Stochastic Processes

MATH4240 (2023/24 Term 2)

Instructor: Renjun DUAN

Course Web Page :
https://www.math.cuhk.edu.hk/course/2324/math4240

Outline I

Outline II

Chapter 0: Review on probability

§0.1: Probability

Perform an experiment:

An outcome: a particular **state** ω

Sample space: the set of all outcomes, Ω

An **event**: a subset of Ω, *e.g.*, $A \subseteq \Omega$

Examples:

1. Toss a coin.

$$\omega_1 = H, \omega_2 = T$$
$$\Omega = \{H, T\} = \{\omega_1, \omega_2\}$$
$$\text{all possible events: } A = \emptyset, \Omega, \{H\}, \{T\}$$

2. Toss 3 coins.

$$\omega_1 = (H, H, H), \omega_2 = ..., \cdots, \omega_8 = ...$$
$$\Omega = \{(H, H, H), (H, H, T), (H, T, H), (T, H, H), (H, T, T), (T, H, T), (T, T, H), (T, T, T)\}$$

Want an event $A \stackrel{\text{def}}{=}$ "exactly 2 heads occur". Then,

$$A = \{(H, H, T), (H, T, H), (T, H, H)\}$$

Probability measure P: a function that assigns real values in $[0, 1]$ to events, satisfying

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

(ii) $0 \leq P(A) \leq 1, \forall A$

(iii) $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i), \forall \{A_i\}_{i=1}^{n}$ which is disjoint
(n finite or infinite)

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

(i) $\mathcal{F}$ is an **event space**, i.e. a collection of events one is interested in, satisfying

(a) $\Omega \in \mathcal{F}$
(b) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
(c) If $A_i \in \mathcal{F}, i = 1, 2, \ldots,$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

$\mathcal{F}$ is a σ-**field** over Ω in measure theoretical term.

(ii) $P : \mathcal{F} \to [0, 1]$ is a probability measure.

Examples:

1. Given Ω, the largest σ-field is the set of all subsets of Ω.

2. Given Ω, the smallest σ-field is $\mathcal{F} = \{\Phi, \Omega\}$.

Conditional probability: A, B are two events, the probability that B happens given that A occurs is

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

Note:

- A, B are **independent** if

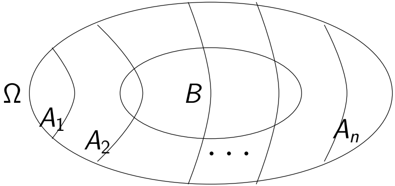
$$P(B|A) = P(B), \text{i.e. } P(A \cap B) = P(A)P(B).$$

- Let A be a fixed event,

$$P_A(\cdot) \stackrel{\text{def}}{=} P(\cdot|A)$$

is called the **conditional probability measure**.

Theorem. Let $\Omega = \bigcup_{i=1}^{n} A_i$ where $A_1, \ldots, A_n$ are disjoint events. Then, for any event B



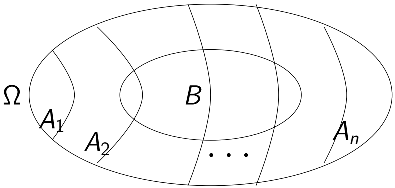
(i) $P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$

(ii) $P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$

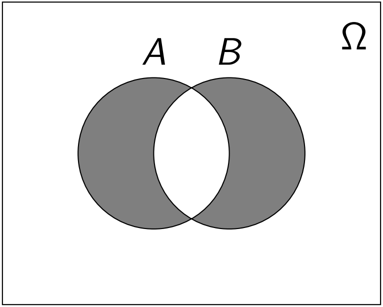
(Bayes' formula)

Note:

In many practical applications, we are given $P(B|A_i)$ and $P(A_i)$, and we want to find $P(A_i|B)$, i.e. to find the probability of the "causes" $A_i (i = 1, 2, \ldots, n)$ subject to the outcome B.



Example: $P(B) = P(B|A)P(A) + P(B|A^C)P(A^C)$
(B is caused by either A or A^c)



Proof: $B = (B \cap A) \cup (B \cap A^C)$

$$P(B) = P(B \cap A) + P(B \cap A^C)$$
$$= P(B|A)P(A) + P(B|A^C)P(A^C). \quad \square$$

One more example:

Suppose that we have 3 cards that are identical in form, except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground.

If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

Sol.: Denote

$\mathrm{RR} \overset{\text{def}}{=}$ the event that the chosen card is red-red
$\mathrm{BB} \overset{\text{def}}{=}$ the event that the chosen card is black-black
$\mathrm{RB} \overset{\text{def}}{=}$ the event that the chosen card is red-black
$\mathrm{R} \overset{\text{def}}{=}$ the event that the upper side of the chosen card is red

Then

$$
\begin{aligned}
P(\mathrm{RB}|\mathrm{R}) &= \frac{P(\mathrm{RB} \cap \mathrm{R})}{P(\mathrm{R})} \\
&= \frac{P(\mathrm{R}|\mathrm{RB})P(\mathrm{RB})}{P(\mathrm{R}|\mathrm{RR})P(\mathrm{RR}) + P(\mathrm{R}|\mathrm{BB})P(\mathrm{BB}) + P(\mathrm{R}|\mathrm{RB})P(\mathrm{RB})} \\
&= \frac{\frac{1}{2} \cdot \frac{1}{3}}{1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}} \\
&= \frac{1}{3}.
\end{aligned}
$$

□

§0.2 Random variables and distributions

Example: Toss a coin n-times.

$$\Omega = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) : \omega_i = H \text{ or } T\}$$

$\sharp$ of $\Omega = 2^n$

$$P(\{\omega\}) = \frac{1}{2^n}$$

Let X denote the number of heads,

then X takes values in $\{0, 1, 2, \ldots, n\}$,

Let $k = 0, 1, \ldots, n,$ then $X = k$ means

the event that we get k number of heads,

$$P(X = k) = \frac{\binom{n}{k}}{2^n}.$$

Random variable: A random variable (r.v.) X on $(\Omega, \mathcal{F}, P)$ is to assign an outcome with a real number

$$X : \Omega \to \mathbb{R}$$

$$\Omega \ni \omega \mapsto X(\omega) \in \mathbb{R}$$

Note: Let $R_X =$ the set of all possible values of X on Ω, then R_X is either "discrete" or "continuous":

Case 1: R_x is a discrete set. In this case, X is called a discrete r.v.

Case 2: R_x is an interval of $\mathbb{R}$ or itself. In this case, X is called a continuous r.v.

Discrete random variable: Assume

$$X(\Omega) = \{k\}_{k=0}^{N} \qquad (N \text{ finite or infinite})$$

Then the values

$$p_k = P(X = k), (k = 0, 1, \dots, N)$$

is called the **probability density function** (p.d.f.).

Note: $\{X = k\} \stackrel{\text{def}}{=} \{\omega \in \Omega : X(\omega) = k\} \in \mathcal{F}$

Examples (Important!)

1. Binomial distribution

We perform n independent trials. At each trial, the prob of **success** is p, and the prob of **failure** is $1 - p$.

Let X denote the *number of successes in n trials*. X has the p.d.f.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, 0 \leq k \leq n.$$
$$\stackrel{\text{def}}{=} B(n, p)$$

2. Poisson distribution:

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

For instance, X counts the number of arrivals in a unit time with rate of arrivals given by $\lambda > 0$.

Theorem: For each $k = 0, 1, \cdots$

$$\lim_{n\to\infty, np=\lambda} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

Note: Therefore, the Poisson distribution can be used to approximate the Binomial distribution when p is small and n is large compared to k.

3. Geometric distribution:

$$P(X = k) = p(1-p)^{k-1}, k = 1, 2, \ldots$$

is the prob that the first occurrence of success requires k independent trials, each with success probability p.

X denotes the number of trials for the first success.

Continuous random variable:

If

$$P(a \leq X \leq b) = \int_a^b f(x)dx,$$

then f is called a **density function** of X.

Note:

the event "$a \leq X \leq b$" $\stackrel{\text{def}}{=} \{\omega \in \Omega : a \leq X(\omega) \leq b\}$

Examples (Important!)

1. Uniform distribution:

$$f(t) = \begin{cases} \frac{1}{b-a}, & a \leq t \leq b \\ 0, & \text{otherwise} \end{cases}$$

2. Exponential distribution:

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

3. Normal distribution:

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \stackrel{\text{def}}{=} N(\mu, \sigma^2)$$

$N(0, 1)$: standard normal density.

Exercise: Assume X, Y are two independent continuous (or discrete) r.v. with densities f, g (or $(p_k), (q_k)$).

Find the density function for the random variable $Z = X + Y$.

§0.3 Expectation and variance

The **expectation** (or **mean**) of X:

$$\mu = E(X) = \sum_k kp_k \text{ or } \int_{-\infty}^{\infty} tf(t)dt$$

The 2^{nd} **moment** of X:

$$E(X^2) = \sum_k k^2 p_k, \text{ or } \int_{-\infty}^{\infty} t^2 f(t)dt$$

The **variance** of X:

$$\sigma^2 \stackrel{\text{def}}{=} Var(X) = E(X-\mu)^2 = E(X^2) - \mu^2$$

(a measurement of how spread the distribution is)

Conditional expectation:

Discrete case: Suppose (X, Y) has a joint density

$$p(x_i, y_i) \stackrel{\text{def}}{=} P(X = x_i, Y = y_i)$$

$$E(Y|X = x_i) = \sum_j y_j P(Y = y_j|X = x_i)$$

$$= \sum_j y_j \frac{p(x_i, y_j)}{p(x_i)}, \quad p(x_i) = \sum_j p(x_i, y_j)$$

Note:

a. $P(Y = y_j|X = x_i)$ is the conditioned density function of Y given $X = x_i$.

b. $E(Y|X = x_i)$ is a function of x_i, and thus regarded as a r.v. on the σ-field generated by X, denoted by $E(Y|X)$.

Continuous case: Let $f(x,y)$ be such that

$$P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, dudv.$$

Then,

$$E(Y|X=x) = \int_{R_Y} y \frac{f(x,y)}{f(x)} dy,$$

$$f(x) = \int_{R_Y} f(x,y) dy.$$

Here $E(Y|X)$ can be understood to be a r.v. on the σ-field generated by X.

§0.4 Sequence of random variables

Repeat a random experiment independently. We obtain a sequence of random variables which are **independent and identically distributed** (i.i.d)

$$\{X_n\}_{n=0}^{\infty}.$$

Two basic theorems are:

- Law of Large Number
- Central Limit Theorem

(<u>Ref</u>: Ross p.389)

However, in many cases $\{X_n\}_{n=0}^{\infty}$ **may not be independent**. There exists dependence in a certain way.

In general, we call

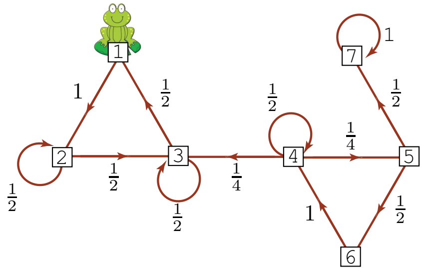
- $\{X_n\}_{n=0}^{\infty}$ a discrete time stochastic process, and
- $\{X_t\}_{t\geq 0}$ a continuous time stochastic process.

We will mainly consider the

"**Markov**" processes

(to be defined) in the discrete time and continuous time.

Example 1.1. A frog hops about on 7 lily pads. The numbers next to arrows show the probabilities with which, at the next jump, he jumps to a neighbouring lily pad (and when out-going probabilities sum to less than 1 he stays where he is with the remaining probability).



$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

There are 7 'states' (lily pads). In the matrix P the element P_{57} $(= 1/2)$ is the prob that, when starting in state 5, the next jump takes the frog t state 7.

Some questions we may want to know:

1. Starting in state 1, what is the prob that we are still in state 1 after 3 steps? after 5 steps? or after 1000 steps?
2. Starting in state 4, what is the prob that we ever reach state 7?
3. Starting in state 4, how long on average does it take to reach either 3 or 7?
4. Starting in state 2, what is the long-run proportion of time spent in state 3?

We can answer those by the end of this course

——End of Chapter 0——

Chapter 1:
Markov Chain

§1.1: Definition & Examples

Example:

- Consider the weather (0=Sunny, 1=Rainy, 2=Cloundy) of days in Hong Kong.
- Let X_0 be a r.v. describing the weather of the 0th day, then
$$X_0 = 0,\ 1,\ \text{or}\ 2,$$
i.e. X_0 takes values in
$$S := \{0, 1, 2\}.$$
- Similarly, for $n \geq 0$ let X_n be a r.v. describing the weather of the nth day, then $X_n = 0$, 1, or 2, i.e. X_n takes values in the same state space S.
- In the end we get a chain $\{X_n\}_{n\geq 0}$.

Definitions:

- Let S be a finite or countably infinite set of integers.

  For instance, $S = \{0, 1, 2, \dots, N\}$, or $S = \{0, 1, 2, \dots\}$, or $S = \{\dots, -1, 0, 1, \dots\}$.

  We call each element of S a **state** and S the **state space**.

- Let $\{X_n\}_{n=0}^{\infty}$ be a sequence of r.v. taking values in S, defined on a *common* probability space $(\Omega, \mathcal{F}, P)$.

Notation for the future:

• For random variables, we use

$$X, Y, Z, \cdots$$

• For states (which are values of random variables), we use

$$x, y, z, \cdots \in S$$

or

$$x_i, y_i, z_i, \cdots \in S,$$

or

$$i, j, k, \cdots \in S.$$

Def: $\{X_n\}_{n=0}^{\infty}$ is a **Markov chain** if

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n). \qquad (*)$$

Note:

- $(*)$ is called the **Markov property** which says that given the *present* state, the *past* states have no influence on the *future*!

- $P(X_{n+1} = y | X_n = x)$ is called the **transition probability**. If it is independent of n, we denote

$$P(x, y) = P(X_{n+1} = y | X_n = x) = P(X_1 = y | X_0 = x)$$

which is the transition probability from state x to state y. In such case, $\{X_n\}_{n=0}^{\infty}$ is called a *time-homogeneous* Markov chain.

It is clear that

(i) $P(x,y) \geq 0$.

(ii) $\sum_{y \in S} P(x,y) = 1$.

Proof:

(i) $P(x,y) = P(X_{n+1} = y | X_n = x) \geq 0$.

(ii) $\sum_{y \in S} P(x,y) = \sum_{y \in S} P(X_{n+1} = y | X_n = x) = 1$.

e.g. for $S = \{0, 1, 2, \dots, N\}$ (N finite or ∞), we may express all the transition probabilities

$$P(x, y), \quad x, y \in S$$

as a *matrix form*:

$$P = [P(x, y)] \text{ (or } [P(i, j)])$$
$$= \begin{bmatrix} P(0,0) & P(0,1) & \cdots & P(0,N) \\ P(1,0) & P(1,1) & \cdots & P(1,N) \\ \vdots & \vdots & \ddots & \vdots \\ P(N,0) & P(N,1) & \cdots & P(N,N) \end{bmatrix}$$

which is called the **Markov matrix** (or transition matrix) (Note: each row vector is a probability vector).

Example 1. Toss a possibly biased coin repeatedly with prob p for H and $1 - p$ for T.

Q.: Set up the model as a Markov chain.

Example 2. Consider a machine that at the start of the day is broken down or in operation. Assume

(i) if it is broken down, the prob that it will be repaired and in operation on the next day is p, $(0 < p < 1)$.

(ii) if it is in operation, the prob that it will be broken down on the next day is q, $(0 < q < 1)$.

Q: Set up the model as a Markov chain. Further,

(a) Find the transition prob.

(b) Find the prob that the machine is broken down on the n^{th} day.

(c) In the long term, what is the prob that the machine is broken down on a day.

Example 3 (Random walk):

Let $\{\xi_i\}_{i=1}^{\infty}$ be i.i.d. r.v. taking values in

$$S = \{\cdots, -1, 0, 1, \cdots\}$$

and having a pdf f, i.e. for each i

$$P(\xi_i = k) = f(k), \quad k = 0, \pm 1, \pm 2, \cdots$$

Let $X_n = X_0 + \xi_1 + \cdots + \xi_n$, where X_0 is the initial position independent of $\{\xi_i\}_{i=1}^{\infty}$. Then,

$$\begin{aligned} P(x, y) &= P(X_{n+1} = y | X_n = x) \\ &= P(\xi_{n+1} = y - x | X_n = x) \\ &= P(\xi_{n+1} = y - x) \\ &= f(y - x). \end{aligned}$$

A simple random walk:

Consider a move to left or right with prob $p, 1 - p$ resp, i.e. $\xi_i = +1$ or -1 with prob $p, 1 - p$ resp.

How does the chain behave as $n \to \infty$?

Example 4 (Gambler's ruin chain)

A gambler starts out with a certain amount and bets against the house.

(i) Each time he wins or loses \$1 with prob p and $q = 1 - p$ resp.

(ii) If he reaches \$0, he is ruined and his amount remain \$0. (he quits playing)

Q.: Set up the model as a Markov chain.

Let X_n denote the amount he has at the n-th stage.
$S = \{0, 1, 2, \cdots\}$.

- For $x = 0$,

$P(0, 0) = 1$,
$P(0, y) = 0, \; y = 1, 2, \cdots$

> **<u>Def:</u>** A state $a \in S$ is absorbing if $P(a, a) = 1$, i.e. $P(a, y) = 0, \forall\, y \neq a$.

$\therefore$ 0 is an absorbing state.

- For $x > 0$,

$$P(x, y) = \begin{cases} p & y = x + 1 \\ 1 - p & y = x - 1 \\ 0 & otherwise \end{cases}$$

$$P = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots \\ 1-p & 0 & p & \cdots & \cdots \\ 0 & 1-p & 0 & p & \cdots \\ \cdots & \cdots & \ddots & \ddots & \ddots \end{bmatrix}$$

A modification of the model: Add a rule

(iii) If he reaches \$N, he quits playing.

Then,

$S = \{0, 1, \cdots, N\}$.

0 and N are absorbing,

$$P(x,y) = \begin{cases} p & y = x+1 \\ 1-p & y = x-1 \\ 0 & otherwise \end{cases} \quad \text{for } 1 \le x \le N-1.$$

$$P = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots & \\ 1-p & 0 & p & \cdots & \cdots & \cdots & \\ 0 & 1-p & 0 & p & \cdots & \cdots & \\ \cdots & \cdots & \ddots & \ddots & \ddots & \cdots & \\ \cdots & \cdots & \cdots & 1-p & 0 & p & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix}$$

Alternative view to the above modified "gambler's ruin chain": Two gamblers start a series of \$1 bets against each other.

(i) The total amount is $\$N$.

(ii) $p =$ prob of the 1^{st} gambler wining
$q = 1 - p =$ prob of the 2^{nd} gambler winning.

(iii) The game is over when one of them losses all.

$X_n \stackrel{\text{def}}{=}$ \$ of the 1^{st} gambler at the n^{th} stage

Q:

- What is the expected value?
- Wo has higher prob of winning?
- How long does the game last?

Remark: The more general form of the chains in examples 3 & 4:

$$P(x,y) = \begin{cases} p_x & y = x - 1 \\ q_x & y = x + 1 \\ r_x & y = x \\ 0 & otherwise \end{cases}$$

which corresponds to the "**birth** & **death**" chain. Here

$$p_x, q_x, r_x \geq 0,$$
$$p_x + q_x + r_x = 1.$$

Example 5 (Queueing chain)

Consider a check out counter at a supermarket.

(i) Let ξ_n denote the number of arrivals in the n^{th} period (say, one minute). Then $\{\xi_n\}_{n=1}^{\infty}$ is i.i.d. r.v. having pdf f (usually Poisson distribution).

(ii) Suppose that if there are any customers waiting for service at the beginning of any given period, then exactly one customer will be served during that period.

Q.: Set up the model as a Markov chain.

- $n = 0$:

$X_0 \overset{\textbf{def}}{=}$ the number of persons on the line initially.

- $n \geq 1$:

$X_n \overset{\textbf{def}}{=}$ the number of persons on the line present at the end of the n^{th} period.

- Then,

$$X_{n+1} = \begin{cases} 0 + \xi_{n+1} & \textit{if} \quad X_n = 0 \\ X_n + \xi_{n+1} - 1 & \textit{if} \quad X_n \geq 1 \end{cases}$$

- For $x = 0$,

$$\begin{aligned} P(0, y) &= P(X_{n+1} = y | X_n = 0) \\ &= P(\xi_{n+1} = y | X_n = 0) \\ &= P(\xi_{n+1} = y) \\ &= f(y). \end{aligned}$$

- For $x \geq 1$,

$$\begin{aligned} P(x, y) &= P(X_{n+1} = y | X_n = x) \\ &= P(\xi_{n+1} = y - x + 1 | X_n = x) \\ &= P(\xi_{n+1} = y - x + 1) \\ &= f(y - x + 1). \end{aligned}$$

For instance, f is Poisson:

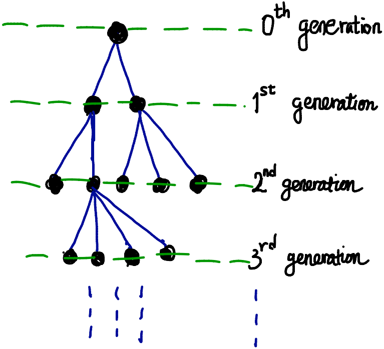
$$f(k) = P(\xi_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \cdots$$

Then

$$P = e^{-\lambda}\begin{bmatrix} 1 & \lambda & \frac{\lambda^2}{2!} & \frac{\lambda^3}{3!} & \cdots \\ 1 & \lambda & \frac{\lambda^2}{2!} & \frac{\lambda^3}{3!} & \cdots \\ 0 & 1 & \lambda & \frac{\lambda^2}{2!} & \cdots \\ \cdots & \ddots & \ddots & \ddots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Example 6 (Branching chain, population growth)

Each individual generates ξ offspring in the next generation independently.



$X_n \stackrel{\text{def}}{=}$ the total NO in the n^{th} generation.

$$P(x, y) = P(\xi_1 + \xi_2 + \cdots + \xi_x = y)$$

Question concerns the extinction or growth of the population!

§1.2 Computations with transition probabilities

Setup:

- $\{X_n\}_{n=0}^{\infty}$: a time-homogeneous Markov chain
- $S = \{0, 1, 2, \ldots, N\}$: state space
  (N: finite or ∞)
- $P = [P(x, y)] = [P(X_{n+1} = y | X_n = x)]$:
  transition prob matrix

Question 1: Given pdf of X_0, can one compute pdf of X_n for any $n \geq 1$?

Let the pdf of X_0 be

$$\pi_k^{(0)} \stackrel{\text{def}}{=} P(X_0 = k), \quad k = 0, 1, \cdots, N,$$

or equivalently we write in the prob row-vector form

$$\pi^{(0)} = [\pi_0^{(0)}, \pi_1^{(0)}, \cdots, \pi_N^{(0)}].$$

- $n = 1$: $P(X_1 = k), k \in S$? or $\pi^{(1)} = [\pi_0^{(1)}, \cdots, \pi_N^{(1)}]$?

$$
\begin{aligned}
P(X_1 = k) &= \sum_{i \in S} P(X_1 = k, X_0 = i) \\
&= \sum_{i \in S} P(X_1 = k | X_0 = i) P(X_0 = i) \\
&= [P(X_0 = 0), P(X_0 = 1), \ldots, P(X_0 = N)] \begin{bmatrix} P(0, k) \\ P(1, k) \\ \vdots \\ P(N, k) \end{bmatrix}
\end{aligned}
$$

Write them for $k = 0, 1, \cdots, N$ in matrix:

$$
\begin{aligned}
&[P(X_1 = 0), P(X_1 = 1), \ldots, P(X_1 = N)] \\
=&[P(X_0 = 0), P(X_0 = 1), \ldots, P(X_0 = N)] \begin{bmatrix} P(0,0) & P(0,1) & \cdots & P(0,N) \\ P(1,0) & P(1,1) & \cdots & P(1,N) \\ \vdots & \vdots & \ddots & \vdots \\ P(N,0) & P(N,1) & \cdots & P(N,N) \end{bmatrix}
\end{aligned}
$$

i.e.

$$
\boxed{\pi^{(1)} = \pi^{(0)} P}
$$

• In general, for $n \geq 1$, setting the pdf of X_n as a probability row-vector in the form

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

Then,

$$\boxed{\pi^{(n)} = \pi^{(n-1)} P}.$$

- Then, by iteration,

$$\begin{aligned}\pi^{(n)} &= \pi^{(n-1)}P \\ &= \pi^{(n-2)}P \cdot P \\ &= \cdots \\ &= \pi^{(0)}\underbrace{P \cdot P \cdot \ldots \cdot P}_{\text{n terms}} \\ &= \pi^{(0)}P^n\end{aligned}$$

where

$$P^n = \underbrace{P \cdot P \cdot \ldots \cdot P}_{\text{product of n terms}}$$

Theorem: $\pi^{(n)} = \pi^{(0)}P^n$, $n = 1, 2, \cdots$

Remark: How to compute the matrix product

$$P^n := \underbrace{P \cdot P \cdot \cdots \cdot P}_{n \text{ terms}}, \quad n = 2, 3, \cdots$$

Indeed, for $x, y \in S$,

$$P^2(x,y) = \sum_{x_1 \in S} P(x, x_1) P(x_1, y)$$

$$P^3(x,y) = \sum_{x_1} \sum_{x_2} P(x, x_1) P(x_1, x_2) P(x_2, y)$$

$$\cdots$$

$$P^n(x,y) = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_{n-1}} P(x, x_1) P(x_1, x_2) \cdots P(x_{n-1}, y).$$

Proof: Left for an exercise. Argument: use induction in n and the formula $P^n = P^{n-1} \cdot P$.

Proposition:

(i) $P(X_n = y) = \sum_x \pi^{(0)}(x) P^n(x, y)$.

(ii) $P(X_n = y | X_0 = x) = P^n(x, y)$.

Proof: (i) is a direct consequence of the formula $\pi^{(n)} = \pi^{(0)} P^n$.
To show (ii),

$$
\begin{aligned}
& P(X_n = y | X_0 = x) \\
&= P(X_n = y, X_{n-1} \in S, \cdots, X_1 \in S | X_0 = x) \\
&= \sum_{x_1 \in S} \cdots \sum_{x_{n-1} \in S} P(X_n = y, X_{n-1} = x_{n-1}, \cdots, X_1 = x_1 | X_0 = x) (\text{tutorial}) \\
&= \sum_{x_1, \cdots, x_{n-1} \in S} \frac{P(X_n = y, X_{n-1} = x_{n-1}, \cdots, X_1 = x_1, X_0 = x)}{P(X_0 = x)} \\
&= \sum_{x_1, \cdots, x_{n-1} \in S} \frac{P(X_0 = x) P(x_0, x_1) \cdots P(x_{n-1}, y)}{P(X_0 = x)} (\text{proof later}) \\
&= \sum_{x_1, \cdots, x_{n-1} \in S} P(x, x_1) \cdots P(x_{n-1}, y) = P^n(x, y). \quad \square
\end{aligned}
$$

Claim:

$$P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) \\ = P(X_0 = x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n).$$

Proof of claim:

$$P(\underbrace{X_0 = x_0, X_1 = x_1, \cdots}_{A}, \underbrace{X_n = x_n}_{B})$$

$$= P(X_n = x_n | X_0 = x_0, \cdots, X_{n-1} = x_{n-1})$$

$$\cdot P(X_0 = x_0, \cdots, X_{n-1} = x_{n-1})$$

$$= P(X_n = x_n | X_{n-1} = x_{n-1}) P(X_0 = x_0, X_1 = x_1, \cdots, X_{n-1} = x_{n-1})$$

$$= P(x_{n-1}, x_n) P(X_0 = x_0, X_1 = x_1, \cdots, X_{n-1} = x_{n-1})$$

$$= \cdots$$

$$= P(x_{n-1}, x_n) P(x_{n-2}, x_{n-1}) \cdots P(x_0, x_1) P(X_0 = x_0). \quad \square$$

Remark: (ii) also immediately implies (i). In fact,

$$P(X_n = y) \overset{\text{by def}}{=} \sum_x P(X_n = y | X_0 = x) P(X_0 = x)$$

$$\overset{\text{by (ii)}}{=} \sum_x P^n(x, y) \pi^{(0)}(x).$$

Definition:

$$P^m(x, y), (m = 0, 1, \cdots)$$

is called the ***m*-step transition function**, which gives the prob of going from state x to state y in m steps. Here we set

$$P^0(x, y) = \delta_{xy} = \begin{cases} 1, & \textit{if} \quad x = y \\ 0, & \textit{otherwise}. \end{cases}$$

Correspondingly, P^m is called the ***m*-step transition matrix**.

Proposition:

$$P(X_{n+m} = y | X_n = x) = P^m(x, y).$$

Proof

$$\begin{aligned}
&P(X_{n+m} = y | X_n = x) \\
&= P(X_{n+m} = y, X_{n+m-1} \in S, \cdots, X_{n+1} \in S | X_n = x) \\
&= \sum_{x_{n+m-1}} \cdots \sum_{x_{n+1}} P(X_{n+m} = y, X_{n+m-1} = x_{n+m-1}, \cdots, X_{n+1} = x_{n+1} | X_n = x) \\
&= \sum_{x_{n+1}, \cdots, x_{n+m-1}} P(X_{n+m} = y, \cdots, X_{n+1} = x_{n+1} | X_0 \in S, \cdots, X_{n-1} \in S, X_n = x) \ (*) \\
&= \sum_{x_{n+1}, \cdots, x_{n+m-1}} P(x, x_{n+1}) \cdots P(x_{n+m-1}, y) \quad \text{(see below)} \\
&= P^m(x, y). \quad \text{(by def of } P^m\text{)}
\end{aligned}$$

Each term in the sum $(*)$ is equal to

$$\begin{aligned}
&P(X_{n+m} = y, \cdots, X_{n+1} = x_{n+1} | X_0 = x_0, \cdots, X_{n-1} = x_{n-1}, X_n = x) \\
&= P(X_{n+m} = y, X_{n+m-1}, \cdots, X_0 = x_0) / P(X_n = x, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0) \\
&= \frac{P(X_0 = x_0) P(x_0, x_1) \cdots P(x, x_{n+1}) P(x_{n+1}, x_{n+2}) \cdots P(X_{n+m-1}, y)}{P(X_0 = x_0) P(x_0, x_1) \cdots P(x_{n-1}, x)} \\
&= P(x, x_{n+1}) \cdots P(x_{n+m-1}, y). \quad \square
\end{aligned}$$

Remark: To compute

$$\begin{aligned} & P(X_n = y|X_0 = x) \\ & = P(X_{1+n} = y|X_1 = x) \\ & = P(X_{2+n} = y|X_2 = x) \\ & = \cdots \\ & = P(X_{m+n} = y|X_m = x), \quad m = 0, 1, 2, \cdots \end{aligned}$$

is equivalent to compute $P^n(x, y)$, that is to find P^n.

For n large, one can reduce P to a diagonal matrix (if possible)

$$P = QDQ^{-1}$$

where $D = \begin{bmatrix} \lambda_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_N \end{bmatrix}$, and Q is the matrix for change basis consisting of eigenvectors. Then

$$P^n = \left[QDQ^{-1}\right]^n = QD^nQ^{-1} = Q \begin{bmatrix} \lambda_0^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_N^n \end{bmatrix} Q^{-1}.$$

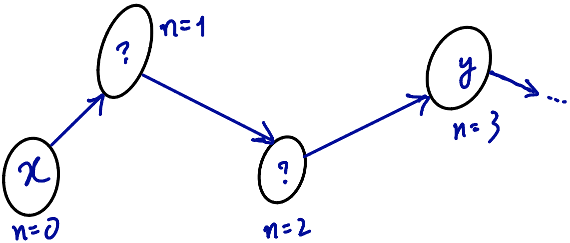
Hence P^n can be calculated in such situation.

Exercise: Do this for the two-state Markov matrix.

Question 2: How to compute the conditional prob that the chain visits y in **finite** time given that it starts from x?

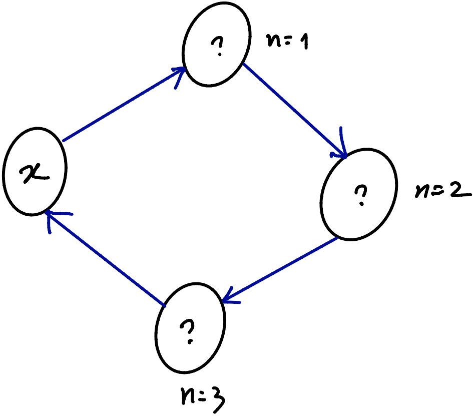
We set it as ρ_{xy}, then

$$\rho_{xy} = P(\exists\, n \geq 1 \text{ such that } X_n = y | X_0 = x).$$



We are interested in a state x such that

$$\rho_{xx} = 1, \text{ or } \rho_{xx} < 1.$$



Def (Hitting Time):
Let $A \subseteq S$. The **hitting time** T_A of A is defined by

$$T_A \stackrel{\text{def}}{=} min\{n \geq 1 : X_n \in A\}.$$

Rks:

- $T_A =$ the first positive time the chain hits A.
  T_A is a r.v. Range of $T_A = \{1, 2, 3, \cdots\} \cup \{\infty\}$.
  Convention: $T_A = \infty$ if $X_n \notin A$ for all $n \geq 1$.
  For $m = 1, 2, \cdots$

$$\{T_A = m\} = \{X_1 \notin A, \cdots, X_{m-1} \notin A, X_m \in A\}.$$

- Convention:

$$T_y \stackrel{\text{def}}{=} T_{\{y\}} = min\{n \geq 1 : X_n = y\}, \quad y \in S,$$

  i.e. the first positive time the chain visits y.

A convenient notion:

$$P_x(\cdot) \stackrel{\text{def}}{=} P(\,\cdot\,|X_0 = x)$$

i.e. the probabilities of various events defined in terms of the Markov chain starting at $x \in S$.

For instance,

$$P_x(A) = P(A|X_0 = x)$$

is the prob of A given that the chain starts at x.

Prop (i) $P_x(T_y = 1) = P(x, y)$.

Proof:

$\because \{T_y = 1\} = \{X_1 = y\}$

$\therefore$

$$
\begin{aligned}
&P_x(T_y = 1) \\
&= P(X_1 = y | X_0 = x) \\
&= P(x, y). \quad \square
\end{aligned}
$$

Prop (ii)

$$P_x(T_y = n+1) = \sum_{z \neq y} P(x,z) P_z(T_y = n), n \geq 1.$$

Proof: Note

$$\{T_y = n+1\} = \bigcup_{z: z \neq y} \{X_1 = z, X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y\}.$$

$\therefore$

$$\begin{aligned} & P_x(T_y = n+1) \\ &= \sum_{z \neq y} P_x(X_1 = z, X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y) \\ &= \sum_{z \neq y} P_x(X_1 = z) P_x(X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y | X_1 = z) \\ &= \sum_{z \neq y} \underbrace{P(X_1 = z | X_0 = x)}_{=P(x,z)} \underbrace{P(X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y | X_0 = x, X_1 = z)}_{=P_z(T_y = n) \text{ why?}} \end{aligned}$$

Recall an **Exercise:**

$$P(X_{n+1} \in B_1, \cdots, X_{n+m} \in B_m | X_0 \in A_0, \cdots, X_{n-1} \in A_{n-1}, X_n = x)$$
$$= P_x(X_1 \in B_1, X_2 \in B_2, \cdots, X_m \in B_m).$$

See the tutorial for the proof.

Rk: It is essentially due to the Markovian property (i.e., given "the present" state, "the past" has no influence on "the future"). So, the prob on the LHS is understood to be the prob in the situation when the chain initially starts at x.

$$\therefore \; P(X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y | X_0 = x, X_1 = z)$$
$$= P_z(X_1 \neq y, X_2 \neq y, \cdots, X_{n-1} \neq y, X_n = y)$$
$$= P_z(T_y = n). \quad \square$$

Prop (iii)

$$P^n(x,y) = \sum_{m=1}^{n} P_x(T_y = m)P^{n-m}(y,y).$$

Proof: Note: $P^n(x,y) = P(X_n = y | X_0 = x) = P_x(X_n = y)$

$\{X_n = y\} = \bigcup_{m=1}^{n} \{T_y = m, X_n = y\}$ (disjoint union)

$$\begin{aligned}
\therefore P^n(x,y) &= P_x(X_n = y) \\
&= \sum_{m=1}^{n} P_x(T_y = m, X_n = y) \\
&= \sum_{m=1}^{n} P_x(T_y = m) P_x(X_n = y | T_y = m) \\
&= \sum_{m=1}^{n} P_x(T_y = m) P(X_n = y | X_0 = x, X_1 \neq y, \cdots, X_{m-1} \neq y, X_m = y) \\
&= \sum_{m=1}^{n} P_x(T_y = m) P^{n-m}(y,y). \quad \square
\end{aligned}$$

Sum:

Proposition:

(i) $P_x(T_y = 1) = P(x, y)$.

(ii) $P_x(T_y = n+1) = \sum_{z \neq y} P(x, z) P_z(T_y = n), \quad n \geq 1$.

(iii) $P^n(x, y) = \sum_{m=1}^{n} P_x(T_y = m) P^{n-m}(y, y)$.

Corollary: If $a \in S$ is absorbing , i.e. $P(a,a) = 1$, then for any $n \geq 1$, $P^n(x,a) = P_x(T_a \leq n)$.

Proof:

$$P^n(x,a) = \sum_{m=1}^{n} P_x(T_a = m) \underbrace{P^{n-m}(a,a)}_{=1 \text{ (to be shown later)}}$$

$$= \sum_{m=1}^{n} P_x(T_a = m)$$

$$= P_x(\cup_{m=1}^{n} \{T_a = m\})$$

$$= P_x(T_a \leq n). \quad \square$$

It remains to show: For any $n \geq 0$, $P^n(a,a) = 1$.
Indeed:

- $n = 0, 1$ is obvious.
- $n \geq 2$:

$$\begin{aligned} P^n(a,a) &= \sum_{x_1,\cdots,x_{n-1}} P(a,x_1)P(x_1,x_2)\cdots P(x_{n-1},a) \\ &= \sum_{x_2,\cdots,x_{n-1}} P(a,x_2)\cdots P(x_{n-1},a) \\ &= \cdots \\ &= \sum_{x_{n-1}} P(a,x_{n-1})P(x_{n-1},a) \\ &= P(a,a) \\ &= 1. \quad \square \end{aligned}$$

Recall: $\rho_{xy} = P_x(T_y < \infty)$ is the prob that the chain starting at x will visit y at some positive time.

In particular,

$$\rho_{yy} = P_y(T_y < \infty)$$

is the prob that the chain starting at y will ever return to y.

Def.:

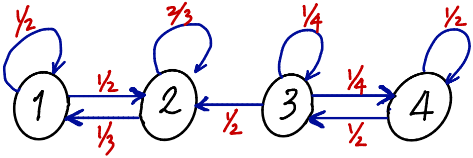
- A state y is called **recurrent** if $\rho_{yy} = 1$, and **transient** if $\rho_{yy} < 1$.
- A chain is called a recurrent (transient) chain if **all** states are recurrent (transient).

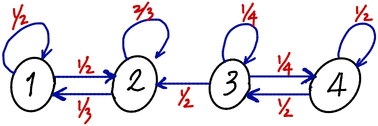
Rk: An absorbing state is recurrent.

Example:

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array}.$$

Q: Find the matrix $[\rho_{xy}]$ from $P = [P(x, y)]$.





Observe:

(i) $0 = \rho_{13} = \rho_{14}, \quad 0 = \rho_{23} = \rho_{24}.$

(ii) $1 = \rho_{11} = \rho_{22}, \quad \therefore$ 1, 2 are recurrent.

(iii) $\rho_{33} < 1, \rho_{44} < 1, \quad \therefore$ 3, 4 are transient.

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & * & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \end{array}.$$

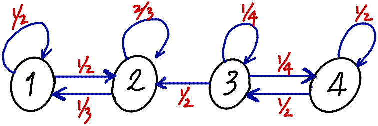
Recalling $\rho_{xy} = P_x(T_y < \infty)$, we have

$$\rho_{xy} = P(x,y) + \sum_{z:z\neq y} P(x,z)\rho_{zy}$$

(Exercise)

Argument: Start at x.

- If $T_y = 1$, i.e. visit y at $n = 1$, prob is $P(x,y)$.
- If it does not visit y at $n = 1$, then it will first visit $z(z \neq y)$ and then start from such z to visit y at some positive time.



$$\begin{cases} \rho_{33} = \cancelto{0}{0 \cdot \rho_{13}} + \cancelto{0}{\frac{1}{2} \cdot \rho_{23}} + \frac{1}{4} + \frac{1}{4} \cdot \rho_{43} \\ \rho_{43} = \cancelto{0}{0 \cdot \rho_{13}} + \cancelto{0}{0 \cdot \rho_{23}} + \frac{1}{2} + \frac{1}{2} \cdot \rho_{43} \end{cases}$$

$$\therefore \rho_{43} = 1, \quad \rho_{33} = \frac{1}{2}$$

Similarly,

$$\rho_{34} = \cancelto{0}{0 \cdot \rho_{14}} + \cancelto{0}{\frac{1}{2} \cdot \rho_{24}} + \frac{1}{4} \cdot \rho_{34} + \frac{1}{4}, \quad \therefore \rho_{34} = \frac{1}{3},$$

$$\rho_{44} = \cancelto{0}{0 \cdot \rho_{14}} + \cancelto{0}{0 \cdot \rho_{24}} + \frac{1}{2} \cdot \rho_{34} + \frac{1}{2} \quad \therefore \rho_{44} = \frac{2}{3}.$$

$$
[\rho_{ij}] = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & \frac{1}{3} \\ 1 & 1 & 1 & \frac{2}{3} \end{bmatrix} \end{array}.
$$

Note: There is a matrix argument for finding $[\rho_{xy}]$.
See *Lawler p.23-27*.

Question 3. Times of visit to a state.

$\{X_n\}_{n=0}^{\infty}$: a time-homogeneous Markov chain

$S = \{0, \cdots, N\}$ (N: finite or ∞): state space

$X_0 = x \in S$

$$\boxed{N(y) \overset{\text{def}}{=} \text{no of times that } X_n (n \geq 1) \text{ visits } y.}$$

Note:

- $N(y) = \sum_{n=1}^{\infty} 1_y(X_n)$, where $1_y(X_n) = \begin{cases} 1, & X_n = y \\ 0, & X_n \neq y \end{cases}$
- $N(y) \in \{0, 1, 2, 3, \cdots\} \cup \{\infty\}$.

$$\{N(y) = 0\} = \text{"}y \text{ is not visited"}$$

$$\{N(y) = k\} = \text{"}y \text{ is visited exactly } k \text{ times"}$$

$$\{N(y) = \infty\} = \text{"}y \text{ is visited infinitely times"}$$

Some **Facts**:

- $$\underbrace{\{N(y) \geq 1\}}_{\text{"}y\text{ is visited at least one time"}} = \underbrace{\{T_y < \infty\}}_{\text{"}y\text{ is visited at a positive finite time"}}.$$

$$\therefore \boxed{P_x(N(y) \geq 1) = P_x(T_y < \infty) = \rho_{xy}}.$$

- $\{N(y) = 0\} = \{N(y) \geq 1\}^c$.

$$\therefore \boxed{P_x(N(y) = 0) = 1 - \rho_{xy}}.$$

Claim: For $m \geq 1$, $P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$.

Case $m = 2$. To show: $P_x(N(y) \geq 2) = \rho_{xy}\rho_{yy}$.
Note:

$$\{N(y) \geq 2\} = \cup_{k\geq 1} \cup_{n\geq 1} \{\text{chain starting at } x \text{ first visits } y \text{ at } k \geq 1 \text{ and next visit } y \text{ again after } n \text{ units of time}\}.$$

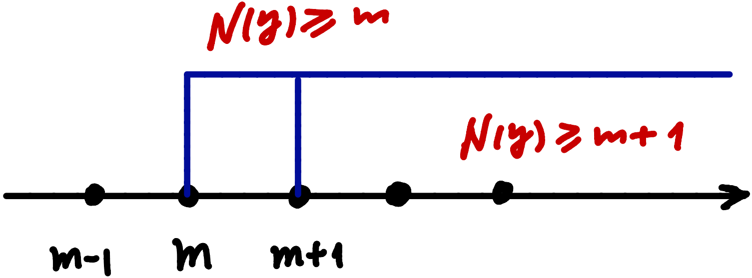
For each $k \geq 1$ and $n \geq 1$, prob $= P_x(T_y = k)P_y(T_y = n)$.
Therefore,

$$\begin{aligned} P_x(N(y) \geq 2) &= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} P_x(T_y = k)P_y(T_y = n) \\ &= \sum_{n=1}^{\infty} P_x(T_y < \infty)P_y(T_y = n) \\ &= \rho_{xy}P_y(T_y < \infty) \\ &= \rho_{xy}\rho_{yy}. \end{aligned}$$

Use the same idea to show $P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$ for $m \geq 2$.

A further **fact:**

$$\{N(y) = m\} = \{N(y) \geq m\} \setminus \{N(y) \geq m+1\}$$



$$\therefore P_x(N(y) = m) = \rho_{xy}\rho_{yy}^{m-1} - \rho_{xy}\rho_{yy}^{(m+1)-1}$$
$$= \rho_{xy}\rho_{yy}^{m-1}(1 - \rho_{yy}).$$

Sum:

Proposition:

(i) $P_x(N(y) \geq 1) = P_x(T_y < \infty) = \rho_{xy}$,

$P_x(N(y) = 0) = 1 - \rho_{xy}$.

(ii) For $m \geq 1$,

$P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$,

$P_x(N(y) = m) = \rho_{xy}\rho_{yy}^{m-1}(1 - \rho_{yy})$.

Proposition: $E_x(N(y)) = \sum_{n=1}^{\infty} P^n(x,y).$

l.h.s.=the expected no of visit to y from x.

Warning: The value can be ∞!

Proof:

$$
\begin{aligned}
E_x(N(y)) &= E_x\Big(\sum_{n=1}^{\infty} 1_y(X_n)\Big) \\
&= \sum_{n=1}^{\infty} E_x(1_y(X_n)) \\
&= \sum_{n=1}^{\infty} P_x(X_n = y) \\
&= \sum_{n=1}^{\infty} P(X_n = y | X_0 = x) = \sum_{n=1}^{\infty} P^n(x,y). \quad \square
\end{aligned}
$$

Theorem (i): y is transient iff $P_y(N(y) = \infty) = 0$.

Proof: Note

$$
\begin{aligned}
P_x(N(y) = \infty) &= \lim_{m \to \infty} P_x(N(y) \geqslant m) \\
&= \lim_{m \to \infty} \rho_{xy} \rho_{yy}^{m-1} \\
&= \begin{cases} 0 & \text{if } \rho_{yy} < 1 \\ \rho_{xy} & \text{if } \rho_{yy} = 1 \end{cases} \qquad (*)
\end{aligned}
$$

$\therefore$ y transient

$\iff \rho_{yy} < 1$

$\overset{(*)}{\iff} P_y(N(y) = \infty) = 0.$ $\square$

Theorem (ii): If y is transient then

$$E_x(N(y)) = \frac{\rho_{xy}}{1-\rho_{yy}} < \infty, \quad x \in S.$$

Proof: For a transient state y,

$$\begin{aligned} E_x(N(y)) &= \sum_{m=0}^{\infty} m P_x(N(y)=m) \\ &= \sum_{m=1}^{\infty} m \rho_{xy} \rho_{yy}^{m-1}(1-\rho_{yy}) \quad (\rho_{yy} < 1) \\ &= \rho_{xy}(1-\rho_{yy}) \cdot \frac{1}{(1-\rho_{yy})^2} \\ &= \frac{\rho_{xy}}{1-\rho_{yy}} < \infty. \quad \square \end{aligned}$$

Theorem (iii):
y is recurrent,
iff $P_y(N(y) = \infty) = 1$,
iff $E_y(N(y)) = \infty$.

Proof: y recurrent

$$\iff \rho_{yy} = 1 \overset{(*)}{\iff} P_y(N(y) = \infty) = 1$$
$$\overset{(**)}{\iff} E_y(N(y)) = \infty.$$

To show $(**)$:

"$\Longrightarrow$": $\because P_y(N(y) = \infty) = 1$
$\therefore E_y(N(y)) = \infty$.

"$\Longleftarrow$": If $E_y(N(y)) = \infty$ then y must be recurrent by Theorem (ii). □

Remark: If y is recurrent, then for $x \in S$,

$$E_x(N(y)) = \begin{cases} 0 & \text{if } \rho_{xy} = 0 \\ \infty & \text{if } \rho_{xy} > 0. \end{cases}$$

WHY? It is heuristically obvious.

Left for an exercise.

Corollary: If S is **finite**, then the chain must have at least one recurrent state.

Proof: Otherwise, all states are transient. Then, for any $x \& y$,

$$\sum_{n=1}^{\infty} P^n(x,y) = E_x(N(y)) = \frac{\rho_{xy}}{1-\rho_{yy}} < \infty.$$

$\therefore \lim_{n\to\infty} P^n(x,y) = 0$. Then

$$\begin{aligned} 0 &= \sum_{y\in S} \lim_{n\to\infty} P^n(x,y) \\ &= \lim_{n\to\infty} \sum_{y\in S} P^n(x,y) \qquad (S: \text{finite}) \\ &= \lim_{n\to\infty} P_x(X_n \in S) = \lim_{n\to\infty} 1 = 1. \quad \square \end{aligned}$$

Question 4. Decomposition of state space.

Def: x **leads to** y (denoted by $x \to y$) if

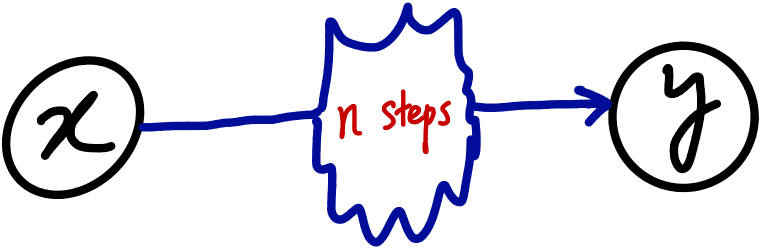
$$\rho_{xy} > 0.$$

Fact 1: $x \to y$ (i.e. $\rho_{xy} > 0$) **iff**

$$P^n(x, y) > 0 \quad \text{for some } n \geqslant 1.$$

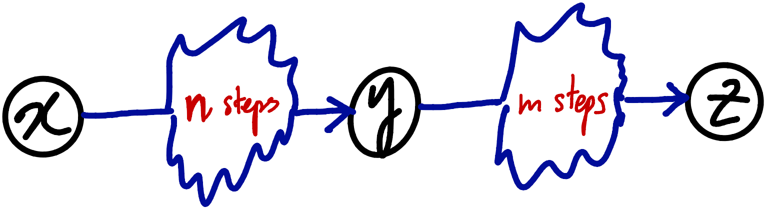
Proof: Note:

- $\rho_{xy} = P_x(T_y < \infty) = P_x(\{\exists\, m \geq 1\, s.t.\ X_m = y\})$.
- $P^n(x, y) = P(X_n = y | X_0 = x) = P_x(X_n = y)$.


x
n steps
y

Fact 2: $\left.\begin{array}{l} x \to y \\ y \to z \end{array}\right\} \Longrightarrow x \to z.$

Proof: Note

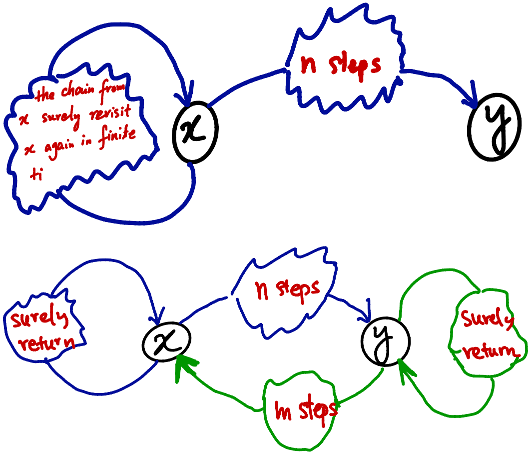


$$P^{n+m}(x,z) = \sum_{i \in S} P^n(x,i)P^m(i,z) \geqslant P^n(x,y)P^m(y,z) > 0.$$

Fact 3:

$$\left.\begin{array}{l} x \text{ recurrent } (\rho_{xx} = 1) \\ x \to y \end{array}\right\} \Longrightarrow \begin{cases} \text{(i)} & y \to x \\ \text{(ii)} & y \text{ recurrent} \\ \text{(iii)} & \rho_{yx} = \rho_{xy} = 1 \end{cases}$$

Proof (Heuristic):



Def.:

(i) $C \subseteq S$ is **closed** if

$$\rho_{xy} = 0, \quad \forall\, x \in C,\ \forall\, y \notin C,$$

i.e. no state in C leads to any state out C.

(ii) A closed set C is **irreducible** if

$$x \to y\,(\text{i.e. } \rho_{xy} > 0), \quad \forall\, x \in C, \forall\, y \in C,$$

namely, any two in C can communicate with each other.

(iii) $\{X_n\}_{n=0}^{\infty}$ is an **irreducible MC** if its state space S is irreducible.

Remark (a): One can claim that

$$
\begin{aligned}
& C \text{ is closed, i.e. } \rho_{xy} = 0, \forall x \in C, \forall y \notin C && (1)\\
\iff & P^n(x, y) = 0, \forall x \in C, \forall y \notin C, \forall n \geqslant 1 && (2)\\
\iff & P(x, y) = 0, \forall n \in C, \forall y \notin C. && (3)
\end{aligned}
$$

- Direct to see: (1) $\iff$ (2) $\implies$ (3).
- To show (3) $\implies$ (2): For $x \in C$ & $y \notin C$,

$$
\begin{aligned}
P^2(x, y) &= \sum_{x_1 \in S} P(x, x_1) P(x_1, y) \\
&= \sum_{x_1 \in C} P(x, x_1) \underbrace{P(x_1, y)}_{0} + \sum_{x_1 \notin C} \underbrace{P(x, x_1)}_{0} P(x_1, y) \\
&= 0.
\end{aligned}
$$

Induction $\implies P^n(x, y) = 0$, $\forall n \geqslant 1$.

Remark (b): If

$$C \text{ is closed, } x \in C,\ P(x, y) > 0$$

then

$$y \in C.$$

Remark (c): If $C \subset S$ is closed, then

$$\{X_n\}_{n=0}^{\infty}$$

can also be regarded as a Markov Chain with the state space C.

Theorem: If C is an irreducible closed set, then **either**

all states in C are recurrent

or

all states in C are transient.

In particular, if C is a **finite** irreducible closed set, then all states in C must be recurrent.

Proof: Two cases in general:

(i) C does NOT contain any recurrent state. In the case, all states in C are transient.

(ii) C contains at least one recurrent state. As C is irreducible, all states in C are recurrent.

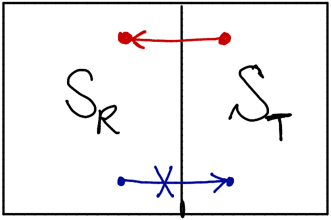
The particular case follows from the fact that any finite closed set must contain at least one recurrent state. □

Set

$$S_R = \{\text{recurrent states}\},$$
$$S_T = \{\text{transient states}\}.$$

Then,

$$S = S_R \cup S_T.$$



$$\therefore S_R \text{ is closed!}$$

A further question: Is S_R irreducible? namely, can any two recurrent states communicate to each other?

Observe: Assume $S_R \neq \phi$, for instance, $\exists\, x_0 \in S_R$. Define

$$C_{x_0} = \{x \in S_R : x_0 \to x\}.$$

Then, C_{x_0} must be **closed & irreducible**.

Proof:

(1) "C_{x_0} closed" $\iff$ "If $x \in C_{x_0}$ & $x \to y \in S$ then $y \in C_{x_0}$"
(Indeed, $y \in S_R, \therefore x_0 \to x \to y \in S_R$)

(2) "C_{x_0} irreducible" $\iff$ " If $x, y \in C_{x_0}$ then $x \to y$".

Indeed, $\left.\begin{array}{l} x_0 \to x \in S_R \\ x_0 \to y \in S_R \end{array}\right\} \Longrightarrow x \to x_0 \to y.$

□

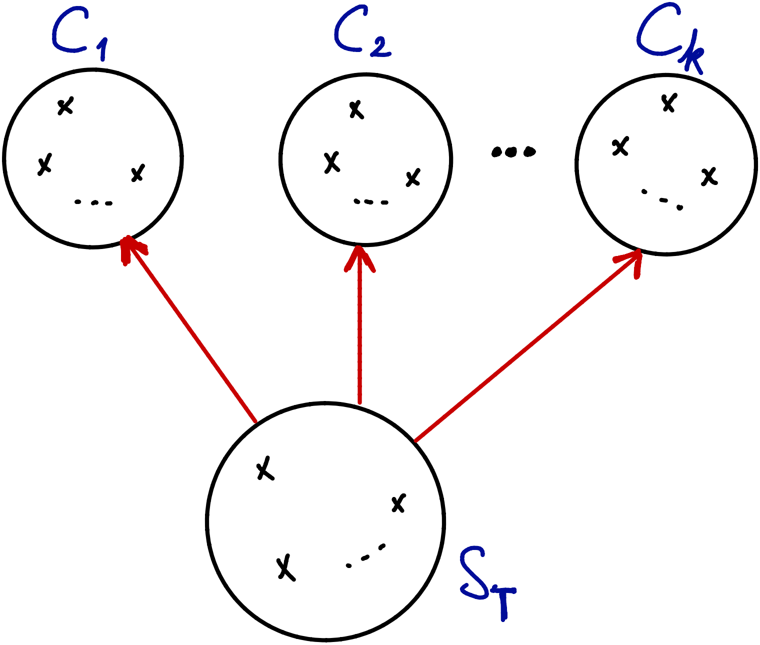
Theorem: Assume $S_R \neq \phi$. Then

$$S_R = \bigcup_{i=1}^{k} C_i \quad (k\text{: finite or infinite}),$$

where C_i, $1 \leqslant i \leqslant k$ are disjoint irreducible closed sets of recurrent states.

Proof: It suffices to show: If $C_1 \& C_2$ are two irreducible & closed sets, then either $C_1 = C_2$ or $C_1 \cap C_2 = \phi$.

Assuming $C_1 \cap C_2 \neq \phi$, we need to show $C_1 = C_2$. In fact, let $y \in C_1$ be arbitrary, we want: $y \in C_2$ $(\therefore C_1 \subseteq C_2 \subseteq C_1)$.Indeed, $\exists x \in C_1 \cap C_2$, then $C_2 \ni x \rightarrow y$. $\therefore y \in C_2$.


C_1
C_2
C_k
S_T

Corollary: If C is an irreducible & closed set, then

$$\text{either } C \subseteq S_R \text{ or } C \subseteq S_T.$$

In particular, if C is a finite, irreducible & closed set, then

$$C \subseteq S_R.$$

In terms of the (*disjoint*) decomposition

$$S = S_R \cup S_T = \left(\cup_{i=1}^{k} C_i\right) \cup S_T,$$

we may rewrite P as the canonical form:

$$P = \begin{array}{c||c|c|c|c|c} & \boxed{C_1} & \boxed{C_2} & \cdots & \boxed{C_k} & \boxed{S_T} \\ \hline \boxed{C_1} & \boxed{*} & 0 & \cdots & 0 & 0 \\ \hline \boxed{C_2} & 0 & \boxed{*} & \ddots & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \boxed{*} & 0 & 0 \\ \hline \boxed{C_k} & 0 & 0 & 0 & \boxed{*} & 0 \\ \hline \boxed{S_T} & \boxed{*} & \boxed{*} & \boxed{*} & \boxed{*} & \boxed{*} \end{array},$$

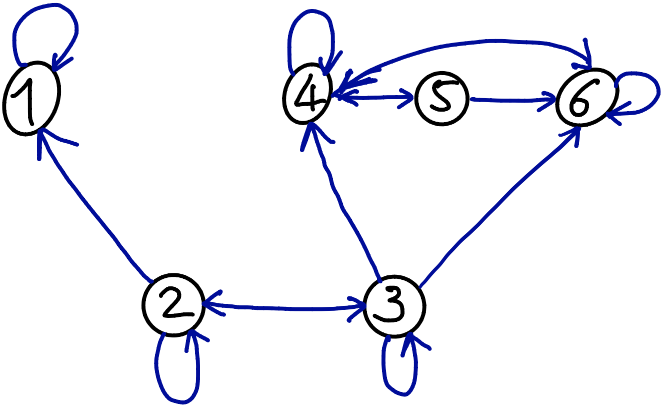
where $\boxed{*}$ denotes the sub-matrix with possible $\neq 0$ entries.

Example:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$$

Q.: Determine $S = S_R \cup S_T = \left(\cup_{i=1}^{k} C_i\right) \cup S_T$.

- $1 \to 1 \ \therefore \ C_1 = \{1\}$.
- $4 \to 5 \to 6 \to 4$ (irreducible), and $4, 5, 6$ do not lead to any other state (closed). $\therefore \ C_2 = \{4, 5, 6\}$.
- $2 \to 1$, $3 \to 4$, $\therefore \ S_T = \{2, 3\}$.



We then reformulate P in the canonical form of

$$P = \begin{array}{c} \\ a=1 \\ b=4 \\ c=5 \\ d=6 \\ e=2 \\ f=3 \end{array} \begin{array}{c} \begin{array}{cccccc} a=1 & b=4 & c=5 & d=6 & e=2 & f=3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix} \end{array}.$$

Final Issue: Assume that C is an irreducible & closed set of recurrent states. Then,

$$\boxed{T_c \stackrel{\text{def}}{=} \min\{n \geqslant 1 : X_n \in C\}}$$

denotes the **hitting time** of C.

We can also consider

$$\boxed{\rho_C(x) \stackrel{\text{def}}{=} P_x(T_C < \infty)}$$

is the prob that the chain starting at x hits C in finite time (or is absorbed by the set C).

NOTE: Once the chain hits C, it remains in C forever. (Why?)

$\therefore \rho_C(\cdot)$ is called the **absorption prob**.

It is clear to see:

$$\rho_C(x) = \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{if } x \text{ is recurrent but } \notin C. \end{cases}$$

Q.: How to compute $\rho_C(x)$, $x \in S_T$?

Indeed, assume **S_T is finite**, then for $x \in S_T$,

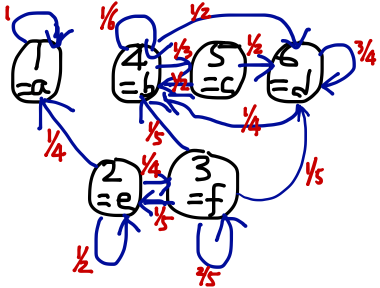
$$\boxed{\rho_C(x) = \sum_{y \in C} P(x,y) + \sum_{y \in S_T} P(x,y)\rho_C(y).} \qquad (*)$$

Assume $d_T \stackrel{\text{def}}{=} \#$ of S_T is finite
$\#$ of unknowns $= d_T$: $\rho_C(x)$, $x \in S_T$
$\#$ of equations $= d_T$
$\therefore$ it is possible to find out $\rho_C(x)$, $x \in S_T$ by solving the linear system of d_T equations.

Theorem. Let S_T be finite. Then $(*)$ admits a unique solution.

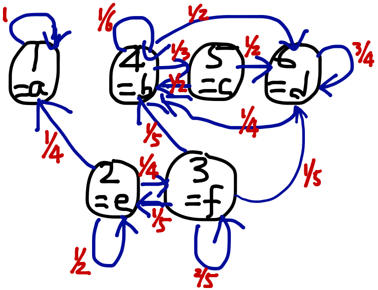
Proof. Omitted.

Example: Find $\underbrace{\rho_{C_2}(e)}_{\stackrel{\text{def}}{=}x}$, $\underbrace{\rho_{C_2}(f)}_{\stackrel{\text{def}}{=}y}$?



$$\begin{cases} x = \rho_{C_2}(e) = \underbrace{[0+0+0]}_{\sum_{j\in C_2=\{b,c,d\}} P(e,j)} + \underbrace{[\frac{1}{2}x + \frac{1}{4}y]}_{\sum_{j\in S_T=\{e,f\}} P(e,j)\rho_{C_2}(j)} \\ y = \rho_{C_2}(f) = \underbrace{[\frac{1}{5} + 0 + \frac{1}{5}]}_{\sum_{j\in C_2=\{b,c,d\}} P(f,j)} + \underbrace{[\frac{1}{5}x + \frac{2}{5}y]}_{\sum_{j\in S_T=\{e,f\}} P(f,j)\rho_{C_2}(j)} \end{cases} \quad \therefore\ x = \frac{2}{5}, y = \frac{4}{5}.$$

Similarly, let $\rho_{C_1}(e) = x$ and $\rho_{C_1}(f) = y$,



then

$$\begin{cases} x = [\frac{1}{4}] + [\frac{1}{2}x + \frac{1}{4}y] \\ y = [0] + [\frac{1}{5}x + \frac{2}{5}y] \end{cases} \Longrightarrow x = \frac{3}{5}, \quad y = \frac{1}{5}.$$

Remark (i): $\sum_i \rho_{C_i}(x) \equiv 1$, $x \in S_T$ (**finite**).

Indeed,

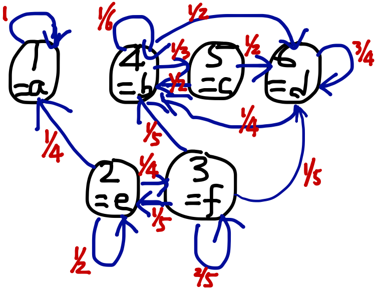
$$\sum_i \rho_{C_i}(x) = \sum_i P_x(T_{C_i} < \infty) = P_x(T_{S_R} < \infty) = 1.$$

Heuristically, it is obvious:

- We totally have finite transient states.
- Each transient state is visited only finite times.
- Surely the chain from x hits a recurrent state in finite time, so the prob $= 1$.

Remark (ii): $\rho_{xy} = \rho_C(x)$, $x \in S_T$, $y \in C$.

Apply it to the previous example:



$$\frac{2}{5} = \rho_{C_2=\{b,c,d\}}(e) = \rho_{eb} = \rho_{ec} = \rho_{ed},$$
$$\frac{4}{5} = \rho_{C_2=\{b,c,d\}}(f) = \rho_{fb} = \rho_{fc} = \rho_{fd}.$$

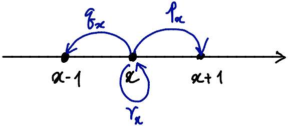
§1.3 More examples

Examples 1: Birth & Death Chain.

- Setting:

$\{X_n\}_{n=0}^{\infty}$, $S = \{0, 1, \cdots, d\}$ (d : finite or ∞)

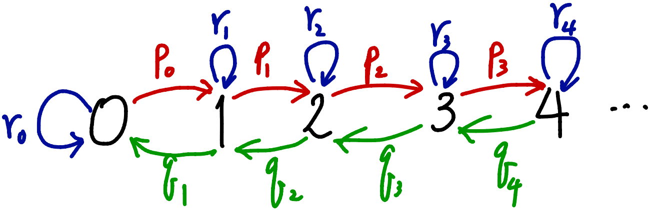
$$P(x,y) = \begin{cases} q_x & \text{if } y = x-1 \\ r_x & \text{if } y = x \\ p_x & \text{if } y = x+1 \\ 0 & \text{otherwise} \end{cases} \qquad \text{where } q_x + \gamma_x + \rho_x = 1.$$



$q_0 = 0$; $p_d = 0$, if d is finite.

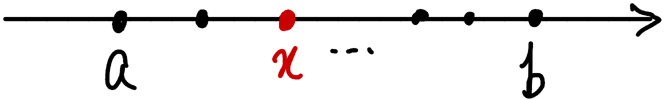
Note: the transition probs are functions of states!

$$
P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ d-1 \\ d \end{array}
\begin{array}{c}
\begin{array}{ccccccc} 0 & 1 & 2 & 3 & \cdots & d-1 & d \end{array} \\
\begin{bmatrix}
r_0 & p_0 & & & & & \\
q_1 & r_1 & p_1 & & & & \\
 & q_2 & r_2 & p_2 & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & & q_{d-1} & r_{d-1} & p_{d-1} \\
 & & & & & q_d & r_d
\end{bmatrix}
\end{array}
$$



A general question: Given $a, b \in S$ with $a < b$, compute

$$u(x) \stackrel{\text{def}}{=} P_x(T_a < T_b), \quad a < x < b,$$
$$v(x) \stackrel{\text{def}}{=} P_x(T_a > T_b), \quad a < x < b.$$



$\{T_a < T_b\}$ = Before the chain hits b, it hits a, (i.e., the chain hits a earlier than b)

$\{T_a > T_b\}$ = Before the chain hits a, it hits b, (i.e., the chain hits b earlier than a)

Claim:

(i) $u(a) = 1$, $u(b) = 0$.

(ii) $u(x) = q_x u(x-1) + r_x u(x) + p_x u(x+1)$ for $a < x < b$.

(iii) $u(x) = \dfrac{\sum_{y=x}^{b-1} \gamma_y}{\sum_{y=a}^{b-1} \gamma_y}$ for $a < x < b$.

(iv) $v(x) = 1 - u(x) = \dfrac{\sum_{y=a}^{x-1} \gamma_y}{\sum_{y=a}^{b-1} \gamma_y}$ for $a < x < b$,

where γ_x are defined by

$$\gamma_x \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } x = 0 \\ \frac{q_1 \cdots q_x}{p_1 \cdots p_x} & \text{if } 1 \leq x \leq d-1. \end{cases}$$

Proof:

(i) is obvious.

(ii) follows by

$$\begin{aligned}
P_x(A) =& P_x(A, X_1 = x-1) + P_x(A, X_1 = x) \\
& + P_x(A, X_1 = x+1) \\
=& P_x(X_1 = x-1) P_x(A | X_1 = x-1) \\
& + P_x(X_1 = x) P_x(A | X_1 = x) \\
& + P_x(X_1 = x+1) P_x(A | X_1 = x+1) \\
=& P(X_1 = x-1 | X_0 = x) P(A | X_0 = x, X_1 = x-1) \\
& + P(X_1 = x | X_0 = x) P(A | X_0 = x, X_1 = x) \\
& + P(X_1 = x+1 | X_0 = x) P(A | X_0 = x, X_1 = x+1) \\
=& q_x P_{x-1}(A) + r_x P_x(A) + p_x P_{x+1}(A).
\end{aligned}$$

Proof of (iii):

$$u(x) = q_x u(x-1) + (1 - p_x - q_x)u(x) + p_x u(x+1)$$
$$(p_x + q_x)u(x) = q_x u(x-1) + p_x u(x+1)$$

$$\begin{aligned} u(x+1) - u(x) &= \frac{q_x}{p_x}[u(x) - u(x-1)] \quad (a < x < b) \\ &= \frac{q_x \cdot q_{x-1}}{p_x \cdot p_{x-1}}[u(x-1) - u(x-2)] \\ &= \cdots \\ &= \left(\frac{q_x}{p_x}\right)\left(\frac{q_{x-1}}{p_{x-1}}\right)\cdots\left(\frac{q_{a+1}}{p_{a+1}}\right)[u(a+1) - u(a)] \\ &= \frac{\gamma_x}{\gamma_a}[u(a+1) - u(a)]. \end{aligned}$$

Note: $\sum_{x=1}^{b-1}(\cdot) \Rightarrow \underbrace{u(b) - u(a)}_{=-1} = \frac{\sum_{x=a}^{b-1}\gamma_x}{\gamma_a}[u(a+1) - u(a)]$

$$\therefore u(x+1) - u(x) = -\frac{\gamma_x}{\sum_{x=a}^{b-1}\gamma_x} \quad (a \leqslant x < b)$$

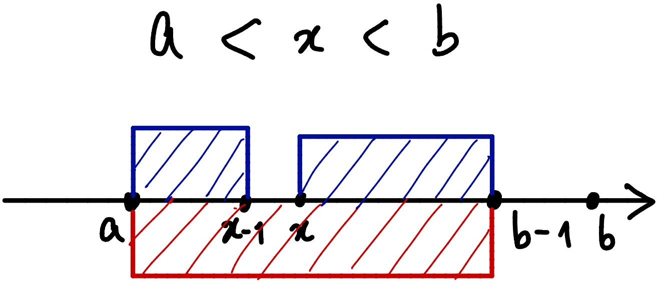
Further, change x to y, $\sum_{y=x}^{b-1} \Rightarrow$

$$\underbrace{u(b)}_{=0} - u(x) = -\frac{\sum_{y=x}^{b-1} \gamma_y}{\sum_{y=a}^{b-1} \gamma_y}, \qquad \therefore u(x) = \sum_{y=x}^{b-1} \gamma_y \Big/ \sum_{y=a}^{b-1} \gamma_y.$$

Reminder:

- $u(x) \stackrel{\text{def}}{=} P_x(T_a < T_b), \quad a < x < b.$
- $\gamma_x \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } x = 0 \\ \frac{q_1 \cdots q_x}{p_1 \cdots p_x} & \text{if } 1 \le x \le d-1. \end{cases}$

Sum:



$$P_x(\underbrace{T_a < T_b}_{\text{“Death faster”}}) = \sum_{y=x}^{b-1} \gamma_y \Big/ \sum_{y=a}^{b-1} \gamma_y,$$

$$P_x(\underbrace{T_a > T_b}_{\text{“Birth faster”}}) = \sum_{y=a}^{x-1} \gamma_y \Big/ \sum_{y=a}^{b-1} \gamma_y.$$

e.g.: Set:

- A gambler bets \$1 each time.
- The prob of winning or losing each bet is $9/19$ and $10/19$, resp.
- The gambler will quit as soon as his net winning is \$25 or his net loss is \$10.

Q.:

(i) Find the prob he quits and wins.

(ii) Find his expected loss.

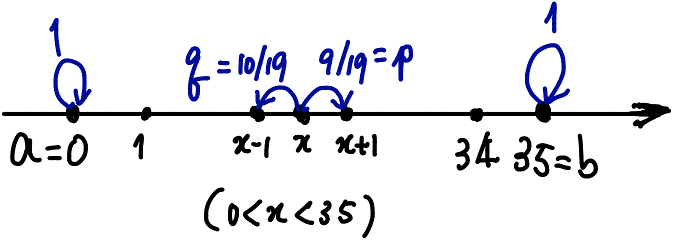
Sol.: Let

$$X_n \overset{\text{def}}{=} \text{the capital of the gambler at time } n = 0, 1, 2, \cdots$$

For simplicity, we choose

$$X_0 = 10, \quad S = \{0, 1, \cdots, 35\}.$$

$\{X_n\}_{n=0}^{\infty}$ forms a birth & death chain on S with



$$\gamma_y = \left(\frac{q}{p}\right)^y = \left(\frac{10}{9}\right)^y, 0 \leqslant y \leqslant 34.$$

(i) **Find the prob he quits and wins:** To find

$$P_{10}(\underbrace{T_{35} < T_0}_{\text{“Birth faster”}}) = \frac{\sum_{y=0}^{9} \gamma_y}{\sum_{y=0}^{34} \gamma_y} = \frac{\sum_{y=0}^{9} (\frac{10}{9})^y}{\sum_{y=0}^{34} (\frac{10}{9})^y} = \frac{(\frac{10}{9})^{10} - 1}{(\frac{10}{9})^{35} - 1}$$

$$= 0.047.$$

(ii) **Find his expected loss:**

gain (+25)	loss (−10)
0.047	1− 0.047

The expected loss is

$$(1 - 0.047)(-10) + (0.047)(25) = -8.36. \quad \square$$

- We are further interested in the below situation:

Assume that $S = \{0, 1, 2, \cdots\}$ is **infinite**, and the birth & death chain is **irreducible**, namely,

$$p_x > 0, \forall\, x \geqslant 0, \quad \text{and} \quad q_x > 0, \forall\, x \geqslant 1.$$

Q.: When such chain is recurrent or transient? (NOT obvious for an irreducible chain with infinite states!)

Proposition: The chain is recurrent **iff**

$$\sum_{k=0}^{\infty} \gamma_k = \infty.$$

Pf.: Since the chain is irreducible, we only need to consider one state, namely, 0. Observe that

$$\rho_{00} = P_0(T_0 < \infty) = r_0 + p_0 P_1(T_0 < \infty), \qquad (*)$$

where

$$\begin{aligned}\rho_{10} = P_1(T_0 < \infty) &= \lim_{n\to\infty} P_1(T_0 < T_n) \\ &= \lim_{n\to\infty}\left[1 - \frac{1}{\sum_{k=0}^{n-1}\gamma_k}\right]. \quad (**)\end{aligned}$$

Therefore,

0 is recurrent, i.e. $\rho_{00} = 1$

$\overset{(*)\&r_0+p_0=1}{\Longleftrightarrow} \rho_{10} = P_1(T_0 < \infty) = 1$

$\overset{(**)}{\Longleftrightarrow} \sum_{k=0}^{\infty}\gamma_k = \infty.$ □

Remark: For instance, let

$$p_x \equiv p > 0,\ q_x \equiv q > 0,\ 0 < \rho + q \leqslant 1.$$

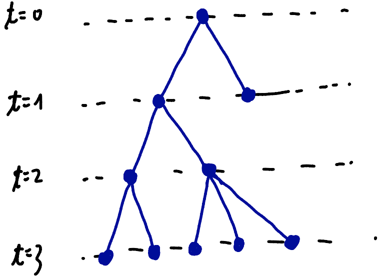
Then,

$$\boxed{\sum_{k=0}^{\infty} \gamma_k = \sum_{k=0}^{\infty} \left(\frac{q}{p}\right)^k.}$$

- If $p > q$, then $\sum_{k=0}^{\infty} \gamma_k$ is finite. The chain is transient .
- If $p = q$ or $p < q$, then $\sum_{k=0}^{\infty} \gamma_k = \infty$. The chain recurrent.

Example 2. **Branching chain.**

Each particle generates ξ particles independently in the next generation.



$X_n \stackrel{\text{def}}{=}$ the total no of particles in the n^{th} generation

$$P(0,0) = 1.$$

$$P(x,y) = P(\xi_1 + \xi_2 + \cdots + \xi_x = y), \quad x \geqslant 1.$$

Q.: Determine

$$\rho \stackrel{\text{def}}{=} \text{the prob that the descendants of a given particle eventually become } \textbf{extinct}.$$

We call ρ to be the **extinction prob** of the chain. Then,

$$\rho = \rho_{10} = P_1(T_0 < \infty).$$

1^{st} **Obervation:** Suppose ξ has the pdf

$$p_k = P(\xi = k), \; k = 0, 1, 2, \cdots$$

Then,

$$P(1, k) = P(\xi_1 = k) = p_k, \; k = 0, 1, 2, \cdots$$

From this we see:

- If $p_0 = 0$, then each individual cannot change to zero, so population never extinct, i.e. $\rho = 0$.
- If $p_0 = 1$, then it extincts for sure, i.e., $\rho = 1$.

To avoid two trivial cases, we always assume

$$0 < p_0 < 1.$$

2^{nd} **Obervation:** Assuming there are x particles, the prob for them to extinct is

$$\rho_{x0} = \rho^x.$$

(**Pf.:** Use **independence**!)

3^{rd} **Obervation:** Let

$$\mu \stackrel{\text{def}}{=} E(\xi) = \sum_{k=0}^{\infty} kp_k = \sum_{k=1}^{\infty} kp_k.$$

Then, $E(X_{n+1}|X_n = k) = E(\xi_1 + \cdots + \xi_k) = k\mu$,

$$\begin{aligned} E(X_n) &= \sum_{k=0}^{\infty} E(X_n|X_{n-1} = k)P(X_{n-1} = k) \\ &= \sum_{k=0}^{\infty}(k\mu)P(x_{n-1} = k) \\ &= \mu E(X_{n-1}) \\ &= \cdots \\ &= \mu^n E(X_0). \end{aligned}$$

Claim: If $\mu < 1$, then population will **extinct for sure**, i.e., $\rho = 1$.

Proof:

$$
\begin{aligned}
P_1(T_0 > n) &\leqslant P_1(X_n \geqslant 1) \quad (\because \{T_0 > n\} \subseteq \{X_n \geqslant 1\}) \\
&= \sum_{k=1}^{\infty} P_1(X_n = k) \leqslant \sum_{k=1}^{\infty} k P_1(X_n = k) \\
&= \sum_{k=0}^{\infty} k P_1(X_n = k) \\
&= E(X_n) = \mu^n E(X_0) \xrightarrow{n\to\infty} 0 \ (\because \mu < 1)
\end{aligned}
$$

Therefore

$$
\underbrace{\rho}_{\text{extinction prob}} = \rho_{10} = P_1(T_0 < \infty) = \lim_{n\to\infty} P_1(T_0 \leqslant n)
$$

$$
= \lim_{n\to\infty} [1 - P_1(T_0 > n)] = 1. \quad \square
$$

What about $\mu \geqslant 1$?

$$\begin{aligned}\rho = \rho_{10} = P_1(T_0 < \infty) \\ = P(1,0) + \sum_{k=1}^{\infty} P(1,k)\rho_{k0} \\ = p_0 + \sum_{k=1}^{\infty} p_k \rho^k = \sum_{k=0}^{\infty} p_k \rho^k,\end{aligned}$$

i.e., ρ solves the equation $\boxed{t = \Phi(t)}$ with

$$\boxed{\Phi(t) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} p_k t^k,}$$

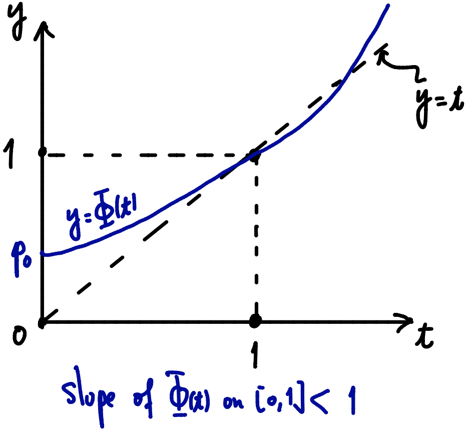
which is called the **moment generating function** of the pdf $(p_k)_{k\geq 0}$ of ξ.

Observe:

- $\Phi'(t) \geqslant 0$, $\Phi''(t) \geqslant 0$, ($\therefore \Phi(t) \uparrow$ & concave upward).
- $\Phi(0) = p_0 \in (0,1)$, $\Phi(1) = \sum_{k=0}^{\infty} p_k = 1$.
- $\Phi'(1) = \sum_{k=1}^{\infty} kp_k = E(\xi) = \mu$.

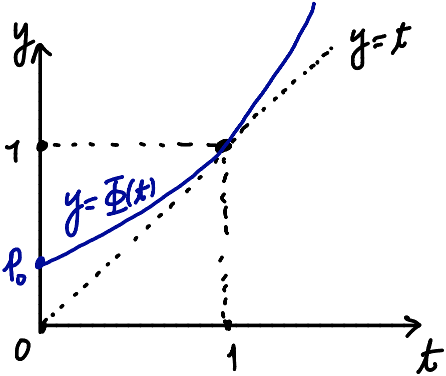
Then, we have **three cases:**

Case (i): $\mu < 1$.



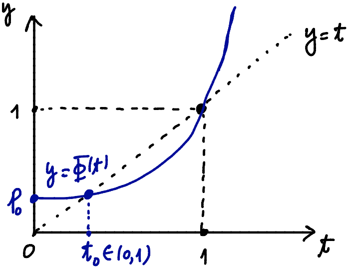
$\therefore\ \rho = 1$ (**extinct for sure**, as proved before)

Case (ii): $\mu = 1$.



$\therefore\ \rho = 1$ (**extinct for sure!**)

Case (iii): $\mu > 1$.



$\Phi(t) = t$ at $t = t_0 \in (0,1)$ or $t = 1$.

Claim: In this case, $P_1(T_0 \leqslant n) \leqslant t_0$ for all $n = 1, 2, \cdots$ (**proved later**).

$$\therefore \rho = \rho_{10} = P_1(T_0 < \infty) = \lim_{n\to\infty} P_1(T_0 \leqslant n) \leqslant t_0$$

$\therefore \rho = t_0$ is the **only** solution.

Proof of Claim: Use induction. Set

$$a_n \stackrel{\text{def}}{=} P_1(T_0 \leqslant n).$$

$n = 0$: $a_0 = P_1(T_0 \leqslant 0) = 0 < t_0$.

Assuming $a_n \leqslant t_0 (n \geqslant 0)$, consider

$$\begin{aligned} a_{n+1} &= P_1(T_0 \leqslant n+1) \\ &= \underbrace{P(1,0)}_{=p_0} + \sum_{k=1}^{\infty} \underbrace{P(1,k)}_{p_k} \underbrace{P_k(T_0 \leqslant n)}_{=[P_1(T_0 \leqslant n)]^k = a_n^k} \\ &= \sum_{k=0}^{\infty} p_k a_n^k \\ &= \Phi(a_n) \leqslant \Phi(t_0) = t_0 \quad (\Phi \text{ is nondecreasing}). \quad \square \end{aligned}$$

e.g.: Every man has 3 kids with prob $1/2$ being boy and $1/2$ being girl. Find the prob that the male live eventually extinct.

Sol.: $p_0 = P(\xi = 0) = \frac{1}{8}$, $p_1 = P(\xi = 1) = \frac{3}{8}$, $p_2 = P(\xi = 2) = \frac{3}{8}$, $p_3 = P(\xi = 3) = \frac{1}{8}$.

$$E(\xi) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2} > 1$$

$$\Phi(t) = \frac{1}{8} + \frac{3}{8}t + \frac{3}{8}t^2 + \frac{1}{8}t^3$$

$$\text{let } \Phi(t) = t, \text{i.e. } t = \frac{1}{8} + \frac{3}{8}t + \frac{3}{8}t^2 + \frac{1}{8}t^3$$

Solutions: $t = 1$, $\sqrt{5} - 2$. Then

$$\rho = \sqrt{5} - 2$$

is the extinct prob. □

Example 3. Queuing chain.

Setting:

- In a queue, let ξ_n denote the no of arrivals in the n-th unit time. $\{\xi_n\}_{n=1}^{\infty}$ are i.i.d.r.v. with pdf:

$$f(k) = p_k, \quad k = 0, 1, 2, \cdots$$

- The service of a customer is exactly one in a unit time.

Let X_n denote the no of customers in the queue.

$$P(x, y) = f(\underbrace{y - (x - 1)}_{\text{no of arrivals}}), \quad x \geqslant 1,$$

$$P(0, y) = f(y).$$

Note: $P(1, y) = P(0, y)$.

Q.: Assuming that the chain is irreducible, check if the chain is **recurrent** or **transient**, i.e. letting

$$\rho = \rho_{00} = P_0(T_0 < \infty),$$

decide

$$\text{if } \rho = 1 \text{ or } \rho < 1.$$

Note. If

$$p_0 > 0 \ \& \ p_0 + p_1 < 1,$$

then the chain is irreducible. (Ex. 37 on Page 46).

Let

$$\Phi(t) \stackrel{\text{def}}{=} p_0 + p_1 t + p_2 t^2 + \cdots$$
$$= \sum_{k=0}^{\infty} p_k t^k$$
$$= \sum_{k=0}^{\infty} f(k) t^k$$

be the moment generating function of f.

Claim: $\rho = \rho_{00}$ solves $\Phi(t) = t$.

Pf.:

- Note

$$\rho_{00} = P(0,0) + \sum_{k=1}^{\infty} P(0,k)\rho_{k0},$$

$$\rho_{10} = P(1,0) + \sum_{k=1}^{\infty} P(1,k)\rho_{k0},$$

$$P(1,k) = P(0,k), \ \forall k \geq 0.$$

Therefore,

$$\rho_{10} = \rho_{00} = \rho.$$

- **To show:** $\rho_{x,x-1} = \rho_{10} = \rho$ **for all** $x > 1$**.**

In fact, we observe that
for the chain starting at $x > 1$ ($\therefore\ x - 1 \geq 1$),
the event $T_{x-1} = n$ means

$$n = \min\{m > 0 : x + (\xi_1 - 1) + \cdots (\xi_m - 1) = x - 1\},$$

i.e.

$$n = \min\{m > 0 : 1 + (\xi_1 - 1) + \cdots (\xi_m - 1) = 0\}.$$

Therefore, $P_x(T_{x-1} = n) = P_1(T_0 = n)$, $\forall n \geq 1$.

$$\therefore \rho_{x,x-1} = \rho_{10} = \rho.$$

- **To show:**

$$\rho_{x,0} = \rho_{x,x-1} \cdot \rho_{x-1,0}, \ \forall x \geqslant 2. \qquad (*)$$

(Ex. 39, P46). If so, then

$$\rho_{x,0} = \rho\rho_{x-1,0} = \cdots = \rho^x,$$

(also true for $x = 1$), and hence

$$\begin{aligned} \rho = \rho_{00} &= P(0,0) + \sum_{k=1}^{\infty} P(0,k)\rho_{k0} \\ &= p_0 + \sum_{k=1}^{\infty} p_k \rho^k \\ &= \Phi(\rho). \end{aligned}$$

Proof of ($*$): Let $x \geq 2$. Note that for $m \geq 2$,

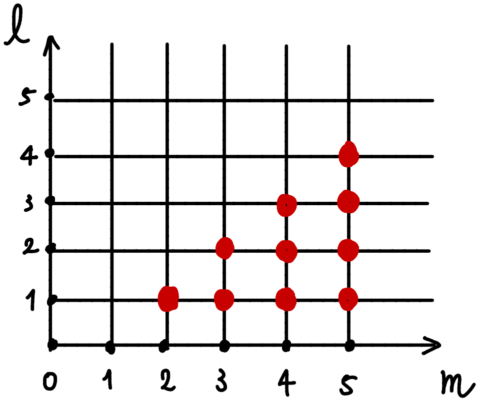
$$P_x(T_0 = m) = \sum_{\ell=1}^{m-1} P_x(T_{x-1} = \ell) P_{x-1}(T_0 = m - \ell).$$

Then,

$$\begin{aligned}
\rho_{x,0} &= P_x(T_0 < \infty) = \sum_{m=1}^{\infty} P_x(T_0 = m) \\
&= \sum_{m=2}^{\infty} P_x(T_0 = m) \quad \text{(Note: } P_x(T_0 = 1) = 0 \text{ for } x \geq 2\text{)} \\
&= \sum_{m=2}^{\infty} \sum_{\ell=1}^{m-1} P_x(T_{x-1} = \ell) P_{x-1}(T_0 = m - \ell) \\
&= \sum_{\ell=1}^{\infty} \sum_{m=\ell+1}^{\infty} P_x(T_{x-1} = \ell) P_{x-1}(T_0 = m - \ell) \ \text{(see later)} \\
&= \sum_{\ell=1}^{\infty} P_x(T_{x-1} = \ell) \rho_{x-1,0} \\
&= \rho_{x,x-1} \rho_{x-1,0}. \quad \square
\end{aligned}$$

Note:

$$\sum_{m=2}^{\infty}\sum_{\ell=1}^{m-1} = \sum_{\ell=1}^{\infty}\sum_{m=\ell+1}^{\infty}.$$



Sum: Let $\mu = E(\xi)$. Then

- If $\mu \leqslant 1$, then $\Phi(\rho) = \rho$ has the only solution $\rho = 1$. The chain is recurrent.

- If $\mu > 1$, then $\Phi(\rho) = \rho$ has two solutions 1 and $t_0 \in (0, 1)$. As in the previous example, one has to take $\rho = t_0$. The chain is transient.

□