$f: \mathbb{R}_+ \rightarrow \mathbb{R} : \mathcal{L}_f(s) = \int_0^{\infty} e^{-sx} f(x) dx$$

$$1) f \text{-density of } \xi, \text{ then } \mathcal{L}_f(s) = \mathbb{E}[e^{-s\xi}]$$

$$2) f_1, f_2 : \mathcal{L}_{\underbrace{f_1 * f_2}_{\text{densities}}}(s) = \mathcal{L}_{f_1}(s) \cdot \mathcal{L}_{f_2}(s)$$

$$3) F \text{-distribution function, } F(0)=0, \quad p = F'$$

$$\mathcal{L}_F(s) = \frac{\mathcal{L}_p(s)}{s}$$

$$\begin{aligned} \text{l.h.s.} &= \int_{\mathbb{R}_+} F(x) \frac{d(e^{-sx})}{s} = - \frac{F(x)e^{-sx}}{s} \Big|_0^{\infty} + \frac{1}{s} \int_{\mathbb{R}_+} p(x) e^{-sx} dx \\ &= \text{r.h.s.} \end{aligned}$$

Ex 1)

$$\begin{aligned} 1) \mathcal{L}_{x^k}(s) &= \int_{\mathbb{R}_+} x^k \frac{d(e^{-sx})}{s} = \frac{n}{s} \int_{\mathbb{R}_+} x^{n-1} e^{-sx} dx \\ &= \frac{n}{s} \cdot \frac{n-1}{s} \cdots \frac{2}{s} \int_{\mathbb{R}_+} e^{-sx} dx = \frac{n!}{s^n} \end{aligned}$$

$$2) \mathcal{L}_{e^{ax}}(s) = \frac{1}{s-a}, \text{ if } a < s$$

1.9 Laplace transform. Calculation of an expectation of a counting process (2)

$$F \rightarrow \mathbb{E}N_t$$

$$\mathbb{E}N_t = U(t) = \sum_{n=1}^{\infty} F^{n*}(t) = F(t) + \left(\sum_{n=1}^{\infty} F^{n*}(t) \right) * F(t)$$

$$\Leftrightarrow U = F + U * F = F + U * p \quad \text{if } F' = p \text{ exists}$$

\downarrow dist. fun. \downarrow densities

$$\int_{\mathbb{R}} U(x-y) dF(y) = \int_{\mathbb{R}} U(x-y) p(y) dy$$

$$\mathcal{L}_U(s) = \mathcal{L}_F(s) + \mathcal{L}_U(s) \mathcal{L}_p(s)$$

$$\mathcal{L}_p(s)$$

$$\boxed{\mathcal{L}_U(s) = \frac{\mathcal{L}_p(s)}{s(1 - \mathcal{L}_p(s))}}$$

$$\textcircled{1} F \rightarrow \mathcal{L}_p$$

$$\textcircled{2} \mathcal{L}_p \rightarrow \mathcal{L}_U$$

$$\textcircled{3} \mathcal{L}_U \rightarrow U$$

1.10 Laplace transform. Calculation of an expectation of a counting process (3)

Example: $S_n = S_{n-1} + \xi_n$, ξ_1, ξ_2, \dots have density $p(x)$

$$p(x) = \frac{e^{-x}}{2} + e^{-2x}, \quad x > 0$$

$$\mathbb{E}N_t = ?$$

$$\begin{aligned} \textcircled{1} p \rightarrow \mathcal{L}_p : \mathcal{L}_p(s) &= \frac{1}{2} \mathcal{L}_{e^{-x}}(s) + \mathcal{L}_{e^{-2x}}(s) \\ &= \frac{1}{2(s+1)} + \frac{1}{s+2} = \frac{3s+4}{2(s+1)(s+2)} \end{aligned}$$

$$\textcircled{2} \mathcal{L}_p \rightarrow \mathcal{L}_u : \mathcal{L}_u(s) = \frac{\mathcal{L}_p(s)}{s(1-\mathcal{L}_p(s))} = \frac{3s+4}{s^2(2s+3)}$$

$$\begin{aligned} \textcircled{3} \mathcal{L}_u \rightarrow u : \mathcal{L}_u(s) &= \frac{A}{s^2} + \frac{B}{s} + \frac{C}{2s+3} \\ &= \frac{A(2s+3) + B(2s^2+3s) + Cs^2}{s^2(2s+3)} \end{aligned}$$

$$3s+4 = (2B+C)s^2 + (2A+3B)s + 3A$$

$$A = \frac{4}{3}, \quad 2A+3B = 3 \Leftrightarrow B = \frac{1}{9}, \quad 2B+C = 0 \Leftrightarrow C = -\frac{2}{9}$$

$$u(t) = \frac{4}{3}t + \frac{1}{9}(1) - \frac{1}{9}e^{-3/2 t}$$

1.11 Limit theorems for renewal processes

$$S_n = S_{n-1} + \xi_n; \quad \xi_1, \xi_2, \dots \text{ iid } > 0 \text{ a.s.}$$

$$\text{Thm 1 } \mu = \mathbb{E}\xi_1 < \infty \Rightarrow \frac{N_t}{t} \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mu} \text{ a.s.}$$

(Analog to SLLN)

$$\text{SLLN: } \frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow[n \rightarrow \infty]{} \mu \text{ a.s.}$$

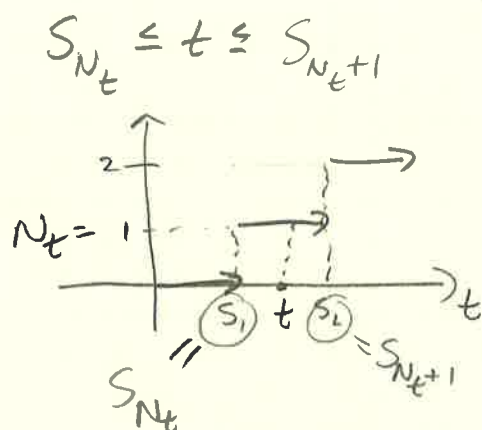
Thm 2: (Analog of CLT) $\sigma^2 = \text{Var } \xi_1 < \infty$

$$\text{Then } Z_t = \frac{N_t - t/\mu}{\sigma \sqrt{t}/\mu^{3/2}} \xrightarrow[t \rightarrow \infty]{} N(0,1)$$

$$P\{Z_t \leq x\} \rightarrow \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

$$\text{CLT: } \frac{\xi_1 + \dots + \xi_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{} N(0,1)$$

proof (thm 1)



$$\frac{N_t}{S_{N_t+1}} \leq \frac{N_t}{t} \leq \frac{N_t}{S_{N_t}}$$

$$\lim_{t \rightarrow \infty} \frac{N_t}{S_{N_t}} = \lim_{n \rightarrow \infty} \frac{n}{S_n} = \frac{1}{\mu} \text{ by SLLN}$$

$$\lim_{t \rightarrow \infty} \frac{N_t}{S_{N_t+1}} = \lim_{t \rightarrow \infty} \frac{N_t}{N_{t+1}} \cdot \lim_{t \rightarrow \infty} \frac{N_{t+1}}{S_{N_t+1}} = \frac{1}{\mu}$$

\parallel \parallel
 1 $1/\mu$

proof (thm 2)

$$P\left\{\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right\} \rightarrow \Phi(x), x \in \mathbb{R}$$

$$P\{S_n \leq n\mu + \sigma\sqrt{n}x\} \rightarrow \Phi(x)$$

$$\Leftrightarrow P\{N_t \geq n\}$$

(set complements)

$$n\mu \approx t, \quad n \approx t/\mu \quad (\text{for } n \text{ large enough})$$

$$n = \frac{t}{\mu} - \frac{\sigma\sqrt{n}}{\mu}x \approx \frac{t}{\mu} - \frac{\sigma\sqrt{t}}{\mu^{3/2}}x$$

$$\Rightarrow P\{Z_t \geq -x\} \rightarrow \Phi(x) \quad \Leftrightarrow P\{Z_t \leq x\} = 1 - P\{Z_t \geq -x\} \rightarrow 1 - \Phi(-x) = \Phi(x)$$

Poisson Processes

2.1 Definition of a Poisson process as a special example of a renewal process. Exact forms of the distributions of the renewal process and the counting process (1)

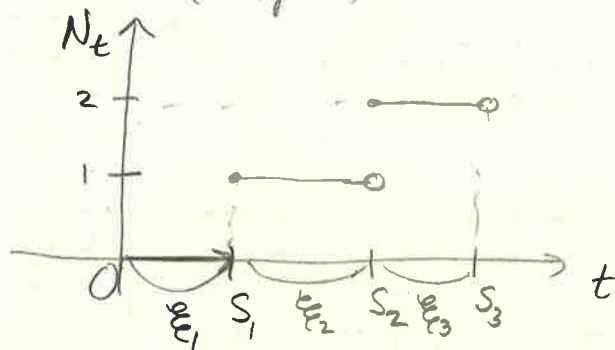
Renewal process

$$S_0 = 0, S_n = S_{n-1} + \xi_n, \xi_1, \xi_2, \dots \text{ i.i.d } > 0 \text{ a.s.}, \xi_i \sim F$$

$$N_t = \arg \max_k \{ S_k \leq t \} \quad (\text{Counting process})$$

$$U(t) = \mathbb{E} N_t = \sum_{n=1}^{\infty} F^{*n}(t)$$

$$L_u(s) = \frac{L_p(s)}{s(1-L_p(s))} : p \rightarrow L_p \rightarrow L_u \rightarrow u \quad (p = F')$$



$$L_u(s) = \int_{\mathbb{R}_+} e^{-sx} U(x) dx$$

2.2 ... (2)

Poisson process

Def 1: A process is a renewal process s.t.

$$\xi_i \sim p(x) = \lambda e^{-\lambda x} \mathbb{I}\{x > 0\}, \lambda - \text{intensity or rate}$$

Thm (i): A distribution function of S_n

$$F_{S_n}(x) = \begin{cases} 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}, & x > 0 \\ 0, & x < 0 \end{cases}$$

$$p_{S_n}(x) = \lambda \frac{(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} \mathbb{I}\{x > 0\}$$

$$(ii) \mathbb{P}\{N_t = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, N_t \sim \text{Poisson}(\lambda t)$$

2.3 ... (3)

Proof (i) $n=1$: $S_1 = \xi_1$
 $p_{S_1}(x) = \lambda e^{-\lambda x}, x > 0$

 $n \rightarrow n+1$

$$\begin{aligned}
 p_{S_{n+1}}(x) &= \int_0^x p_{S_n}(x-y) p_{\xi_{n+1}}(y) dy \\
 &= \int_0^x \frac{\lambda^n (x-y)^{n-1}}{(n-1)!} e^{-\lambda(x-y)} \lambda e^{-\lambda y} dy \\
 &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \int_0^x (x-y)^{n-1} dy = \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \frac{x^n}{n} \\
 &= \lambda \frac{(\lambda x)^n}{n!} e^{-\lambda x} \quad \square
 \end{aligned}$$

2.4 ... (4)

proof (ii)

$$P\{N_t = n\} = P\{S_n \leq t\} - P\{S_{n+1} \leq t\} \quad (=)$$

$$\{N_t = n\} = \underbrace{\{S_n \leq t\}}_A \cap \underbrace{\{S_{n+1} > t\}}_B$$

$$A \cap B = A \setminus B^c \quad \Rightarrow P\{A \cap B\} = P\{A\} - P\{B^c\}$$

Here: $B^c \subset A$

$$\begin{aligned}
 &= \left(1 - e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}\right) - \left(1 - e^{-\lambda t} \sum_{k=0}^n \frac{(\lambda t)^k}{k!}\right) \\
 &= e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad \square
 \end{aligned}$$

2.5 Memoryless property

A r.v. X possesses the memoryless property iff

$P\{X > u+v\} = P\{X > u\} P\{X > v\}$. If $P\{X > v\} > 0$, then

$$P\{X > u+v \mid X > v\} = P\{X > u\}$$

Thm 2: Let X be a r.v. with density $p(x)$, then
 X -memoryless $\Leftrightarrow p(x) = \lambda e^{-\lambda x}$

Ex buses arrive every 20 ± 2 minutes

$$v = 19 \text{ min}, u = 10 \text{ min}$$

$$(l.h.s.) P\{X > 29 | X > 19\} = 0 \text{ given the data}$$

$$(r.h.s.) P\{X > 10\} = 1$$

Thus, Poisson process is not appropriate

2.6. Other definitions of Poisson processes (1)

Def 2 N_t - an integer value process s.t.

$$0) N_0 = 0 \text{ a.s.}$$

$$1) N_t \text{ has independent increments: } \forall t_0 < t_1 < \dots < t_n, \\ N_{t_1} - N_{t_0}, \dots, N_{t_n} - N_{t_{n-1}} \text{ are independent}$$

$$2) N_t \text{ has stationary increments} \\ N_t - N_s \stackrel{d}{=} N_{t-s}$$

$$3) N_t - N_s \sim \text{Poisson}(\lambda(t-s)), t > s$$

$$3) \Rightarrow 2)$$

2.7 Other definitions of Poisson processes (2)

$$P\{N_{t+h} - N_t = 0\} = 1 - \lambda h + o(h), h \rightarrow 0$$

$$P\{N_{t+h} - N_t = 1\} = \lambda h + o(h), h \rightarrow 0$$

$$P\{N_{t+h} - N_t \geq 2\} = o(h), h \rightarrow 0$$

$$\lim_{h \rightarrow 0} \frac{1 - P\{N_{t+h} - N_t = 0\}}{h} = \lim_{h \rightarrow 0} \frac{1 - e^{-\lambda h}}{h} = \lambda$$

Def 3 N_t is a Poisson process, if

$$0) N_0 = 0$$

$$1) N_t \text{ has independent increments}$$

$$2) N_t \text{ has stationary increments}$$

$$3') \lim_{h \rightarrow 0} \frac{P\{N_{t+h} - N_t \geq 2\}}{P\{N_{t+h} - N_t = 1\}} = 0$$

2.8 Non-homogeneous Poisson processes (1)

$$N_t \sim \text{Pois}(\lambda t) \Rightarrow \mathbb{E}N_t = \lambda t$$

Def: Let $\Lambda(t)$ be a differentiable, increasing function s.t. $\Lambda(0) = 0$. Then, $X_t = N_t$ is a non-homogeneous Poisson process if,

if, 0) $N_0 = 0$

1) N_t has independent increments

2) $N_t - N_s \sim \text{Pois}(\Lambda(t) - \Lambda(s))$

2.9 Non-homogeneous Poisson processes (2) (NHPP)

$\lambda(t) = \Lambda'(t)$ - intensity function

Properties NHPP:

1) $\mathbb{E}N_t = \Lambda(t)$

$\Lambda(t) = \alpha t^\beta, \alpha > 0, \beta > 0$ (for example)

2) if $\lambda(t) = \text{const} \Rightarrow \Lambda(t) = \text{const} \cdot t$

3) $\Lambda(t)$ - differentiable $\Rightarrow \Lambda(t)$ - continuous \Rightarrow
 $\Lambda(t)$ - increasing

$\Rightarrow \exists \Lambda^{-1}(t)$. If Image $\Lambda(t) = \mathbb{R}_+$, $N_{\Lambda^{-1}(t)}$ - homogeneous P.P.

2.10 Relation between renewal theory and NHPP (1)

$$S_n = \arg\min_t \{N_t = n\}, \quad \xi_n = S_n - S_{n-1}$$

ξ_1, ξ_2, \dots - i.i.d.?

1) $p_{\xi}(x) = \lambda(x) e^{-\Lambda(x)}$

$$P\{\xi_1 \leq x\} = P\{S_1 \leq x\} = P\{N_x \geq 1\} = 1 - P\{N_x = 0\} \Leftrightarrow$$

$$\{S_1 > x\} = \{N_x < 1\}$$

$\Leftrightarrow 1 - e^{-\Lambda(x)}$

Take derivatives of both sides to finish proof \square .

2.11 ... (2)

$$S_k = \argmin_t \{N_t = k\}$$

$$\xi_k = S_k - S_{k-1}$$

$$1) p_{\xi_1}(t) = \lambda(t) e^{-\Lambda(t)}$$

$$2) p_{\xi_2|\xi_1}(t|s) = \lambda(t+s) e^{-\Lambda(t+s) + \Lambda(s)}$$

$$F_{(\xi_1, \xi_2)}(s, t) = P\{\xi_1 \leq s, \xi_2 \leq t\} = \int_0^s P\{\xi_1 \leq s, \xi_2 \leq t \mid \xi_1 = y\} p_{\xi_1}(y) dy$$

Since $y \leq s$

$$= \int_0^s P\{N_{t+y} - N_y \geq 1 \mid \xi_1 = y\} p_{\xi_1}(y) dy$$

independent

$$= \int_0^s (1 - e^{-\Lambda(t+y) + \Lambda(y)}) \lambda(y) e^{-\Lambda(y)} dy$$

$$p_{(\xi_1, \xi_2)}(s, t) = \frac{\partial}{\partial t} \left(\frac{\partial}{\partial s} F_{(\xi_1, \xi_2)}(s, t) \right)$$

$$= \frac{\partial}{\partial t} \left((1 - e^{-\Lambda(t+s) + \Lambda(s)}) \lambda(s) e^{-\Lambda(s)} \right)$$

$$= \lambda(t+s) e^{-\Lambda(t+s) + \Lambda(s)} \lambda(s) e^{-\Lambda(s)}$$

Then $p_{\xi_2|\xi_1}(t|s) = \frac{p_{(\xi_1, \xi_2)}(s, t)}{p_{\xi_1}(s)}$ finishes the proof \square .

2.12 ... (3)

ξ_1, ξ_2, \dots - i.i.d. ? (NHPP can be obtained from renewal process iff NHPP is homogeneous PP)

$$p_{\xi_1}(t) = p_{\xi_2|\xi_1}(t|s), \quad \forall t, s > 0$$

$$\lambda(t) e^{-\Lambda(t)} = \lambda(t+s) e^{-\Lambda(t+s) + \Lambda(s)}$$

$$\left(\int_0^T \dots dt \right) : e^{-\Lambda(0)} - e^{-\Lambda(T)} = e^{-\Lambda(s)} - e^{-\Lambda(T+s) + \Lambda(s)}$$

$$\Lambda(T) = \Lambda(T+s) - \Lambda(s), \quad \forall s, T > 0 \Rightarrow \Lambda(t) = \lambda t$$

correct

$\Lambda(t)$ - increasing

2.13 Elements of queuing theory. $M/G/k$ systems (1)

$$\begin{aligned}
 P\{N_{t+h} - N_t = 0\} &= 1 - \lambda h + o(h) \\
 P\{N_{t+h} - N_t = 1\} &= \lambda h + o(h) \\
 P\{N_{t+h} - N_t \geq 2\} &= o(h)
 \end{aligned}$$

 $M/G/k$

I) Arrival Process: M - memoryless (Poisson)
 D - deterministic
 G - general

II) Service time (M, D, G)

III) A number of services ($1, 2, \dots, \infty$)

 $M/G/\infty$ $\tau > 0$ (time moment)

$N(t)$
 Customer arrivals \rightarrow $N_1(t)$ - still being served at $\tau : \lambda_1(t) = \lambda(1 - G(\tau - t))$
 \rightarrow $N_2(t)$ - already completed by $\tau : \lambda_2(t) = \lambda G(\tau - t)$

$$\begin{aligned}
 P\{N_1(t+\delta) - N_1(t) = 1\} &= P\{N(t+\delta) - N(t) = 1\} \cdot (P\{Y > \tau - t\} + o(\delta)) \\
 &= (\delta\lambda + o(\delta)) (1 - G(\tau - t) + o(\delta)) \\
 &= \boxed{\lambda\delta(1 - G(\tau - t) + o(\delta))}
 \end{aligned}$$

2.14 ... (2)

$$P\{N_1(t) = n_1, N_2(t) = n_2\} = P\{N_1(t) = n_1, N_2(t) = n_2 \mid N(t) = n_1 + n_2\} \cdot P\{N(t) = n_1 + n_2\}$$

$$= C_{n_1+n_2}^{n_1} (1 - G(\tau - t))^{n_1} G(\tau - t)^{n_2} \cdot e^{-\lambda t} \frac{(\lambda t)^{n_1+n_2}}{(n_1+n_2)!}$$

$$= \frac{\lambda t (1 - G(\tau - t))^{n_1}}{n_1!} e^{-\lambda(1 - G(\tau - t))} \cdot \frac{\lambda t (G(\tau - t))^{n_2}}{n_2!} e^{-\lambda G(\tau - t)}$$

$$= P\{N_1(t) = n_1\} \cdot P\{N_2(t) = n_2\}$$

Therefore $N_1 \perp N_2$

2.15 Compound Poisson Processes (1)

$$X_t = \sum_{k=1}^{N_t} \xi_k, \quad \xi_1, \xi_2, \dots - \text{i.i.d.}, \quad N_t - \text{P.P. with intensity } \lambda$$

and ξ_1, ξ_2, \dots and N_t are independent

ξ_1, ξ_2, \dots claim sizes

N_t - amount of claims until time t (Insurance interpretation)

X_t - aggregated claim amount

1) Probability generating function (PGF)

ξ - integer, ≥ 0 values

$$\boxed{\phi_\xi(u) = \mathbb{E}[u^\xi], \quad |u| < 1}$$

$$\xi_1 \perp \xi_2 \Rightarrow \phi_{\xi_1 + \xi_2}(u) = \phi_{\xi_1}(u) \phi_{\xi_2}(u)$$

2) Moment-generating function (MGF)

$$\boxed{L_\xi(u) = \mathbb{E}[e^{-u\xi}], \quad \xi \geq 0, u > 0}$$

2.16 ... (2)

3) Characteristic function

$$\phi_\xi(u) = \mathbb{E}[e^{iu\xi}], \quad u \in \mathbb{R}, \forall \xi, \quad \phi_\xi: \mathbb{R} \rightarrow \mathbb{C}, \quad \xi_1 \perp \xi_2 \Rightarrow \phi_{\xi_1 + \xi_2}(u) = \phi_{\xi_1}(u) \phi_{\xi_2}(u)$$

$$\text{Thm } \boxed{\phi_{X_t - X_s}(u) = e^{\lambda(t-s)(\phi_\xi(u) - 1)}}$$

$$\begin{aligned} \text{Proof: } \text{lhs} &= \mathbb{E} e^{iu(X_t - X_s)} = \sum_{k=0}^{\infty} \mathbb{E} \left[e^{iu(X_t - X_s)} \mid N_t - N_s = k \right] P\{N_t - N_s = k\} \\ &= \sum_{k=0}^{\infty} (\phi_\xi(u))^k e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^k}{k!} \end{aligned}$$

$\xi_1 + \dots + \xi_k$ from 11

□

2.17 ... (3)

$$X_t = \sum_{k=1}^{N_t} \xi_k, \quad \xi \text{ can be any random variable}$$

$$\xi: \phi_\xi(u) = E[e^{iu\xi}]$$

$$\phi: \mathbb{R} \rightarrow \mathbb{C}$$

$$\xi_1 \perp \xi_2 \Rightarrow \phi_{\xi_1 + \xi_2}(u) = \phi_{\xi_1}(u) \phi_{\xi_2}(u)$$

$$\text{Thm: } \phi_{X_t - X_s}(u) = e^{\lambda(t-s)(\phi_\xi(u) - 1)}, \quad t > s \geq 0$$

proof:

$$\text{lhs} = E[e^{iu(X_t - X_s)}]$$

$$= \sum_{k=0}^{\infty} E[e^{iu(X_t - X_s)} | N_t - N_s = k] \cdot P\{N_t - N_s = k\}$$

\downarrow $\text{d} \parallel$ $\text{since } \xi_1, \xi_2, \dots \perp N_t \sim \text{Pois}(\lambda(t-s))$
 ξ_1, \dots, ξ_k

$$= \sum_{k=0}^{\infty} [\phi_\xi(u)]^k \cdot e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^k}{k!} \quad \square$$

2.18 ... (4)

$$\text{Corollary: } \begin{cases} EX_t = \lambda t E\xi, \\ \text{Var } X_t = \lambda t E\xi^2 \end{cases}$$

proof: $E[\xi^r] < \infty \Rightarrow \phi(u)$ is r -times differentiable at 0
 and $\phi^{(r)}(0) = i^r E\xi^r$

$$EX_t = \frac{\phi'_{X_t}(0)}{i} = \frac{\lambda t (\phi'_\xi(0) \cdot \phi_{X_t}(0))}{i} = \lambda t E\xi, \quad \square$$

\downarrow
 $i = E\xi$

3.1 Definition of a Markov chain. Some examples

Def: A Markov chain - S_n , $n=0,1,2,\dots$
 S' - state space (countable)

$$P\{S_n = j \mid S_{n-1} = i_{n-1}, \dots, S_0 = i_0\} = P\{S_n = j \mid S_{n-1} = i_{n-1}\}$$

$$i_0, \dots, i_{n-1}, j \in S' \text{ and } P\{S_{n-1} = i_{n-1}, \dots, S_0 = i_0\} \neq 0$$

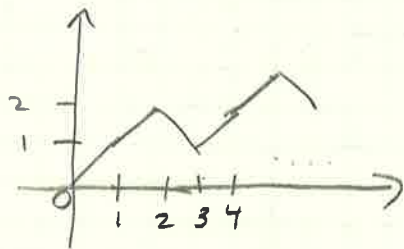
$$P\{S_n = i_n, S_{n-1} = i_{n-1}, \dots, S_0 = i_0\} = P\{S_n = i_n \mid S_{n-1} = i_{n-1}, \dots, S_0 = i_0\} \\ \cdot P\{S_{n-1} = i_{n-1}, \dots, S_0 = i_0\}$$

$$= P\{S_n = i_n \mid S_{n-1} = i_{n-1}\} \cdot P\{S_{n-1} = i_{n-1}, \dots, S_0 = i_0\}$$

$$= P\{S_n = i_n \mid S_{n-1} = i_{n-1}\} \cdot P\{S_{n-1} = i_{n-1} \mid S_{n-2} = i_{n-2}\} \\ \cdot \dots \cdot P\{S_1 = i_1 \mid S_0 = i_0\} \cdot P\{S_0 = i_0\}$$

Ex ① Random walk (not a renewal process)

$$S_0 = 0, S_n = S_{n-1} + \xi_n, \xi_1, \xi_2, \dots \text{ i.i.d. } \sim \begin{cases} 1, \text{ w.p. } p \\ -1, \text{ w.p. } 1-p \end{cases}$$



$$P\{S_n = j \mid S_{n-1} = i_{n-1}\} = \begin{cases} p, & j = i_{n-1} + 1 \\ 1-p, & j = i_{n-1} - 1 \\ 0, & \text{otherwise} \end{cases}$$

② Taxis in the airport

1 taxi at any 1 moment, $n=1,2,3,\dots$

X_k = # people waiting for a taxi at time k

Y_k = # people arriving at k

$$X_k = Y_k + (X_{k-1} - 1)_+ = \begin{cases} Y_k, & \text{if } X_{k-1} = 0 \\ Y_k + X_{k-1} - 1, & \text{if } X_{k-1} - 1 > 0 \end{cases}$$

③ X_n : $P\{X_n = j \mid X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} = P\{X_n = j \mid X_{n-1} = i_{n-1}, \dots, X_{n-m} = i_{n-m}\}$
 $m \in \mathbb{N}$, fixed (X_n is not a Markov chain)

$S_n = (X_n, \dots, X_{n-m-1})$, $n = (m-1), m, \dots$ S_n is a Markov chain

3.2 Matrix representation of a Markov chain. Transition matrix. Chapman-Kolmogorov equation.

3.2

Matrix representation

$$S = (1, 2, \dots, M)$$

$$P\{X_n = j | X_{n-1} = i\} = p_{ij} \text{ - homogeneous (no dependence on } n\text{)}$$

$$P = (p_{ij})_{i,j=1}^M \text{ - transition matrix}$$

$$\sum_{j=1}^M p_{ij} = 1, \forall i; \quad p_{ij} \geq 0 \quad \} \text{ - stochastic matrix}$$

$$p_{ij}^{(m)} = P\{X_{n+m} = j | X_n = i\}$$

$$P^{(m)} = (p_{ij}^{(m)}) \text{ - } m\text{-step transition matrix}$$

$$\text{Thm: } \boxed{P^{(m)} = P^m}$$

$$\text{proof: } p_{ij}^{(m)} = \sum_{k=1}^M P\{X_{n+m} = j | X_{n+m-1} = k, X_n = i\} \\ = \sum_{k=1}^M P\{X_{n+m} = j | X_{n+m-1} = k | X_n = i\}$$

$$\begin{aligned} & \stackrel{(\text{Markov property})}{=} \sum_{k=1}^M P\{X_{n+m} = j | X_{n+m-1} = k\} P\{X_{n+m-1} = k | X_n = i\} \\ & = \sum_{k=1}^M p_{kj} p_{ik}^{(m-1)} \Rightarrow P^{(m)} = P \cdot P^{(m-1)} = \dots = P^m \quad \square \end{aligned}$$

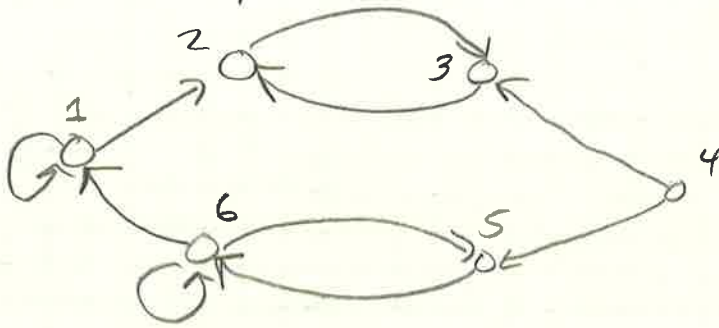
$$P\{X_k = j\} := \pi_j^{(k)}, \quad (\pi_1^{(k)}, \dots, \pi_m^{(k)}) := \vec{\pi}^{(k)}$$

$$\pi_j^{(k)} = \sum_{i=1}^M P\{X_k = j | X_{k-1} = i\} P\{X_{k-1} = i\}$$

$$= \sum_{i=1}^M p_{ij} \pi_i^{(k-1)} \Rightarrow \vec{\pi}^{(k)} = \vec{\pi}^{(k-1)} \cdot P = \vec{\pi}^{(0)} P^k$$

$$\vec{\pi}^* \text{ - stationary distribution for Markov chain if } \boxed{\vec{\pi}^* P = \vec{\pi}^*}$$

Graphical representation



1 node = 1 state

 $i, j\text{-arc} \Leftrightarrow p_{ij} \neq 0$

$$P = \begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}_{6 \times 6}$$

Def (1) j is accessible from $i \exists$ walk (path) from i to j ($i \rightarrow j$)

$$1 \rightarrow 3 ; 1 \nrightarrow 4$$

(2) i and j communicate if $i \rightarrow j$ and $j \rightarrow i$ ($i \leftrightarrow j$)

$$2 \leftrightarrow 3$$

(3) \mathcal{Y} -set, and relation \sim called an equivalence relation $a \sim a, a \in \mathcal{Y}$ - reflexivity $a \sim b \Rightarrow b \sim a, a, b \in \mathcal{Y}$ - symmetry $a \sim b, b \sim c \Rightarrow a \sim c, a, b, c \in \mathcal{Y}$ - transitivity $\mathcal{Y} = \sqcup B_i$ (\sqcup - disjoint union), B_i - equivalence classes

$$\Leftrightarrow \begin{array}{c} \nearrow \\ \searrow \end{array} \Leftrightarrow P \begin{pmatrix} \vdots \\ 0 \\ \vdots \end{pmatrix}$$

 B_1, B_2, \dots - equivalence classes

$$\forall j \in B_i, \forall k \in S \begin{cases} k \in B_i, k \leftrightarrow j \\ k \notin B_i, k \nleftrightarrow j \end{cases}$$

 $2 \leftrightarrow 3, 5 \leftrightarrow 6, 1, 4$ are the four equivalence classes

3.4 ... (2)

Def: i is recurrent, $\forall j: i \rightarrow j \Rightarrow j \rightarrow i$ i is transient if it's not recurrent $\Leftrightarrow \exists j: i \rightarrow j, j \nrightarrow i$ ex: ①, ④, ⑤, ⑥ - transient

②, ③ - recurrent

Thm: In 1 class of equivalence, all states are either recurrent or transient.

proof k -transient: $\exists j: k \rightarrow j, j \nrightarrow k$

$i, k \in 1 \text{ class} \Rightarrow i \rightarrow k \rightarrow j$, but $j \nrightarrow i: j \rightarrow i \rightarrow k$ is a contradiction \square

3.5 ... (3)

Def: Period of a state i is $\text{gcd}\{n: p_{ii}(n) \neq 0\} =: d(i)$

$d(i) = 1 \Rightarrow i$ -aperiodic

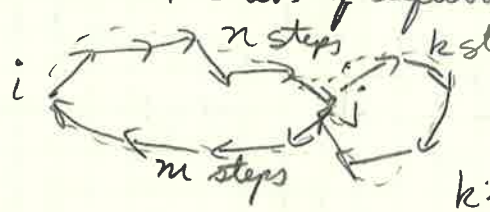
(ex) $d(1) = 1 = d(4) = d(5) = d(6)$
 $d(2) = 2 = d(3)$

④ has no return, so $d(4) = 1$ by convention

Thm: All elements in 1 class of equivalence have the same period

proof:

$p_{ii}(n+m+k) \neq 0$



$k: p_{jj}(k) \neq 0 \Rightarrow n+m+k \mid d(i)$

$\Rightarrow k \mid d(i) \Rightarrow \left. \begin{matrix} d(i) \mid d(j) \\ d(j) \mid d(i) \end{matrix} \right\} \Rightarrow d(i) = d(j)$

\square

3.6 Ergodic chains. Ergodic Theorem (1)

Matrix representation

$$\underline{P}, \underline{\pi}^{(k)}$$

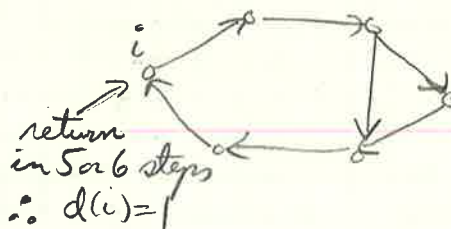
$$\underline{P}^{(n)} = \underline{P}^n$$

Ergodic Markov chains:

- 1 class of equivalence
- recurrent
- $d(i) = 1$ (aperiodic)

Graphical representation

classes of equivalence.
 recurrent/transient
 $d(i)$ - period



Prop: Markov chain is ergodic $\Leftrightarrow \exists m \in \mathbb{N} : p_{ij}(m) \neq 0, \forall i, j \in S$ (*)

If chain is ergodic, then (*) hold $\forall m \geq (M-1)^2 + 1$.

3.7 ... (2)

Ergodic theorem: Let X_t -ergodic Markov chain, i.e. X_t has 1 class of equivalence, recurrent and aperiodic. Then,

$$\boxed{\exists \lim_{n \rightarrow \infty} p_{ij}(n) = \pi_j^* > 0 \text{ (doesn't depend on } i)}$$

$$\sum_{j=1}^M \pi_j^* = 1 \quad \vec{\pi}^* = (\pi_1^*, \dots, \pi_M^*)$$

Corr(i) $\vec{\pi}^*$ -stationary distribution: $\vec{\pi}^* \underline{P} = \vec{\pi}^*$

(ii) $\lim_{n \rightarrow \infty} P\{X_n = j\} = \pi_j^* \quad [\pi_j^{(0)} \text{ is arbitrary}]$

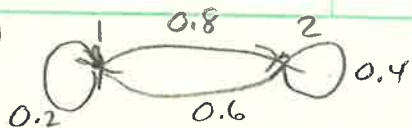
proof (i) $i = 1, \dots, M$

$$\begin{aligned} (\vec{\pi}^* \underline{P})_i &= \sum_{j=1}^M \pi_j^* p_{ji} = \sum_{j=1}^M \lim_{n \rightarrow \infty} p_{kj}(n) p_{ji} \quad (k \in 1, \dots, M) \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^M \underbrace{p_{kj}(n) p_{ji}}_{\underline{P}^{(n)} \underline{P} = \underline{P}^{(n+1)}} \\ &= \lim_{n \rightarrow \infty} p_{ki}(n+1) = \pi_i^* \quad \square \end{aligned}$$

proof (ii) $\lim_{n \rightarrow \infty} \pi_j^{(n)} = \lim_{n \rightarrow \infty} \sum_{k=1}^M \pi_k^{(0)} p_{kj}(n) \quad \pi_j^{(0)} \text{ is arbitrary}$

$$\begin{aligned} &\quad \vec{\pi}^{(n)} = \vec{\pi}^{(0)} \underline{P}^{(n)} \\ &= \sum_{k=1}^M \pi_k^{(0)} \underbrace{\lim_{n \rightarrow \infty} p_{kj}(n)}_{= \pi_j^*} = \pi_j^* \sum_{k=1}^M \pi_k^{(0)} = \pi_j^* \quad \square \end{aligned}$$

(ex)



$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix}$$

$$\vec{\pi}^* = (a, b); \quad \vec{\pi}^* P = \vec{\pi}^*$$

$$(a \ b) \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix} = (a \ b)$$

$$\left. \begin{aligned} 0.2a + 0.6b &= a \\ 0.8a + 0.4b &= b \end{aligned} \right\} \Rightarrow a = \frac{3}{7}, \ b = \frac{4}{7}$$

$$P\{X_n = 1\} \rightarrow \frac{3}{7}$$

$$P\{X_n = 2\} \rightarrow \frac{4}{7}$$

4.1 Random vector. Definition and main properties

$$\xi \sim N(\mu, \sigma^2), \quad p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \sigma > 0, \mu \in \mathbb{R}$$

$$\phi(u) = e^{iu\mu - \frac{1}{2}u^2\sigma^2}$$

$$\left. \begin{array}{l} X_1, X_2 \sim N(0, 1) \\ \text{cor}(X_1, X_2) = 0 \end{array} \right\} \not\Rightarrow X_1 \perp\!\!\!\perp X_2$$

$$P\{X = \mu\} = 1 \Rightarrow \sigma = 0$$

Def: A random vector $\vec{X} = (X_1, \dots, X_n)$ is Gaussian iff
 $\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n, \quad \sum_{k=1}^n \lambda_k X_k \sim N$

4.2 Gaussian vector. Definition and main properties

Thm: \vec{X} - Gaussian iff any of the following holds:

$$(i) \quad \phi_{\vec{X}}(\vec{u}) = \mathbb{E}[e^{i\langle \vec{u}, \vec{X} \rangle}] = \exp\left\{i\langle \vec{u}, \vec{\mu} \rangle - \frac{1}{2}\vec{u}^T C \vec{u}\right\}$$

$\vec{\mu} \in \mathbb{R}^n$; C - symmetric, positive semidefinite (size $n \times n$)

$$(ii) \quad \vec{X} = A\vec{X}^0 + \vec{\mu}; \quad A \in \text{Mat}(n \times n), \quad \vec{X}^0 - \text{standard normal vector}$$

Remark: $\vec{\mu} = (\mathbb{E}X_1, \dots, \mathbb{E}X_n)$

$$C = (C_{jk})_{j,k=1}^n; \quad C_{jk} = \text{cov}(X_j, X_k)$$

$$\sum_{k,j=1}^n u_k C_{kj} u_j \geq 0, \quad \forall u \in \mathbb{R}^n \Leftrightarrow u^T C u \geq 0, \quad \forall u \in \mathbb{R}^n$$

$$\sum_{k,j=1}^n u_k \text{cov}(X_j, X_k) u_j = \text{cov}\left(\sum_{j=1}^n u_j X_j, \sum_{k=1}^n u_k X_k\right)$$

$$= \text{var}\left(\sum_{j=1}^n u_j X_j\right) \geq 0$$

$$A = C^{1/2}: \quad AA^T = C \Rightarrow \exists U: U^{-1} = U^T: C = U^T \begin{pmatrix} d_1 & 0 \\ & \ddots \\ 0 & d_n \end{pmatrix} U$$

$$A = U^T \begin{pmatrix} \sqrt{d_1} & 0 \\ & \ddots \\ 0 & \sqrt{d_n} \end{pmatrix} U \Rightarrow C = AA^T$$

Proof: Def $\Leftrightarrow (i) \Leftrightarrow (ii)$

Def $\Rightarrow (i)$: $\langle \vec{u}, \vec{x} \rangle \sim N$

$$\phi_{\vec{x}}(\vec{u}) = \mathbb{E} e^{i \langle \vec{u}, \vec{x} \rangle} = \phi_{\vec{x}}(1) = e^{i \mu_{\vec{x}} - \frac{1}{2} \sigma_{\vec{x}}^2}$$

$$\mu_{\vec{x}} = \mathbb{E} \left[\sum_{k=1}^n u_k X_k \right] = \sum_{k=1}^n u_k \mathbb{E} X_k = \sum_{k=1}^n u_k \mu_k = \langle \vec{\mu}, \vec{u} \rangle$$

$$\sigma_{\vec{x}}^2 = \text{cov} \left(\sum_{k=1}^n u_k X_k, \sum_{j=1}^n u_j X_j \right)$$

$$= \sum_{k=1}^n \sum_{j=1}^n u_k \text{cov}(X_k, X_j) u_j = \vec{u}^T C \vec{u}$$

(i) \Rightarrow Def By definition of ϕ for Gaussian

(ii) \Rightarrow (i)

$$\vec{X}^0\text{-Gaussian} \Rightarrow \phi_{\vec{X}^0}(\vec{u}) = \exp \left\{ -\frac{1}{2} \vec{u}^T \vec{u} \right\}$$

$$\begin{aligned} \phi_{\vec{x}}(\vec{u}) &= \mathbb{E} \left[e^{i \langle \vec{u}, A \vec{X}^0 + \vec{\mu} \rangle} \right] = \mathbb{E} \left[e^{i \langle \vec{u}, \vec{\mu} \rangle + \langle \vec{u}, A \vec{X}^0 \rangle} \right] \\ &= e^{i \langle \vec{u}, \vec{\mu} \rangle} \phi_{\vec{X}^0}(A^T \vec{u}) = e^{i \langle \vec{u}, \vec{\mu} \rangle} e^{-\frac{1}{2} \vec{u}^T \underbrace{A A^T}_C \vec{u}} \\ &= e^{i \langle \vec{u}, \vec{\mu} \rangle} e^{-\frac{1}{2} \vec{u}^T C \vec{u}} \end{aligned}$$

(i) \Rightarrow (ii) $A = C^{1/2} \quad \square$

4.3 Connection between independence of normal random variates and absence of correlation

Thm: Let $X_1, X_2 \sim N(0, 1)$ and $\text{cov}(X_1, X_2) = 0$, then $X_1 \perp\!\!\!\perp X_2 \Leftrightarrow (X_1, X_2)$ - Gaussian vector

Proof (\Rightarrow) $\lambda_1 X_1 + \lambda_2 X_2 \sim N \Rightarrow (X_1, X_2)$ - Gaussian vector

$$(\Leftarrow) C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow A = C^{1/2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = C$$

$$(ii) \Rightarrow \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = A \vec{X}^0 + \vec{\mu} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_1^0 \\ X_2^0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} X_1^0 \\ X_2^0 \end{pmatrix} \quad \square$$

ex $X_1 \sim N(0,1)$, $X_2 := |X_1| \cdot \xi$, $\xi = \begin{cases} 1 & \text{w.p. } 1/2 \\ -1 & \text{w.p. } 1/2 \end{cases}$, $\xi \perp\!\!\!\perp X_1$

1) $X_2 \sim N(0,1)$, $x > 0$

$$\begin{aligned} P\{X_2 \leq x\} &= P\{|X_1| \leq x \mid \xi = 1\} P\{\xi = 1\} \\ &\quad + P\{|X_1| \geq x \mid \xi = -1\} P\{\xi = -1\} \\ &\stackrel{\xi \perp\!\!\!\perp X_1}{=} P\{|X_1| \leq x\} \cdot \frac{1}{2} + P\{|X_1| \geq x\} \cdot \frac{1}{2} \\ &= \frac{1}{2} [1 + P\{|X_1| \leq x\}] = P\{X_1 \leq x\} \end{aligned}$$

2) $\text{cov}(X_1, X_2) = 0$

$$\begin{aligned} E[X_1 X_2] - E X_1 E X_2 &= E[X_1 |X_1| \xi] \\ &= E[X_1 |X_1|] \underbrace{E \xi}_{=0} - 0 \cdot E X_2 \\ &= 0 - 0 = 0 \quad \square \end{aligned}$$

3) X_1, X_2 are dependent

Assume $X_1 \perp\!\!\!\perp X_2 \Rightarrow (X_1, X_2) - \text{Gaussian}$

$$Y = X_1 - X_2 = X_1 - |X_1| \xi \sim N$$

$\{Y > 0\}$ when $X_1 > 0$ and $\xi = -1$

$$\begin{aligned} P\{Y > 0\} &\geq P\{X_1 > 0 \cap \xi = -1\} = P\{X_1 > 0\} P\{\xi = -1\} \\ &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \end{aligned}$$

$$P\{Y = 0\} \geq \frac{1}{4} \text{ since } \text{var}(Y) \neq 0$$

Thus Y is not Gaussian

4.4 Definition of a Gaussian process. Covariance function (1)

Def: A Gaussian process X_t is a stochastic process s.t.

$\forall t_1, t_2, \dots, t_n: (X_{t_1}, \dots, X_{t_n})$ - Gaussian vector

$m(t) = \mathbb{E}X_t$ - mathematical expectation $m: \mathbb{R}_+ \rightarrow \mathbb{R}$

$K(t, s) = \text{cov}(X_{t_1}, X_{t_2})$, $K: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$

$K(t, t) = \text{Var } X_t$ and $K(t, s) = K(s, t)$

K - a positive semidefinite function (p.s.d.)

$\forall (t_1, \dots, t_n) \in \mathbb{R}_+^n$,

$\forall (u_1, \dots, u_n) \in \mathbb{R}^n$

$$\sum_{k=1}^n \sum_{j=1}^n u_k u_j K(t_k, t_j) \geq 0$$

$$\Leftrightarrow \text{cov}\left(\sum_{k=1}^n u_k X_{t_k}, \sum_{j=1}^n u_j X_{t_j}\right) = \text{var}\left(\sum_{k=1}^n u_k X_{t_k}\right) \geq 0$$

4.5 Definition of a Gaussian process. Covariance function (2)

Thm: $m: \mathbb{R}_+ \rightarrow \mathbb{R}$, $K: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$, K - symmetric and p.s.d.

Then, \exists Gaussian process X_t : $\mathbb{E}X_t = m(t)$, $\text{cov}(X_t, X_s) = K(t, s)$

Gaussian r.v.: $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$

Gaussian vector: $\vec{\mu} \in \mathbb{R}^n$, $C \in \text{Mat}(n, n)$ and symmetric and p.s.d.

Gaussian process: $m: \mathbb{R}_+ \rightarrow \mathbb{R}$, $K: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ (sym. and p.s.d.)

Ex: $K(t, s) = |t - s|$ is not p.s.d.

Assume p.s.d. $\Rightarrow \exists X_t: \text{cov}(X_t, X_s) = |t - s|$

let $t = s$, $\text{var } X_t = 0 \Rightarrow X_t = f(t)$ - deterministic

$$\text{cov}(X_t, X_s) = \mathbb{E}[X_t X_s] - \mathbb{E}X_t \mathbb{E}X_s = f(t)f(s) - f(t)f(s) = 0 \neq |t - s|$$

Contradiction \square

$$\lambda_1(B_b - B_a) + \lambda_2(B_d - B_c) = \lambda_1 B_b - \lambda_1 B_a + \lambda_2 B_d - \lambda_2 B_c \sim N(0,0)$$

$$\Rightarrow \begin{bmatrix} B_b - B_a \\ B_d - B_c \end{bmatrix} \sim \text{Gaussian vector}$$

$$\Rightarrow B_b - B_a \perp B_d - B_c$$

$$(2) (B_t, B_s) \sim \text{Gaussian vector} \Rightarrow B_t - B_s \sim N$$

$$\mathbb{E}[B_t - B_s] = \mathbb{E}B_t - \mathbb{E}B_s = m(t) - m(s) = 0 - 0 = 0$$

$$\begin{aligned} \text{Var}[B_t - B_s] &= \text{cov}(B_t - B_s, B_t - B_s) = \text{cov}(B_t, B_t) - 2\text{cov}(B_t, B_s) + \text{cov}(B_s, B_s) \\ &= \min(t, t) - 2\min(t, s) + \min(s, s) \\ &= t - 2s + s = t - s \quad \square \end{aligned}$$

$$\text{Def 2} \Rightarrow \text{Def 1} \quad t_1 < t_2 < \dots < t_n$$

$$\begin{aligned} \sum_{k=1}^n \lambda_k B_{t_k} &= \lambda_n (B_{t_n} - B_{t_{n-1}}) + (\lambda_n + \lambda_{n-1}) B_{t_{n-1}} + \sum_{k=1}^{n-2} \lambda_k B_{t_k} \\ &= \sum_{k=1}^n d_k (B_{t_k} - B_{t_{k-1}}) \sim N_{\dots} \quad (t_0 = 0) \end{aligned}$$

$$\Rightarrow (B_{t_1}, \dots, B_{t_n}) \sim \text{Gaussian vector} \Rightarrow B_t \sim \text{Gaussian}$$

$$B_t \sim N(0, t) \Rightarrow m(t) = \mathbb{E}B_t = 0 \quad (t, s \text{ } t > s)$$

$$K(t, s) = \text{cov}(B_t, B_s) = \text{cov}(B_t - B_s + B_s, B_s)$$

$$= \text{cov}(B_t - B_s, B_s) + \text{cov}(B_s, B_s)$$

$$= \underbrace{\text{cov}(B_t - B_s, B_s - B_0)}_{\parallel} + \text{cov}(B_s, B_s)$$

$$= \underbrace{0}_{\parallel} = \text{var}(B_s) = s$$

$$\text{If } s > t, K(t, s) = t$$

$$\therefore K(t, s) = \min(t, s) \quad \square$$

4.7 Modification of a process. Kolmogorov continuity theorem

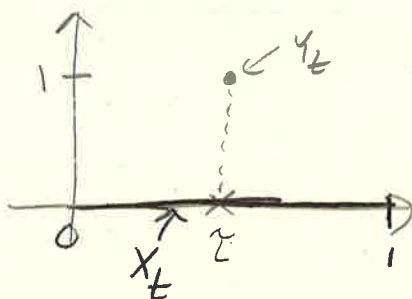
Kolmogorov continuity theorem

Def X_t, Y_t are stochastically equivalent if
 $P\{X_t = Y_t\} = 1, \forall t \geq 0$

Ex: $X_t = 0, \forall t \in [0, 1]$

$Y_t = \mathbb{1}\{\tau = t\}, \tau \sim \text{Unif}(0, 1)$

$$P\{X_t = Y_t\} = P\{Y_t = 0\} = P\{t \neq \tau\} = 1$$



Thm If $\exists C, \alpha, \beta > 0$ s.t. $E[|X_t - X_s|^\alpha] \leq C|t-s|^{1+\beta}$
 $\forall t, s \in [a, b]$, then $\exists Y_t$ that is stochastically
 equivalent to X_t s.t. Y_t has continuous trajectories,
 i.e., X_t has a continuous modification.

$$\text{Ex: } E[|B_t - B_s|^4] = (t-s)^2 \underbrace{E[\xi^4]}_3 = 3(t-s)^2$$

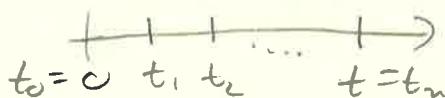
$N(0, t-s)$

$$B_t - B_s = \sqrt{t-s} \xi, \xi \sim N(0, 1)$$

$$\Rightarrow C = 3, \beta = 1, \alpha = 4$$

4.8 Main properties of Brownian Motion

① Quadratic variation



$$\lim_{n \rightarrow \infty} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 = t$$

$\nearrow S_n$

$\nwarrow \max |t_i - t_{i-1}| \rightarrow 0 \text{ as } n \rightarrow \infty$

$$\mathbb{E}(S_n - t)^2 \xrightarrow{n \rightarrow \infty} 0 \iff \lim_{n \rightarrow \infty} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 = t \quad \text{quadratic variation}$$

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n |B_{t_k} - B_{t_{k-1}}| = \infty \quad \text{variation}$$

② B_t - everywhere continuous, but nowhere differentiable

$$B_{t+h} \xrightarrow[h \rightarrow 0]{P} B_t, \quad \forall t \geq 0$$

③ $\lim_{t \rightarrow \infty} \frac{B_t}{t} = 0 \quad \text{a.s.}$

$$\lim_{t \rightarrow \infty} \frac{B_t}{\sqrt{t}} = \infty \quad \text{a.s.}$$

Law of iterated logarithm:

$$\lim_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log(\log t)}} = 1$$

5.1 Two types of stationarity (1)

StationarityDef 1: X_t is (strictly) stationary if (S.S.)

$$(X_{t_1+h}, \dots, X_{t_n+h}) \stackrel{d}{=} (X_{t_1}, \dots, X_{t_n}), \forall (t_1, \dots, t_n) \in \mathbb{R}_+^n, \forall h > 0$$

Def 2: X_t is (weakly) stationary if (wide-sense stationarity) (W.S.)

$$1. \quad m(t) = \mathbb{E}X_t = \text{const}$$

$$2. \quad K(t, s) = \text{cov}(X_t, X_s) = k(t+h, s+h), \quad \forall (t, s) \in \mathbb{R}_+^2, \forall h > 0$$

$$\Leftrightarrow \gamma: \mathbb{R} \rightarrow \mathbb{R} : K(t, s) = \gamma(t-s)$$

Properties of $\gamma(\cdot)$:

$$1) \quad \boxed{\gamma(0) \geq 0}$$

$$\text{cov}(X_t, X_t) = \text{Var}(X_t)$$

$$2) \quad \boxed{|\gamma(t)| \leq \gamma(0)}$$

$$|\text{cov}(X_t, X_0)| \leq \sqrt{\text{Var} X_t} \cdot \sqrt{\text{Var} X_0} \quad \text{Cauchy-Schwarz}$$

$$= \sqrt{\gamma(0)} \cdot \sqrt{\gamma(0)} = \gamma(0)$$

$$3) \quad \boxed{\gamma \text{ is even}}$$

$$\gamma(t) = \text{cov}(X_t, X_0) = \text{cov}(X_0, X_t) = \gamma(-t)$$

Ex:

$$(i) \quad \mathbb{E}X_t^2 < \infty$$

 X_t is strictly stationary $\Rightarrow X_t$ is weakly stationary(ii) X_t - Gaussian process $\Rightarrow X_t$ - S.S. $\Leftrightarrow X_t$ - W.S.① white noise process: $X_t, t=0, \pm 1, \pm 2, \pm 3, \dots$

$$\boxed{\mathbb{E}X_t = 0}, \text{Var} X_t = \sigma^2 \quad [\text{WN}(0, \sigma^2)]$$

$$K(t, s) = 0, \text{ if } t \neq s$$

$$K(t, s) = \sigma^2 \mathbb{1}_{\{t=s\}} = \gamma(t-s), \quad \boxed{\gamma(\kappa) = \sigma^2 \mathbb{1}_{\{\kappa=0\}}} \Rightarrow \text{W.S.}$$

- a) X_1, X_2, \dots - iid. noise } X_t - S.S.
 b) X_t - Gaussian

② Random walk

$$S_n = S_{n-1} + \xi_n, \quad \xi_1, \xi_2, \dots \text{ - iid. } \begin{cases} 1, \text{ w.p. } p \\ -1, \text{ w.p. } 1-p \end{cases}$$

$$S_0 = 0$$

$$S_n = \xi_1 + \dots + \xi_n$$

$$\mathbb{E}S_n = n\mathbb{E}\xi_1 = n(2p-1) \Rightarrow \text{not W.S.} \Rightarrow \text{not S.S.}$$

a) $p = 1/2 \Rightarrow \mathbb{E}S_n = 0$

$$K(n, m) = \text{cov}(S_m + \xi_{m+1} + \dots + \xi_n, S_m) \quad (n > m)$$

$$= \text{cov}(S_m, S_m) + \text{cov}(\xi_{m+1} + \dots + \xi_n, S_m)$$

//

$$\text{Var } S_m = m \text{Var } \xi_1 = \min(n, m) \text{Var } \xi_1$$

$$\Rightarrow \text{not W.S.} \Rightarrow \text{not S.S.}$$

③ Brownian motion

$$\mathbb{E}B_t = 0$$

$$\text{Var } B_t = t \quad [B_t - B_s \underset{\substack{\parallel \\ 0}}{\sim} N(0, \underset{\substack{\parallel \\ t}}{t-s})]$$

$$\text{Var } B_t = t \neq \gamma(0) \Rightarrow \text{not W.S.} \Rightarrow \text{not W.S.}$$

$$K(t, s) = \min(t, s), \quad t > s$$

$$K(t+h, s+h) = s+h \neq s$$

S.2 Two types of stationarity (2)

④ Moving average process: $MA(q)$

$$Y_t = a_0 X_t + a_1 X_{t-1} + \dots + a_q X_{t-q} \quad (a_1, \dots, a_q) \in \mathbb{R}^q; a_0 = 1$$

$$X_t \sim WN(0, \sigma^2)$$

$$\mathbb{E}Y_t = 0$$

$$K(t, s) = \text{cov}\left(\sum_{j=0}^q a_j X_{t-j}, \sum_{k=0}^q a_k X_{t-k}\right)$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_j a_k \operatorname{cov}(X_{t-j}, X_{s-k}) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_j a_k \sigma^2 \mathbb{1}\{t-s=j-k\}$$

$$\text{MA}(1): \gamma(x) = C \mathbb{1}\{|x|=1\} + D \mathbb{1}\{x=0\}$$

$\text{MA}(\xi)$ is W.S.

⑤ Autoregressive $\text{AR}(p)$

$$Y_t = b_1 Y_{t-1} + \dots + b_p Y_{t-p} + \varepsilon_t$$

$\varepsilon_t \sim \text{WN}(0, \sigma^2)$

$$\operatorname{cov}(\varepsilon_t, Y_s) = 0, \forall t > s$$

$$\text{AR}(1): Y_t - b Y_{t-1} = \varepsilon_t, \quad b \in \mathbb{R}$$

$$Y_t = \sum_{j=0}^{\infty} b^j \varepsilon_{t-j}, \quad \mathbb{E} Y_t = 0$$

$$K(t, s) = \sum_{j, k=0}^{\infty} b^{j+k} \sigma^2 \mathbb{1}\{t-s=j-k\}$$

$$\text{If } t-s=0 \Rightarrow K(t, s) = \sum_{k=0}^{\infty} b^{2k} < \infty \Leftrightarrow |b| < 1$$

$$\text{If } |b| < 1, Y_t - \text{W.S.}$$

5.3 Spectral density of a wide-sense stationary process (1)

Spectral density

Bochner - Khintchine theorem

$\phi(u)$ - char. fun. \Leftrightarrow 1) ϕ - continuous

$\phi: \mathbb{R} \rightarrow \mathbb{C}$

2) ϕ - positive semi-definite

$$\sum_{j, k=1}^n \bar{z}_j \bar{z}_k \phi(u_j - u_k) \geq 0$$

$\forall (z_1, \dots, z_n) \in \mathbb{C}^n$
 $\forall (u_1, \dots, u_n) \in \mathbb{R}^n$

$$\phi(u) = \mathbb{E} e^{iu\xi}$$

$$3) \phi(0) = 1$$

$$1), 2) \text{ properties only } \Rightarrow \exists \mu: \phi(u) = \int e^{iu\lambda} \mu(d\lambda)$$

$$1), 2), \int |\phi(u)| du < \infty \Rightarrow \phi(u) = \int e^{iu\lambda} s(\lambda) d\lambda$$

X_t - w.s. $\Rightarrow \gamma: K(t,s) = \gamma(t-s)$

If γ is continuous, $\int |\gamma(u)| du < \infty$ Fourier transform

$$\Rightarrow \exists g(x): \gamma(u) = \int e^{iux} g(x) dx = \mathcal{F}[g](u)$$

$$g(x) = \frac{1}{2\pi} \int e^{-iux} \gamma(u) du \quad - \text{spectral density}$$

$$g(x) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} e^{-ihx} \gamma(h) \quad - \text{discrete time spectral density}$$

5.4: Spectral density of a wide-sense stationary process (2)

(Ex) 1) $WN(0, \sigma^2) \Rightarrow \gamma(u) = \sigma^2 \mathbb{1}_{\{u=0\}}$

$$g(x) = \frac{\sigma^2}{2\pi}$$

2) $MA(1)$

$$\gamma(u) = \begin{cases} 0, & |u| > 1 \\ a\sigma^2, & |u| = 1 \\ a(1+u^2)\sigma^2, & u=0 \end{cases}$$

$$g(x) = \frac{\sigma^2}{2\pi} (1 + a^2 + 2a \cos(x))$$

Prop: A real-valued function $g(x)$ is a spectral density of a stochastic process X_t iff γ on $[-\pi, \pi]$

1) $g(x) \geq 0$

2) g -even

3) $\int_{-\pi}^{\pi} g(x) dx < \infty$

5.5 Moving-average filters (1)

Filter: $X_t \rightarrow Y_t$

$$Y_t = a_0 X_t + a_1 X_{t-1} + \dots + a_n X_{t-n}$$

$$Y_t = \int_{\mathbb{R}} e^{-\beta(t-s)} X_s ds$$

1) Linearity: $\left. \begin{matrix} X_t^{(1)} \rightarrow Y_t^{(1)} \\ X_t^{(2)} \rightarrow Y_t^{(2)} \end{matrix} \right\} \Rightarrow c_1 X_t^{(1)} + c_2 X_t^{(2)} \rightarrow c_1 Y_t^{(1)} + c_2 Y_t^{(2)}$

2) time-invariance

$$[X_t \rightarrow Y_t] \Rightarrow [X_{t+h} \rightarrow Y_{t+h}], \forall h > 0$$

$$\int_{\mathbb{R}} e^{-\beta(t-s)} X_{s+h} ds = \int_{\mathbb{R}} e^{-\beta((t+h)-(s+h))} X_{s+h} ds = Y_{t+h}$$

$$Y_t = \int_{\mathbb{R}} g(s) X_{t-s} ds \quad (\text{continuous time})$$

$$Y_t = \sum_{h=-\infty}^{\infty} g(h) X_{t-h} \quad (\text{discrete time})$$

5.6 Moving-average filters (2)

Thm X_t - W.S. process with $EX_t = 0$, $g_X(\cdot)$ and

$$Y_t = \int_{\mathbb{R}} g(s) X_{t-s} ds, \text{ then}$$

(i) Y_t - W.S. process

$$(ii) g_Y(x) = g_X(x) \cdot |F[g](x)|^2$$

$$F[g](x) = \int_{\mathbb{R}} e^{iux} g(u) du$$

Proof: (i) $EY_t = \int_{\mathbb{R}} g(s) \underbrace{EX_{t-s}}_0 ds = 0$

$$K_Y(t_1, t_2) = E \left[\int_{\mathbb{R}} g(s_1) X_{t_1-s_1} ds_1 \cdot \int_{\mathbb{R}} g(s_2) X_{t_2-s_2} ds_2 \right]$$

$$= \iint_{\mathbb{R} \times \mathbb{R}} g(s_1) g(s_2) E \underbrace{[X_{t_1-s_1} X_{t_2-s_2}]}_{\gamma_X(t_2-t_1-(s_2-s_1))} ds_1 ds_2$$

$$\Rightarrow \gamma_Y(x) = \iint_{\mathbb{R} \times \mathbb{R}} g(s_1) g(s_2) \gamma_X(x-(s_2-s_1)) ds_1 ds_2$$

$$(ii) \gamma_Y(x) = \int_{\mathbb{R}} g(s_1) \underbrace{\int_{\mathbb{R}} \gamma_X(x+s_1-s_2) g(s_2) ds_2}_{[\gamma_X * g](x+s_1)} ds_1$$

$$= \int_{\mathbb{R}} [Y_X * g](x+s_1) g(s_1) ds_1$$

$$\text{let } g^o(x) := g(-x)$$

$$= \int_{\mathbb{R}} [Y_X * g](x-s_1) \underset{g^o(s_1)}{g(-s_1)} ds_1 = Y_X * g * g^o(x)$$

$$\left(g_Y(x) = \frac{1}{2\pi} \mathcal{F}[Y_Y](-x) \right)$$

$$\frac{1}{2\pi} \mathcal{F}[Y_Y](x) = \frac{1}{2\pi} \mathcal{F}[Y_X](x) \cdot \mathcal{F}[g](x) \cdot \mathcal{F}[g^o](x)$$

$$\Rightarrow \frac{1}{2\pi} \mathcal{F}[Y_Y](-x) = \frac{1}{2\pi} \mathcal{F}[Y_X](-x) \cdot \underbrace{\mathcal{F}[g](-x) \cdot \mathcal{F}[g^o](-x)}_{\text{complex conjugates}}$$

$\underset{g_Y(x)}{\parallel}$
 $\underset{g_X(x)}{\parallel}$

5.7 Moving-average filters (3)

X_n - W.S. (discrete time), g_X

$Y_n = a_1 X_{n-1} + a_2 X_{n-2}$. What are a_1, a_2 such that

$\mathbb{E}[(X_n - Y_n)^2]$ is minimum?

$Z_n = X_n - Y_n = X_n - a_1 X_{n-1} - a_2 X_{n-2} \Rightarrow Z_n$ - W.S. from then

$$g_Z(x) = g_X(x) / |\mathcal{F}[g](x)|^2 \text{ from then}$$

$$g(x) = \mathbb{1}\{x=0\} - a_1 \mathbb{1}\{x=1\} - a_2 \mathbb{1}\{x=2\}$$

$$\mathcal{F}[g](x) = 1 - a_1 e^{ix} - a_2 e^{2ix}$$

$$\underbrace{\text{Var } Z_n}_{\parallel} \rightarrow \min_{a_1, a_2}$$

$$K_Z(n, n) = \gamma_Z(0) = \int_{\mathbb{R}} \underbrace{e^{i \cdot 0 \cdot x}}_{\parallel} g_Z(x) dx$$

$$= \int_{\mathbb{R}} g_X(x) \underbrace{|1 - a_1 e^{ix} - a_2 e^{2ix}|^2}_{\parallel} dx$$

$$= \sum_{i,j=1}^2 \beta_{ij} a_i a_j + \sum_{i=1}^2 c_i a_i + 0 \quad \left(\begin{array}{c} \text{minimized wr.t} \\ a_1, a_2 \end{array} \right) \Big|_D$$