NPTEL Stochastic Processes Lec 01 1.1.1

Probability Theory Refresher (Module 01)

E: The possible outcomes

Ω: The set of all possible outcomes

↳ countably infinite, countably finite, uncountably many

event - subset of Ω

ϕ, Ω to name a couple

σ-algebra over Ω: $\mathcal{F}$

$\mathcal{F}$: the collection of all possible subsets of Ω s.t.

1) $\phi \in \mathcal{F}$
2) If $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$
3) If $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

$\mathcal{F} = \{\phi, \Omega\}$ = the trivial σ-algebra

Probability Space: $(\Omega, \mathcal{F}, P)$

P is a set function s.t.

(1) $P(A) \geq 0$ for $A \in \mathcal{F}$

(2) $P(\Omega) = 1$

(3) If $A_1, A_2, \ldots$ are mutually exclusive events, then

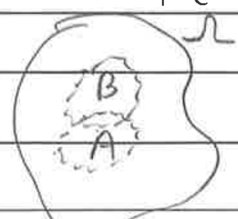
$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

$\boxed{P\text{: probability measure}}$

Conditional Probability and Independence of Events

Let $(\Omega, \mathcal{F}, P)$ be a probabilty space. If $P(A) > 0$, let B be the event

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



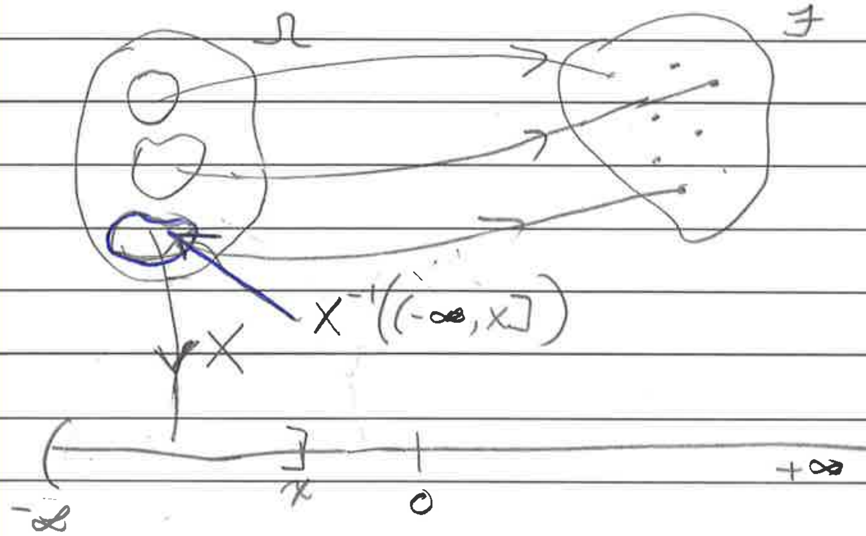
If A & B are independent, then $P(B|A) = P(B)$

Random Variables

$(\Omega, \mathcal{F}, P)$

$X: \Omega \to \mathbb{R}$ s.t. $X^{-1}((-\infty, x]) \in \mathcal{F}, \ \forall x \in \mathbb{R}$

X - random variable (dependent on choice of $\mathcal{F}$)



Cumulative Distribution Function

$F_X(x) = P(\{X \le x\}), \quad -\infty < x < \infty$

CDF of the r.v. X

$\{X \le x\} = \{\omega \mid X(\omega) \le x, \omega \in \Omega\} \in \mathcal{F}$ (event)

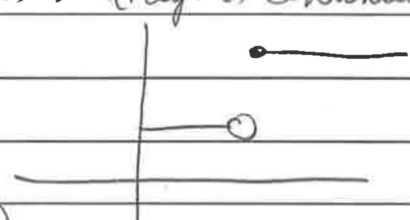
(1) $0 \le F_X(x) \le 1, \quad -\infty < x < \infty$

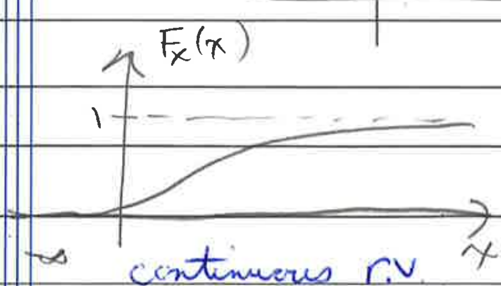
(2) $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$

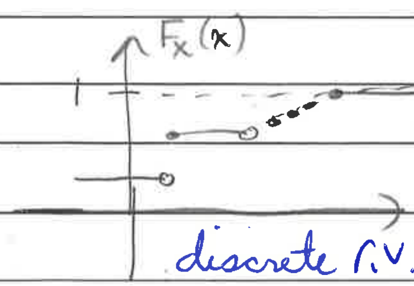
(3) $F_X(x)$ is monotonically increasing in x

for $x < y$, $F_X(x) \le F_X(y)$

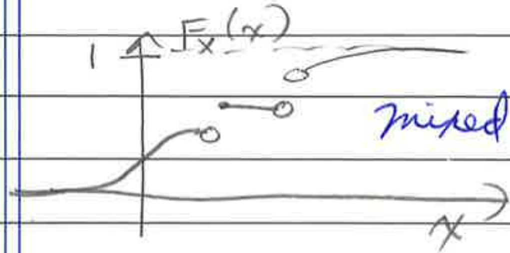
(4) $F_X(x)$ is (right) continuous in x







countably infinite jumps are possible



mixed r.v. can have countably infinite jumps

Standard Discrete Random Variable

1. Discrete uniform

$X \sim U(x_1, x_2, \dots, x_n)$

$P(X = x_i) = \frac{1}{n}$, $i = 1, \dots, n$ (pmf)

2. Binomial

$X \sim B(n, p)$

$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, 2, \dots, n$

3. Geometric

$X \sim Geo(p)$

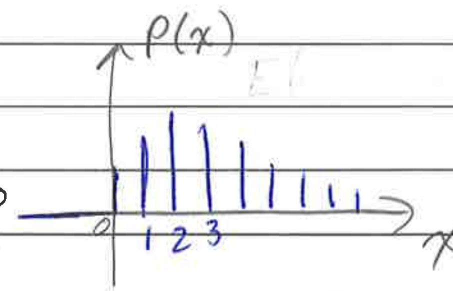
$P(X = r) = (1-p)^{r-1} p$, $x = 1, 2, \dots$

4. Poisson

$X \sim P(\lambda)$

$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \mathbb{1}(\{x \geq 0\})$, $\lambda > 0$

$E(X) = \lambda = Var(X)$

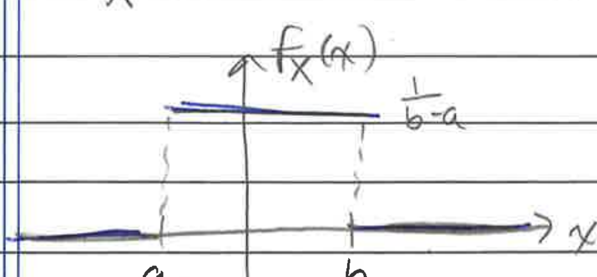


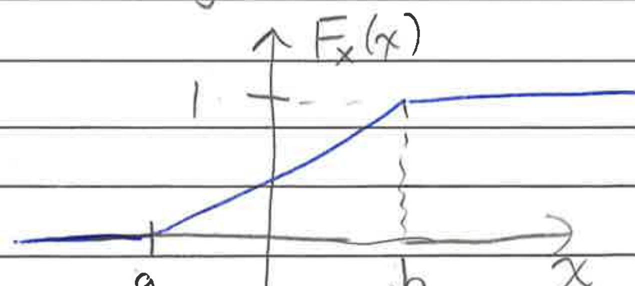
Standard Continuous Random Variables

1. Continuous uniform

$X \sim U(a, b)$

$f_X(x) = \frac{1}{b-a} \mathbb{1}(\{a \leq x \leq b\})$ (density)

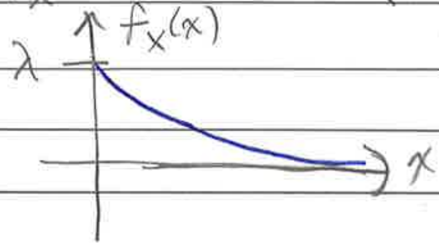


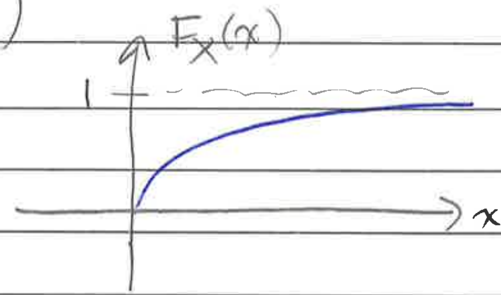


2. Exponential

$X \sim Exp(\lambda)$

$f_X(x) = \lambda e^{-\lambda x} \, \mathbb{1}(\{x > 0\})$

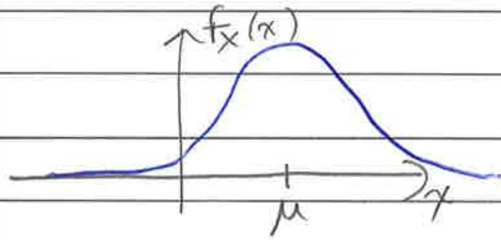




3. Normal

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

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < \mu, x < \infty, \quad \sigma > 0$$



$Z = \frac{X-\mu}{\sigma} \Rightarrow Z \sim N(0,1)$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

Lec 02

Joint Distribution of Random Variables

$(X_1, X_2, \ldots, X_n)$ is an n-dim random vector

Joint distribution

- joint probability mass function
- " " density function

(X, Y) both discrete

$P_{X,Y}(x,y) = Prob(\{X = x, Y = y\})$

$\{X = x, Y = y\} = \{\omega \mid X(\omega) = x, Y(\omega) = y, \omega \in \Omega\}$

(X, Y) - 2 dim. continuous r.v

$F_{X,Y}(x,y) = P(\{X \le x, Y \le y\})$

$$= \int_{-\infty}^{x} \int_{-\infty}^{y} \underbrace{f_{X,Y}(r,s)}_{\text{joint pdf}} \, dr \, ds$$

ex 1. (X,Y) discrete

$$P_{X,Y}(i,j) = \frac{1}{2^{i+j}}, \quad i,j = 1,2,\ldots$$

$$P_X(i) = \sum_j P_{X,Y}(i,j) = \sum_j \frac{1}{2^{i+j}} = \frac{1}{2^i}$$

$$P_Y(j) = \sum_i P_{X,Y}(i,j)$$

$(X_1, X_2, \ldots, X_5)$ all discrete

$$P_{X_1}(x) = \sum_{x_2,\ldots,x_5} P_{X_1,X_2,\ldots,X_5}(x_1,x_2,x_3,x_4,x_5)$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

Independent R.V. (works for n R.V.s)

(X,Y) are independent iff

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \quad \forall x,y$$

$$P_{X,Y}(x,y) = P_X(x)P_Y(y), \quad \forall x,y \qquad \text{(discrete)}$$

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \quad \forall x,y \qquad \text{(continuous)}$$

Covariance and Correlation Coeffecient

X

$$E(X) = \int_{-\infty}^{\infty} x\,dF_X(x) \quad \text{if integral exists}$$

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx \quad \text{if } X \text{ is continuous)}$$

provided $E(|X|) < \infty$

$$X \geq 0 \Rightarrow E(X) \geq 0 \qquad X \geq Y$$

$$E(aX+b) = aE(X)+b \qquad E(X) \geq E(Y)$$

X

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

$\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$

$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{X,Y}(x,y)\,dx\,dy$ if X,Y continuous

$X \perp Y \Rightarrow E(XY) = E(X)E(Y)$

$\Rightarrow \text{Cov}(X,Y) = 0$ (not iff) $\text{Cov}(X,Y)=0 \not\Rightarrow X \perp Y$

Correlation Coefficient

$$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \qquad X \perp Y \Rightarrow \rho_{X,Y} = 0$$

$|\rho| \leq 1$

Conditional Distribution

Given (X,Y)

$$P_{X|Y=y_j}(X=x_i \mid Y=y_j) = \frac{\text{Prob}(\{X=x_i \cap Y=y_j\})}{\text{Prob}(Y=y_j)}$$

provided $P(Y=y_j) > 0$

- on reduced sample space given $Y=y_j$, find $P(X=x_i)$

$(X_1, \dots, X_n)$ discrete

$P_{X_n \mid X_1, \dots, X_{n-1}}$ or $P_{X_1, X_2, X_7 \mid X_3, X_4, X_5, X_6}$

(X,Y) - continuous

$$f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{provided } f_Y(y) > 0$$

$X|Y=y$ is a r.v. called the conditional distribution

Conditional Expectation

$X|Y$ is a r.v.

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

$\quad$ L is a function of Y and a r.v.

$$E(E(X|Y)) = E(X)$$

If $X \perp Y \Rightarrow E(X|Y) = E(X)$

$$(X_1, X_2, \dots, X_n)$$

$$E(X_n \mid X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1})$$

$$(X, Y)$$

1) $p_{X,Y}(x,y) = \dfrac{1}{2^{x+y}}, \quad x, y = 1, 2, \dots$ discrete

2) continuous

$$f_{X,Y}(x,y) = \lambda \mu e^{-\lambda x - \mu y}\, \mathbb{1}(\{x>0, y>0\}), \quad \lambda, \mu > 0$$

$$f_X(x) = \lambda e^{-\lambda x}, \quad f_Y(y) = \mu e^{-\mu y}$$

Bivariate Normal

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}$$

$$X \sim N(\mu_1, \sigma_1^2); \quad Y \sim N(\mu_2, \sigma_2^2)$$

Covariance Matrix of $(X_1, X_2, \dots, X_n)$

$$\begin{array}{c|c} & \begin{matrix} X_1 & X_2 & \cdots & \end{matrix} \\ \begin{matrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{matrix} & \begin{pmatrix} Var(X_1) & & & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & & \\ & & \ddots & \\ Cov(X_n, X_1) & & & Var(X_n) \end{pmatrix}_{n \times n} \end{array}$$

$(i,j) \rightarrow Cov(X_i, X_j) = Cov(X_j, X_i)$

Symmetric matrix

Probability generating function

X-discrete, $x_i = 0, 1, 2, \ldots$

$$G_X(z) = \sum_i z^i P(X=i) = E(z^X)$$

$$G_X(1) = 1, \quad P(X=i) = \frac{G_X^{(i)}(0)}{i!}$$

$$X \sim B(n,p) \Rightarrow G_X(z) = (1-p+pz)^n$$

$$X \sim P(\lambda) \Rightarrow G_X(z) = e^{\lambda(z-1)}$$

Moment generating function

$M_X(t) = E(e^{Xt})$ provided expectation exists

$$e^{Xt} = 1 + \frac{Xt}{1!} + \frac{X^2t^2}{2!} + \cdots$$

$$M_X(0) = 1$$

$$X \sim B(n,p) \Rightarrow M_X(t) = (1-p+pe^t)^n$$

$$X \sim P(\lambda) \Rightarrow M_X(t) = e^{\lambda(e^t-1)}$$

$$X \sim N(\mu,\sigma^2) \Rightarrow M_X(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

$(X_1, \ldots, X_n)$, X_i iid $\forall i$

$$X \stackrel{d}{=} Y \Rightarrow F_X(x) = F_Y(y)$$

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

$$M_{S_n}(t) = \prod_{i=1}^n M_{X_i}(t) = (M_X(t))^n$$

Characteristic Function

$\phi_X(t) = E(e^{iXt})$, $i = \sqrt{-1}$ (always exists)

$$= \int_{-\infty}^{\infty} e^{itx}\, dF_X(x)$$

$$= \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx \quad \text{if } X \text{ is absolutely continuous}$$

This is the Fourier transform of f_X

$\phi_X(-it) = M_X(t)$

$S_n = \sum_{i=1}^{n} X_i$, X_i are iid $\Rightarrow \phi_{S_n}(t) = (\phi_X(t))^n$

Convergence of Sequence of Random Variables

$X_1, X_2, \ldots, X_n, \ldots$ in $(\Omega, \mathcal{F}, P)$

1) Convergence in Probability $X_n \xrightarrow{P} X$

for $\varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$

2) Convergence almost surely $X_n \xrightarrow{a.s.} X$

$P(\lim_{n\to\infty} X_n = X) = 1$

$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X$

$\not\Leftarrow$

3) Convergence in distribution $X_n \xrightarrow{d} X$

$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$

$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

$\not\Leftarrow \qquad \not\Leftarrow$

weak law of large numbers (WLLN)

$\overline{X}_n \xrightarrow{P} \mu$

Strong law of large numbers (SLLN)

$\overline{X}_n \xrightarrow{a.s.} \mu$

$X_1, X_2, \ldots$, $X_i \sim$ iid

$E(X_i) = \mu$, $Var(X_i) = \sigma^2$

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

<u>CLT</u> $\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \Phi(x)$

$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}z^2} dz$, $\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$

Lecture 3:

Problems in Random Variables and Distributions

1) Let X be a random variable having geometric distribution with parameter p.

Prove: $P(X=n+k \mid X>n) = P(X=k)$ (Memoryless property)

$$P(X=n+k \mid X>n)$$
$$= \frac{P(X=n+k \cap X>n)}{P(X>n)}$$
$$= \frac{P(X=n+k)}{P(X>n)} = \frac{(1-p)^{n+k-1}p}{\sum_{i=n+1}^{\infty}(1-p)^{i-1}p}$$
$$= \frac{p(1-p)^{n+k-1}}{p(1-p)^n\left(1+(1-p)+(1-p)^2+\dots\right)}$$
$$= \frac{(1-p)^{n+k-1}}{(1-p)^n\left(\frac{1}{1-(1-p)}\right)} = (1-p)^{k-1}p = P(X=k)$$

2) Let X be a r.v. having gamma distribution with parameter n (+ve integer) and λ. Then, the CDF of X is given by

$$F_X(x) = 1 - \sum_{i=0}^{n-1}\frac{(\lambda x)^i e^{-\lambda x}}{i!}$$

$$f_X(x) = \frac{\lambda^n x^{n-1}e^{-\lambda x}}{\Gamma(n)}$$

Now,

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x \frac{\lambda^n t^{n-1}e^{-\lambda t}}{\Gamma(n)}\,dt$$

Let $\lambda t = \mu$

$$F_X(x) = \int_0^{\lambda x}\frac{\mu^{n-1}e^{-\mu}}{\Gamma(n)}\,d\mu = 1 - \int_{\lambda x}^{\infty}\frac{\mu^{n-1}e^{-\mu}}{\Gamma(n)}\,d\mu$$
$$= 1 - \frac{1}{(n-1)!}\left[\frac{\mu^{n-1}e^{-\mu}}{-1}\Bigg|_{\lambda x}^{\infty} - \int_{\lambda x}^{\infty}\frac{(n-1)\mu^{n-2}e^{-\mu}}{-1}\,d\mu\right]$$

$$= 1 - \frac{1}{(n-1)!}(\lambda x)^{n-1} e^{-\lambda x} - \frac{(\lambda x)^{n-2} e^{-\lambda x}}{(n-2)!} - \cdots - \frac{(\lambda x)^0 e^{-\lambda x}}{0!}$$

by repeated integration by parts

$$= 1 - \sum_{i=0}^{n-1} \frac{(\lambda x)^i e^{-\lambda x}}{i!}$$

3) Let X, Y be independent exponential distributed r.v.s with parameter λ and μ, respectively.

Define

$$U = \min\{X, Y\}$$

$$V = \max\{X, Y\}$$

$$N = \begin{cases} 0, & X \le Y \\ 1, & X > Y \end{cases}$$

Find

(1) $P(\{N=0\})$ and $P(\{N=1\})$

(2) $P(\{N=0 \cap U>t\})$

(2) $\{N=0 \cap U>t\} = \{t < X \le Y\}$

$$\therefore P(\{N=0 \cap U>t\}) = P(\{t<X\le Y\})$$

$$= \iint_{t<X\le Y} \lambda e^{-\lambda x} \mu e^{-\mu y}\, dy\, dx$$

$$= \int_t^\infty \left(\int_x^\infty \mu e^{-\mu y} dy \right) \lambda e^{-\lambda x} dx$$

$$= \int_t^\infty e^{-\mu x} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda+\mu} \int_t^\infty (\lambda+\mu) e^{-(\lambda+\mu)x} dx$$

$$= \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} = P(\{N=0 \cap U>t\})$$

Similarly,

$$P(\{N=1 \text{ and } U>t\}) = \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

$$P(\{U>t\}) = P(\{N=0 \cap U>t\}) + P(\{N=1 \cap U>t\})$$
$$= e^{-(\lambda+\mu)t}$$

(1) $P(\{N=0\}) = P(\{N=0 \cap U>0\}) = \frac{\lambda}{\lambda+\mu}$

$$P(\{N=1\}) = 1 - P(\{N=0\}) = \frac{\mu}{\lambda+\mu}$$

$$P(N=0 \text{ and } U>t) = P(N=0)\, P(U>t)$$
$$P(N=1 \cap U>t) = P(N=1)\, P(U>t)$$

4) Let $X \sim B(N, p)$ where $N \sim P(\lambda)$. Find pmf of X.

Given $N \sim P(\lambda)$, $P(N=n) = \frac{e^{-\lambda}\lambda^n}{n!}$, $n=0,1,2,\ldots$

$$P(X=k) = \sum_{n=0}^{\infty} P(X=k \mid N=n)\, P(N=n)$$

$$= \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \frac{\lambda^n e^{-\lambda}}{n!}$$

$$= \frac{\lambda^k e^{-\lambda} p^k}{k!} \sum_{n=k}^{\infty} \frac{[\lambda(1-p)]^{n-k}}{(n-k)!}$$

$$= \frac{(\lambda p)^k e^{-\lambda}}{k!} e^{-\lambda(1-p)}$$

$$P(X=k) = \frac{(\lambda p)^k e^{-\lambda p}}{k!}, \quad k=0,1,2,\ldots$$

$$X \sim P(\lambda p)$$

5) Let the number of offspring of a given species be a r.v. having pmf $p(x)$ for $x=0,1,2,\ldots$ with mean μ and variance σ^2. A population begins with a single parent who produces a random number N of progeny, each of which independently produces offspring according to $p(x)$ to form

a second generation. Find the mean and variance of the number of descendants in the second generation.

Let $X = \xi_1 + \xi_2 + \cdots + \xi_N$

where ξ_i is the number of progeny of the i^{th} offspring of the original parent.

Given $p(x) = p(\xi_i = x)$, $E(\xi_i) = \mu$, $Var(\xi_i) = \sigma^2$

Find $E(X)$

$$E(X) = \sum_{n=0}^{\infty} E(X \mid N=n)\, P(N=n)$$

$$= \sum_{n=0}^{\infty} E(\xi_1 + \xi_2 + \cdots + \xi_n \mid N=n)\, P(N=n)$$

$$= \mu \sum_{n=0}^{\infty} n P(N=n)$$

$$= \mu \cdot \mu = \mu^2$$

$$Var(X) = E[(X-\mu^2)^2] = E[(X - N\mu + N\mu - \mu^2)^2]$$

$$= E[(X-N\mu)^2] + E[\mu^2(N-\mu)^2] + 2E[\mu(X-N\mu)(N-\mu)]$$

$$E[(X-N\mu)^2] = \sum_{n=0}^{\infty} E[(X-N\mu)^2 \mid N=n]\, P(N=n)$$

$$= \sigma^2 \sum_{n=1}^{\infty} n P(N=n) = \mu\sigma^2$$

$$E[\mu^2(N-\mu)^2] = \mu^2 Var(N) = \mu^2\sigma^2$$

$$E[\mu(X-N\mu)(N-\mu)] = \mu \sum_{n=0}^{\infty} E[(X-N\mu)(N-\mu) \mid N=n]\, P(N=n)$$

$$= \mu \sum_{n=0}^{\infty} (n-\mu)\, E(X - N\mu \mid N=n)\, P(N=n)$$

$= 0$ because $E(X - N\mu \mid N=n) = E(\xi_1 + \xi_2 + \cdots + \xi_n - n\mu)$

$$Var(X) = \mu\sigma^2(\mu+1)$$

Lecture 4

Problems in Sequences of Random Variables

1) Let $Z_1, Z_2, \ldots$ be a sequence of r.v.s each having Poisson distribution with parameter n.

$$Z_n \sim P(n), \quad n = 1, 2, 3, \ldots$$

Find limiting distribution of $Y_n = \frac{Z_n - n}{\sqrt{n}}$

$$M_{Z_n}(t) = E(e^{Z_n t}) = \sum_{k=0}^{\infty} e^{kt} \frac{e^{-n} n^k}{k!}$$

$$= e^{-n} \sum_{k=0}^{\infty} \frac{(e^t n)^k}{k!}$$

$$= e^{-n} e^{ne^t} = e^{n(e^t - 1)}$$

$$M_{Y_n}(t) = M_{\frac{Z_n - n}{\sqrt{n}}}(t) = E\left(e^{\frac{Z_n - n}{\sqrt{n}} t}\right)$$

$$= e^{-t\sqrt{n}} M_{Z_n}\left(\frac{t}{\sqrt{n}}\right)$$

$$= e^{-t\sqrt{n}} e^{n\left(e^{\frac{t}{\sqrt{n}}} - 1\right)}$$

$$= e^{-t\sqrt{n}} e^{n\left(\cancel{1} + \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \frac{t^3}{3!\, n^{3/2}} + \cdots - \cancel{1}\right)}$$

$$= e^{-t\sqrt{n}} e^{t\sqrt{n}} e^{t^2/2} e^{\left[\frac{t^3}{3!\, n^{1/2}} + \cdots\right]}$$

$$= e^{\left(t^2/2 + \frac{t^3}{3!\, n^{1/2}} + \cdots\right)}$$

As $n \to \infty$ $\quad M_{Y_n}(t) \to e^{t^2/2} \Longleftrightarrow$ MGF of standard normal

limiting distribution of $\frac{Z_n - n}{\sqrt{n}}$ is $N(0,1)$

2) Let $X_1, X_2, \ldots$ be a sequence of r.v.'s each having CDF

$$F_{X_n}(x) = \begin{cases} 0, & -\infty < x < 0 \\ 1-\left(1-\frac{x}{n}\right)^n, & 0 \le x < n \\ 1, & n \le x < \infty \end{cases}$$

As $n \to \infty$

$$F_{X_n}(x) = \begin{cases} 0, & -\infty < x < 0 \\ 1-e^{-x}, & 0 \le x < \infty \end{cases}$$

Suppose X is r.v. with CDF

$$F_X(x) = \begin{cases} 0, & -\infty < x < 0 \\ 1-e^{-x}, & 0 \le x < \infty \end{cases}$$

Then $X_n \xrightarrow{d} X$; $X \sim Exp(1)$.

3) Suppose we choose at random n numbers from $[0,1]$ with the uniform distribution. Let X_i be a r.v. describing the i^{th} choice. Then, for $i = 1, 2, \ldots$

$$E(X_i) = \int_0^1 x\,dx = \frac{1}{2}$$

$$Var(X_i) = \int_0^1 x^2\,dx - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Let $S_n = \sum_{i=1}^{n} X_i$. $E\left(\frac{S_n}{n}\right) = \frac{1}{2}$

$$Var\left(\frac{S_n}{n}\right) = \frac{1}{12n}$$

For any $\varepsilon > 0$, using Chebyshev's Inequality

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \ge \varepsilon\right) \le \frac{1}{12n\varepsilon^2}$$

As $n \to \infty$

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \ge \varepsilon\right) \to 0 \iff \frac{S_n}{n} \xrightarrow{P} \frac{1}{2}$$

The X_n obey WLLN.

4) Consider a repairman who replaces a light bulb the instant it burns out. Suppose the first light bulb is put in a $t=0$, and let X_i be the lifetime of the i^{th} lightbulb. Let $T_n = X_1 + X_2 + \dots + X_n$ where X_i are iid be the time the n^{th} light bulb burns out.

Assume that $X_i \sim Exp(\lambda) \Rightarrow E(X_i) = \frac{1}{\lambda}$

$$\frac{T_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{n\to\infty} \frac{1}{\lambda} \text{ a.s.}$$

$\{X_1, X_2, \dots\}$ obeys SLLN

5) Let $X_1, X_2, \dots$ be a sequence of r.v. each having student t distribution with n degrees of freedom. The pdf of X_n is

$$f_{X_n}(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1+\frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$$

For large n,

$$\lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} = \frac{1}{\sqrt{2\pi}} \quad \text{(Sterling's approximation)}$$

Also $$\lim_{n\to\infty}\left(1+\frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} = e^{-x^2/2}$$

$$\Rightarrow \lim_{n\to\infty} f_{X_n}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$\Rightarrow \quad X_n \xrightarrow{d} Z \sim N(0,1)$$

6) Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.s each having pmf

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}, \quad i = 1, 2, 3, \ldots$$

Define $M_k = \sum_{j=1}^{k} X_j$, $k = 1, 2, \ldots$

For a fixed integer n, define

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt} \text{ for all } t \geq 0 \text{ s.t. } nt \text{ is an integer}$$

Now, for $0 \leq s \leq t$,

$$E\left(W^{(n)}(t) - W^{(n)}(s)\right) = 0 \text{ since } E(M_k) = 0$$

$$Var\left(W^{(n)}(t) - W^{(n)}(s)\right) = t - s$$

Fix $t \geq 0$, as $n \to \infty$

$$W^{(n)}(t) \xrightarrow{d} X \sim N(0, t) \quad \text{(using CLT)}$$

Used for model of Brownian Motion

Module 02 Definition and Simple Stochastic Processes

Lecture 01: Definition, Classification and Examples

A. What is a Stochastic Process?

Def: Let $(\Omega, \mathcal{F}, P)$ be a given probability space. A collection of random variables $\{X(t), t \geq 0\}$ defined on the probability space $(\Omega, \mathcal{F}, P)$ is called a stochastic process.

Definition:

A stochastic process is also defined as a function of two arguments $X(\omega, t)$, $\omega \in \Omega$, $t \in T$. A stochastic process is also called a chance process or random process.

B. Parameter and State Spaces

The set T is called the parameter space where $t \in T$ may denote time, length, distance or any other quantity.
The set S is the set of all possible values of $X(t)$, $\forall t$ and is called the state space, where $X(t): \Omega \to A_t$, $A_t \subseteq \mathbb{R}$ and $S = \bigcup_{t \in T} A_t$

T - finite, countably infinite, uncountably infinite
S - " " " " "
set of intervals of $\mathbb{R}$ or $\mathbb{R}$ itself

C. Two approches to creating stochastic processes $\{X(\omega, t), \omega \in \Omega, t \in T\}$

Case (1): a family of r.v.s $\Leftrightarrow \{X(\cdot, t), t \in T\}$

Case (2): a set of functions on $T \Leftrightarrow \{X(\omega, \cdot), \omega \in \Omega\}$
- realization of the process, or trajectory, or sample path or sample function.

$\{X(t), t \in T\}$ - 1D, 2D or n-dim

e.g. $X(t) = (X_1(t), X_2(t))$ (2D random vector)

$X_1(t)$ – maximum temperature; $X_2(t)$ – minimum temperature

$\{X(t), t \in T\}$ – 2D stochastic process

$X(t) = (X_1(t), X_2(t), \dots, X_n(t))$ n-dim r.v. for each t

$\{X(t), t \in T\}$ – n-dim stochastic process

$X(t) = X_1(t) + i X_2(t)$ complex valued r.v. for fixed t

$\{X(t), t \in T\}$ – complex valued stochastic process

D. Classification of stochastic processes

T – parameter space

S – state space

S	$\{X(t), t \in T\}$
$\{0, 1, \dots\}$	integer valued or discrete state
$\mathbb{R}$	real-valued
Euclidean k-dim	k-vector

T	$\{X(t), t \in T\}$
$\{0, 1, \dots\}$	discrete parameter, stochastic sequence
$\mathbb{R}$	continuous, or simply a stochastic process

T \ S	discrete	continuous
discrete	discrete time discrete state	discrete time continuous state
continuous	continuous time discrete state	continuous time continuous state

(1) discrete time discrete state stochastic process

ex 1 X_n = # of customers in the barber shop after n^{th} customer departure from the shop

$T = \{1, 2, \dots\}$ $\{X_n, n = 1, 2, \dots\}$
$S = \{0, 1, 2, \dots\}$

ex 2 X_n = # of packets waiting in the buffer at the n^{th} time unit in the communication router.

$S = \{0, 1, \dots\}$
$T = \{1, 2, \dots\}$
$\{X_n, n = 1, 2, 3, \dots\}$

(2) continuous time discrete state stochastic process

ex 1: $X(t)$ = # of customers in the barber shop at any time t.

$T = \{t \mid t \geq 0\}$
$S = \{0, 1, 2, \dots\}$
$\{X(t), t \in T\}$

ex 2: $X(t)$ = # of customers eating food in a restaurant at any time t

$T = \{t \mid t \geq 0\}$
$S = \{0, 1, 2, \dots\}$
$\{X(t), t \in T\}$

(3) discrete time continuous state stochastic process

ex 1: X_n = the content of a dam observed at the n^{th} time unit

$T = \{1, 2, \dots\}$, $S = \{x \mid x \geq 0\}$, $\{X_n, n = 1, 2, \dots\}$

ex 2: X_n = the amount of 1 U.S. dollar in rupees at the n^{th} time unit in a day

$S = \{x \mid x \geq 0\}$, $T = \{1, 2, \dots\}$, $\{X_n, n = 1, 2, \dots\}$

(4) continuous time continuous state stochastic process

ex 1: $X(t)$ = temperature of a particular city at any time t

$S = \{x \mid -50 \le x \le 60\}$ degrees C

$T = \{t \mid t \ge 0\}$

$\{X(t) \mid t \in T\}$

ex 2: $X(t)$ = the content of a dam observed at any time t

$S = \{x \mid x \ge 0\}$, $T = \{t \mid t \ge 0\}$, $\{X(t), t \in T\}$

E. Summary

- Stochastic process is a collection of random variables
- Simple stochastic processes can be observed from a current real world problem
- We will describe the probability distribution of a stochastic process in future lectures

Lecture 02: Simple Stochastic Processes 2.2.1

A. Arrival Process

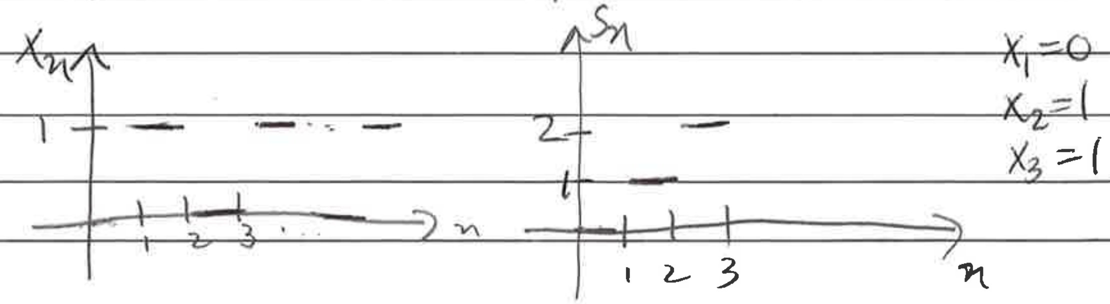
Bernoulli Process (discrete time)

$\{X_i, i=1,2,\dots\}$, $X_i \overset{iid}{\sim} \text{Bernoulli}(p) = B(1,p)$

Define $S_n = \sum_{i=1}^{n} X_i$, $P(X_i = k) = \begin{cases} 1-p, & k=0 \\ p, & k=1 \end{cases}$

$S_n \sim B(n,p)$ = # of arrivals in n trials

$\{S_n, n=1,2,\dots\}$ = Binomial process



$E(S_n) = np$, $Var(S_n) = np(1-p)$

$T =$ # of trials up to and including to 1st success or 1st arrival

$$P(T=n) = (1-p)^{n-1}p$$

$T \sim \text{Geometric}(p)$, $E(T) = \frac{1}{p}$, $Var(T) = \frac{1-p}{p^2}$

- $P(T-n=m \mid T>n) = P(T=m)$

This property is called the "memoryless property"

Poisson Process (aka counting process)

Process of arrival of customers at a barbershop.

$N(t) =: N_t$: # of arrivals occur during the interval $[0,t]$

$\{N(t), t \geq 0\}$ - Continuous time discrete state stochastic process

Assume that

(1) In $(t, t+\Delta t)$, Prob of one arrival $= \lambda\Delta t + O(\Delta t)$, $\lambda > 0$

(2) Prob of more than arrival in $(t, t+\Delta t) = O(\Delta t)$

(3) events occuring in non-overlapping intervals are mutually independent

partition $[0,t)$ into n equal parts of length t/n

Apply binomial distribution

$$P(N(t)=k) = \binom{n}{k}\left(\lambda\frac{t}{n}\right)^k\left(1-\lambda\frac{t}{n}\right)^{n-k}, \quad k=0,1,2,\dots,n$$

As $n \to \infty$

$$P(N(t)=k) = \lim_{n\to\infty}\binom{n}{k}p^k(1-p)^{n-k}, \quad p=\lambda\frac{t}{n}, \quad k=0,1,2,\dots$$

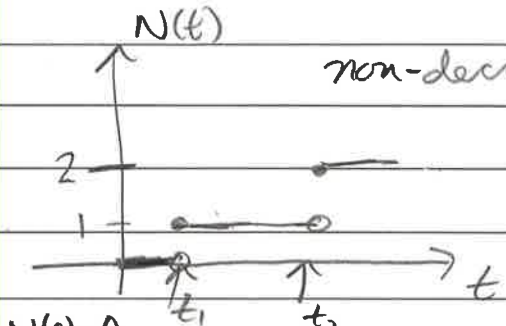
$$= \lim_{n\to\infty}\frac{n!}{(n-k)!\,k!}\left(\lambda\frac{t}{n}\right)^k\left(1-\lambda\frac{t}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty}\underbrace{\frac{n!}{n^k(n-k)!}}_{1}\frac{(\lambda t)^k}{k!}\underbrace{\left(1-\lambda\frac{t}{n}\right)^n}_{e^{-\lambda t}}\underbrace{\left(1-\lambda\frac{t}{n}\right)^{-k}}_{1}$$

$$= \frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

$$P(N(t)=k) = e^{-\lambda t}\frac{(\lambda k)^k}{k!}, \quad k=0,1,2,\dots$$

For fixed t, $N(t) \sim \text{Poisson}(\lambda t)$ r.v., $\{N(t), t\geq 0\}$ is a homogeneous poisson Stochastic Process



non-decreasing, right continuous, jumps of size 1

$N(0)=0$

$N(t_1) = N(t_1^+) = 1$

$N(t_1^-) = 0$

$N(t_2^{-1}) = 1$

$N(t_2) = N(t_2^+) = 2$

Let T_k be the time s.t. k^{th} arrival, $k = 1, 2, \ldots$

Let X_k be the successive inter arrival times of the k^{th} customer for $k = 1, 2, \ldots$: $X_1 = T_1$

$X_2 = T_2 - T_1$

$X_k = T_k - T_{k-1}, \; k = 1, 2, \ldots, \; T_0 = 0$

What is the distribution of $X_k, k = 1, 2, \ldots$?

$$P(X_1 > t) = P(N(t) = 0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

$P(X_1 > t) = e^{-\lambda t} \iff P(X_1 \le t) = 1 - e^{-\lambda t}$ (CDF of X_1)

$X_1 \sim Exp(\lambda)$

Similarly, $X_2 \sim Exp(\lambda), \ldots, X_i \sim Exp(\lambda), \; i = 1, 2, 3, \ldots$

$\{X_i, i = 1, 2, 3, \ldots\}$ discrete time continuous state stochastic process

$X_i \overset{iid}{\sim} Exp(\lambda)$ r.v.s for each $i = 1, 2, 3, \ldots$

Renewal process

Matlab Code

```
1) lambda = input('Enter the arrival rate :');
2) Tmax = input('Enter maximum time :');
3) T(1) = 0;
4) i = 1;
```

5) while T(i) < Tmax
   U(i) = rand(1,1);
   T(i+1) = T(i) - (1/lambda)*(log(U(i)));
   i = i+1;
end

Simple Random Walk (SRW)

For $(\Omega, \mathcal{F}, P)$

X_i, $i = 1,2,3,\dots$ are iid, integer-valued r.v.s

As a special case

$$P(X_i = k) = \begin{cases} p, & k=1 \\ 1-p, & k=-1 \end{cases}, \quad 0<p<1$$

Define $S_n := \sum_{i=1}^{n} X_i$, $\{S_n, n=1,2,\dots\}$ is a SRW

In general k can take any value, called a general random walk

If $p = 1/2$ for $k \in \{1,-1\}$, this is a symmetric random walk

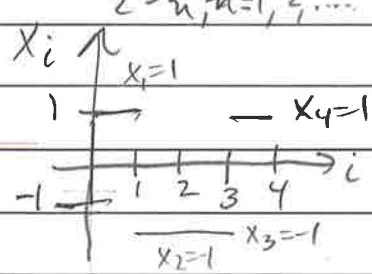
Two person coin tossing game: If A wins, B give 1 to A.

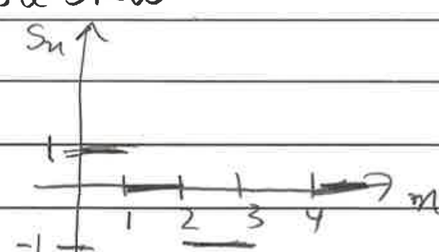
If A loses, A gives 1 to B.

X_i, $i=1,2,\dots$ amount of person A earnings at the i^{th} game

$S_n = \sum_{i=1}^{n} X_i$, ~~total amount earned at the end~~

$\{S_n, n=1,2,\dots\}$ is a SRW





$$E(X_i) = p + (-1)(1-p) = 2p-1$$

For $p = 1/2$, $E(X_i) = 0$, $E(X_i^2) = 1$, $E(S_n) = 0$, $Var(S_n) = n$

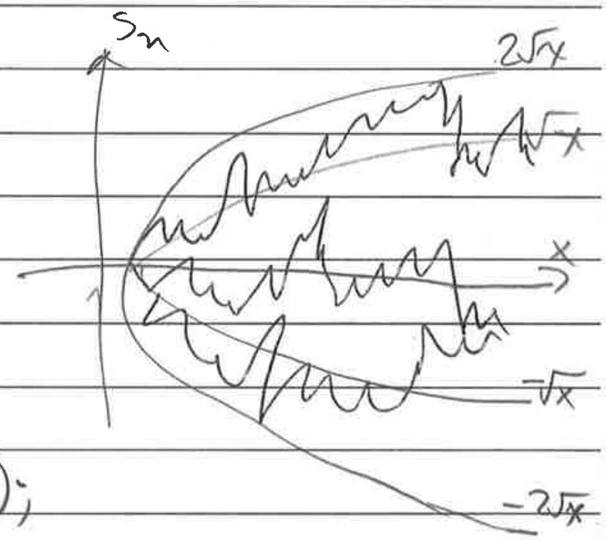
$$E\left(\left(\frac{S_n}{\sqrt{n}}\right)^2\right) = 1$$

Using CLT, $\dfrac{\frac{S_n}{\sqrt{n}} - 0}{1} \xrightarrow{d} N(0,1)$

i.e., $\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$

Matlab code

```
1) x0 = input('Enter the initial position :');
2) nsteps = input ('Enter the number of steps :');
3) p = input ('Probability of FORWARD move success in any step:');
4) S(1: nsteps ) =0;
5) S(1) = x0;
6) for istep = 2: nsteps
       if (rand() < 1-p)
          x = -1;
       else
          x = 1;
       end
       S(istep) = S(istep-1)+x;
   end
7) stairs (0:(istep-1), S(1:(istep)));
```

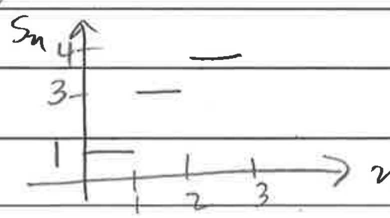


Population Processes

Consider the population of tigers in India. At the end of its lifetime produces a random number X of offspring with pmf

$$P(X=k) = a_k,\ k=0,1,2,\dots,\ a_k \ge 0,\ \sum_{k=0}^{\infty} a_k = 1$$

$\{S_n, n=0,1,2,\dots\}$ population size for tigers at the end of n^{th} generation – discrete time discrete state stochastic process



$S_0 = 0$

$S_1 = X_1$

$X_1 = 3$

$S_2 = X_1 + X_2 + X_3$

$X_1 = 3, X_2 = 0, X_3 = 1 \Rightarrow S_2 = 4$

Module 3: Stationary and Autoregressive Processes

Lecture 1: Stationary Processes

A. Introduction

Time series $\{X(t), t \geq 0\}$ (special case of stochastic processes)

B. Definitions

1. Mean function : $m(t) = E[X(t)]$
2. Second-order stochastic process : $E[X^2(t)] < \infty$, $\forall t$
3. Covariance function : $c(s,t) = \text{cov}(X(s), X(t))$
$$= E[X(s)X(t)] - E[X(s)]E[X(t)]$$

It satisfies: (1) $c(s,t) = c(t,s)$, $\forall s,t \in T$

(2) Using Schwarz inequality:
$$c(s,t) \leq \sqrt{c(s,s)c(t,t)} = \sqrt{Var(X(s))\,Var(X(t))}$$

(3) It is non-negative definite: $\forall a_1, a_2, \dots, a_n \in R$ and $t_i \in T$
$$\sum_{j=1}^{n}\sum_{k=1}^{n} a_j a_k c(t_j, t_k) = E\left[\left(\sum_j a_j X(t_j)\right)^2\right] \geq 0$$

(4) Sum & Product of covariance functions are also covariance functions, not always, but assumed for this course.

4. Autocorrelation function: $R(s,t) = \dfrac{E[X(t)X(s)] - E[X(t)]E[X(s)]}{\sqrt{Var(X(t))}\sqrt{Var(X(s))}}$
$$= \frac{Cov(X(s), X(t))}{\sqrt{Var(X(t))\,Var(X(s))}}$$

Assume $R(s,t)$ depends only on $|t-s|$, then
$$R(\tau) = \frac{E[(X(t)-\mu)(X(t+\tau)-\mu)]}{\sigma^2}$$

where $m(t) = E[X(t)] = \mu$ and $\sigma^2 = Var[X(t)]$

$R(\tau) = R(-\tau)$

5. Independent increments: If $\forall\, t_1 < t_2 < \dots < t_n$, $X(t_2)-X(t_1)$, $X(t_3)-X(t_2)$, …, $X(t_n)-X(t_{n-1})$ are mutually independent r.v.s $\forall n$, then $\{X(t), t \in T\}$ has the independent increments property.

6. Ergodic property: to be defined later.

C. Strict-sense stationary processes

If for arbitrary $t_1, t_2, \dots, t_n$, the joint distribution of the random vector $(X(t_1), X(t_2), \dots, X(t_n))$ and $(X(t_1+h), X(t_2+h), \dots, X(t_n+h))$ are the same for all $h>0$, then $\{X(t), t\in T\}$ is strict-sense stationary of order n.
If this is satisfied $\forall n$, then this is a strict-sense stationary process.

D. Wide-sense stationary process or weakly stationary or covariance stationary process

(1) $E(X(t)) = m(t)$ is independent of t

(2) $E[X^2(t)] < \infty$

(3) $c(s,t)$ is a function only on the time difference $|t-s|$, $\forall s,t$
i.e. $c(s,t) = f(|t-s|)$

Ex 1: let $X_i \overset{iid}{\sim} B(1,p)$, $\{X_i, i=1,2,\dots\}$ - stochastic process

(1) $m(i) = E(X_i) = p$

(2) $E[X_i^2] = p$

(3) $c(i,j) = E[X_iX_j] - E[X_i]E[X_j] = \begin{cases} 0, & i \neq j \\ p(1-p), & i=j \end{cases}$ from iid

$\{X_i, i=1,2,\dots\}$ - W.S.S.P

$(X_{i_1}, X_{i_2}, X_{i_3}, \dots, X_{i_n}) \sim \prod$ of n $B(1,p)$
$(X_{i_1+h}, X_{i_2+h}, \dots, X_{i_n+h})$ are distributed identically as above $\Rightarrow$ S.S.S.P. also

Ex 2: Let $\{X(t), t\in T\}$ be a S.S.S.P. with finite second-order moments, and let $Y(t) := a+bt+X(t)$.

$\{Y(t), t\in T\}$ is a S.P. Do it S.S.S.P.?

$m(t) = E(Y(t))$ is a function of t, thus $\{Y(t), t\in T\}$ is not WSSP and not SSSP. since $E(Y(t+h)) \neq E(Y(t))$

Ex 3: Let $\{X_n, n=1,2,\dots\}$ be uncorrelated r.v.s with $E[X_n]=k, \forall n$ and $E[X_m X_n] = \begin{cases} \sigma^2, & m=n \\ 0, & m\neq n \end{cases}$

$\{X_n, n=1,2,\dots\}$ is W.S.S.P. called a white noise process.

Additionally, one can assume that the ergodic property is satisfied along with WSSP.

Ergodicity

In a Markov chain, a state is said to be ergodic if it is aperiodic and positive recurrent. If all states in a Markov chain are ergodic, then the chain is said to be ergodic.

$\Rightarrow \mu = E[X(t)]$ and $R(t,s) = R(\tau)$.

$\Rightarrow$ The mean can be estimated from the time average

$$\hat{\mu}_T = \frac{1}{2T}\int_{-T}^{T} x(t)\,dt \xrightarrow{\text{mean square}} \mu$$

Lecture 2: Autoregressive Processes 3.2.1

Ex 4: $\{X(t), t\geq 0\}$

Define $X(t) = A\cos\theta t + B\sin\theta t$, where A & B are uncorrelated r.v.s with $E(A)=E(B)=0$ and $Var(A)=Var(B)=\sigma^2$

(1) $E(X(t)) = 0 \quad \forall t$

(2) $C(X(s),X(t)) = E[X(t)X(s)] - E[X(t)]E[X(s)]$

$= \sigma^2 \cos\theta(t-s)$

(3) $E[X^2(t)] = \sigma^2 \quad \Rightarrow \{X(t), t\geq 0\}$ is WSSP

Ex 5: Time Series Forecasting

$X_0, X_1, \ldots, X_n$ - time series data

X_{n+1} to be forecasted

$E[X_{n+1} \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n]$

If $X_i \sim N(\mu, \sigma^2) \Rightarrow E[X_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n]$ is a linear combo of $X_1, X_2, \ldots, X_n$

$\hat{X}_{n+1} = a_0 x_0 + a_1 x_1 + \ldots + a_n x_n$, where $a_0, a_1, \ldots$ are parameters to be estimated. If we assume WSSP, $a_0, a_1, \ldots$ are time independent

Ex 6: $\{W(t), t \geq 0\}$ is a Gaussian process

$W(t) \sim N(\mu, \sigma^2)$, $\forall t$

$(W(t_1), W(t_2), \ldots, W(t_n)) \sim$ multivariate normal distribution

$[E(W(t_1)), E[W(t_2)], \ldots, E[W(t_n)]]$ - mean vector

+ covariance matrix, $[cov(W(t_i), W(t_j))]_{i,j}$, $i,j = 1, \ldots, n$

Suppose $\{W(t), t \geq 0\}$:

(1) $E[W(t)]$ is independent of t

(2) $E[W^2(t)] < \infty$

(3) $Cov(W(t_i), W(t_j))$ is a function of $|t_i - t_j|$

Then, $\{W(t), t \geq 0\}$ - WSSP

$\{W(t), t \geq 0\}$ is also SSSP

Definition: Pure Random Process or White Noise Process

$\{X_t, t \geq T\}$ has $E[X_t] = m$

$$E[X_t X_{t+k}] = \begin{cases} \sigma^2, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

It is covariance stationary process

Definition : Moving Average (MA) Process

It is represented as $\{X_t, t \in T\}$, $X_t = a_0 e_t + a_1 e_{t-1} + \dots + a_h e_{t-h}$

where a_i - real constants

$\{e_t\}$ - pure random process (mean 0 and variance σ^2)

When $a_h \neq 0$, $\{X_t, t \in T\}$ is called MA of order h.

$$C_k := E(X_t X_{t+k}) \Rightarrow C_k = \begin{cases} (a_0 a_k + a_1 a_{k+1} + \dots + a_{h-k} a_h)\sigma^2, & k \le h \\ 0, & k > h \end{cases}$$

Define
$$\rho_k := \frac{C_k}{C_0} = \begin{cases} \dfrac{a_0 a_k + a_1 a_{k+1} + \dots + a_{h-k} a_h}{a_0^2 + a_1^2 + \dots + a_h^2}, & k \le h \\ 0, & k > h \end{cases}$$

The first order Markov process is defined as:

$X_t + a_t X_{t-1} = e_t$, where $|a_t| < 1$ and $\{e_t, t \in T\}$ is a pure random process with mean 0 and variance 1.

$$X_t = \sum_{k=0}^{\infty} g_k e_{t-k} \text{ where } g_k = a^{|k|}, \forall k \in \mathbb{Z}$$

Autoregressive Process

Definition : AR Process $\{X_t, t \in T\}$

$$X_t + b_1 X_{t-1} + b_2 X_{t-2} + \dots + b_h X_{t-h} = e_t, \quad b_h \neq 0$$

$\Rightarrow \{X_t, t \in T\}$ - AR Process of order h

When $X_t = \sum_{\gamma=0}^{\infty} b_\gamma e_{t-\gamma} \Rightarrow \{X_t, t \in T\}$ - AR process of infinite order

Special Case: AR process of order 2

$$X_t + b_1 X_{t-1} + b_2 X_{t-2} = e_t$$

$\{X_t, t \in T\}$ is called Yule process

Autoregressive Moving Average Process

Defn: ARMA process

$$\sum_{\gamma=0}^{p} b_\gamma X_{t-\gamma} = \sum_{s=0}^{q} a_s e_{t-s} \quad , b_0 = 1 \qquad ARMA(p,q)$$

Defn Covariance function

$$C_k = E(X_t X_{t+k}) - E(X_t)E(X_{t+k})$$

Defn Correlation function

$$\rho_k = \frac{C_k}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+k})}} = -\int_{\pi}^{\pi} e^{ix\omega} dF_1(\omega)$$

$F_1(\omega)$ - integrated spectrum

$dF_1(\omega) = f_1(\omega)d\omega$

Inverse representation of ρ_k:

$$f_1(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \rho_k e^{-ik\omega}$$

The representation of c_k:

$$c_k = \int_{-\pi}^{\pi} e^{ik\omega} dF(\omega)$$

$$dF(\omega) = c_0 dF_1(\omega)$$

The inverse representation of c_k:

$$f(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} c_k e^{-ik\omega}$$

time domain	Frequency domain
$\{X_t, t\in T\}$, c_k, ρ_k	$\{X_t, t\in T\}$, $f(\omega)$, $f_1(\omega)$, c_k, ρ_k

Example: $\{X_t, t\in T\}$ - white noise process

$$E[X_t] = 0 \quad , \quad E[X_t X_s] = \begin{cases} \sigma^2, & t=s \\ 0, & t\neq s \end{cases}$$

$c_k = 0,\ k \neq 0$ | $\rho_k = 0,\ k = 0$

$c_0 = \sigma^2$ | $\rho_0 = 1$

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} c_k e^{-ik\omega} = \frac{\sigma^2}{2\pi},\quad -\pi \leq \omega \leq \pi$$

$$F(\omega) = \int f(\omega)\,d\omega = \begin{cases} \dfrac{\sigma^2 \omega}{2\pi}, & -\pi \leq \omega \leq \pi \\ 0, & \text{otherwise} \end{cases}$$

$$dF(\omega) = \frac{\sigma^2}{2\pi}\,d\omega$$

$$c_k = \int_{-\pi}^{\pi} e^{ik\omega}\,dF(\omega) = \begin{cases} 0, & k \neq 0 \\ \sigma^2, & k = 0 \end{cases}$$

Summary

- Importance of stationary processes is discussed
- Few properties of stochastic processes are discussed.
- Autoregressive and moving average processes are explained
- Finally, study of power spectrum are also discussed.

Module 4: Discrete-time Markov Chain

Lecture 1: Introduction, Definition and Transition Probability Matrix

$P(\text{Head}) = p$, $P(\text{Tail}) = 1-p$, $0 < p < 1$

For the n^{th} trial, $X_n \in \{0, 1\}$

$$P(X_n = 0) = 1-p$$
$$P(X_n = 1) = p \quad \Rightarrow \{X_i, i = 1, 2, 3 \ldots\} \text{ is a stochastic process}$$

X_i are mutually independent r.v.s, $i = 1, 2, \ldots$

$$S_n := X_1 + X_2 + \cdots + X_n$$
$$S_{n+1} = S_n + X_{n+1}$$

$\{S_n, n = 1, 2, \ldots\}$ is a S.P.

$$P(S_{n+1} = k+1 \mid S_n = k) = p$$
$$P(S_{n+1} = k \mid S_n = k) = 1-p$$

$$P(S_{n+1} = k+1 \mid S_1 = i_1, S_2 = i_2, \ldots, S_n = k)$$
$$= \frac{P(S_{n+1} = k+1, S_n = k, \ldots, S_2 = i_2, S_1 = i_1)}{P(S_1 = i_1, S_2 = i_2, \ldots, S_n = k)}$$
$$= \frac{P(S_{n+1} = k+1 \mid S_n = k)\, P(S_n = k, \ldots, S_2 = i_2, S_1 = i_1)}{P(S_1 = i_1, S_2 = i_2, \ldots, S_n = k)}$$
$$= P(S_{n+1} = k+1 \mid S_n = k) = p, \quad n \geq 1$$

$$P(S_{n+1} = k \mid S_1 = i_1, S_2 = i_2, \ldots, S_n = k)$$
$$= P(S_{n+1} = k \mid S_n = k) = 1-p, \quad n \geq 1$$

$\{S_n, n=1,2,\ldots\}$ satisfies the "memoryless property" or "Markov property"

$\{S_n, n = 1, 2, \ldots\}$ - Markov process
- discrete time discrete state stochastic process
- Markov chain if state space is discrete (MC)

S: state space, discrete type

T: parameter space, discrete

If T is discrete, $\{S_n, n=1,2,..\}$ is a discrete time MC
" continuous " is a continuous time MC

<u>Definition</u> Discrete time Markov Chain (DTMC)

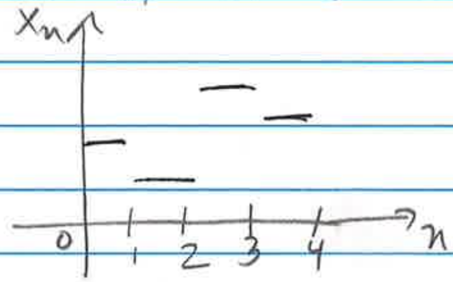
$\{X_n, n=0,1,2,\ldots\}$ - discrete time discrete state S.P.

$S = \{0,1,2,\ldots\}$

Suppose $P(X_{n+1}=j \mid X_0=i_0, X_1=i_1, \ldots, X_n=i\}$

$= P(X_{n+1}=j \mid X_n = i\}$, $\forall$ states $i_0, i_1, \ldots, i, j$ and $\forall n \geq 0$

Then $\{X_n, n=0,1,2..\}$ is a DTMC



X_n: state of the system at n^{th} step

Our interest:

(1) distribution of X_n, $n \geq 1$

(2) distribution of X_n as $n \to \infty$

Need:

(1) distribution of X_0

(2) transition distribution from X_n to X_{n+1} for any n

$P_j(n) = P(X_n = j)$ - pmf of r.v. X_n, $j \in S$

$P_{jk}(m,n) \triangleq P(X_n = k \mid X_m = j)$, $0 \leq m \leq n$, $j,k \in S$

When DTMC is time homogeneous, i.e., $P_{jk}(m,n)$ depends on $n-m$

$P_{jk}(n) = P(X_{m+n} = k \mid X_m = j)$, $\forall n$, $j,k \in S$

-Called n-step transition probability function

$P_{jk}(1) = p_{jk} = P(X_{n+1}=k \mid X_n = j)$, $\forall n \geq 1$, $j,k \in S$

$\forall j,k \in S$, $P_{jk}(0) = \begin{cases} 1 & , j=k \\ 0 & , \text{otherwise} \end{cases}$

4/13

$P = [p_{ij}]$, where $p_{ij} = P(X_{n+1}=j \mid X_n = i)$, $i,j \in S$, $\forall n \geq 1$

This satisfies: (1) $p_{ij} \geq 0$, $\forall i,j \in S$

(2) $\sum_j p_{ij} = 1$, $i \in S$ (row sum $=1$)

P is a stochastic matrix

<u>Transition Probability Matrix</u>

State transition diagram or directed graph or stochastic graph

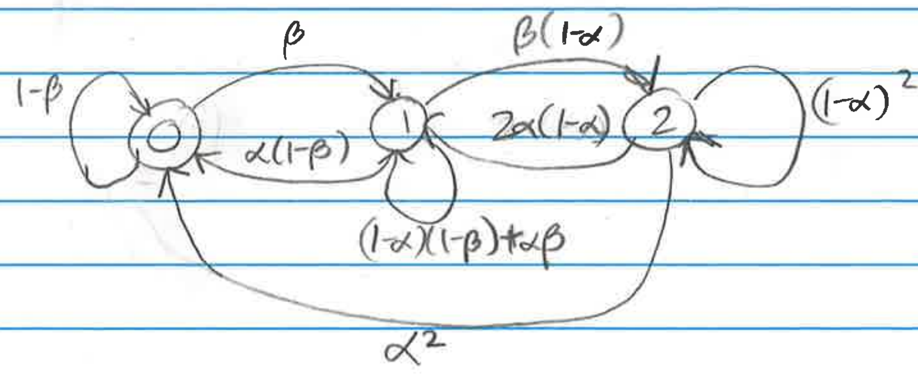
<u>Example 1</u>: A factory has two machines and one repair crew. Assume that the probability of any one machine breaking down on a given day is α. Assume that if the repair crew is working on a machine, the probability that they will complete the repairs in a day is β. For simplicity, ignore the probability of a repair completion or a breakdown occurring except at the end of the day. Let X_n be the number of machines in operation at the end of the n^{th} day. Assume that X_n can be modeled as a Markov chain.

$S = \{0, 1, 2\}$ $\{X_n, n=1,2,\dots\}$ — DTMC

Assume $\{X_n, n=1,2,\dots\}$ is time-homogeneous.

$$P = \begin{array}{cc} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \; X_{n+1} \\ \begin{array}{c} 0 \\ 1 \\ X_n \;\; 2 \end{array} & \begin{pmatrix} 1-\beta & \beta & 0 \\ \alpha(1-\beta) & (1-\alpha)(1-\beta)+\alpha\beta & \beta(1-\alpha) \\ \alpha^2 & 2\alpha(1-\alpha) & (1-\alpha)^2 \end{pmatrix} \end{array}$$

$p_{00}^{(1)} = 1-\beta$, $p_{01}^{(1)} = \beta$, $p_{02}^{(1)} = 0$ (Can't repair 2 machines in a day)



Example 2: The owner of a local one-chair barber shop is thinking of expanding the shop capacity because there seems to be too many people that are turned away. Observations indicate that in the time required to cut one person's hair there may be 0, 1, and 2 arrivals with probability 0.3, 0.4 and 0.3, respectively. The shop has a fixed capacity of six people including the one whose hair is being cut. Any new arrival who finds six people in the barber shop is denied entry. Let X_n be the number of people in the shop at the completion of the n^{th} person's hair cut. $\{X_n\}$ is a Markov chain assuming i.i.d. arrivals.

$\{X_n\}$ is a time homogeneous DTMC

$$P = \begin{array}{c} \boxed{X_n} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ \\ \end{array} \begin{pmatrix} 0.3 & 0.4 & 0.3 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 & 0 & 0 \\ 0 & 0.3 & 0.4 & 0.3 & 0 & 0 \\ 0 & 0 & 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0 & 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0 & 0 & 0 & 0.3 & 0.7 \end{pmatrix} \qquad S = \{0,1,2,3,4,5\}$$

$$\begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \\ & & \boxed{X_{n+1}} & & & \end{array}$$

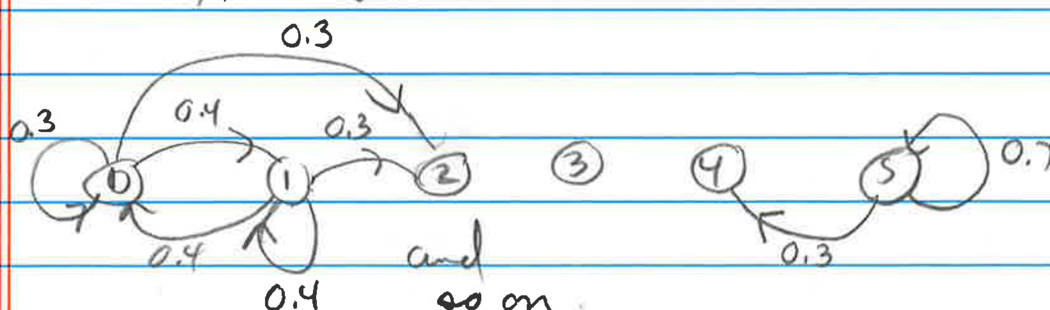
$p_{00}^{(1)} = P(X_{n+1}=0 \mid X_n=0) = 0.3$

$p_{01}^{(1)} = P(X_{n+1}=1 \mid X_n=0) = 0.4$

$p_{02}^{(1)} = P(X_{n+1}=2 \mid X_n=0) = 0.3$

$p_{54}^{(1)} = 0.3$, $p_{55}^{(1)} = 0.3+0.4 = 0.7$

$S = \{0,1,2,3,4,5\}$



Lecture 2: Chapman-Kolmogorov Equations

$$p_{jk}^{(n)} = P(X_{m+n}=k \mid X_m=j), \quad n \geq 0, \; j,k \in S$$

$$p_{jk}^{(1)} = P(X_{n+1}=k \mid X_n=j), \quad n \geq 0; \; j,k \in S$$

$$p_j^{(n)} = P(X_n=j), \; j \in S, \; n=1,2,\ldots$$

$$p_j^{(n)} = \sum_{i \in S} P(X_0=i)\,P(X_n=j \mid X_0=i)$$

$P(0) = [P(X_0=0) \;\; P(X_0=1) \;\; P(X_0=2) \;\; \cdots]$ Initial probability vector

$P(X_n=j \mid X_0=i)$ - ?

$$p_{ij}^{(n)} = P(X_n=j \mid X_0=i)$$

Chapman-Kolmogorov equations

Let $p_{ij}^{(n)} = P(X_{m+n}=j \mid X_m=i)$ time homogeneous DTMC

2-steps

$$p_{ij}^{(2)} = P(X_{n+2}=j \mid X_n=i) = \sum_{k \in S} P(X_{n+2}=j, X_{n+1}=k \mid X_n=i)$$

$$= \sum_{k \in S} \frac{P(X_{n+2}=j, X_{n+1}=k, X_n=i)}{P(X_n=i)}$$

$$= \sum_k \frac{P(X_{n+2}=j \mid X_{n+1}=k, X_n=i)\,P(X_{n+1}=k, X_n=i)}{P(X_n=i)}$$

$$\overset{\text{Markov Property}}{=} \sum_k \frac{P(X_{n+2}=j \mid X_{n+1}=k)\,P(X_{n+1}=k \mid X_n=i)\,\cancel{P(X_n=i)}}{\cancel{P(X_n=i)}}$$

$$\therefore \; p_{ij}^{(2)} = \sum_k p_{ik}^{(1)} p_{kj}^{(1)}$$

$$p_{ij}^{(m+1)} = \sum_k p_{ik}\, p_{kj}^{(m)} = \sum_k p_{ik}^{(m)} p_{kj}$$

$$p_{ij}^{(m+n)} = \sum_k p_{ik}^{(m)} p_{kj}^{(n)}$$ — CK eqn for time homogeneous DTMC

$$\underline{P} = [p_{ij}^{(1)}] \Rightarrow \underline{P}^{(2)} = \underline{P} \cdot \underline{P} = \underline{P}^2 \Rightarrow \underline{P}^{(n)} = \underline{P}^n, \ n \geq 1$$

n-step

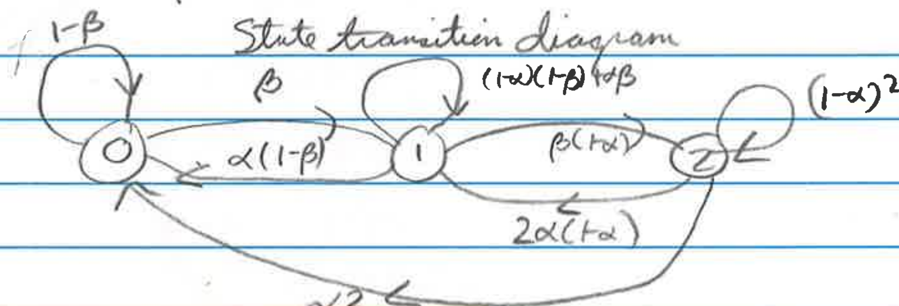
$$P(n) = [P(X_n=0) \ P(X_n=1) \ P(X_n=2) \ \cdots]$$
$$= P(0) P^{(n)} = P(0) P^n$$

Example 1: Repair crew with two machines from last lecture

$S = \{0, 1, 2\}$

$$P = \begin{array}{cc} & X_{n+1} \\ & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ X_n \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 1-\beta & \beta & 0 \\ \alpha(1-\beta) & (1-\alpha)(1-\beta)+\alpha\beta & \beta(1-\alpha) \\ \alpha^2 & 2\alpha(1-\alpha) & (1-\alpha)^2 \end{pmatrix} \end{array}$$

State transition diagram



Q) If the system starts out with both machines operating, what is the probability that both will be in operation two days later?

$$P(X_0=2) = 1$$

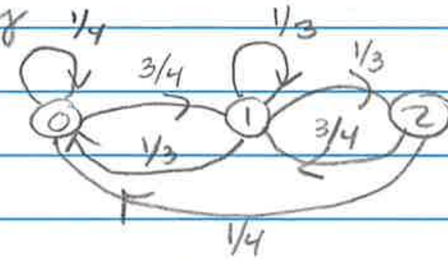
$$P(0) = [P(X_0=0) \ P(X_0=1) \ P(X_0=2)] = [0 \ 0 \ 1]$$

$$P(X_2=2) = ?$$
$$= \sum_i P(X_0=i) P(X_2=2 \mid X_0=i)$$
$$= P(X_0=2) P(X_2=2 \mid X_0=2) + 0 + 0$$
$$= P(X_2=2 \mid X_0=2) = p_{2,2}^{(2)}$$

$$p_{2,2}^{(2)} = [P^2]_{(3,3)}$$

Example 2: Let $\{X_n, n=0,1,2,\dots\}$ be a time homogeneous Markov chain with $S=\{0,1,2\}$, $P(0)=(1/4, 1/2, 1/4)$ and one-step transition probability matrix P given by

$$P = \begin{array}{c} 0 \\ 1 \\ 2 \\ \\ \end{array} \begin{pmatrix} 1/4 & 3/4 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/4 & 3/4 \end{pmatrix}$$
$$\quad\quad\quad 0 \quad\;\; 1 \quad\;\; 2$$



$$P(X_0=0, X_1=1, X_2=1) = P(X_2=1 \mid X_1=1)\, P(X_1=1 \mid X_0=0)\, P(X_0=0)$$
$$= \frac{1}{3}\cdot\frac{3}{4}\cdot\frac{1}{4} = \frac{1}{16}$$

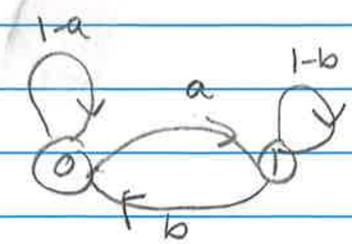
$$P(X_2=1) = \sum_{i\in S} P(X_0=i)\, P(X_2=1 \mid X_0=i)$$

$$P^2 = \begin{bmatrix} 1/4 & 3/4 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/4 & 3/4 \end{bmatrix} \begin{bmatrix} 1/4 & 3/4 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/4 & 3/4 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$P(X_7=0 \mid X_5=0) = p_{0,0}^{(2)} = [P^2]_{(1,1)}$$

Example 3: Consider a communication system which transmits the two digits 0 or 1 through several stages. Let X_0 be the digit transmitted initially at the 0^{th} stage and X_n, $n=1,2,\dots$ be the digit leaving the n^{th} stage. The transition probability matrix of the corresponding system is given by

$S=\{0,1\}$



$$\begin{array}{cc} & X_{n+1} \\ & \begin{array}{cc} 0 & \quad 1 \end{array} \end{array}$$
$$P = \begin{array}{c} \\ X_n \end{array} \begin{array}{c} 0 \\ 1 \end{array} \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \qquad \begin{array}{l} 0<a<1 \\ 0<b<1 \end{array}$$

$$P^{(n)} = \begin{array}{c|c} & \begin{array}{cc} \quad 0 \quad & \quad 1 \quad \end{array} \\ \begin{array}{c} 0 \\ \\ 1 \end{array} & \begin{bmatrix} \dfrac{b+a(1-a-b)^n}{a+b} & \dfrac{a-a(1-a-b)^n}{a+b} \\ \dfrac{b-b(1-a-b)^n}{a+b} & \dfrac{a+b(1-a-b)^n}{a+b} \end{bmatrix} \end{array} \quad \text{for } |1-a-b|<1$$

As $n \to \infty$: $\quad p_0 = \dfrac{b}{a+b}$, $\quad p_1 = \dfrac{a}{a+b}$

Example 4: Let $\{Y_n, n=1,2,\ldots\}$ be a sequence of independent r.v.'s with $P(Y_n=1) = p = 1-P(Y_n=-1)$, $n=1,2,\ldots$, $0<p<1$

Let X_n be defined by $X_0=0$, $X_{n+1}=X_n+Y_{n+1}$, $n=1,2,\ldots$

Check if $\{X_n, n=1,2,\ldots\}$ is a DTMC. If DTMC, find $P(X_n=k)$

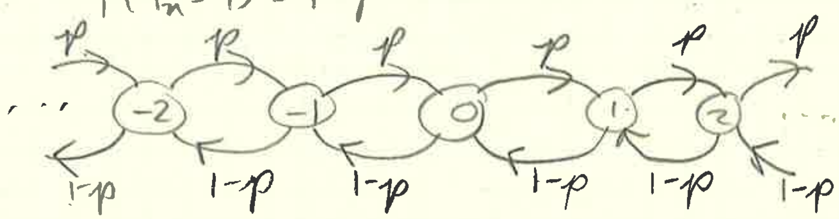
$S=\{0,\pm1,\pm2,\ldots\}$

$X_0=0$, $X_{n+1}=X_n+Y_{n+1}$, $n=1,2,\ldots$

$Y_n \in \{-1,1\}$

$P(Y_n=1)=p$ $\quad 0<p<1$

$P(Y_n=-1)=1-p$



X_{n+1} only depends on X_n, not $X_{n-1}, X_{n-2}, \ldots$

$\{X_n, n=1,2,\ldots\}$ satisfies the Markov property, thus is a DTMC.

$P(X_n=k)$ is found by C-K equations

Summary

- Chapman-Kolmogorov equation is introduced
- n-step Transition Probability Matrix is explained
- Finally, simple examples are also illustrated.

Lecture 3: Classification of State and Limiting Distributions 4.3.1

A. Introduction

Accessible: State i is accessible from state j if

$p_{ji}^{(n)} > 0$ for some $n \geq 0$.

$P(\text{ever enter state } i \mid \text{initially in state } j)$

$= P\left(\bigcup_{n=0}^{\infty} \{X_n = i\} \mid X_0 = j\right)$

Communicate: Two states communicate if state i is accessible from state j and state j is accessible from state i.

$i \leftrightarrow j$

(1) $i \leftrightarrow i \Leftrightarrow p_{ii}^{(0)} = P(X_0 = i \mid X_0 = i) = 1, \forall i$

(2) If $i \leftrightarrow j \Rightarrow j \leftrightarrow i$ (symmetry)

(3) If $i \leftrightarrow j$ and $j \leftrightarrow k \Rightarrow i \leftrightarrow k$ (transitivity)

Communication is an equivalence relation, and hence partitions the set of states into communicating classes.

Class: A class of states is a subset of the state space S such that every state of the class communicates with every state and there is no other state outside the class which communicates with all other states in the class.

Class Property:

- All states belonging to a particular class share the same properties.

Periodicity: State i is a return state if $p_{ii}^{(n)} > 0$ for some $n \geq 1$. The period d_i of a return state i is defined as the greatest common divisor of all m such that $p_{ii}^{(m)} > 0$.

$$d_i = \gcd(\{m \mid p_{ii}^{(m)} > 0\})$$

If $d_i = 1$, then the state is aperiodic.

Within a class, all states have the same period

Closed Set of States:

If C is a set of states s.t. no state outside C can be reached from any state in C, then C is said to be closed.

If only one element in C, state i called an absorbing state

$\Leftrightarrow p_{ii}^{(1)} = 1$

Irreducible: - If a Markov chain does not contain any other proper closed subset of the state space S, other than the state space S itself, then the Markov chain is said to be an irreducible Markov chain.

- The states of a closed communicating class share same class properties. Hence, all the states in the irreducible chain are of the same type.

First Visit:

$f_{jk}^{(n)} = P(\text{state } k \text{ for the first time at the } n^{th} \text{ time step} \mid \text{system starts a state } j \text{ initially})$

pmf

$p_{jk}^{(n)} = P(\text{state } k \text{ at the } n^{th} \text{ time step} \mid \text{state } j \text{ initially})$

$$p_{jk}^{(n)} = \sum_{r=0}^{n} f_{jk}^{(r)} p_{kk}^{(n-r)}, \quad n \geq 1$$

$$p_{kk}^{(0)} = 1, \quad f_{jk}^{(0)} = 0, \quad f_{jk}^{(1)} = p_{jk}^{(1)}$$

First Passage Time

- $F_{jk} = P(\text{the system starting at state } j \text{ will ever reach state } k)$

$$= \sum_{n=1}^{k} f_{jk}^{(n)}$$

We have two possibilities:

- $F_{jk} < 1$ - w.p. $(1-F_{jk})$ you will never reach state k when starting at j
- $F_{jk} = 1$ - w.p. 1 you will reach state k when starting at state j

Mean Recurrence Time (or mean first passage time)

$$\mu_{jk} = \sum_{n=1}^{\infty} n f_{jk}^{(n)}$$

When $k=j \Rightarrow f_{jj}^{(n)}$ - distribution of the recurrence time of the state j

If $F_{jj} = 1 \Rightarrow$ the return to state j is certain.

μ_{jj} - mean recurrence time for state j

$$\mu_{jj} = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$$

$$T_j = \inf\{n \geq 1 : X_n = j \mid X_0 = j\}$$ - first return time to state j

$$\mu_j := \mu_{jj} = E(T_j)$$

B. Classification of States

Recurrent State (or persistent): if $F_{jj} = 1$

$$F_{jj} = \sum_{n=1}^{\infty} f_{jj}^{(n)}$$

Null recurrent : $\mu_{jj} = \infty$

positive recurrent : $\mu_{jj} < \infty$

If positive recurrent and aperiodic, it is called *ergodic*. 4.3.4

In a MC, if all states are ergodic, then the MC is called ergodic.

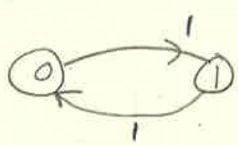
Transient State: If $F_{jj} < 1$. With some probability $(1-F_{jj})$, the system will not return to state j.

state j is transient iff $P(T_j = \infty) > 0$

- State
  - $(F_{jj}=1)$ recurrent
    - positive $(\mu_j < \infty)$
      - aperiodic $(d_j = 1)$ (ergodic)
      - periodic $(d_j > 1)$
        - absorbing (when $p_{jj} = 1$)
    - null $(\mu_j = \infty)$
      - aperiodic
      - periodic
  - transient $(F_{jj} < 1)$

Lecture 4: Limiting and Stationary Distributions 4.4.1

Example 1:



$$P = \begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

(rows: X_n, columns: X_{n+1})

$S = \{0, 1\}$

$f_{00}^{(1)} = 0$, $f_{00}^{(2)} = 1$, $f_{00}^{(n)} = 0$, $n \geq 3$ only for first visit

$$F_{00} = \sum_{n=1}^{\infty} f_{00}^{(n)} = 1 \Rightarrow \{0\}\text{-recurrent state } (+ve)$$

Similarly, $F_{11} = 1 \Rightarrow \{1\}$-recurrent state $(+ve)$

$$\mu_{00} = \sum_{n=1}^{\infty} n f_{00}^{(n)} = 2 f_{00}^{(2)} = 2 \cdot 1 = 2$$

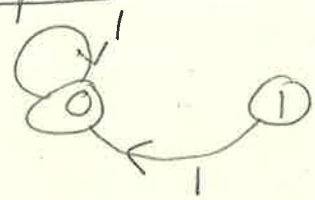
$$\mu_{11} = 2$$

This MC is irreducible.

$d_0 = \gcd\{2, 4, 6, \ldots\} = 2 \Rightarrow$ period for state $\{0\}$ is 2

$d_1 = \gcd\{2, 4, 6, \ldots\} = 2 \Rightarrow$ period for state $\{1\}$ is 2

Example 2:



$S = \{0, 1\}$

$$P = \begin{array}{cc} & X_{n+1} \\ & \begin{array}{cc} 0 & 1 \end{array} \\ X_n \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \end{array}$$

1st closed communicating class $C_1 = \{0\}$

$p_{00}^{(1)} = 1 \Rightarrow \{0\}$ - absorbing state

$f_{11}^{(1)} = 0$, $f_{11}^{(n)} = 0, n \geq 1 \Rightarrow F_{11} = \sum_{n=1}^{\infty} f_{11}^{(n)} = 0$

$\{1\}$ - transient state $\Rightarrow T = \{1\} \leftarrow$ class of all transient state

This is a reducable MC

Example 3:

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

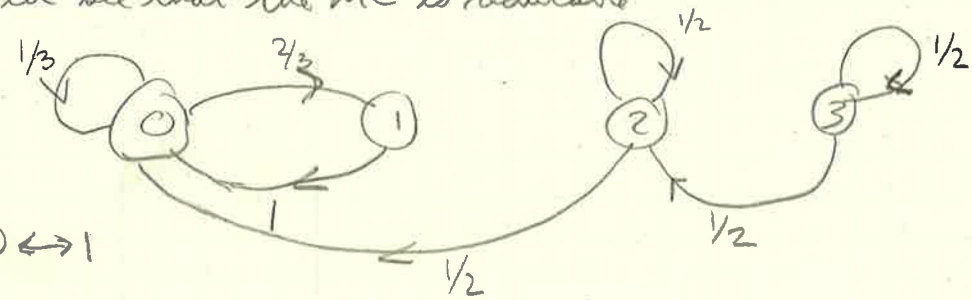
$f_{00}^{(1)} = \frac{1}{3}$, $f_{00}^{(2)} = \frac{2}{3} \cdot 1$, $f_{00}^{(n)} = 0, n \geq 3$

$f_{22}^{(1)} = \frac{1}{2}$; $f_{22}^{(n)} = 0$, $n \geq 2$ (Same for 3)

$F_{00} = 1$, $F_{22} = \frac{1}{2}$

Hence, $\{0, 1\}$ are recurrent states and $\{2, 3\}$ are transient states. The period of state 0 is 1 since $d_0 = \gcd\{1, 2, 3, \ldots\}$. Hence also we see that the MC is reducable

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



$0 \leftrightarrow 1$

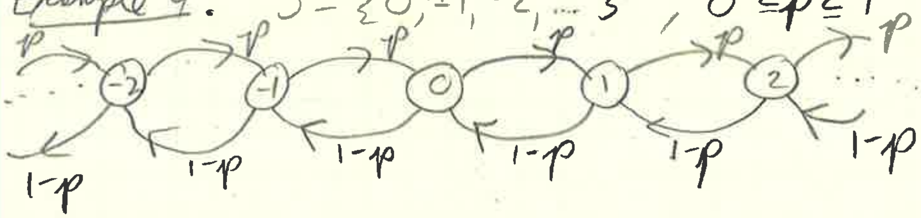
$\mu_{00} = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} = \frac{5}{3} \Rightarrow \{0\}$ - +ve recurrent state

$\mu_{11} < \infty \Rightarrow \{1\}$ - is +ve recurrent state

$d_0 = 1 \Rightarrow$ aperiodic +ve recurrent $\Leftrightarrow$ ergodic

$d_1 = 1 \Rightarrow \{0, 1\}$ — ergodic state

Example 4: $S = \{0, \pm 1, \pm 2, \ldots\}$, $0 \leq p \leq 1$



case 1: $p=0$, all states transition to next state below

$\Rightarrow$ all states are transient states

case 2: $p=1$, all states transition to next forward state

$\Rightarrow$ all states are transient states

Case 3: $0 < p < 1$

$\{0, \pm1, \pm2, \ldots\}$ - recurrent states after even # of steps

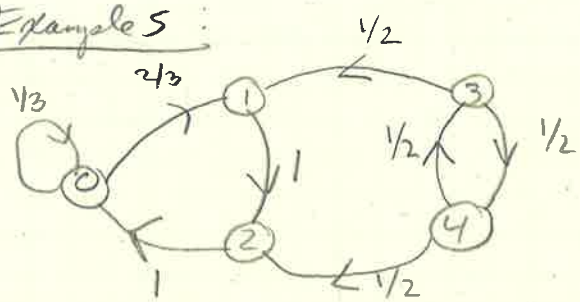
$d_i = \gcd\{n : p_{ii}^{(n)} > 0\} = \gcd\{2,4,6,\ldots\} = 2, \ \forall i \in S$

$\mu_{11} = \sum_{n=1}^{\infty} n f_{ii}^{(n)} = ?$ $\qquad \{0, \pm1, \pm2, \ldots\}$ - not ergodic

all states communicate $\Rightarrow$ one communicating class including all state

$\Rightarrow$ Irreducible MC

Example 5:



$\{3,4\}$ $\quad 1 \not\leftrightarrow 3, \ 2 \not\leftrightarrow 4$

$\{0,1,2\}$ – form loop

$0 \leftrightarrow 1, \ 1 \leftrightarrow 2, \ 2 \leftrightarrow 0$

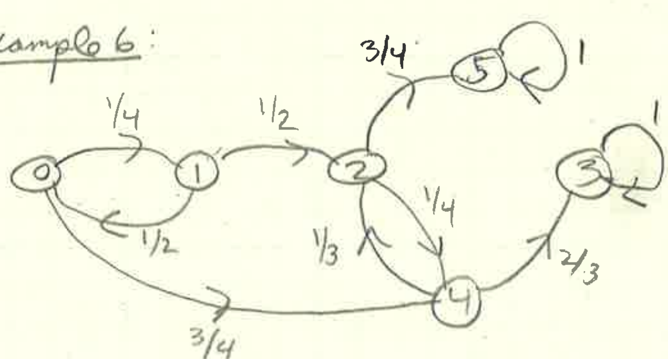
$\{0,1,2\}$ - closed communicating class $\qquad d_0 = \gcd\{1,3,4,5,\ldots\} = 1$

– +ve recurrent states, aperiodic $\Rightarrow$ ergodic

$\{3,4\}$ – transient states

This is a reducible MC

Example 6:



$C_1 = \{3\}$ – absorbing

$C_2 = \{5\}$ - absorbing

$T = \{0,1,2,4\}$ - transient states

Reducible MC

Example 7: Reducible MC

$0 < p < 1$

$\{0, N\}$ - absorbing state

$\{1, 2, \ldots, N-1\}$ - transient states

Lecture 5: Limiting Distributions, Ergodicity and Stationary Distributions

Consider Doeblin's formula

$$F_{jk} = \lim_{m\to\infty} \frac{\sum_{n=1}^{m} p_{jk}^{(n)}}{1+\sum_{n=1}^{m} p_{kk}^{(n)}}$$

In particular, $F_{jj} = 1 - \lim_{m\to\infty} \frac{1}{1+\sum_{n=1}^{m} p_{jj}^{(n)}}$

$$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$$

$$F_{jk} = \sum_{n=1}^{\infty} f_{jk}^{(n)}$$

Results

(1) State j is recurrent iff. $\sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty$

(2) State j is transient ($F_{jj}<1$) if $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$. Also, if j is transient then $p_{jj}^{(n)} \to 0$ as $n\to\infty$

Theorem: Basic limit theorem of renewal theory

If state j is positive recurrent, then as $n\to\infty$:

(i) $p_{jj}^{(n)} \to \frac{t}{\mu_{jj}}$, state j is periodic with period t.

(ii) $p_{jj}^{(n)} = \frac{1}{\mu_{jj}}$, state j is aperiodic

(iii) $p_{jj}^{(n)} \to 0$, state j is transient

If state j is null recurrent,

$$\mu_{jj} = \infty,\ p_{jj}^{(n)} \to 0 \text{ as } n\to\infty$$

Some important results - time homogeneous DTMC

(1) For an irreducible MC, all states are of the same type
(2) For a finite MC, at least one state must be positive recurrent.
(3) For an irreducible finite MC, all states are positive recurrent.

A. Limiting State Probabilities

$$\lim_{n\to\infty} P(X_n=j \mid X_0=i) \quad , i,j = 0,1,2,\ldots$$

Suppose the limiting probabilities are independent of the initial state of the process. Then,

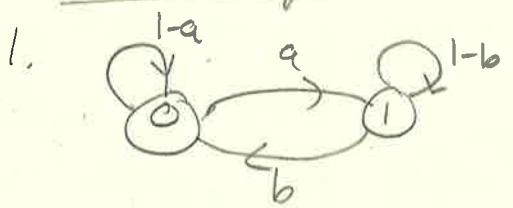
$$V_j := \lim_{n\to\infty} p_{ij}^{(n)} \quad , V=[V_0\ V_1\ \cdots]$$

$$V_k = \sum_j V_j p_{jk} = \sum_j \left(\sum_i V_i p_{ij}\right) p_{jk}$$

$$= \sum_i V_i p_{ik}^{(2)}$$

$$V_k = \sum_i V_i p_{ik}^{(n)} \quad , n \geq 1$$

Some examples

1. 

$0 < a, b < 1$

limiting distribution matrix : $\lim_{n\to\infty} P^{(n)}$

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-a & a \\ 1 & b & 1-b \end{array} \quad , F_{00}=1 \quad , F_{11}=1$$

$\mu_{00} < \infty , \mu_{11} < \infty$

$\Bigg\}$ positive recurrent

$P^{(n)} = P^n$ Use eigenvalue/eigenvectors to find P^n

$$P^n = \begin{pmatrix} \frac{b+a(1-a-b)^n}{a+b} & \frac{a-a(1-a-b)^n}{a+b} \\ \frac{b-b(1-a-b)^n}{a+b} & \frac{a+b(1-a-b)^n}{a+b} \end{pmatrix} \quad , n=2,3,\ldots \quad |1-a-b|<1$$

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$$

Limiting distribution vector, $\Pi = (\Pi_0\ \Pi_1)$. Note that Π_0, Π_i are independent of initial state 'i'.

$\Pi_0 = \frac{b}{a+b} \quad , \quad \Pi_1 = \frac{a}{a+b}$

probabilities

In general, the three distributions, limiting, stationary, and equilibrium are different. They are the same for some special cases of DTMC.

B. Ergodicity

It is a necessary and sufficient condition for the existence of $\{v_j\}$ satisfying

$$v_j = \sum_i v_i P_{ij} \ ; \ \sum v_j = 1$$

in case of an irreducible and aperiodic MC for the MC to be ergodic.

C. Stationary Distribution

The vector π is called a stationary distribution of the DTMC if $\pi = (\pi_0, \pi_1, \dots)$ satisfies

(1) $\pi_j \ge 0, \ \forall j$

(2) $\sum \pi_j = 1$

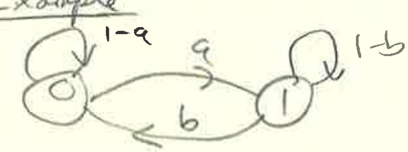
(3) $\pi = \pi P$

For an irreducible, aperiodic, $\underset{\text{positive recurrent}}{\vee}$ MC $\Rightarrow$ limiting distribution $\Leftrightarrow$ stationary distribution $\Leftrightarrow$ equilibrium distribution

Some Important Results

1) For an irreducible, aperiodic, positive recurrent MC, the stationary distribution exists and is unique. The π is uniquely determined by $\pi = \pi P$, i.e., $\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}$

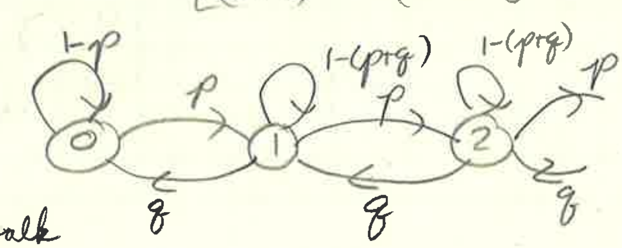
Example

1) 

$0 < a, b < 1$

$$\pi = \left(\frac{b}{a+b}, \frac{a}{a+b}\right) \Rightarrow \left[\frac{b}{a+b} \ \frac{a}{a+b}\right]\begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix} = \left[\frac{b^2+ab}{(a+b)^2} \ \frac{a^2+ab}{(a+b)^2}\right]$$

$$= \left[\frac{b(a+b)}{(a+b)^2} \ \frac{a(a+b)}{(a+b)^2}\right] = \left[\frac{b}{a+b} \ \frac{a}{a+b}\right] \Leftrightarrow \pi = \pi P$$

2) 

.... $0 < p, q < 1$

1D random walk

- Each states communicate with the others so this is irreducible.
- Aperiodic — all states are the same type.
- Is $\mu_{00} < \infty$? Assume that all states are positive recurrent.

$\rightarrow$ the stationary distribution exists and is unique.

$$\Rightarrow \pi = \pi \underbrace{P}_{\text{one-step transition matrix}}, \quad \sum_i \pi_i = 1 \quad \text{where } \pi = [\pi_0 \; \pi_1 \; \pi_2 \; \dots]$$

$$\pi_0 = \pi_0(1-p) + \pi_1 q \;\Rightarrow\; \pi_1 = \frac{p}{q}\pi_0$$

$$\pi_1 = \pi_0 p + \pi_1(1-p-q) + \pi_2 q \;\Rightarrow\; \pi_2 = \frac{p^2}{q^2}\pi_0$$

$$\vdots$$

$$\pi_3 = \frac{p^3}{q^3}\pi_0$$

$$\vdots$$

$$\pi_n = \frac{p^n}{q^n}\pi_0, \quad n = 1, 2, 3, \dots$$

$$\sum_{i=0}^{\infty} \pi_i = 1 : \quad \pi_0\left(1 + \frac{p}{q} + \frac{p^2}{q^2} + \dots\right) = 1$$

$$\Rightarrow \pi_0 = \frac{1}{1 + \frac{p}{q} + \frac{p^2}{q^2} + \dots} \quad \text{converges if } \frac{p}{q} < 1$$

$$\pi_n = \left(\frac{p}{q}\right)^n \pi_0 \quad \text{if } \underbrace{\frac{p}{q} < 1}_{\text{condition for positive recurrent}}$$

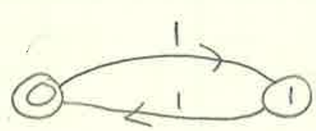
3) Consider a DTMC with $P = [P_{ij}]$ s.t. $\sum_{i=1} P_{ij} = 1$, $j = 1, 2, \dots, n$

Here, P is called a doubly stochastic matrix. If a finite irreducible DTMC has a doubly stochastic matrix, then all of the stationary probabilities are equal, i.e.,

$$\pi_j = \frac{1}{n}, \quad j = 1, 2, \dots, n$$

and all states will be positive recurrent.

4)



$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \sim \text{doubly stochastic}$$

$d_0 = 2$ $\quad \{0, 1\}$ – positive recurrent

$d_1 = 2$ $\quad$ irreducible

$$\pi = \pi P, \quad \sum_i \pi_i = 1, \quad \pi = (\pi_0, \pi_1) = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$\lim_{n \to \infty} P^n$ does not exist — limiting distribution does not exist, but the stationary distribution does exist.

Lecture 6: Time Reversible Markov Chain, Application of Irreducible Markov Chain in Queueing Models

A. Time Reversible Markov Chain

- Consider a DTMC $\{\ldots, X_{n-2}, X_{n-1}, X_n, X_{n+1}, X_{n+2}, \ldots\}$
- Trace the DTMC backwards

$$\{\ldots, X_{n+2}, X_{n+1}, X_n, X_{n-1}, X_{n-2}, \ldots\}$$

Is $\{X_{n-i}, i=0,1,2,\ldots\}$ a DTMC? Yes, the reversible process is also a DTMC.

$Q = [Q_{ij}]$ - one step transition probability matrix

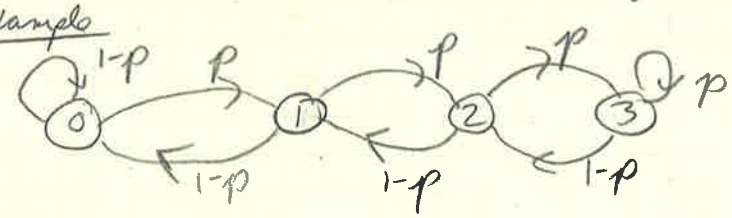
$$Q_{ij} = P[X_n = j \mid X_{n+1} = i]$$
$$= \frac{P[X_n = j]\, P[X_{n+1} = i \mid X_n = j]}{P[X_{n+1} = i]} = \frac{\pi_j p_{ji}}{\pi_i}$$ If the stationary distribution exists

Definition: A DTMC is said to be time reversible if $Q_{ij} = P_{ij}$, i.e., the reverse DTMC has the same transition probability matrix as the original DTMC.

Since, $Q_{ij} = \frac{\pi_j p_{ji}}{\pi_i} \Leftrightarrow \pi_j p_{ji} = \pi_i p_{ij}$ time-reversibity equations

Result: For an irreducible DTMC, if there exists a probability solution π satisfying the time-reversiblity equations $\pi_i p_{ij} = \pi_j p_{ji}, \forall i,j$, then the DTMC has positive recurrent states, is time-reversible and the solution π is the unique stationary distribution.

Example



$$\pi_0 p_{01} = \pi_0 p = \pi_1 (1-p) = \pi_1 p_{10}$$

✓ Irreducible: $\pi_1 p_{12} = \pi_1 p = \pi_2 (1-p) = \pi_2 p_{21}$

$$\pi_2 p_{23} = \pi_2 p = \pi_3 (1-p) = \pi_3 p_{32}$$

Solve

$$\pi_n = \frac{p^n}{(1-p)^n} \pi_0, \quad n = 1,2,3$$

$$\pi_0 = \frac{1}{1 + \frac{p}{1-p} + \frac{p^2}{(1-p)^2} + \frac{p^3}{(1-p)^3}}$$

$$\sum_i \pi_i p_{ij} = \sum_i \pi_j p_{ji} = \pi_j \underbrace{\sum_i p_{ji}}_{=1} = \pi_j$$

$$\sum_i \pi_i p_{ij} = \pi_j \iff \pi P = \pi$$

B. Irreducible MC in queueing models

Consider a barber shop with fixed capacity of six people including the one whose hair is being cut. Observations indicate that in the time required to cut one person's hair there may be 0, 1, and 2 arrivals with probability 0.3, 0.4, and 0.3 respectively. Any new arrival is denied entry when the shop is occupied with six people. Find the long run proportion of time that the shop has six people in it.

$X(t)$: # of customers in the barber shop at time t
$\{X(t), t \geq 0\}$ is a stochastic process (continuous time, discrete-state)

The corresponding queueing model is

M | G | 1 | 6 system

- 6 — capacity
- 1 — # of servers
- G — service time distribution
- M — interarrival time distribution

Markov property / Exponential distribution

General distribution

e.g., M|M|1, M|M|1|4, M|M|∞, M|M|3|7

capacity is ∞

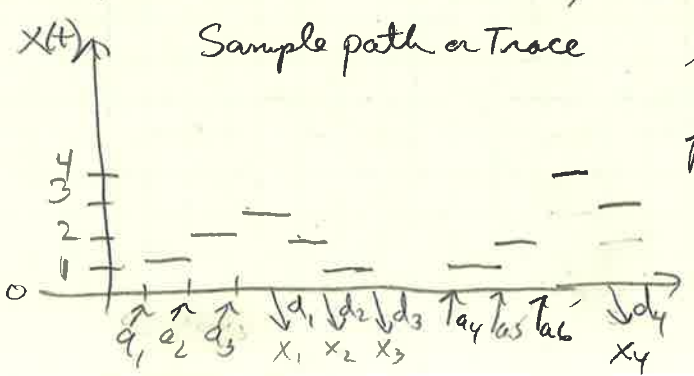
X_n = # of customers in the shop at the completion of the n^{th} customer's hair cut.
$S = \{0,1,2,3,4,5\}$

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0.3 & 0.4 & 0.3 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 & 0 & 0 \\ 0 & 0.3 & 0.4 & 0.3 & 0 & 0 \\ 0 & 0 & 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0 & 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0 & 0 & 0 & 0.3 & 0.7 \end{pmatrix} \end{array}$$

Assuming iid arrivals, $\{X_n, n \geq 1\}$ is a DTMC

$\{X_n\}_{n=0}^{6}$ is embedded in $\{X(t), t \geq 0\}$ called an embedded MC and DTMC.

Sample path or Trace



$p_{00}^{(1)} = P(X_{n+1}=0 \mid X_n = 0), \ \forall n$

$p_{01}^{(1)} = P(X_{n+1}=1 \mid X_n = 0), \ \forall n$

✓ Irreducible
✓ Aperiodic
✓ positive recurrent

$\pi = [\pi_0\ \pi_1\ \pi_2\ \pi_3\ \pi_4\ \pi_5]$ exists

Solve $\pi = \pi P$; $\sum_{i=0}^{5} \pi_i = 1$

What is π_5 ?

Embedded MC in M/G/1 queueing model

Let $\{X_n, n=1,2,\ldots\}$ be an irreducible, aperiodic DTMC with state space $S=\{0,1,2,\ldots\}$ and one-step transition probability matrix

$$P = \begin{pmatrix} k_0 & k_1 & k_2 & k_3 & \cdots \\ k_0 & k_1 & k_2 & k_3 & \cdots \\ 0 & k_0 & k_1 & k_2 & \cdots \\ 0 & 0 & k_0 & k_1 & \cdots \\ \vdots \end{pmatrix} \qquad \sum_{i=0}^{\infty} k_i = 1$$

Define $p(s) = \sum_i k_i s^i$ and $\pi(s) = \sum_i \pi_i s^i$ (generating functions)

If $p'(1) > 1$, states are transient

$\frac{dp}{ds}(1) = p'(1) = 1$, " " null recurrent

$p'(1) < 1$, " " positive recurrent

When $p'(1) < 1$, the unique stationary distribution $\pi = [\pi_0\ \pi_1\ \pi_2 \ldots]$ exists.

Solve $\pi = \pi P$, we get

$\pi_0 = k_0\pi_0 + k_0\pi_1$

$\pi_1 = k_1\pi_0 + k_1\pi_1 + k_0\pi_2$

$\vdots$

$$\pi(s) = \frac{\pi_0(1-s)\,p(s)}{p(s)-s}$$

Use $\sum_i \pi_i = 1$ to find π_0

$$\lim_{s\to 1} \frac{\pi(s)}{\pi_0} = \lim_{s\to 1} \frac{(1-s)\,p(s)}{p(s)-s} \Rightarrow \pi_0 = 1 - p'(1)$$

Hence, $$\pi(s) = \frac{[1-p'(1)]\,(1-s)\,p(s)}{p(s)-s}$$

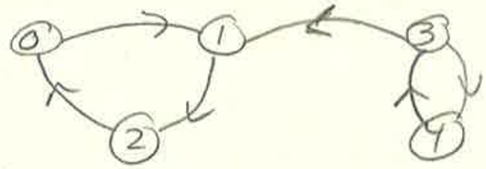
Lecture 7: Reducible Markov Chains

A. Closed Communicating Class of States

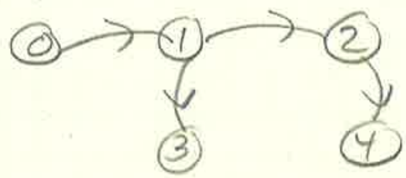
- If C is a set of states such that no such state outside C can be reached from any state in C, then C is said to be a closed communicating class of states
- Only one element in C, state i is called an absorbing state: $p_{ii} = 1$

Definition: If the MC does not contain any other closed communicating class of states other than the state space S, then the MC is called an irreducible MC. Otherwise, it is called a reducible MC.

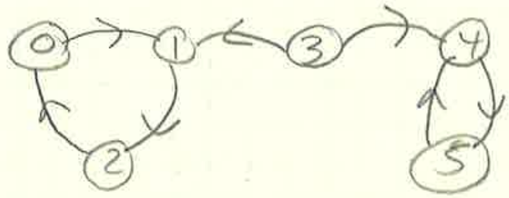
Examples

(1) 

$C_1 = \{0, 1, 2\}$
$T = \{3, 4\}$

(2) 

$C_1 = \{3\}$
$C_2 = \{4\}$
$T = \{0, 1, 2\}$

(3) 

$C_1 = \{0, 1, 2\}$
$C_2 = \{4, 5\}$
$T = \{3\}$

Types of Reducible MC

1. With one closed communicating class of states
2. With one or more absorbing states
3. With more than one closed communicating class of states

1. With one closed communicating class (C)

- Further, assume finite states and that the states of C are aperiodic
- Canonical form:

$$P = \begin{array}{c c} & \begin{array}{cc} C & T \end{array} \\ \begin{array}{c} C \\ T \end{array} & \begin{pmatrix} P_1 & 0 \\ R_1 & Q \end{pmatrix} \end{array} \qquad S = C \cup T$$

P_1 - stochastic submatrix
C - the set of closed communicating class states
T - the set of transient states

B. Stationary Distribution for Reducible MC

- For a reducible finite MC with a closed communicating class and aperiodic states, the stationary distribution exists and is given by $V = (V_1, 0)$. (Ergodic Theorem)

$$P^n = \begin{pmatrix} P_1^n & 0 \\ R_n & Q^n \end{pmatrix}$$

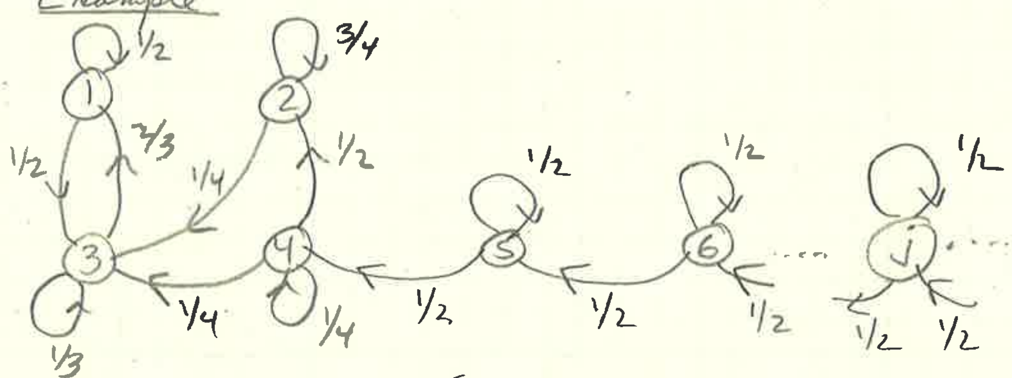
As $n \to \infty$, $P_1^n \to eV_1$ and $Q^n \to 0$, where $e = (1,1,1,\dots,1)$

- Further, assume that positive recurrent and the states of C are aperiodic.

$$\lim_{n \to \infty} p_{ij}^{(n)} = \begin{cases} V_j & , j \text{ is positive recurrent} \\ 0 & , j \text{ is transient} \end{cases}$$

– independent of the initial state 'i'.

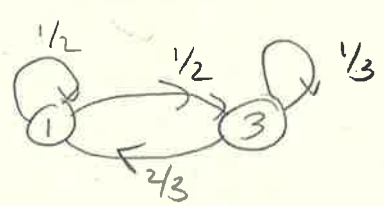
<u>Example</u>



As $n \to \infty$: $p_{i1}^{(n)} \to 4/7$

$p_{i3}^{(n)} \to 3/7$

$p_{ij}^{(n)} \to 0$, $j = 2,4,5,6,7,8,\dots$



$$P_1 = \begin{array}{c|cc} & 1 & 3 \\ \hline 1 & 1/2 & 1/2 \\ 3 & 2/3 & 1/3 \end{array}$$

$$\pi P_1 = \pi$$

$$(\pi_1 \;\; \pi_3)\begin{pmatrix} 1/2 & 1/2 \\ 2/3 & 1/3 \end{pmatrix} = (\pi_1 \;\; \pi_3)$$

$$\tfrac{1}{2}\pi_1 + \tfrac{2}{3}\pi_3 = \pi_1 \Rightarrow \pi_3 = \tfrac{3}{4}\pi_1$$

$$\pi_1 + \pi_3 = 1 \Rightarrow \pi_1 = 4/7,\ \pi_3 = 3/7$$

2. With one or more absorbing states

· Further, assume finite state and that all recurrent states are absorbing states

· Canonical form

$$P = \begin{array}{c|cc} & A & T \\ \hline A & I & 0 \\ T & R & Q \end{array} \quad , \quad S = A \cup T$$

A: The set of absorbing states
T: The set of transient states

· Here,

$$P^n = \begin{array}{c|cc} & A & T \\ \hline A & I & 0 \\ T & R_n & Q^n \end{array}$$

As $n \to \infty$, $Q^n \to 0$ and $p_{ij}^{(n)} \to 0, \forall i,j \in T$

Probability of Absorption

$$p_{ik}^{(n+1)} = \sum_{j \in S} p_{ij} p_{jk}^{(n)}, \quad i \in T, \; k \in S-T$$

Use $p_{kk} = 1$

$$p_{ik}^{(n+1)} = p_{ik} + \sum_{j \in T} p_{ij} p_{jk}^{(n)} \quad \text{— (1)} \quad (\text{element form})$$

Define

a_{ik} = P(system starting at state 'i' eventually gets absorbed into an absorbing state k)

As $n \to \infty$ in (1),

$$a_{ik} = p_{ik} + \sum_{j \in T} p_{ij} a_{jk}$$

In matrix form,

$$A = [a_{ik}], \quad A = R + QA$$

$$A = [I-Q]^{-1} R$$

$[I-Q]^{-1}$ - fundamental matrix

The probability of absorption is <u>not</u> independent of initial state i

C. Time to absorption from a transient state to an absorbing state

- Let T_i denote the number of steps, including the starting state i, in which the MC remains in a transient state before entering an absorbing state.

- It is a discrete r.v. with possible values $1, 2, 3, \ldots$

$$P(T_i = n) = P(T_i \geq n-1) - P(T_i \geq n)$$

$$= P(X_{n-1} \in T\} - Prob(X_n \in T)$$

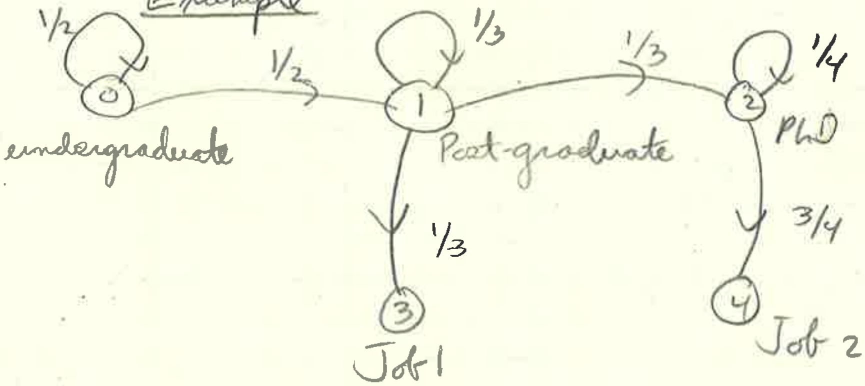
$$= \sum_{j \in T} \left(P_{ij}^{(n-1)} - P_{ij}^{(n)} \right)$$

$$\forall i, j \in T, \quad P_{ij}^{(n)} = Q^n$$

Hence, $P(T_i = n) = Q^{n-1}(I-Q)e$

Mean time : $\mu = (I-Q)^{-1} e$

Example



Canonical form

$$P = \begin{array}{cc} & \begin{array}{ccccc} 3 & 4 & 0 & 1 & 2 \end{array} \\ \begin{array}{c} A \begin{array}{c} 3 \\ 4 \end{array} \\ T \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \end{array} & \left(\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 \\ 0 & 3/4 & 0 & 0 & 1/4 \end{array} \right) \end{array}$$

Here, $I - Q = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 0 & 2/3 & -1/3 \\ 0 & 0 & 3/4 \end{pmatrix}$, $\mu = (I-Q)^{-1} = \begin{array}{cc} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 2 & 3/2 & 2/3 \\ 0 & 3/2 & 2/3 \\ 0 & 0 & 4/3 \end{pmatrix} \end{array}$

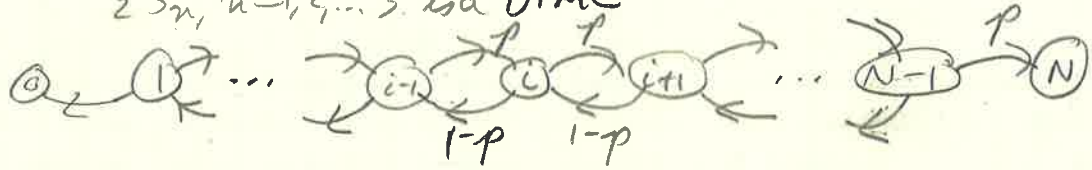
$$A = \begin{array}{cc} & \begin{array}{cc} 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \\ 0 & 1 \end{pmatrix} \end{array}$$

D. Gambler's Ruin Problem

- Consider a gambler who starts with an initial fortune of i, and then on each gamble, either wins 1 or loses 1, independent of the past with probabilities p and $q = 1-p$, respectively.

- Let S_n denote the total fortune after the n^{th} gamble. The gambler's objective is to reach a total fortune of N, without first getting ruined.

$\{S_n, n=1,2,\dots\}$ is a DTMC



Probability that the Gambler Wins

Let p_i denote the probability that the gambler wins when $S_0 = i$.

Clearly $P_0 = 0$, $P_N = 1$

for $1 \le i \le N-1$, $p_i = p P_{i+1} + q P_{i-1}$

Solving

$$P_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)^N}, & p \ne q \\ \dfrac{i}{N}, & p = q = 1/2 \end{cases}$$

$1 - P_i$ = probability of ruin.

Mean number of games

Let M_i denote the mean number of games that must be played until the gambler either goes broke or wins the complete fortune N in the game given that it starts with i.

Clearly, $M_0 = M_N = 0$,

for $1 \le i \le N-1$, $M_i = 1 + p M_{i+1} + q M_{i-1}$

Solving: $$M_i = \begin{cases} i(N-i), & p = q = 1/2 \\ \dfrac{i}{q-p} + \dfrac{N}{q-p} \cdot \dfrac{1-(q/p)^i}{1-(q/p)^N}, & p \ne q \end{cases}$$