26/10-10 Chapter 1 - Probability

Random experiment (doesn't know the outcome)
Different possible outcomes ξ which belong to sample space S of all possible outcomes
An event A is a subset of the sample space

Algebra of sets
$\cup, \cap$, complement $\bar{A}$ (A^c), difference $A \setminus B = A \cap \bar{B}$
A and B disjoint if $A \cap B = \emptyset$ empty set.

Probability messure
$P(A)$ for $A \subseteq S$ gives probabilities to events so that:

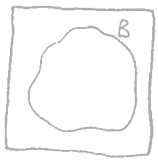
Three axioms are fullfilled
$P(A) \geq 0$
$P(S) = 1$ (sample space set of all outcomes)
$P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$

consequences

Proof.
$P(\bar{A}) = 1 - P(A)$ $(1 = P(S) = P(A \cup \bar{A}) = P(A) + P(\bar{A}))$
$P(\emptyset) = 0$
$P(A) \leq P(B)$ if $A \subseteq B$
$P(A) \leq 1$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(A \setminus B) = P(A) - P(A \cap B)$
$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ for A_i $i = 1, \dots n$ disjoint.

Conditional probability
of an event A given that we know that B has happened

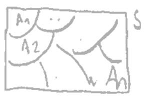
DEFINITION (1)
$P(A|B) = \frac{P(A \cap B)}{P(B)}$



$P(A \cap B) = P(A|B) \cdot P(B)$ follows from (1)

Total probability

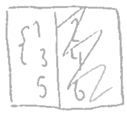
Law of total probability



If $A_1, \dots, A_n$ are disjoint and $\bigcup_{i=1}^{n} A_i = S$
$P(B) = \sum_{i=1}^{n} P(B|A_i) P(A_i)$ (if it is easier to calculate conditional probabilities)

Independence of events
A and B are independent if $P(A \cap B) = P(A) \cdot P(B)$
$A_1, \dots, A_n$ are independent if $P(A_{i_1} \cap \dots \cap A_{i_n}) = P(A_{i_1}) \cdots P(A_{i_n})$ for any selection of different indices $i_1, \dots, i_n \in \{1, \dots, n\}$

example Dice
$S = \{1,2,3,4,5,6\}$ Throw one dice $P = ?$ $P(1) = P(2) = P(3) = \dots = P(6) = \frac{1}{6}$ $P(A) = \frac{\#A}{6} = \frac{\text{nbr of members of}}{6}$
$A = \{1,2,3\}$ $P(A|B) = P(\leq 3 \mid \text{odd}) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{1,3\})}{P(\{1,3,5\})} = \frac{2/6}{3/6} = \frac{2}{3}$
$B = \{1,3,5\}$



intuitively
2/3

Are A and B independent? $P(A \cap B) = P(A) \cdot P(B)$? NO!

2. Random variables (stochastic variables)

Sample space S of all possible outcomes ξ of random experiment.
'robability measure P assigns probabilities $P(A)$ to events $A \subseteq S$

Random variable X
is a function from S to $\mathbb{R}$, $X: S \to \mathbb{R}$
$X(\xi)$ for $\xi \in S$ (random), which is random because ξ is random (the input/argument is random, not the function
$P(X \leq x) = P(\xi \in S : X(\xi) \leq x) \equiv F_X(x)$ called the cumulative distribution function, (CDF), for the random variable X for $x \in \mathbb{R}$

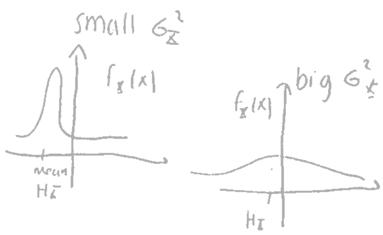
$0 \leq F_X(x) \leq 1$ (it is a probability)
$F_X(x_1) \leq F_X(x_2)$ if $x_1 \leq x_2$ (growing)
$\lim_{x \to \infty} F_X(x) = 1$ and $\lim_{x \to -\infty} F_X(x) = 0$
$P(a < X \leq b) = P(X \in]a, b]) = F_X(b) - F_X(a)$

Discrete random variables
A rv X is discrete if it has finitely or at most countably infinite many possible different values
'robability mass function (PMF) $p_X(x) = P(X = x)$ for possible values x of X.
$0 \leq p_X(x) \leq 1$
$\sum_{\text{all } x} p_X(x) = 1$
$P(X \in A) = \sum_{x \in A} p_X(x)$

Continous random variables
A rv X is continous if $F_X(x)$ is (continuous) and differentiable.
probability density function (PDF) $f_X(x) = F_X'(x)$
$f_X(x) \geq 0$ (because $F_X(x)$ is increasing)
$\int_{-\infty}^{\infty} f_X(x)\,dx = F_X(\infty) - F_X(-\infty) = 1 - 0 = 1$
$P(X \in A) = \int_A f_X(x)\,dx$

Mean, variance and other moments
$\mu_X = E(X)$ (expectation) the average value of X (center of gravity of the PDF and CDF)

$$= \begin{cases} \sum_{\text{all } x} x\, p_X(x), & X \text{ discrete} \\ \int_{-\infty}^{\infty} x f_X(x)\,dx, & X \text{ continous.} \end{cases}$$



n^{th} moment of X $E(X^n) = \begin{cases} \sum_{\text{all } x} x^n p_X(x), & X \text{ discrete} \\ \int_{-\infty}^{\infty} x^n f_X(x)\,dx, & X \text{ continous} \end{cases}$

Variance of X $\sigma_X^2 = Var(X) = E((X - \mu_X)^2) = \begin{cases} \sum_{\text{all } x} (x - \mu_X)^2 p_X(x) \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\,dx \end{cases}$

$Var(X) = E(X^2) - \mu_X^2$) follows from expanding, linearity etc.

examples of r.v. section 2.7

Discrete examples

Bernoulli distribution rv X with possible values $\{0,1\}$ $P(X=1) = p_X(1) = p$ $P(X=0) = p_X(0) = 1-p$

Binomial distribution.

Geometric distribution

Poisson distribution rv. X with possible values $0, 1, 2, \ldots$

$p_X(x) = e^{-\lambda} \dfrac{\lambda^x}{x!}$ for $x = 0, 1, 2 \ldots$ $\mu_X = \sigma_X^2 = \lambda$ (put it into the definitions)

- Binomial distribution

with parameters n, p

Count the number of one's (1's) when we perform n independent Bernoulli random experiments.

$P(X=h) = p_X(h) = \binom{n}{h} p^h (1-p)^{n-h}$ (binomial theorem says sums to one) for $h = 0, \ldots, n$

- Geometric distribution

The number X of the times we have to perform independent Bernoulli experiments to get the first 1. $P(X=h) = p_X(h) = (1-p)^{h-1} p$ $h = 1, 2$

Continous examples

- Uniform distribution over $[a,b]$
- Exponential distribution with parameter $\lambda > 0$
- Gamma distribution with parameter $\lambda > 0$ and $\alpha > 0$
- Normal distribution $N(\mu, \sigma^2)$

8/10 Chapter 3 Multiple random variables

2-dim case

n-dim case. $\xi \in S$

(X,Y) random variable with values in $\mathbb{R}^2$. (the same probability space $X(\xi)$, $Y(\xi)$)

$(X_1, \dots X_n)$ — " — $\mathbb{R}^n$.

cdf. $F_X(x) = P(X \le x)$, $F_Y(y) = P(Y \le y)$ $\not\Rightarrow$ $F_{XY}(x,y) = P(X \le x, Y \le y)$ $\Leftarrow$

no information about their dependence

X and Y are independent if $P(\underbrace{X \le a}_{]-\infty, x]}, \underbrace{Y \le b}_{]-\infty, y]}) = P(X \le a) \cdot P(Y \le b)$

$F_{X,Y}(x,y) = P(X \le x, Y \le y) = P(X \le x) \cdot P(Y \le y) = F_X(x) \cdot F_Y(y)$ (independent)

- $0 \le F_{X,Y}(x,y) \le 1$
- $F_{XY}(x_1,y_1) \le F_{XY}(x_2,y_2)$ for $x_1 \le x_2$, $y_1 \le y_2$ (increasing in both variables)

$\lim_{x,y \to \infty} F_{XY}(x,y) = 1$, $\lim_{x,y \to -\infty} F_{XY}(x,y) = 0$.

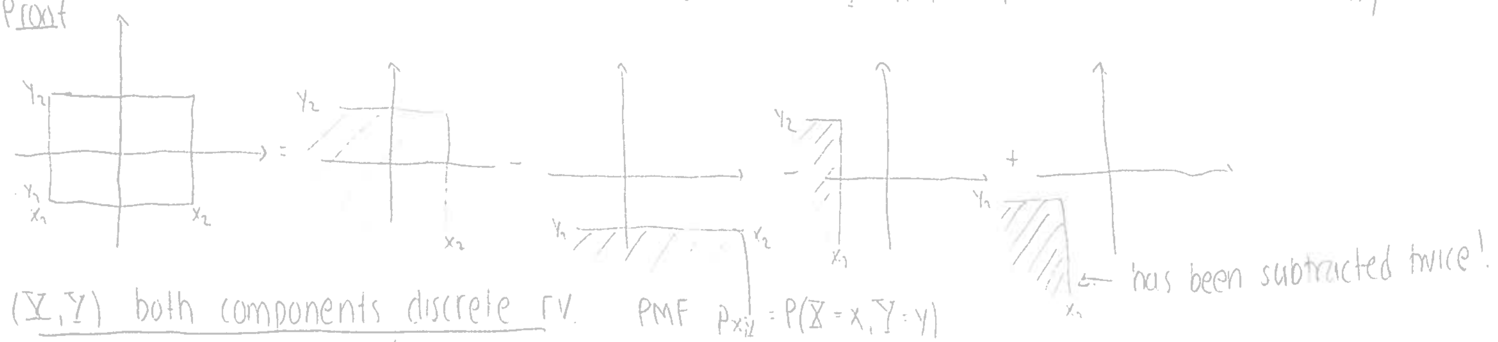
$\lim_{x \to \infty} F_{X,Y}(x,y) = F_Y(y)$, $\lim_{y \to \infty} F_{XY}(x,y) = F_X(x)$

have subtracted this one twice!

one dimension.

$P(x_1 < X \le x_2, y_1 < Y \le y_2) = F_{XY}(x_2,y_2) - F_{XY}(x_1,y_2) - F_{X,Y}(x_2,y_1) + F_{X,Y}(x_1,y_1)$ $\left(P(x_1 < X \le x_2) = F_X(x_2) - F_X(x_1) \right)$

Proof



(X,Y) both components discrete rv. PMF $p_{X,Y} = P(X = x, Y = y)$

- $0 \le p_{XY}(x,y) \le 1$
- $P(X \in A, Y \in B) = \sum_{(x,y) \in A \times B} p_{XY}(x,y)$

(X,Y) both components continuous rv PDF $f_{XY}(x,y) = \dfrac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$

- $f_{XY}(x,y) \ge 0$ (because the cdf is increasing in both arguments)
- $P(X \in A, Y \in B) = \iint_{(x,y) \in A \times B} f_{XY}(x,y)\,dx\,dy$

$A =]a,b]$ $B =]c,d]$

- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$, $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$ (1)
- X, Y independent $\Leftrightarrow$ $f_{XY}(x,y) = f_X(x) f_Y(y)$

proof (1)

$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx}(P(X \le x)) = \frac{d}{dx} F_{XY}(x, \infty) = \frac{d}{dx} P(X \le x, Y < \infty) = \frac{d}{dx} \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(\hat{x}, y)\,d\hat{x}\,dy = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$

example: we get a Poisson(λ) distributed number N of lottery tickets. Each lottery ticket has winning chance p. What is the expected number of total wins on the N lottery tickets?

$$E[\text{total number of wins}] = \sum_{n=0}^{\infty} \underbrace{E[\text{total number of wins} \mid N=n]}_{np} \cdot \underbrace{P(N=n)}_{\frac{\lambda^n e^{-\lambda}}{n!}} \ldots = p\lambda$$

Conditional distributions

(X,Y) discrete conditional PMF $\underset{P(X=x \mid Y=y)}{p_{X|Y}(x|y)} \overset{(1)}{=} \frac{p_{XY}(x,y)}{p_Y(y)} = \frac{P(X=x, Y=y)}{P(Y=y)}$

$$\underset{\substack{\uparrow \\ \text{total number of wins} \\ \text{in earlier example}}}{P(X=x)} = \sum_{\text{all } y} \underbrace{P(X=x \mid Y=y)}_{\binom{n}{x} p^x (1-p)^{n-x}} \underbrace{P(Y=y)}_{\frac{\lambda^n}{n!}e^{-\lambda}} = \sum_{\text{all } y} \underbrace{p_{X|Y}(x|y)\, p_Y(y)}_{p_{X,Y}(x,y) \text{ according to (1)}}$$

(N = n)

$$E[X] = \sum_{\text{all } y} E[X \mid Y=y]\, p_Y(y)$$

(X,Y) continous

$f_{X|Y}(x|y) \overset{\text{DEF}}{\underset{(2)}{=}} \frac{f_{X,Y}(x,y)}{f_Y(y)}$ motivation: $f_{X|Y}(x|y) = \frac{d}{dx} \lim_{\varepsilon \to 0} P(X \le x \mid Y \in]y-\varepsilon, y]) = \frac{d}{dx} \lim_{\varepsilon \to 0} \frac{P(X \le x, Y \in]y-\varepsilon, y])}{P(Y \in]y-\varepsilon, y])} =$

$$= \frac{d}{dx} \lim_{\varepsilon \to 0} \underbrace{\frac{F_{X,Y}(x,y) - F_{X,Y}(x,y-\varepsilon)}{\varepsilon}}_{\to \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)} \underbrace{\frac{\varepsilon}{F_Y(y) - F_Y(y-\varepsilon)}}_{\to \frac{1}{F_Y'(y)}} = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$P(X \le x) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X|Y}(\hat{x} \mid \hat{y})\, f_Y(\hat{y})\, d\hat{y}\, d\hat{x}$$

$$E[X \mid Y=y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, dx$$

$E[X] = \int_{-\infty}^{\infty} E[X \mid Y=y]\, f_Y(y)\, dy$ Proof $\int_{-\infty}^{\infty} E[X \mid Y=y] f_Y(y)\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, \underbrace{f_{X|Y}(x|y)\, dx\, f_Y(y)}_{=(2)= f_{X,Y}(x,y)} dy = \int_{-\infty}^{\infty} x \Big(\underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy}_{f_X(x) \text{ shown earlier today.}} \Big) dx =$

$$= \int_{-\infty}^{\infty} x f_X(x)\, dx = E[X] \quad \text{qed}$$

N-variate normal (Gauss) distribution X n-dimensional r.v.

$$f_X(x) = \frac{1}{(2\pi)^{n/2} |\det K|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1} (x-\mu)\right) \qquad \mu = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{bmatrix} \qquad K_{ij} = \mathrm{Cov}(X_i, X_j)$$

$f_{X_1, \ldots, X_n}$

Covariance and variance (not 106-107 in the book)

$$\mathrm{Cov}(X,Y) = E[(X - \underset{E[X]}{\mu_x})(Y - \underset{E[Y]}{\mu_y})] = E[XY] - \mu_x E[Y] - \mu_y E[X] - \mu_x\mu_y = E[XY] - E[X]E[Y]$$

$\mathrm{Var}(X) = V[X] = E[(X-\mu_x)^2] = \mathrm{Cov}[X,X] = E[X^2] - E[X]^2$ X and Y independent if $\mathrm{Cov}[X,Y] = 0$.

X and Y independent $\underset{\not\Leftarrow}{\Rightarrow}$ $\mathrm{Cov}(X,Y) = 0$ since $E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \underbrace{f_X(x) f_Y(y)}_{f_{XY}(x,y)}\, dx\, dy = E[X]E[Y]$

$\mathrm{Corr}[X,Y] = \rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{V[X]V[Y]}}$, $|\rho_{XY}| \le 1$ (big correlation ≈ 1) (small correlation ≈ 0)

$$\mathrm{cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = [a_i, b_i \in \mathbb{R}] = \sum_{i=1}\sum_{j=1} a_i b_j \mathrm{Cov}[X_i, Y_j]$$

$$\mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \mathrm{Cov}\left[\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{cov}[X_i, X_j]$$

$$\mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \mathrm{Var}[X_i] \quad \text{if } X_1, \dots X_n \text{ independent}$$

29/10-10 Chapter 4 Functions of random variables

n-variate normal distribution

$$f_X(x) = \frac{1}{(2\pi)^{n/2}|\det K|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right)$$ PDF of n-variate normal distributed rv. $X = (X_1, \dots, X_n)$, $x = (x_1, \dots, x_n)$

2-variate case

$$\mu = \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad K = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \begin{bmatrix} \mathrm{var}(X_1) & \mathrm{cov}[X_1, X_2] \\ \mathrm{cov}[X_1, X_2] & \mathrm{var}[X_2] \end{bmatrix} \quad |\rho| < 1$$

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}[x_1 \; x_2]\frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right)$$

$$[x_1 \; x_2]\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - \rho x_1 x_2 - \rho x_1 x_2 + x_2^2$$

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)} = \left[f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\,dx_1 = \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right) dx_1 = \right.$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x_1-\rho x_2)^2}{2(1-\rho^2)}\right) \exp\left(-\frac{(1-\rho^2)x_2^2}{2(1-\rho^2)}\right) dx_1 = \frac{1}{\sqrt{2\pi}} e^{-x_2^2/2} \quad \left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}\right) \quad \underbrace{\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{(x_1-\rho x_2)^2}{2(1-\rho^2)}\right)}_{f_{N(\rho x_2, 1-\rho^2)}(x_1)} \quad \int_{-\infty}^{\infty} \dots =$$

$$= \frac{\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right)}{\frac{1}{\sqrt{2\pi}} e^{-x_2^2/2}} = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{(x_1-\rho x_2)^2}{2(1-\rho^2)}\right) \quad f_{N(\rho x_2, 1-\rho_2)}$$

$Y = g(X)$ $\quad g: \mathbb{R} \to \mathbb{R}$ $\quad X$ is continuous, g is invertible. increasing

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} P(Y \le y) = \frac{d}{dy} P(g(X) \le y) = \frac{d}{dy} P(X \le g^{-1}(y)) = \frac{d}{dy} F_X(g^{-1}(y)) = f_X(g^{-1}(y)) \frac{1}{g'(g^{-1}(y))}$$

$(Z, W) = j(X, Y)$ $\quad j: \mathbb{R}^2 \to \mathbb{R}^2$ $\quad X, Y$ continuous $\quad j$ invertible

$$f_{Z,W}(z,w) = f_{X,Y}(x,y)(j^{-1}(z,w)) \frac{1}{\begin{vmatrix} \frac{dj_1}{dx} & \frac{dj_1}{dy} \\ \frac{dj_2}{dx} & \frac{dj_2}{dy} \end{vmatrix} (j^{-1}(z,w))}$$

$$f_{Z,W}(z,w) = \frac{\partial^2}{\partial z \partial w} P(Z \le z, W \le w) = \frac{\partial^2}{\partial z \partial w} \iint_{\{j(x,y) \in]-\infty,z] \times]-\infty,w]\}} f_{X,Y}(x,y)\,dx\,dy$$

4.5 expectation $Y = g(X)$

$$E[Y] = E[g(X)] \begin{cases} = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx & , X \text{ continous} \\ = \sum_{\text{all } x} g(x) p_X(x) & , X \text{ discrete.} \end{cases}$$

$$E[g(X_1, \dots X_n)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \dots x_n) f_{X_1 \dots X_n}(x_1, \dots x_n)\,dx_1 \cdots dx_n \quad \text{or } \int_{\mathbb{R}^n} g f_X$$

$$= \sum_{x_1} \cdots \sum_{x_n} g(x_1, \dots, x_n) p_{X_1 \dots X_n}(x_1, \dots x_n)$$

$E[\sum_{i=1}^{n} a_i X_i] = \sum_{i=1}^{n} a_i E[X_i]$ $\quad$ $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ if X and Y independent, g, h, functions.

$a_i \in \mathbb{R}$ X_i rv.

conditional expectation

$E[Y \mid X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy$ $\quad$ Y cont

$\sum$. $\quad$ Y discrete.

$E[Y \mid X] = \int_{-\infty}^{\infty} y f_{Y|X}(y|X)\,dy$ $\quad$ (keep X random)

$E[Y] = E[E[Y \mid X]] = \int_{-\infty}^{\infty} E[Y \mid X = x] f_X(x)\,dx$

Jensens inequality

$E[g(x)] \geq g(E(x))$ for g a convex function.

example $\quad E[X^2] \geq (E(X))^2 \quad x^2$ convex

$E[|X|] \geq |E[X]| \quad |x|$ convex

Cauchy-Schwarz inequality

$|E[XY]| \leq \sqrt{E[X^2]E[Y^2]}$ $\quad$ $|\langle x, y \rangle| \leq |x| \cdot |y|$

$\int fg \leq \sqrt{\int f^2 \int g^2}$

TRANSFORMS!

Probability generating function

$G_X(z) = E[z^X] = \sum_{x=0}^{\infty} z^x p_X(x)$ for $|z| < 1$ for $\mathbb{N}$-valued rv. X

$G_X^{(n)}(z)\Big|_{z=0} = n!\,p_X(n) \Rightarrow p_X(n) = \dfrac{G_X^{(n)}(0)}{n!}$

Moment generating function (MGF)

$$M_X(t) = E[e^{tX}] \begin{cases} \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx & , X \text{ continous} \\ \sum_{\text{all } x} e^{tx} p_X(x) & , X \text{ discrete.} \end{cases}$$

$$M_X(j\omega) = \Psi_X(\omega)$$

$$M_X^{(n)}(t)\Big|_{t=0} = \frac{d^n}{dt^n} E[e^{tX}]\Big|_{t=0} = E\left[\frac{d^n}{dt^n} e^{tX}\right]\Big|_{t=0} = E[X^n e^{tX}]\Big|_{t=0} = E[X^n]$$

$$= M_X^{(n)}(0)$$

$$M_{X,Y}(t_1, t_2) = E[e^{t_1 X + t_2 Y}]$$

Characteristic function (CHF) $\boxed{j^2 = -1}$

$$\Psi_X(\omega) = E[e^{j\omega X}] = \begin{cases} \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)\,dx \\ \sum_{\text{all } x} e^{j\omega x} p_X(x) \end{cases} \qquad \omega \in \mathbb{R}$$

always convergent! always well defined.

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega x} \Psi_X(\omega)\,d\omega$$

$$\Psi_X^{(n)}(0) = \frac{d^n}{d\omega^n} E[e^{j\omega X}]\Big|_{\omega=0} = E[(j\omega)^n e^{j\omega X}]\Big|_{\omega=0} = E[(jX)^n] \qquad E[(X^n)] = \frac{\Psi_X^{(n)}(0)}{j^n}$$

2-d fourier-transform

$$\Psi_{XY}(\omega_1, \omega_2) = E[e^{j\omega_1 X + j\omega_2 Y}]$$

X and Y independent $\iff F_{X,Y}(x,y) = F_X(x) F_Y(y) \iff f_{X,Y}(x,y) = f_X(x) f_Y(y)$

$\iff \Psi_{X,Y}(\omega_1, \omega_2) = \Psi_X(\omega_1) \Psi_Y(\omega_2)$.

$\left(\iff M_{X,Y}(t_1, t_2) = M_X(t_1) \cdot M_Y(t_2)\right)$

example

$X = N(\mu, \sigma^2)$

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

$$\Psi_X(\omega) = \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)\,dx = \int_{-\infty}^{\infty} e^{j\omega x} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \underbrace{\cancel{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu-j\omega\sigma^2)^2}{2\sigma^2}}\,dx}}_{1} \cdot e^{\mu j\omega} e^{-\frac{\omega^2\sigma^2}{2}}$$

$$e^{\mu j\omega - \frac{\omega^2\sigma^2}{2}}$$

1/11-10

lids-Beta $\int_0^1 (\sin(\frac{1}{x}))^2 dx = \int_1^\infty \frac{\sin(x)^2}{x^2} dx \approx 0.671563$ 0.673427 ± 0.001 (convergence error)
0.673824 (simulation) $n = 10^6$ simulations.

Chapter 4.

Section 4.9 Law of Large Numbers, CLT (central Limit Theorem)

$X_1, X_2, \ldots X_n$ independent identically distributed (IID) random variables.

$\frac{1}{n}\sum_{i=1}^{n} X_i \to E[X_1]$, $n \to \infty$

Chapter 5 Random processes

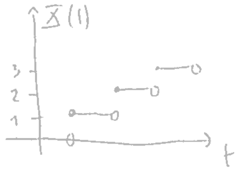
Random process (stochastic) family of rv. indexed by time.

$\{X(t)\}_{t\in T}$ (bunch of rv) is a function $X : \Omega \times T \to \mathbb{R}$ defined on a sample space Ω and parameter set T. (probability space, time)

$= \{X(\omega, t)\}_{t \in T}$ Most commonly T = time, $T = \mathbb{R}$, $T = [0, \infty[$, $T = \mathbb{Z}$, $T = \mathbb{N}$, $T = [0,1]$, $T = \{0, \ldots, n\}$

* continous time
\+ discrete time.

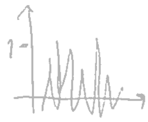
ex) Poisson process



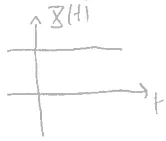
the value at each time t random (poisson)

ex) completely "wild" white noise

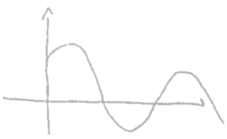
Each $X(t)$ is independent $N(0,1)$-distributed.



ex) $X(t) = \xi$ a single $N(0,1)$ random variable



ex) $X(t) = A\cos(\omega t) + B\sin(\omega t)$ where $\omega \in \mathbb{R}$, A and B are zero-mean uncorrelated r.v. with variance σ^2.



5.3 Characterization of random processes

$(F_{X(t)}(x) = P(X(t) \le x))$ (very incomplete as a description of a process) no information about dependence among the rv

$F_{X(t_1), \ldots, X(t_n)}(x_1, \ldots, x_n) = P(X(t_1) \le x_1, \ldots X(t_n) \le x_n)$ for each choice of $n \in \mathbb{N}$, $t_1, \ldots, t_n \in T$.

finite dimensional distributions (fidi's)

PDF when process value are continuous: $f_{X(t_1), \ldots X(t_n)}(x_1, \ldots, x_n) = \frac{\partial^n F(x_1, \ldots x_n)}{\partial x_1 \cdots \partial x_n}$ (know everything about the process if this is known)

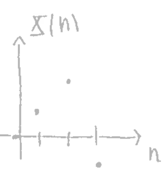
PMF discrete case (are discrete.) $p_{X(t_1), \ldots X(t_n)}(x_1, \ldots x_n) = P(X(t_1) = x_1, \ldots X(t_n) = x_n)$.

ex) discrete time.

$X(n) = \sum_{i=1}^{n} Y_i$ for $n = 0, 1, 2 \ldots$ where $\{Y_i\}_{i=1}^{\infty}$ are IID rv.

$(X(\omega_0, t))_{t \in T}$ for a fixed "typical" $\omega_0 \in \Omega$ is a realization of the process.

random walk



(everytime one sees a picture of a process it is a realization of a process.

Mean function

$\mu_{X(t)} = E[X(t)]$

these say interesting things about the process, but not everything.

Autocorrelation function

$R_X(s,t) = E[X(s)X(t)]$

auto: because of covariance etc. between the process and itself.

Autocovariance function

$K_X(s,t) = Cov[X(s), X(t)]$ $\quad K_X(s,t) = R_X(s,t) - \mu_X(s)\mu_X(t)$.

process values may differ, but not the law that determines the ...

Stationary processes $(X(t))_{t\in\mathbb{R}}$ is such that $[X(t_1+h), \dots X(t_n+h)]$ same CDF $[X(t_1), \dots X(t_n)]$

finite dimensional distributions are invariant under time translations h.

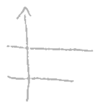
Have to be able to calculate it's CDF for all n-dimensional vectors to decide wether it is a stationary process.

Normal processes (Gaussian) $(X(t))_{t\in T}$ process such that $(X(t_1), \dots X(t_n))$ is n-dimensional normally distributed for each $t_1, \dots, t_n \in T$. That is:

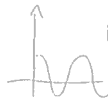
$$\Psi_{X(t_1),\dots X(t_n)}(\omega_1, \dots \omega_n) = E[e^{j(\omega_1 X(t_1) + \dots + \omega_n X(t_n))}] = e^{j\sum_{i=1}^{n} \omega_i E[X(t_i)] - \frac{1}{2}\sum_{i=1}^{n}\sum_{h=1}^{n} \omega_i \omega_h Cov[X(t_i), X(t_h)]}$$

$$f_{X(t_1),\dots X(t_n)}(x_1, \dots x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{|\det K|}} e^{-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)} \qquad \mu = \begin{bmatrix} E[X(t_1)] \\ \vdots \\ E[X(t_n)] \end{bmatrix} \quad K = (Cov(X(t_i), X(t_h)))_{i,h}$$

ex) $X(t)$



single normal distr r.v. = normal process $\qquad X(t) = A\cos(\omega t) + B\sin(\omega t) \quad A,B \sim N(0,\sigma^2)$



normal process

Ergodic processes $(X(t))_{t\in\mathbb{R}}$ for each ω - doesn't depend on ω

$X(t)_{t\in\mathbb{R}}$ is a process such that $\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} X(t)dt = E[X(0)]$. Assuming that $E[X(t)] = E[X(0)]$ is constant. (doesn't depend on t)

$E[X(1)] \leftarrow \frac{1}{n}\sum_{i=1}^{n} X_i \;, n\to\infty. \quad (X_i)_{i=1}^{\infty}$ iid.

important in practice because there is only one way of getting real results (only one world)

ex) If $X(t)$ are iid. ("built on basic LLN")

WSS, weak stationary processes $(X(t))_{t\in\mathbb{R}}$ such that
- $E[X(t+h)] = E[X(t)]$
- $Cov[X(s+h), X(t+h)] = Cov[X(s), X(t)]$

very much easier for a process to be WSS than stationary

most used in practice.

ex) $X(t) = A\cos(\omega t) + B\sin(\omega t)$ where A, B are zero-mean, uncorrelated (zero-covariance) with common variance σ^2

$Cov[X(s), X(t)] = [\text{autocovariance}] = Cov[A\cos(\omega s) + B\sin(\omega s), A\cos(\omega t) + B\sin(\omega t)] =$

$\underbrace{V[A]}_{\sigma^2}\cos(\omega s)\cos(\omega t) + 0 + 0 + V[B]\sin(\omega s)\sin(\omega t) = \sigma^2\cos(\omega(s-t))$

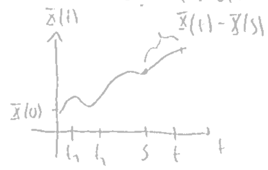
because of uncorrelation

5/11-10 Processes with stationary independent increments - Levy processes

$\{X(t)\}_{t\geq 0}$ $\quad X(0), X(t_1)-X(0), X(t_2)-X(t_1), \ldots, X(t_n)-X(t_{n-1})$ independent for $0<t_1<\ldots<t_n$
(independent increments)

$X(t+h)-X(t+s)$ same distribution as $X(t)-X(s)$ for $0\leq s<t$ and $h>0$ (stationary increments)

ex) Poisson process (ex) independent process)



continuous time version of a discrete time random walk

$X(n)=\sum_{i=1}^{n} Y_i$ for $n=0,1,2\ldots$ where $\{Y_i\}_{i=1}^{\infty}$ are iid r.v's.

Markov processes

now the future depends on the history just on the value right now just right now important

$(X(t))_{t\geq 0}$ $\quad (X(n))_{n\in\mathbb{N}}$ $\quad P(X(t_{n+1})\leq x_{n+1} \mid X(t_1)=x_1, \ldots X(t_n)=x_n) = P(X(t_{n+1})\leq x_{n+1} \mid X(t_n)=x_n)$

for $0\leq t_1<\ldots<t_n<t_{n+1}$ (can skip previous part of the history)

every stationary independent increment process is a Markov process.

ex) SIIP

$P(X(t_{n+1})\leq x_{n+1} \mid X(t_1)=x_1 \ldots X(t_n)=x_n) = P(\underbrace{X(t_{n+1})-X(t_n)}_{\text{independent of history}} + \underbrace{X(t_n)}_{\substack{=x_n \\ \text{determined} \\ \text{by } X(t_n)}} \leq x_{n+1} \mid \cancel{X(t_1)=x_1}, \ X(t_n)=x_n) =$

$= P(X(t_{n+1})\leq x_{n+1} \mid X(t_n)=x_n)$

Markov chains

Discrete time Markov processes with discrete values

$F_{X(t_1),\ldots X(t_n)}(x_1,\ldots,x_n) = F_{X(t_1)}(x_1)\prod_{i=2}^{n} P(X(t_n)\leq x_n \mid X(t_{n-1})=x_{n-1})$ (528)

$F_{X(t_1)}(x_1)\prod_{i=2}^{n} P(X(t_n)\leq x_n \mid X(t_{n-1})\leq x_{n-1})$ Prob. 5.25

solution

$F_{X(t_1),\ldots X(t_n)}(x_1,\ldots,x_n) = P(X(t_n)\leq x_n \mid X(t_1)\leq x_1, \ldots X(t_n)\leq x_n)\cdot P(X(t_1)\leq x_1, \ldots, X(t_{n-1})\leq x_{n-1})$

wrong $\neq P(X(t_n)\leq x_n \mid X(t_{n-1})\leq x_{n-1})P(X(t_1)\leq x_1, \ldots, X(t_{n-1})\leq x_{n-1})$

Because send $x_{n-1}\to\infty$ $\quad P(X(t_1)\leq x_1, \ldots, X(t_{n-2})\leq x_{n-2}, X(t_n)\leq x_n) = P(X(t_n)\leq x_n)P(X(t_1)\leq x_1, \ldots, X(t_{n-2})\leq x_{n-2})$

5.5 Discrete parameter Markov chains

$E=\{0,1,\ldots N\}$ ==> finite matrix

$\{X_n\}_{n=0}^{\infty}$ $\quad X_0, X_1, X_2, \ldots$ possible values $E=\{0,1,2,\ldots\}$ => infinite matrix.

Markov property

$P(X_{n+1}=j \mid X_0=i_0, \ldots, X_n=i) = P(X_{n+1}=j \mid X_n=i) = p_{ij}$ doesn't depend on n.

Time homogeneous

To characterize such processes we need two quantities

$p_0=[P(X_0=0)\ P(X_0=1) \ldots P(X_0=n)] = [P(X_0=i)]_i$ initial state probabilities

$P=(p_{ij})_{ij}$ transition matrix $\quad P=\begin{bmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & \cdot & \\ & \cdot & \\ & \cdot & \end{bmatrix} = \begin{bmatrix} P(X_{n+1}=0 \mid X_n=0) & P(X_{n+1}=1 \mid X_n=0) & \cdots \\ P(X_{n+1}=0 \mid X_n=1) & P(X_{n+1}=1 \mid X_n=1) & \\ P(X_{n+1}=0 \mid X_n=2) & P(X_{n+1}=1 \mid X_n=2) & \cdots \\ \vdots & \vdots & \end{bmatrix}$

ex) $E=\{0,1\}$ $(N=1)$ $\quad P = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$

tossing a coin for example

X_n = "value" of coin after n'th toss.

$\sum_{j=1}^{\infty} p_{ij} = 1 \quad \forall i$ (since some value has to be attained)

$= \sum_{j=1}^{\infty} P(X_{n+1}=j \mid X_n=i) = 1 \quad \forall i$, i.e. row sums are one.

$P^{(m)} = P(X_{n+m}=j \mid X_n=i)$ the m-step transition matrix.

$P^{(m)} = P^m$ proof for $m=2$ $\quad P(X_{n+2}=j \mid X_n=i) = \frac{P(X_{n+2}=j, X_n=i)}{P(X_n=i)} \overset{\text{law of total probability}}{=} \sum_{h=0}^{\infty} \frac{P(X_{n+2}=j, X_{n+1}=h, X_n=i)}{P(X_n=i)} =$

$= \sum_{h=0}^{\infty} \frac{P(X_{n+2}=j, X_{n+1}=h, X_n=i)}{P(X_{n+1}=h, X_n=i)} \cdot \frac{P(X_{n+1}=h, X_n=i)}{P(X_n=i)}$ $\qquad$ chapman-kolmogorov theorem

$= \sum_{h=0}^{\infty} \underbrace{P(X_{n+2}=j \mid X_{n+1}=h, \cancel{X_n=i})}_{p_{hj}} \underbrace{P(X_{n+1}=h \mid X_n=i)}_{p_{ih}} = \sum_{h=0}^{\infty} p_{hj} p_{ih} = (P^2)_{ij}$ $\quad$ markov

correct version of formula 5.28

$P(X_0=i_0, \ldots, X_n=i_n) = P(X_0=i_0) \prod_{k=1}^{n} P(X_k=i_k \mid X_{k-1}=i_{k-1})$

Proof: $P(X_0=i_0, \ldots, X_n=i_n) \overset{\text{markov}}{=} P(X_n=i_n \mid \cancel{X_0=i_0}, \ldots, X_n=i_n)\, P(X_0=i_0, X_1=i_1, \ldots, X_{n-1}=i_{n-1}) =$

$= P(X_0=i_0, \ldots, X_{n-1}=i_{n-1})\, P(X_n=i_n \mid X_{n-1}=i_{n-1})$

$\vdots$

$=$ ______ $\qquad$ ______

in n steps move from i to j $= p_{ij}^n$

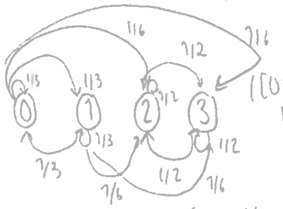
$P(X_n=j) = \sum_{i=0}^{\infty} P(X_n=j \mid X_0=i) P(X_0=i) = \sum_{i=0}^{\infty} p_{ij}^{(n)} p(0)_i = (p(0)P^n)_j$

$p(n) = [P(X_n=0), P(X_n=1), \ldots] = p(0)P^n$

ex) $E=\{0,1,2,3\}$ $\quad p(0) = \ldots$ (starting distribution) $\quad (X_n)_{n=0}^{\infty}$

qualitatively different to 0,1

$P = \begin{bmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/3 & 1/3 & 1/6 & 1/6 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$



(coin tossing behaviour) between 2/3.

one can never get back to zero and one when once left.

transient (eventually you can move from them and never come back)

$\{0,1,2,3\}$

recurrent (can never escape from them)

Classification of states

T_j = the number of steps to the first visit of state j given starting at state j at time 0

j is RECURRENT if $P(T_j<\infty)=1$

j is TRANSIENT if $P(T_j<\infty)<1$

$i \to j$ if $p_{ij}^{(n)} = P(X_n=j \mid X_0=i) > 0$ for some n. j is ACCESSIBLE from i.

$i \leftrightarrow j$ if $i \to j$ and $j \to i$ $\quad$ i and j communicate.

ex) $2 \leftrightarrow 3$
$0 \to 1,2,3$
$1 \to 0,2,3$
$2 \not\to 0,1$
$3 \not\to 0,1$

11/10-10

Markov chains (special case of Markov process)

State space of possible values $E=\{0,1, \ \}$ $E=\{0,1,\ldots,N\}$

$(X_n)_{n=0}^{\infty}$ random process with values in E posseses the Markov property.

$P(X_{n+1}=x_{n+1} \mid X_n=x_n, \ldots, X_0=x_0) = P(X_{n+1}=x_{n+1} \mid X_n=x_n)$

Initial distribution $p(0)=[P(X_0=0) \ \ P(X_0=1) \ldots\ldots]$

Time homogeneity $P(X_{n+1}=j \mid X_n=i)$ doesn't depend on n. (the probability law of how to move around doesn't change)

Transition probability matrix

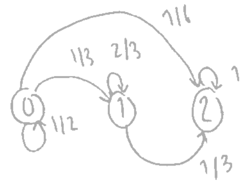
$$P=\begin{bmatrix} P(X_{n+1}=0 \mid X_n=0) & P(X_{n+1}=1 \mid X_n=0) & \ldots \\ P(X_{n+1}=0 \mid X_n=1) & P(X_{n+1}=1 \mid X_n=1) & \ldots \\ P(X_{n+1}=0 \mid X_n=2) & P(X_{n+1}=1 \mid X_n=2) & \ldots \\ \vdots & & \ddots \end{bmatrix}$$

computer problem

$E=\{0,1,2\}$ $p(0)=[1 \ 0 \ 0]$ starts in 0 for sure.

$$P=\begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 0 & 2/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}$$

(Geometric distribution)



aperiodic (non periodic, $p=1$)

$T=\min\{n\in N : X_n=2\}$

$E(T)=4$.

average in moving from zero

average in moving from 1 to 2

$\frac{2}{3}\cdot(2+3)+\frac{1}{3}\cdot 2$ (different ways) added up.

simulate T n times with result $T_1, \ldots, T_n$

$\bar{T} \pm 2\cdot S_T$ (99%-confidence interval)

states 0,1 transient - "temporary"

state 2 reccurent, even absorbing.

classification of states $\quad T_i=\min\{n\geq 1 : X_n=i \text{ given } X_0=i\}$

State i is TRANSIENT if $P(T_i=\infty \mid X_0=i)>0$ (might have to wait infinitely long to get back to i) $\iff P(T_i<\infty \mid X_0=i)<1$ (temporary)

—"— RECCURENT if $P(T_i<\infty \mid X_0=i)=1$

—"— ABSORBING if $\begin{cases} P(X_{n+1}=j \mid X_n=i)=0 & \text{for } j\neq i \\ P(X_{n+1}=i \mid X_n=i)=1 \end{cases}$

(special of reccurent)

$i\to j$ j is accessible from i if $P(X_n=j \mid X_0=i)>0$ for some n. ?

i and j communicate if $i\to j$ and $j\to i$

State i is periodic with period p if $\gcd\{n\geq 1 : P_{ii}(n)=P\{X_n=i \mid X_0=i\}>0\}=p>1$

ex) periodic chain. $E=\{0,1\}$ $P=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$\{4,6,8,\ldots\}$

$\gcd\{n\geq 1 : P(X_n=i \mid X_0=i)>0\}=2$ for $i=0,1$.

$X(t_i)=X_{t_i}$

PMF $p_{X(t_1),\ldots,X(t_n)} = P(X(t_1)=x_1, \ldots, X(t_n)=x_n)$ $\quad 0\leq t_1\leq \ldots \leq t_n$

$= \frac{P(X(t_1)=x_1,\ldots,X(t_n)=x_n)}{P(X(t_1)=x_1,\ldots,X(t_{n-1})=x_{n-1})} P(X(t_1)=x_1,\ldots,X(t_{n-1})=x_{n-1}) = P(X(t_n)=x_n \mid X(t_{n-1}))\,P(X(t_1)=x_1,\ldots,X(t_{n-1})=x_{n-1})$

$= \ldots\ldots = P(X(t_n)=x_n \mid X(t_{n-1})=x_{n-1})\,P(X(t_{n-1})=x_{n-1} \mid X(t_{n-2})=x_{n-2})\cdot P(X(t_1)=x_1,\ldots,X(t_{n-2})=x_{n-2}) =$

$= \ldots\ldots = P(X(t_n)=x_n \mid X(t_{n-1})=x_{n-1})\,P(X(t_{n-1})=x_{n-1} \mid X(t_{n-2})=x_{n-2})\ldots\ldots P(X(t_1)=x_1)$

p. 588 in the book $\quad (P^{t_n-t_{n-1}})_{x_{n-1},x_n}\times(P^{t_{n-1}-t_{n-2}})_{x_{n-2},x_{n-1}}\times\ldots\times(p(0)P^{t_1})_{x_1}$

stationary distribution $\hat{\mathbb{p}}$ given by $\begin{cases} \hat{\mathbb{p}} = \hat{\mathbb{p}} P \\ \hat{\mathbb{p}} \text{ probability distribution} \end{cases}$

ex) $P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$ $\alpha, \beta \in [0,1]$ real numbers $E = \{0,1\}$

$[p_0 \ \ p_1] \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} = \begin{bmatrix} p_0 \\ p_1 \end{bmatrix}$

$\begin{cases} (1-\alpha) p_0 + \beta p_1 = p_0 \\ \alpha p_0 + (1-\beta) p_1 = p_1 \\ p_0 + p_1 = 1 \end{cases}$

(overdetermined - reason why probability distribution doesn't always exist. may or may not exist, depends on P.)

$\left. \begin{array}{l} p_0, p_1 \geq 0 \\ p_0 + p_1 = 1 \end{array} \right\}$ probability distribution

$\begin{cases} -\alpha p_0 + \beta p_1 = 0 \\ \alpha p_0 - \beta p_1 = 0 \\ p_0 + p_1 = 1 \end{cases}$

$\alpha = \beta = 0 \Rightarrow p_0 + p_1 = 0$ $\quad p_0 = p \quad p_1 = 1 - p. \ \forall p \in [0,1]$

$\alpha = 0, \beta > 0 \Rightarrow p_0 = 1, p_1 = 0.$

$\alpha > 0, \beta = 0 \Rightarrow p_0 = 0, p_1 = 1$

$\alpha, \beta > 0. \quad p_0 = \frac{\beta}{\alpha} p_1 \quad (\frac{\beta}{\alpha} + 1) p_1 = 1 \quad p_1 = \frac{1}{\frac{\beta}{\alpha} + 1} = \frac{\alpha}{\alpha + \beta} \quad p_0 = \frac{\beta}{\alpha + \beta}$

If the chain is started according to the stationary distribution then it will always have the stationary distribution.

$p(n) = [P(X_n = 0) \ \ P(X_n = 1) \ \ldots \ ] = \hat{\mathbb{p}}$

proof: $\mathbb{p}(n) = \mathbb{p}(0) P^n = \hat{\mathbb{p}} P^n = \underbrace{(\hat{\mathbb{p}} P)}_{\hat{\mathbb{p}}} P^{n-1} = \ldots = \hat{\mathbb{p}}$

$P(X_n = i) \to \hat{\mathbb{p}}_i$ as $n \to \infty$ (sort of a steady state) volt-meter.

In the book

$\lim_{n \to \infty} P^n = \hat{\mathbb{P}} = \begin{bmatrix} \hat{\mathbb{p}} \\ \hat{\mathbb{p}} \\ \vdots \end{bmatrix}$

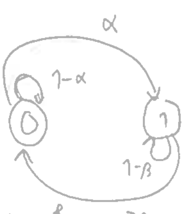
Proof that Hsu's formula implys Albin's

$P(X_n = i) = [\mathbb{p}(0) P^n]_i \longrightarrow \left(\mathbb{p}(0) \begin{bmatrix} \hat{\mathbb{p}} \\ \hat{\mathbb{p}} \\ \vdots \end{bmatrix} \right)_i = \left(\mathbb{p}(0) \begin{bmatrix} \hat{p}_1 & \hat{p}_2 & \hat{p}_3 \\ \hat{p}_1 & \hat{p}_2 & \hat{p}_3 \\ \vdots & \vdots & \vdots \end{bmatrix} \right)_i = [\hat{\mathbb{p}}]_i$

$T_i = \min\{n \geq 1 : X_n = i \text{ if } X_0 = i\}$

$[\hat{\mathbb{p}}]_i = \frac{1}{E[T_i]}$ the longer time it goes in coming back to i is dependent of how likely it is to be in i

ex) $\hat{\mathbb{p}} = [\frac{\beta}{\alpha + \beta} \quad \frac{\alpha}{\alpha + \beta}]$ $\quad E = \{0, 1\}$ $\quad P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$ $\quad \alpha, \beta \in]0,1[$



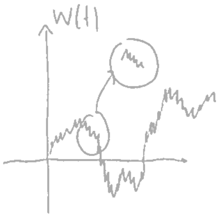
$E[T_0]$ = expected value of smallest time greater than one that you come back to zero when starting in zero

$E[T_0] = 1 \cdot (1-\alpha) + \alpha (1 + \frac{1}{\beta}) = \frac{\alpha + \beta}{\beta} = \frac{1}{\hat{\mathbb{p}}_0}$

↑ geometric distribution

2/11-10 Final words about Markov chains
Poisson process
Wiener process (Brownian motion) - most important process in the world.

Wiener process
continuous process
but nowhere
differentiable



- $W(0)=0$ starts at zero
- independent increments $W(t_1), W(t_2)-W(t_1), \ldots W(t_n)-W(t_{n-1})$ independent for $0 \le t_1 \le \ldots \le t_n$
- stationary increments with $W(t+s)-W(s) \sim N(0, \sigma^2 t)$
- continuous sample paths.

$$\begin{cases} E[W(t)] = 0 \\ Cov[W(s), W(t)] = \underbrace{V[W(1)]}_{\sigma^2} \min(s,t) \end{cases}$$

$\{W(t)\}_{t \ge 0}$ normal process (Gaussian process)

$K = (Cov(X(t_k), X(t_l)))_{k,l}$

DEF: $\{X(t)\}_{t \in T}$ normal iff $\Psi_{X(t_1), \ldots X(t_n)}(\theta\omega_1, \ldots \theta\omega_n) = E[e^{\theta j \sum_{l=1} \omega_l X(t_l)}] = e^{\theta j \omega^T H - \frac{1}{2}\omega^T K \omega \theta^2}$

$\omega = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}$

$H = \begin{bmatrix} E[X(t_1)] \\ \vdots \\ E[X(t_n)] \end{bmatrix}$

Thus $\{X(t)\}_{t \in T}$ normal $\Longleftrightarrow \sum_{i=1}^{n} \omega_i X(t_i)$ is one-dimensional normally distributed for any $n \in \mathbb{N}, \omega_1, \ldots, \omega_n \in \mathbb{R}, t_1, \ldots, t_n \in T$

$E(e^{j\theta N(m,s^2)}) = e^{j\theta m - \frac{1}{2}\theta^2 s^2}$ characteristic function of normal distribution

$$\sum_{i=1}^{n} \omega_i W(t_i) = \omega_n (W(t_n) - W(t_{n-1})) + (\omega_n + \omega_{n-1})(W(t_{n-1}) - W(t_{n-2})) + \ldots + (\omega_n + \ldots + \omega_1) W(t_1)$$

one dimensional normal distributed.

$t_1 \le \ldots \le t_n$ → normally distributed
$t_0 = 0$

Reason that $(W(t))_{t \ge 0}$ is so important is that the non-existing derivative $(W'(t))_{t \ge 0}$ is continuous (completely) white noise!

Verify this empirically $\left(\frac{W(t+\Delta) - W(t)}{\Delta}\right)_{t \ge 0}$ for very small Δ.

noise that disturbs electrical signals for example.

$K_{W'}(s,t) = Cov[W'(s), W'(t)] = \sigma^2 \delta(t-s)$

ex) Nobel prize awarded Black Sholes model for stock-price.

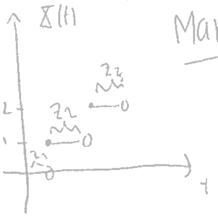
$(S(t))_{t \ge 0}$ $\quad \frac{dS(t)}{dt} = S(t) \frac{dW(t)}{dt}$, $S(0) = S_0$ $\quad S'(t) = S(t) W'(t)$

Wiener process is Markov process $\quad t_1 \le \ldots \le t_n \le t_{n+1}$

$P(W(t_{n+1}) \le x_{n+1} | W(t_n) = x_n, \ldots, W(t_1) = x_1) = P(W(t_{n+1}) - W(t_n) \le x_{n+1} - x_n | W(t_n) = x_n, \ldots \cancel{W(t_1) = x_1}) =$

$= P(W(t_{n+1}) \le x_{n+1} | W(t_n) = x_n)$

Poisson process   discrete valued process, continuous time $(X(t))_{t\geq 0}$



Markov process with continuous time.   most convenient for computer simulations

$T_n = Z_1 + Z_2 + \dots + Z_n$ time for n'th jump.

$X(t) = \max\{n : T_n \leq t\}$

$(Z_n)_{n=1}^{\infty}$ interarrival time, independent, positive random variables that are $\exp(\lambda)$ distributed.

Renewal process is such process where $(Z_n)_{n=1}^{\infty}$ can be any sequence of iid positive random variables.

Renewal process = counting process (counting arrivals)

Alternative definition Poisson process $(X(t))_{t\geq 0}$ with intensity λ is given by:

- $X(0) = 0$
- $X(t)$ has independent increments.

  $X(t+s) - X(s)$ is $Po(\lambda t)$-distributed.

Another alternative definition of Poisson process $(X(t))_{t\geq 0}$ with intensity λ is

- $X(0) = 0$
- $X(t)$ independent increments
- $X(t)_{t\geq 0}$ integer valued. ($\mathbb{N}$-valued), non-decreasing with $P(X(t+\Delta t) - X(t) = 0) = 1 - \lambda\Delta t + o(\Delta t)$

$P(X(t+\Delta t) - X(t) = 1) = \lambda\Delta t + o(\Delta t)$

$P(X(t+\Delta t) - X(t) \geq 2) = o(\Delta t)$

$E[X(t)] = E\{X(n)\}t = \lambda t$

$Cov[X(s), X(t)] = V[X(1)]\min(s,t) = \lambda\min(s,t)$

can use characteristic fcn to show that a sum of two poisson is poisson!

Note: $Y(t) = W(t) + \sigma^2 t$

$E[Y(t)] = 0 + \sigma^2 t$

$Cov[Y(s), Y(t)] = Cov[W(s), W(t)] = \sigma^2\min(s,t)$

If $(X_1(t))_{t\geq 0}$ and $(X_2(t))_{t\geq 0}$ are independent poisson processes with intensity parameters λ_1 and λ_2 respectively, then $Y(t) = X_1(t) + X_2(t)$ is a poisson process with intensity $\lambda_1 + \lambda_2$.

$M_{X_1(t)}(s) = E[e^{sX_1(t)}] = E[e^{sPo(\lambda t)}] = e^{\lambda_1(s-1)}$

$M_{X_2(t)}(s) = \dots = e^{\lambda_2(s-1)}$

$M_{X_1(t)+X_2(t)}(s) = E[e^{s(X_1(t)+X_2(t))}] = E[e^{sX_1(t)}]E[e^{sX_2(t)}] = e^{\lambda_1(s-1)}e^{\lambda_2(s-1)} = e^{(\lambda_1+\lambda_2)(s-1)}$

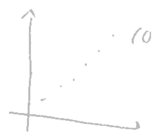
Final words about Markov chains

not communicating

$(X_n)_{n=0}^{\infty}$ $E = \{0,1,2,\ldots\} = T \cup \begin{pmatrix}\text{communicating} \\ \text{recurrent} \\ \text{states 1}\end{pmatrix} \cup \begin{pmatrix}\text{communicating} \\ \text{recurrent} \\ \text{states 2}\end{pmatrix} \ldots \cup \begin{pmatrix}\text{communicating} \\ \text{recurrent states } n\end{pmatrix} \cup$

T: transient states

$\begin{pmatrix}\text{absorbing} \\ \text{state 1}\end{pmatrix} \cup \ldots \cup \begin{pmatrix}\text{absorbing} \\ \text{state } m\end{pmatrix}$ can't get out of any of the sets except the transient set.



could be transient states only.

infinitely many transient states

if finitely many transient states will after a certain time be in another set w certainty.

Stationary distribution

$\hat{p} = \hat{p} P \quad \lim_{n\to\infty} P(X_n = i) = \lim_{n\to\infty} P(X_n = i \mid X_0 = j) = (\hat{p})_i = \frac{1}{E[T_i]}$

For a chain where all states communicate (irreducible chain) and are aperiodic.
(all states must have the same period)

ex) $E = \{0,1\}$ $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ can start it randomly, otherwise it is not random

$\hat{p} = [\frac{1}{2} \ \frac{1}{2}]$ $\quad [\frac{1}{2} \ \frac{1}{2}]\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [\frac{1}{2} \ \frac{1}{2}]$ $\quad \hat{p}(0) = [1 \ 0]$

$\lim_{n\to\infty} P(X_n = i) = \frac{1}{2} \quad P(X_{2n} = 0) = 1 \quad P(X_{2n+1} = 0) = 0$

it's periodic - that's why this isn't true.

18/11-10

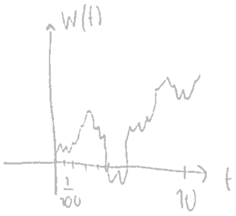
Computer problem 3

Wiener process $\{W(t)\}_{t\geq 0}$ independent stationary increment process with $W(0) = 0$ and

$W(t+s) - W(s) \sim N(0, \sigma^2 t) \quad \sigma^2 = 1 \qquad \varepsilon = \frac{1}{100}$ $\qquad \varepsilon_1 \quad \varepsilon_1 + \varepsilon_2$ $\quad t_1 \quad t_2$

Find $\mathrm{cov}(\Delta_\varepsilon(s), \Delta_\varepsilon(t))$ where $\Delta_\varepsilon(t) = \frac{W(t+\varepsilon) - W(t)}{\varepsilon}$ $\quad t > 0$ $\quad$ plot $\Delta_\varepsilon(t)$ $\quad \varepsilon_2$

plot $W(t)$



$\{W(\frac{i}{100})\}_{i=0}^{1000} \quad W(\frac{i}{100}) = \sum_{j=1}^{i} \varepsilon_j \quad \varepsilon_j \sim N(0, \frac{1}{100})$

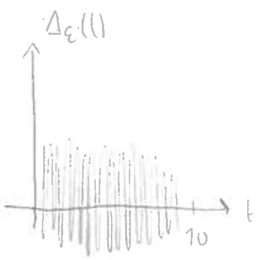
Need a vector of ε:s

$[\varepsilon_1, \ldots \varepsilon_{1000}]$ $\qquad \varepsilon_3 + \varepsilon_2$

$W = [0, \varepsilon_1, \varepsilon_1 + \varepsilon_2, \ldots \sum_{i=1}^{1000} \varepsilon_i]$

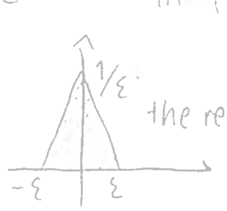
$W_{t+\varepsilon} = [\varepsilon_1, \varepsilon_1 + \varepsilon_2, \ldots, \sum_{i=1}^{1000} \varepsilon_i]$ $\qquad t_1, t_2, t_3$

$\frac{W_{t+\varepsilon} - W}{1/100}$ $\qquad t_3 - t_1$



$\mathrm{Cov}[\Delta_\varepsilon(s), \Delta_\varepsilon(t+s)] = \mathrm{Cov}\left[\frac{W(s+\varepsilon) - W(s)}{\varepsilon}, \frac{W(t+s+\varepsilon) - W(t+s)}{\varepsilon}\right] = \frac{\min(s+\varepsilon, t+s+\varepsilon) - \min(s+\varepsilon, t+s) - \min(s, t+s+\varepsilon) + \min(s, t+s)}{\varepsilon^2}$

$\mathrm{Cov}[W(s), W(t+s)] = \min[s, t+s]$



the result of the plot of covariance converges to $\delta(t)$ as $\varepsilon \to 0$.

5.8 Martingales one of the most important classes of processes

$\{X_n\}_{n=0}^{\infty}$ discrete time stochastic process

$E[X_{n+1} | X_1 = x_1, \ldots, X_n = x_n] = x_n$ DEFINITION OF a MARTINGALE PROCESS. wants to model a fair game is exactly fair.

$E[X_{n+1} | X_1, \ldots, X_n] = X_n$ $g(X_1, \ldots, X_n)$ $E[X_{n+1} | F_n] = X_n$.

ex) $Y_1, Y_2, \ldots$ independent, zero-mean r.v's

$X_n = \sum_{i=1}^{n} Y_i$ $n = 0, 1, 2$ is a martingale (the most basic one)

accumulated gain after n turns of the game.

each X a sum of Y's.

$E[X_{n+1} | X_1 = x_1, \ldots X_n = x_n] = E[\underbrace{Y_{n+1}}_{\text{independent of } Y_1, \ldots, Y_n} + \underbrace{\sum_{i=1}^{n} Y_i}_{X_n = x_n} | Y_1 = x_1 \; Y_1 + Y_2 = x_2, \ldots \sum_{i=1}^{n} Y_i = x_n) = 0 + x_n = x_n$

can forget about the past, because the r'v are independent

The process that is zero all the time is also a martingale but not a very interesting one.

conditioning (continued)

$E[Y | X = x] = g(x)$ $\left[\int_{-\infty}^{\infty} y \underbrace{f_{Y|X}(y|x)}_{\frac{f_{Y,X}(y,x)}{f_X(x)}} dy = g(x) \right]$ when they are continuous.

$E[Y | X] = g(X)$ more convenient because the notation is more compact. means the same thing but X is not determined yet.

$E[Y | X_1 = x_1, \ldots, X_n = x_n] = g(x_1, \ldots, x_n)$

$E[Y | X_1, \ldots, X_n] = g(X_1, \ldots, X_n)$ the conditional expectation is random.

the information given by knowing $X_1, \ldots, X_n$ is denoted F_n

$\boxed{E[Y | X_1, \ldots, X_n] \equiv E[Y | F_n]}$

What are the properties of the conditional expectation $E[Y | F_n]$?

$E[aY_1 + bY_2 | F_n] = a E[Y_1 | F_n] + b E[Y_2 | F_n]$ (proof same as ordinary ...)

$Y \geq 0 \Rightarrow E[Y | F_n] \geq 0$ • Y is F_n-measurable $\Rightarrow E[Y | F_n] = Y$ (because we know the value of Y)

Random variable Z is F_n-measurable (explainable) if we can say what the value of Z is knowing the information F_n.

ex) $Z = X_1 + \ldots + X_n \Rightarrow Z$ is F_n-measurable.

• Z is F_n-measurable $\Rightarrow E[ZY | F_n] = Z E[Y | F_n]$ we know Z

Y independent of $F_n \Rightarrow E[Y | F_n] = E[Y]$

Towering: $E[E[\underbrace{Y | F_m}_{X_1, \ldots X_m}] | F_n] = E[Y | F_n]$ when $m > n$. version (generalization) of $E[Y] = \int_{-\infty}^{\infty} E[Y | X = x] f_X(x) dx$

Important!

$X_1, \ldots, X_n, X_{n+1}, \ldots X_m$

$E[\underbrace{E[Y | F_n]}_{E[Y | X_1, \ldots, X_n]}] = E[Y]$ $\left(= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} E[Y | X_1 = x_1, \ldots X_n = x_n] f_{X_1, \ldots X_n}(x_1, \ldots x_n) dx_1 \ldots dx_n \right)$

· Jensens inequality

option which has the correct value, a slot machine that gives 100% back ...

$E[g(Y)|F_n] \geq g(E[Y|F_n])$

Consequences of Martingale definition

If $\{X_n\}_{n=0}^{\infty}$ is martingale then a) $E[X_n] = E[X_0]$: is constant (same all the time)

b) $E[X_{m+n}|F_n] = X_n$ for $m \geq 0$.

Proof

b) $E[X_{m+n}|F_n] = E[\underbrace{E[X_{m+n}|F_{\overset{\text{greater than } n}{m+n-1}}]}_{X_{m+n-1}}|F_n] = \begin{bmatrix}\text{Martingale} \\ \text{definition}\end{bmatrix} = E[X_{m+n-1}|F_n] = \dots = E[X_{n+1}|F_n] = X_n$

a) $n=0$ in b) $E[X_m|F_0] = X_0 \Rightarrow \underbrace{E[E[X_m|F_0]]}_{= E[X_m]} = E[X_0]$

$\{X_n\}_{n=0}^{\infty}$ submartingale if $E[X_{n+1}|F_n] \geq X_n$

$\{X_n\}_{n=0}^{\infty}$ supermartingale if $E[X_{n+1}|F_n] \leq X_n$

ex) submartingale/supermartingale

$Y_1, \dots, Y_n$ independent r.v with common expectation $E[Y_n] = c$

$X_n = \sum_{i=0}^{n} Y_i$ martingale if $c=0$

— " — submartingale if $c>0$

— " — supermartingale if $c<0$

$E[X_{n+1}|X_1 = x_1, \dots X_n = x_n] = E[Y_{n+1} + \sum_{i=1}^{n} Y_i | Y_1 = x_1, \dots, \sum_{i=1}^{n} Y_i = x_n] = c + x_n$

Theorem

$\{X_n\}_{n=0}^{\infty}$ supermartingale $\Rightarrow \{-X_n\}_{n=0}^{\infty}$ submart.

— " — submart $\Rightarrow$ — " — supermart

$\{X_n\}_{n=0}^{\infty}$ sub- and supermart $\Rightarrow \{X_n\}_{n=0}^{\infty}$ martingale

Normally distributed $\Rightarrow$ every linear comb normally distributed.

$\sum_{i=1}^{n} a_i(c_i + X_i) = \underbrace{\sum_{i=1}^{n} a_i c_i}_{\text{constant}} + \sum_{i=1}^{n} a_i X_i$

Normally distributed with shifted expectation

19/11-10

$\{X_n\}_{n=0}^{\infty}$ time discrete random process

$E[X_{n+1} | X_n, \ldots X_1] = E[X_{n+1} | F_n] = X_n$ martingale

$\geq$ submartingale

$\leq$ supermartingale

Theorem

Doob decomposition

If $\{X_n\}_{n=0}^{\infty}$ is a submartingale then there exists a martingale $\{M_n\}_{n=0}^{\infty}$ and an increasing process $\{A_n\}_{n=0}^{\infty}$ such that $X_n = M_n + A_n$

ex) $X_n = \sum_{i=1}^{n} Y_i$ where $Y_1, Y_2, \ldots$ are independent with common expected value $E[Y_i] = c > 0$.

independent of history!

$E[X_{n+1} | X_1, \ldots, X_n] = E[Y_{n+1} + X_n | X_1, \ldots, X_n] = E[Y_{n+1}] + X_n = c + X_n > X_n$

fixed

$X_n = \underbrace{\sum_{i=1}^{n} (Y_i - c)}_{M_n \text{ martingale}} + nc$

zero-mean indep. r.v.

Stopping-times

A stopping time is a $\mathbb{N}$-valued r.v. T such that $\{T=n\}$ is F_n-measurable, that is, one can tell wether the event $\{T=n\}$ happens (ed) or not using $S_1, \ldots, S_n$

ex 51 A gambler has \$100 starting capital and repeatingly plays a simple \$1 slot machine where at each play the player earns \$5 with probability 1/5 and louses his \$1 with probability 4/5

(fair slot machine) gets back \$1 on average.

Fortune after round n is $S_n = S_0 + \sum_{i=1}^{n} X_i$ where $S_0 = 100$ and X_i is gain. (The fair game Martingale) same proof as last lecture but starts at 100.

T_1 = first time at which fortune is \$200 = $\min\{n : S_n = 200\}$ stopping time with respect to S_n

T_2 = time for ruin = $\min\{n : S_n = 0\}$

$T_3 = \min\{T_1, T_2\}$ = first time at which either fortune is \$200 or ruin $= \min\{n : S_n = 200 \text{ or } S_n = 0\}$

$\{S_n\}_{n=0}^{\infty}$ is martingale $E[S_{n+1} | S_0, S_1, \ldots S_n] = S_n$?

$\{T_1\}, \{T_2\}, \{T_3\}$ can all be cheched wheter they hold by inspection of $S_1, \ldots, S_n$

$T = \max\{S_n : n\} = \max_{n \geq 0} S_n$ is NOT stopping time. (can't tell wheter it is the maximum because one doesn't know the future)

Optional stopping theorem

If $\{X_n\}_{n=0}^{\infty}$ is a martingale and T a stopping time, then $E[X_T] = E[X_0]$ $\left[\text{For a martingale } E[X_n] = E[X_0]\right]$

Under technical conditions

⋮

1) $E[T] < \infty$

2) $E[|X_T|] < \infty$

3) $\lim_{n\to\infty} E[|X_n| I_{\{T>n\}}] = 0$ $\quad I_{\{T>n\}} = \{1, T>n;\ 0, T \not> n\}$

Application to ex. 5.1

T_3 = first time at which $S_n = 0$ or $S_n = 200$

$E[S_{T_3}] = E[S_0] = 100$

$= P(\text{ruin}) \cdot 0 + P(200) \cdot \underset{\approx}{200}$ (200-204) (can only have two values)

$P(\text{ruin}) = 1 - p(\text{success})$ $\Rightarrow P(\text{success getting } 200) \approx \frac{1}{2}$

1) $E[T] < \infty$ (by some arguments)

2) is satisfied. $E[|S_{T_3}|] \leq 200 < \infty$

3) $E[\underbrace{|S_n|}_{\leq 200} \underbrace{J_{\{t>n\}}}_{\to 0}] \to 0$

$E[S_{T_1}] = 200 \neq E[S_0] = 100$ (not really correct?! can have 198, and earn 5 $\Rightarrow$ 203)

$E[S_{T_2}] = 0 \neq E[S_0] = 100$

Chapter 6 Analysis and processing of Random processes.

continuous time martingales

$\{X(t)\}_{t\geq 0}$ $\quad E[X(t)|F_s] = X(s)$ for $s \leq t$ where F_s is the information available from $\{X(r)\}_{r\leq s}$

ex) wiener process

$\{W(t)\}_{t\geq 0}$ is mart, because $E[W(t)|\{W(r)\}_{r\leq s}] = E[W(t) - W(s) + W(s)|\{W(r)\}_{r\leq s}] =$

$= E[\underbrace{W(t) - W(s)}_{0}] + W(s)$

T = first time $W(t) = 200$ or $W(t) = -100$ (stopping time)

$0 = E[W(0)] = E[W(T)] = 200 \cdot p(\text{reach 200 first}) + (-100) \cdot p(\text{reach -100 first})$ $\quad p(\text{reach -100 first}) = 1 - p(\text{reach 200 first})$ $\quad P(\text{reach 200 first}) = 1/3$

Continuity, derivation and integration of random processes. $\{X(t)\}_{t\in\mathbb{R}}$

Continuity: $X(t+\varepsilon) - X(t) \to 0$ as $\varepsilon \to 0$ $\qquad E[(X(t+\varepsilon) - X(t))^2] \to 0$ as $\varepsilon \to 0$

Derivative: $\frac{X(t+\varepsilon) - X(t)}{\varepsilon} - X'(t) \to 0$ as $\varepsilon \to 0$ $\qquad E[(\frac{X(t+\varepsilon) - X(t)}{\varepsilon} - X'(t))^2] \to 0$ as $\varepsilon \to 0$

Integration (riemann): $\sum_{i=1}^{n} X(s_i)(t_i - t_{i-1}) - \int_a^b X(s)ds \to 0$, $n \to \infty$ when $a = t_0 < t_1 < \ldots < t_n = b$ satisfies $\max_{1\leq i\leq n} t_i - t_{i-1} \to 0$

as $n \to \infty$, $s_i \in [t_{i-1}, t_i]$ $\qquad E[(\sum_{i=1}^{n} X(s_i)(t_i - t_{i-1}) - \int_a^b X(s)ds)^2] \to 0$, as $n \to \infty$

Mean square convergence $\quad$ There are processes converging in first mode but not in the second! $\wedge^2$

Autocorrelation function $\quad R_{XX}(s,t) = E[X(s)X(t)]$

Autocovariance function $\quad K_{XX}(s,t) = Cov[X(s)X(t)] = R_{XX}(s,t) - \mu_X(s)\mu_X(t)$

Mean function $\quad \mu_X(t) = E[X(t)]$

say interesting things, but NOT everything $\quad$ dependence among function values etc.

$F_{X(t_1),\dots,X_n(t)}(x_1,\dots,x_n) = P(X(t_1)\le x_1,\dots,X(t_n)\le x_n)$

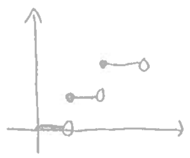
in general, to calculate this is impossible $\Rightarrow$ use the above functions to describe the distribution
For normal / Gaussian processes it is enough!

Continuity

$$E[(X(t+\varepsilon)-X(t))^2] = R_{XX}(t+\varepsilon,t+\varepsilon)-2R_{XX}(t,t+\varepsilon)+R_{XX}(t,t)$$

X is continuous iff R_{XX} is continuous

ex) Poisson process



if $\varepsilon>0$, because of stationary increments

$$E[(X(t+\varepsilon)-X(t))^2] = E[X(|\varepsilon|)^2] = \lambda|\varepsilon|+\lambda^2\varepsilon^2 \to 0 \text{ as } \varepsilon\to 0$$

"explanation": probability of jumping in a specific point $=0$

X is differentiable $\Leftrightarrow$ R_{XX} is two times differentiable.

If X is differentiable with derivative process X'

$\mu_{X'}(t)=?$

$R_{X'X'}(s,t)=?$

$$\boxed{\mu_{X'}(t)} = E[X'(t)] = E\left[\lim_{\varepsilon\to 0}\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right] = \lim_{\varepsilon\to 0} E\left[\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right] =$$

$$\lim_{\varepsilon\to 0}\frac{\mu_X(t+\varepsilon)-\mu_X(t)}{\varepsilon} = \boxed{\mu'_X(t)} \qquad \mu_{X'}(t)=\mu'_X(t)$$

$$R_{X'X'}(s,t) = E[X'(s)X'(t)] = E\left[\lim_{\varepsilon\to 0}\left(\frac{X(s+\varepsilon)-X(s)}{\varepsilon}\right)\left(\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right)\right] = \lim_{\varepsilon\to 0}\frac{R_{XX}(s+\varepsilon,t+\varepsilon)-R_{XX}(s+\varepsilon,t)-R_{XX}(s,t+\varepsilon)+R_{XX}(s,t)}{\varepsilon^2}$$

$$f(t)+f'(t)\varepsilon+\frac{f''(t)}{2}\varepsilon^2-2f(t)+f(t)+f'(t)(-\varepsilon)+\frac{1}{2}f''(t)(-\varepsilon)^2$$

$$= \frac{\partial^2 R_{XX}(s,t)}{\partial s\partial t} \qquad \frac{f(t+\varepsilon)-2f(t)+f(t-\varepsilon)}{\varepsilon^2}\to f''(t)$$

Integration $E\left[\int_a^b X(s)ds\right] \approx E\left[\sum_{i=1}^n X(s_i)(t_i-t_{i-1})\right] = \sum_{i=1}^n E[X(s_i)](t_i-t_{i-1}) \approx \int_a^b E[X(s)]ds$

$$\operatorname{cov}\left[\int_a^b X(s)ds,\int_a^b X(t)dt\right] = \int_a^b\int_a^b \operatorname{cov}[X(s),X(t)]\,ds\,dt$$

$$E\left[\int_a^b X(s)ds\int_a^b X(t)dt\right] = \int_a^b\int_a^b R_{XX}(s,t)\,ds\,dt$$

25/11-10

p: probability success

$S_0 = 100 \quad S_n = \sum_{i=1}^{n} X_i \quad X_i = \begin{cases} 4 & \text{wp } 1/5 \\ -1 & \text{wp } 4/5 \end{cases}$

$T = \min\{n : S_n \geq 200 \text{ or } S_n = 0\}$

$100 = E[S_0] = E[S_T] \begin{matrix} \geq p \cdot 200 \\ \leq p \cdot 203 \end{matrix} + (1-p)\cdot 0$ divides: $p \in [\frac{100}{203}, \frac{1}{2}]$

$\hat{p}$ proportion of successes from simulations.

$p = \hat{p} \pm \underbrace{\sqrt{\hat{p}(1-\hat{p})}}_{\leq 0.5} / \sqrt{n} \cdot 4$ (standard deviation) $p = \hat{p} \pm \frac{2}{\sqrt{n}}$ $\quad \frac{2}{\sqrt{n}} \approx \frac{1}{1000}$ to get a better estimate.

Section 6.3-6.5

wss process $\underset{X_t}{\{X(t)\}_{t\in\mathbb{R}}}$ $\underset{X_n}{\{X(n)\}_{n\in\mathbb{Z}}}$ linear "device" (invariant in time)

$\underset{\text{in signal}}{\xrightarrow{X}} \boxed{T} \underset{\text{out signal}}{\xrightarrow{Y}}$ $\quad Y(t) = (TX)(t) \quad Y_n = (TX)_n$

linearity $\quad T(\alpha X_1 + \beta X_2) = \alpha T(X_1) + \beta T(X_2)$

time-homogeneity (invariant in time) $\quad T(\underbrace{X(t-t_0)}_{\text{delayed signal}}) = TX(t-t_0)$

impulse response $h(t) = (T\delta)(t)$ $\quad h(n) = (T\delta)_n$

Dirac function (distribution)

$\delta_n = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases}$

$\begin{cases} 0, & x \neq 0 \\ \infty, & x = 0 \end{cases}$ $\quad \int_{-\infty}^{\infty} \delta \, dx = 1$

Any in signal X_n in discrete time can be written $X_n = \sum_{k=-\infty}^{\infty} X_k \delta(n-k)$

$(TX)_n = T\left(\sum_{k=-\infty}^{\infty} X_k \delta(n-k)\right) = \sum_{k=-\infty}^{\infty} X_k \underbrace{T(\delta(n-k))}_{\substack{(T\delta)(n-k) \\ h(n-k)}} = \sum_{k=-\infty}^{\infty} X_k h(n-k) = (X * h)(n)$ discrete case

$(TX)(t) = \int_{-\infty}^{\infty} X(s) h(t-s) ds = (X * h)(t)$ continuous case

wss process is such that $\cdot\, R_{XX}(\tau) = E[X(t)X(t+\tau)]$ doesn't depend on t.

$\cdot\, \mu_X = E[X(t)] = $ constant.

For the in-signal

is it a wss to?) yes!

often assumes the mean to be zero

out-signal $R_{YY}(\tau)$

$\mu_Y = E[Y(t)] = E[\int_{-\infty}^{\infty} X(s) h(t-s) ds] = \int_{-\infty}^{\infty} \underbrace{E[X(s)]}_{\mu_X} h(t-s)\, ds = \mu_X \int_{-\infty}^{\infty} h(s) ds$ variable substitution

$= h * h(-\cdot) * R_{XX}$

$R_{YY}(\tau) = E[\int_{-\infty}^{\infty} X(r) h(t-r) dr \int_{-\infty}^{\infty} X(s) h(t+\tau-s) ds] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \underbrace{E[X(r)X(s)]}_{\substack{R_{XX}(r-s) \\ \hat{r}-\hat{s}}} h(t-r) h(t+\tau-s)\, dr\, ds$

$-\hat{r} \quad \tau-\hat{s}$

$r = \hat{r}+t \quad s = \hat{s}+t \quad d\hat{r}\, d\hat{s}$

because convolutions are involved, one might suspect Fourier transforms to be involved.

$S_{\underline{XX}}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{\underline{XX}}(\tau)\, d\tau$ Fourier transform of $R_{\underline{XX}}$ called power spectral density of $R_{\underline{XX}}^{\sigma}(\underline{X})$

$R_{\underline{XX}}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega\tau} S_{\underline{XX}}(\omega)\, d\omega$

all frequencies are equally important in white noise!

ex (white noise)

$R_{\underline{XX}}(\tau) = \sigma^2 \delta(\tau)$

$S_{\underline{XX}}(\omega) = \sigma^2$

Delta function, Dirac's d, dirac function, delta distribution, dirac distribution, continuous time version of Kronecker's delta

$\delta: \mathbb{R} \to \mathbb{R}$ $\quad \delta(t) = 0, \ t \neq 0$

$\int_{-\infty}^{\infty} g(t)\delta(t)\, dt = g(0)$

$\Theta =$  δ

$\delta(t) = \Theta'(t)$

filter function

$S_{\underline{YY}}(\omega) = H(\omega)\overline{H(\omega)}\, S_{\underline{XX}}(\omega) = |H(\omega)|^2 S_{\underline{XX}}(\omega)$ $\quad H(\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} h(t)\, dt.$

$\overline{H(\omega)} = \int_{-\infty}^{\infty} e^{j\omega t} h(t)\, dt = \int_{-\infty}^{\infty} e^{-j\omega t} h(-t)\, dt.$

$S_{\underline{XX}}(\Omega) = \sum_{h=-\infty}^{\infty} e^{-j\Omega h} R_{\underline{XX}}(h)$

$R_{\underline{XX}}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\Omega h} S_{\underline{XX}}(\Omega)\, d\Omega$ $\quad H(\Omega) = \sum_{h=-\infty}^{\infty} e^{-j\Omega h} h(h)$

$S_{\underline{YY}}(\Omega) = |H(\Omega)|^2 S_{\underline{XX}}(\Omega)$

two WSS processes

$R_{\underline{XY}}(\tau) = E[\underline{X}(t)\underline{Y}(t+\tau)]$ $\quad \{\underline{X}(t)\}_{t\in\mathbb{R}}, \{\underline{Y}(t)\}_{t\in\mathbb{R}}$ WSS (says something about dependence between $\underline{X}$ and $\underline{Y}$)

$S_{\underline{XY}}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{\underline{XY}}(\tau)\, d\tau$

$R_{\underline{X}, h*\underline{X}} = E[\underline{X}(t)\int_{-\infty}^{\infty} \underline{X}(s) h(t+\tau-s)\, ds] = \int_{-\infty}^{\infty} h(t+\tau-s) R_{\underline{XX}}(t-s)\, ds = \begin{bmatrix} \hat{t} = t-s & s=+\infty \to \hat{t} = -\infty \\ d\hat{t} = -ds & s=-\infty \to \hat{t} = \infty \end{bmatrix} =$

$\underline{Y} = h * \underline{X}$

$= -\int_{+\infty}^{-\infty} h(\tau+\hat{t}) R_{\underline{XX}}(\hat{t})\, d\hat{t} = \int_{-\infty}^{\infty} h(\tau+\hat{t}) R_{\underline{XX}}(\hat{t})\, d\hat{t} = \begin{bmatrix} \hat{t} = -\hat{\hat{t}} & R_{\underline{XX}}(\hat{t}) = R_{\underline{XX}}(-\hat{t}) \\ R_{\underline{XX}} \text{ symmetric} & \end{bmatrix} = \int_{-\infty}^{+\infty} h(\tau-\hat{\hat{t}}) R_{\underline{XX}}(\hat{\hat{t}})\, d\hat{\hat{t}} =$

$= (h * R_{\underline{XX}})(\tau)$ $\quad S_{\underline{X}, h*\underline{X}}(\omega) = H(\omega) S_{\underline{XX}}(\omega)$

[26/11-10]

Convolution (faltning)

Continuous time $f, g: \mathbb{R} \to \mathbb{R}$ $\quad f * g(t) = \int_{-\infty}^{\infty} f(s) g(t-s) ds = \int_{-\infty}^{\infty} f(t-s) g(s) ds$

Discrete time $f, g: \mathbb{Z} \to \mathbb{R}$ $\quad f * g(n) = \sum_{h=-\infty}^{\infty} f(h) g(n-h) = \sum_{h=-\infty}^{\infty} f(n-h) g(h)$

ex) X and Y independent random variables with pdf f_X, f_Y

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(r) f_Y(z-r) dr = f_X * f_Y(z)$$

Fourier transforms

Continuous time $f: \mathbb{R} \to \mathbb{R}$ $\quad \hat{f}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega x} f(x) dx$

inverse transform $\quad f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega x} \hat{f}(\omega) d\omega$

Discrete time : $f: \mathbb{Z} \to \mathbb{R}$ $\quad \hat{f}(\omega) = \sum_{h=-\infty}^{\infty} e^{-j\omega h} f(h)$

inverse transform $\quad f(h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega h} \hat{f}(\omega) d\omega$

$$\widehat{(f * g)}(\omega) = \hat{f}(\omega) \hat{g}(\omega)$$

Proof (continuous time)

$$\widehat{(f * g)}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega x} \left[\int_{-\infty}^{\infty} f(y) g(x-y) dy \right] dx = \int_{-\infty}^{\infty} e^{-j\omega y} \left[\int_{-\infty}^{\infty} e^{-j\omega(x-y)} g(x-y) dx \right] f(y) dy = \hat{g}(\omega) \hat{f}(\omega)$$

Impulses - delta function

Continuous time $\delta: \mathbb{R} \to \mathbb{R}$ $\quad \delta(x) = 0 \quad x \neq 0 \quad \int_{-\infty}^{\infty} g(x) \delta(x) dx = g(0)$

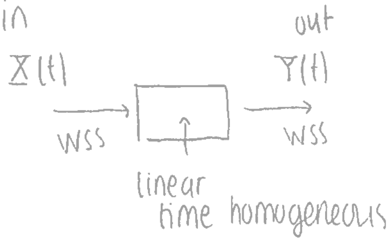
Discrete time $\delta: \mathbb{Z} \to \mathbb{R}$ $\quad \delta(h) = \begin{cases} 1, & h = 0 \\ 0, & h \neq 0 \end{cases}$

ex) white noise

continuous time: random process $\{W(t)\}_{t \in \mathbb{R}}$ that is zero-mean, wss with $R_{WW}(\tau) = \sigma^2 \delta(\tau)$

Discrete time: random process $\{W(h)\}_{h \in \mathbb{Z}}$ that is zero-mean wss with $R_{WW} = \sigma^2 \delta(h)$

Linear system (filter)



$Y(t) = (TX)(t)$

$h(t) = (T\delta)(t) \Rightarrow Y(t) = (h * X)(t)$

impulse response

$R_{YY}(\tau) = h * h(-x) * R_{XX}(\tau)$

$H_Y = H_X \int_{-\infty}^{\infty} h(t) dt.$

PSD (power spectral density) $S_{XX}(\omega) = (\hat{R}_{XX})(\omega)$

filter function $H(\omega) = \hat{h}(\omega)$

$S_{YY}(\omega) = (\hat{R}_{YY})(\omega) = \hat{h}(\omega)(\hat{h}(-)) (\omega) S_{XX}(\omega) = |H(\omega)|^2 S_{XX}(\omega)$

$H(\omega) \cdot \overline{H}(\omega)$

ex) Send random WSS signal $X(t)$ on noisy channel. We receive $Y(t) = X(t) + \underbrace{N(t)}_{\text{noise}}$

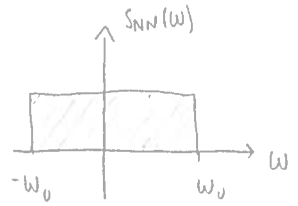
N is WSS zero-mean, independent of X

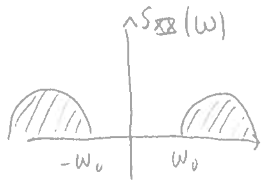
we want to construct a filter that removes noise "as good as possible"

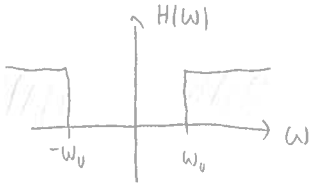
$Y(t) \longrightarrow$ [Filter $h(t)$ $H(\omega)$] $\longrightarrow Z(t)$ which is as like $X(t)$ as possible

X and N zero-mean WSS $R_{XX}(\tau)$ $R_{NN}(\tau)$

$S_{XX}(\omega)$ $S_{NN}(\omega)$







<u>Wiener filter</u> (the best filter for the example)

$$H(\omega) = \frac{S_{XX}(\omega)}{S_{XX}(\omega) + S_{NN}(\omega)}$$

X is WSS process with autocorrelation function R_{XX}

$R_{XX}(\tau) = E[X(t) X(t+\tau)]$

1) $R_{XX}(\tau) = R_{XX}(-\tau)$
2) $R_{XX}(0) = E[X(t)^2] \geq 0$
3) $|R_{XX}(\tau)| \leq R_{XX}(0)$

Proof 3

$0 \leq E[(X(t) \pm X(t+\tau))^2] = 2(R_{XX}(0) \pm R_{XX}(\tau))$

$S_{XX}(\omega) = (\hat{R}_{XX})(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{XX}(\tau)\, d\tau$

$S_{XX}(\omega)$ real valued non-negative

$S_{XX}(\omega) = S_{XX}(-\omega)$

$R_{XX}(0) = E[X(t)^2] = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(\omega)\, d\omega$

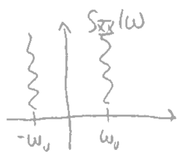
$R_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} S_{XX}(\omega)\, d\omega$

ex) $R_{XX}(\tau) = \sigma^2 \delta(\tau)$ (white noise) $\Rightarrow S_{XX}(\omega) = \sigma^2$
equal amount of all frequences

$X(t) = Y\cos(\omega_0 t) \Rightarrow R_{XX}(\tau) = \cos(\omega_0 \tau)$

zero-mean
unit variance

$S_{XX}(\omega) \overset{?}{=} \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$



ex) $R_{XX}(\tau) = e^{-|\tau|}$

$S_{XX}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{XX}(\tau)\, d\tau =$

$\int_0^{\infty} e^{-j\omega\tau} e^{-\tau}\, d\tau + \int_{-\infty}^{0} e^{-j\omega\tau} e^{\tau}\, d\tau = \left[\frac{e^{-(1+j\omega)\tau}}{-(1+j\omega)}\right]_0^{\infty}$

$+ \left[\frac{e^{(1-j\omega)\tau}}{1-j\omega}\right]_{-\infty}^{0} = \frac{1}{1+j\omega} + \frac{1}{1-j\omega} = \frac{2}{1+\omega^2}$

$R_{XX}(\tau) = \frac{2}{1+\tau^2} \Rightarrow S_{XX}(\omega) = 2\pi e^{-|\omega|}$

If $X(t)$ and $Y(t)$ are WSS, then $R_{XY}(\tau) = E[X(t)Y(t+\tau)]$

$S_{XY}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega\tau} R_{XY}(\tau)\, d\tau$

$R_{XY}(\tau) = (h * R_{XX})(\tau) \qquad S_{XY}(\omega) = H(\omega) S_{XX}(\omega)$

Derivation of Wiener filter (not in book)

Select to minimize $E[(Z(t) - X(t))^2] = E[((h * Y(t)) - X(t))^2] = E[(h * X(t) + h * N(t) - X(t))^2] =$

$R_{h*X, h*X}(0) + R_{H*N, H*N}(0) + R_{XX}(0) + \underbrace{2R_{h*X, h*N}}_{=0} - 2R_{h*X, X}(0) - \underbrace{2R_{h*N, X}(0)}_{0} =$

$= E[\int_{-\infty}^{\infty} h(t-s)N(s)ds\, X(t)]$ could lift out the integral $= \frac{1}{2\pi}\int_{-\infty}^{\infty}(|H(\omega)|^2 S_{XX}(\omega) + |H(\omega)|^2 S_{NN}(\omega) + S_{XX}(\omega) - 2H(\omega)S_{XX}(\omega))\, d\omega$

Differentiate with respect to H: $2HS_{XX} + 2HS_{NN} - 2S_{XX} = 0 \Rightarrow H(\omega) = \frac{S_{XX}(\omega)}{S_{XX}(\omega) + S_{NN}(\omega)}$

The reason for skipping the absolute value is that $E[(Z(t) - X(t))^2]$ is always real

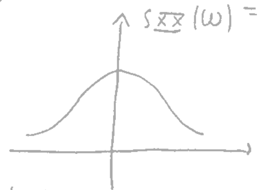
2/12-10 Queues

Computational problem

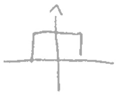
$\{X(t)\}_{t\in\mathbb{R}}$ OU-process Gaussian process with zero-mean, $R_{XX}(\tau)=e^{-|\tau|}$

$\{N(t)\}_{t\in\mathbb{R}}$ Gaussian noise zero-mean

$S_{XX}(\omega)=2/(1+\omega^2)$

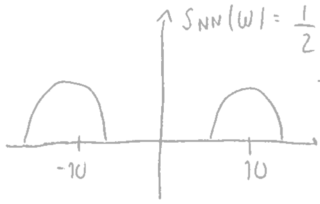


$Y(t)=X(t)+N(t)$

$H(\omega)$ filter, = 

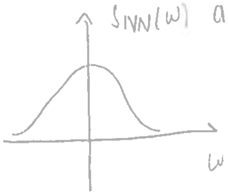
$S_{NN}(\omega)=\frac{1}{2}\left(S_{XX}(\omega-10)+S_{XX}(\omega+10)\right)$

$S_{NN}(\omega)=\frac{1}{2}\left(S_{XX}(\omega-10)+S_{XX}(\omega+10)\right)$



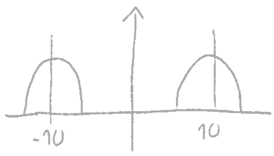
easier problem

$S_{NN}(\omega)$ also OU process



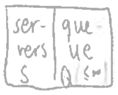
"fool" the noise move the frequency content

$\hat{X}(t)=\cos(\Theta+\omega(t))\,X(t)$



High pass filter.

$n=Q+S\ (\le\infty)$

customers arrive →  → exit.

$X(t)$ = number of customers in queue system (served customers + those queuing)

- interarrival times — FIFO (first in first out) fair queue

times between customers.

$\exp(\lambda)$-distributed (existing in nature - the most important example)

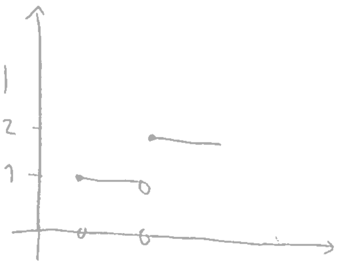
- service times

$\exp(\mu)$-distributed

- $X(0)$

- Required to specify the queuing system

ex)



$X(0)=0$ can either "jump" up one step or down.

queue process $\{X(t)\}_{t\ge0}$ (Markov process, continuous time, discrete valued)

Markov chain in continuous time.

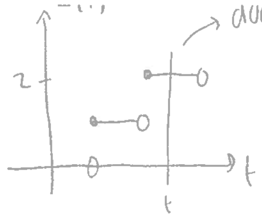
ex) MC in continuous time is the POISSON PROCESS $\{X(t)\}_{t\ge0}$ with intensity λ.

$P(X(t)\le x\mid X(t_1),\ldots,X(t_n)) = P(\underbrace{X(t)-X(t_n)}\le x-X(t_n)\mid \cancel{X(t_1)}\ X(t_n)) = P(X(t)\le x\mid X(t_n))$

lack of memory property. — independent increment doesn't depend on history and $X(t_n)$ depends only on itself.

for $0\le t_1<\ldots<t_n<t$

Theorem A continuous r.v. $T>0$ has the lack of memory property $P(T>t+s\mid T>s)=P(T>t)$ iff T is exponentially distributed.



doesn't matter how long time has elapsed of the interval.

Proof

$\boxed{\Leftarrow}$ The exponential distribution satisfies the property.

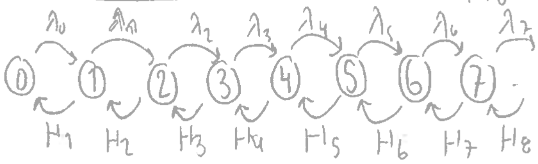
$$P(T>t+s \mid T>s) = \frac{P(T>t+s, T>s)}{P(T>s)} = \frac{\exp(-\lambda(t+s))}{\exp(-\lambda s)} = \exp(-\lambda t) = P(T>t)$$

$\boxed{\Rightarrow}$ $P(T>t+s) = \ldots = P(T>t)P(T>s)$ the only solution to this equation is the exponential distribution.

(whatever Markov chain one has, the time inbetween happenings has to be exponentially distributed)

$\{X(t)\}_{t\geq 0}$ with discrete values is MC iff it stays independent exp-distributed times at its different discrete values.

ex) Birth-death process $\{X(t)\}_{t\geq 0}$ (the by far most important MC)



very useful

Poisson process special case, $\lambda_i = \lambda \quad \mu_i = 0$

could be finite if one of the λ_is equals to 0.

If $X(t) = n$, then we wait $\min(\underbrace{\exp(\lambda_n)}_{B}, \underbrace{\exp(\mu_n)}_{D})$

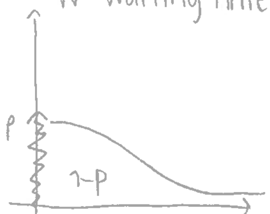
distribution of minimum of two independent exponential distributions.

$$P(\min(B,D) > t) = P(B>t)\cdot P(D>t) = e^{-\lambda_n t} e^{-\mu_n t} = P(\exp(\lambda_n + \mu_n) > t)$$

$n \to n-1$ wp $\frac{\mu_n}{\lambda_n+\mu_n}$

$n \to n+1$ wp $\frac{\lambda_n}{\lambda_n+\mu_n}$

W - waiting time for service

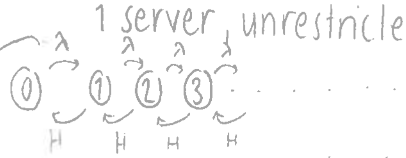


Markov arrivals and service time.

1 server

M/M/1/(∞)

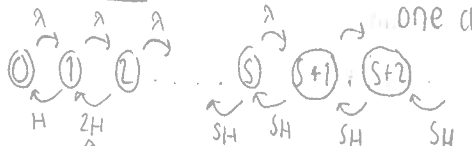
1 server, unrestricted queue.



Because the arrival process with intensity λ

s servers

M/M/S/(∞) ← unrestricted queue

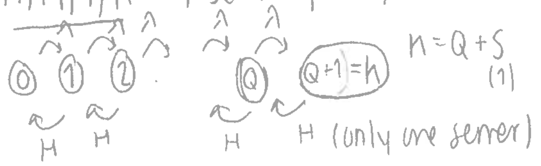


one and the same arrival process.

because more servers are active

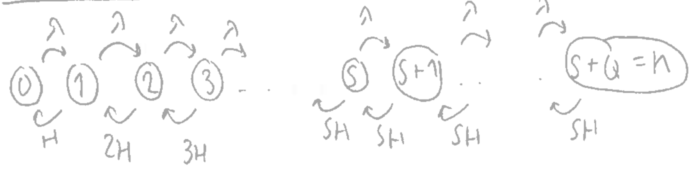
M/M/1/K 1 server, Q queue places

In the book K is total capacity of system, that is the total number of servers + queuing places



$K = Q + S$
(1)

M/M/S/K S servers, Q queue places $K = Q + S$



$p_n(t) = P(X(t) = n)$ → how the probability of being in n changes.

$$0 = \begin{cases} p_n'(t) = -(\lambda_n + \mu_n)p_n(t) + \mu_{n+1}p_{n+1}(t) + \lambda_{n-1}p_{n-1}(t) & n \geq 1 \\ p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t) \end{cases}$$

comes from above ↗ ↖ comes from above

Stationary distribution. (the probability of being in n shouldn't change in time)

$$p_n(t) = p_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} \cdot C$$

^ any constant times this works, the sum over p_n has to be 1

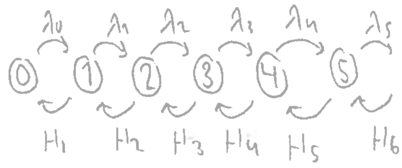
verification

$$-(\lambda_n + \mu_n)\frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} + \mu_{n+1}\frac{\lambda_0 \cdots \lambda_n}{\mu_1 \cdots \mu_{n+1}} + \lambda_{n-1}\frac{\lambda_0 \cdots \lambda_{n-2}}{\mu_1 \cdots \mu_{n-1}} = 0$$

$$-\frac{\lambda_0 \cdots \lambda_n}{\mu_1 \cdots \mu_n} - \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_{n-1}} + \frac{\lambda_0 \cdots \lambda_n}{\mu_1 \cdots \mu_n} + \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_{n-1}}$$

$\boxed{3/12-10}$ Chapter 9, Queues

Birth-death process $\{X(t)\}_{t \geq 0}$



spend time in state (n) according to $\min(\exp(\lambda_n), \exp(\mu_n)) = \exp(\lambda_n + \mu_n)$

go to $n-1$ wp $\frac{\mu_n}{\lambda_n + \mu_n}$

— " — $n+1$ wp $\frac{\lambda_n}{\lambda_n + \mu_n}$

Usually either $X(0) = 0$ or $X(0)$ has the stationary distribution (steady-state)

$$P(X(0) = n) = \frac{\prod_{i=1}^{n} \frac{\lambda_{i-1}}{\mu_i}}{\sum_{n=0}^{\infty} \prod_{i=1}^{n} \frac{\lambda_{i-1}}{\mu_i}} = \frac{\frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}}{\sum \cdots}$$

M/M/1 one server, infinitely many queue places

$\lambda_n = \lambda$ $\mu_n = \mu$

under steady-state.

$L = E\{$number of customers in $\underbrace{\text{queue}}_{\text{served + queuing}}$ – whole system$\}$

$L_q = E\{$ — " — waiting in queue$\}$

$L_s = E\{$ — " — being served$\}$

Balance equal eq (9.2)

" L/λ

$W = E\{$Amount of time a customer spends in whole system$\}$

$W_q = E\{\text{amount of time a customer spends queuing}\} = W - W_s$ (total time in system = queuing time + waiting time)

$W_s = E\{$ — " — being served$\} = E[\exp(\mu)] = \frac{1}{\mu}$

$L = E[X(t)] = \sum_{n=0}^{\infty} n p_n$ ← stat. distribution. $L_q = \lambda W_q$ $L_s = \lambda W_s$

M/M/1 1 server, infinitely many queue places
$\lambda_n = \lambda$ $\mu_n = \mu$

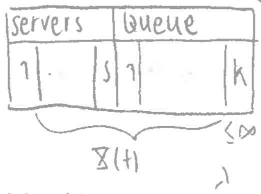
M/M/S s servers, infinitely many queue places
$\lambda_n = \lambda$, $\mu_n = \min(n,s)\mu = \begin{cases} n\mu, & n \le s \\ s\mu, & n > s \end{cases}$

M/M/S/K s servers, $K = s + Q$ total number of places in queue, Q = number of queue places $< \infty$.

$\lambda_n = \begin{cases} \lambda, & n \le K-1 \\ 0, & n \ge K \end{cases}$ $\mu_n = \min(n,s)\mu$

queue system

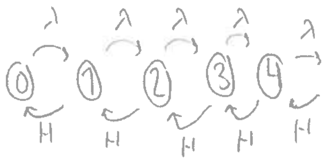
poisson process with intensity λ



$X(t)$ number of busy places. (total, servers + queue)

$L = L_q + L_s$

M/M/1



$p_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$ $\lambda < \mu$ $L = \frac{\mu}{\lambda}$

$L = \sum_{n=0}^{\infty} n \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right) = \frac{\lambda/\mu}{(1-\lambda/\mu)^2}\left(1 - \frac{\lambda}{\mu}\right) = \frac{\lambda}{\mu - \lambda}$

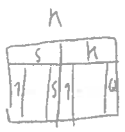
$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x} - 1\right) = \frac{d}{dx}\sum_{n=1}^{\infty} x^n = \sum_{n=1}^{\infty} n x^{n-1}$

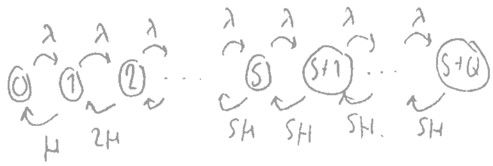
$W = L/\lambda = \frac{1}{\mu - \lambda}$ $W_q = \frac{L}{\lambda} - \frac{1}{\lambda} = \frac{1}{\mu-\lambda} - \frac{1}{\mu} = \frac{\lambda}{(\mu-\lambda)\mu}$

stochastic differential equations
$dX(t) = \mu(X(t))\,dt + \sigma(X(t))\,dW(t)$

M/M/S/K
exp(λ)-times between arrivals
exp(μ) - service times
s servers
$K = Q + S$ total places in queue system where Q = number of places in queue.

Poisson with intensity λ





$\sum_{n=0}^{N} x^n = \frac{1 - x^{N+1}}{1-x}$

$p_n = \begin{cases} \frac{\lambda^n}{\mu^n (n!)} \Big/ \Sigma & n \le s \\ \frac{\lambda^n}{\mu^n s!\, s^{n-s}} \Big/ \Sigma & n = s+1, \ldots, s+Q \end{cases}$

$\Sigma = \sum_{n=0}^{s-1} \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} + \underbrace{\sum_{n=s}^{s+Q} \left(\frac{\lambda}{\mu}\right)^n \frac{1}{s!\, s^{n-s}}}$ → $\left(\frac{\lambda}{\mu}\right)^s \frac{1}{s!} \underbrace{\sum_{n=s}^{s+Q} \left(\frac{\lambda}{\mu s}\right)^{n-s}}$ $\frac{1 - \left(\frac{\lambda}{\mu s}\right)^{Q+1}}{1 - \left(\frac{\lambda}{\mu s}\right)}$