# *Lecture 10. More on PP and DTMC*

Sim, Min Kyu, Ph.D.
mksim@seoultech.ac.kr

서울과학기술대학교 데이터사이언스학과

BK21 FOUR 데이터사이언스와 비즈니스포텐셜

## Minimum of exponential random variables

- Minimum of exponentially distributed random variables is again exponential random variable with parameter that is the sum of parameters of original exponential random variables.
- Thm. $X_1 \sim exp(\lambda_1), X_2 \sim exp(\lambda_2)$, $X_1$ and $X_2$ are indep.
  $\Rightarrow min(X_1, X_2) \sim exp(\lambda_1 + \lambda_2)$

# Doubly stochastic matrix

### Definition - Doubly stochastic matrix

- A matrix is said to be *stochastic* if each row sums up to 1.
  - Every legit transition probability matrix in DTMC is *stochastic*.
- A stochastic matrix is said to be *doubly stochastic* if each column sums up to 1 as well.
  - Ex) The first example of periodic matrix.

### Theorem

- $n$ by $n$ doubly stochastic matrix for finite states DTMC has stationary distribution $\pi_i = 1/n$ for every state $i \in S$.
  - Ex1) The above example.
  - Ex2) Ring structure DTMC.

## *Chapman-Kolmogrov Equation for DTMC*

- n-step probability review
- We want to have (m+n)-step transition probability (matrix) using m-step and n-step transition matrix.
- $\mathbf{P}_{ij}^{n+m} = \sum_{k \in S} \mathbf{P}_{ik}^{n} \mathbf{P}_{kj}^{m}$
- Perspective of ''path''.
- pf) $\mathbf{P}_{ij}^{n+m} =$

- Matrix Algebra point of view

-

# Previous Exercise

$$\mathbf{P} = \begin{array}{c} (C,C) \\ (NC,C) \\ (C,NC) \\ (NC,NC) \end{array} \begin{pmatrix} .2 & & .8 & \\ .4 & & .6 & \\ & .6 & & .4 \\ & .8 & & .2 \end{pmatrix}$$

Q1) $\mathbb{P}[Y_{n+2} = (NC,C) | Y_n = (C,C)]$=?

## *Interarrival times in PP*

- In LN( ), p( ). We let $T_1$ :=*Time to first arrival* in $PP(\lambda)$, and showed $T_1 \sim exp(\lambda)$.
- Furthermore, for $n > 1$, let $T_n$ denote the elapsed time between $(n-1)$st and the $n$th event. The sequence $\{T_n, n = 1, 2, ...\}$ is called *sequence of interarrival times.*
- ex) What does $T_1 = 5$ and $T_2 = 10$ mean?

- Theorem: $T_n \sim exp(\lambda)$, for all $n = 1, 2, ...$.
- Remark. The PP from any point on is independent of all that has previously occurred (by independent increments), and also has the same distribution as the original process (by stationary increments).
- Remark. In other words, the process has *no memory*, and hence exponential interarrival times are naturally expected.

## *blank*

## *Extended applications of DTMC and PP*

- We can model various things.
- Sports Analysis
- Gambling
- Financial Market
- and more…

- This lecture could help your project as well as the exam.

## *Gambler's ruin probability*

- Suppose you have \$3(=x), and bet \$1 with winning probability $p = 18/38$ until your wealth becomes \$0(=a) or your wealth becomes \$8(=b). What is chance of you will leave Casino with \$8?

$$
\mathbf{P} = \begin{array}{c} 0(lose) \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8(won) \end{array}
\begin{pmatrix}
1 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
.53 & \bullet & .47 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & .53 & \bullet & .47 & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & .53 & \bullet & .47 & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & .53 & \bullet & .47 & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & .53 & \bullet & .47 & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & .53 & \bullet & .47 & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & .53 & \bullet & .47 \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & 1
\end{pmatrix}
$$

- Result of $a = 0, b = 8, p = 18/38$

$$
\mathbf{P}^{\infty} = \begin{array}{c} 0(lose) \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8(won) \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .92 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .08 \\ .82 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .18 \\ .72 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .28 \\ .60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .40 \\ .48 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .52 \\ .33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .67 \\ .18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .82 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
$$

- Result of $a = 0, b = 1000, p = 18/38, x = 100.$
  - What is the quantity for $\mathbf{P}^{\infty}_{100\$ \to win}$?

- Result of $a = 0, b = 1000, p = 19/38, x = 100.$
  - What is the quantity for $\mathbf{P}^{\infty}_{100\$ \to win}$?

- Result of $a = 0, b = 10 \times 100$\$, $p = 18/38, x = 1 \times 100$ \$ (bet 100\$ for each)
  - What is the quantity for $\mathbf{P}^{\infty}_{1 \times 100\$ \to win}$?

## Appendix - MATLAB code for Gambler's ruin

```
>>
>>
>>
>>
>> p=18/38;
q=1-p;
a=0; % lose
b=1000; % won
n=b-a+1;
P=diag(p*ones(n-1,1),1)+diag(q*ones(n-1,1),-1);
P(1,:)=[1' zeros(n-1,1)']';
P(n,:)=[zeros(n-1,1)' 1']';
P;
P50000=P^500000;
P50000(101,n)
```

## *blank*

# *Squash*

- Racket sports (court number 5 in CRC)
- Rules
  - Two players, three or five games.
  - Only the server scores points.
  - The server, on winning a rally, scores a point
  - The receiver, on winning a rally, becomes the server.
  - The player who scores nine points wins the game

- Rules (cont'd)
  - Suppose A and B are playing for the first set and $8 : \overline{7}$ now. (A's score is 8, B's score is 7, and B is serving)
  - Suppose B wins this play so that it becomes $8 : \overline{8}$.
  - Because A got to 8 first, A can decide either
    - *i)* This set ends at 9
    - *ii)* This set ends at 10

- Questions
  - Suppose the chance of A winning a play is 0.6, then should A choose i) or ii)?

- Suppose A decides “i) This set ends at 9”.
- DTMC
  - Transition diagram and matrix

  - Classification of states

  - What is the chance of A winning this game?

- Suppose A decides "ii) This set ends at 10".
- DTMC

$$
\mathbf{P} = \begin{array}{c} lose \\ 8:\overline{8} \\ \overline{8}:8 \\ 8:\overline{9} \\ \overline{8}:9 \\ 9:\overline{8} \\ \overline{9}:8 \\ 9:\overline{9} \\ \overline{9}:9 \\ win \end{array}
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet
\end{pmatrix}
$$

- What is the chance of A winning this game?

- What if the chance of A winning a rally is not 0.6, but for general $p$?

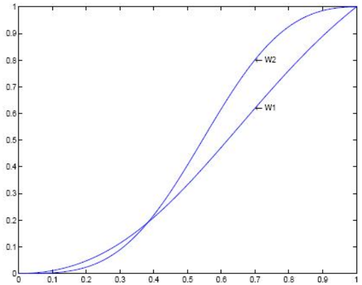



Figure 1: Probability of winning

- optimal decision
  - If $p \leq$ , then choose i) ends at 9
  - Otherwise, choose ii) ends at 10

- Reference
  - Optimal Decision for the Squash Player
  - Jan Vecer, Columbia University, Department of Statistics
  - Journal of Chinese Statistical Association, 2004.
  - www.stat.columbia.edu/∼vecer/squash.ps

## *Stock price - binomial tree*

- Let $X_n$ be the closing price of the stock at $n$th day.
- Let $p = \mathbb{P}(X_{n+1} = x + 1 | X_n = x)$, and $1 - p = \mathbb{P}(X_{n+1} = x - 1 | X_n = x)$
- Consider future evolution, starting with $X_0 = 100$.

- Consider an European call option which matures at day 5 with exercise price 101.
- (If you possess one unit of the call option, then at the day 5, you have a right to buy the stock at 101 dollars.)
- If $X_5 = 103$, then you can buy the stock at 101 and sell at 103. In this case, you earn 2 dollar.
- If $X_5 = 99$, then you still can buy the stock at 101. But you would not do it because you can buy a stock at 99 dollars. (Possessing call option is the "right" not the "obligation")
- i.e., the payoff of a call option is $(X_5 - 101)^+$

- What is the expected payoff for the option, when $p = 0.6$?

## blank

## *Soccer scoring by PP*

- Sports like soccer can be modelled by Poisson process because of many trials with small probability to make goal.
- Suppose the team A has goal-ability of $PP(1 \ per \ 30 \ minutes)$, and the team B has goal-ability of $PP(1 \ per \ 25 \ minutes)$.
- Assume the two Poisson processes are independent. (very absurd assumption...)
- What are the chance for A to win/draw/lose?
- Let $X_A :=$ # of goals that A team makes during 90 minutes, and define $X_B$ same way.
- By Poisson Process principle, $X_A \sim poi(3)$, $X_B \sim poi(90/25)$.
- $\mathbb{P}(X_A > X_B)$, $\mathbb{P}(X_A = X_B)$, $\mathbb{P}(X_A < X_B)$
- Why PP is not a good model for Soccer game scoring?

## blank

## Previous Exercise

$$\mathbf{P} = \begin{pmatrix} .2 & .8 & 0 & 0 & 0 \\ .5 & .5 & 0 & 0 & 0 \\ 0 & .25 & 0 & .75 & 0 \\ 0 & 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

## *blank*

## *Previous Exercise*

- Suppose you are modelling your economic status using DTMC.
  - There are five possible economic status: Trillionaire (T), Billionaire (B), Millionaire (M), Thousandaire (Th), Bankrupt (Bk).
  - Your economic status changes every month.
  - Once you are bankrupt, there is no chance of coming back to other states.
  - Also, once you become a billionaire or a trillionaire, you will never be millionaire, thousandaire, or bankrupt.
  - Other chances of transitions are as follows. (Chances for transitions to itself are omitted.)

| Transition | Prob. |
|---|---|
| T $\rightarrow$ B | 0.9 |
| B $\rightarrow$ T | 0.3 |
| M $\rightarrow$ B | 0.1 |
| M $\rightarrow$ Th | 0.5 |
| Th $\rightarrow$ M | 0.2 |
| Th $\rightarrow$ Bk | 0.3 |

## *blank*

## *Previous Exercise*

- Nortel in Canada operates a call center for customer service.
  - Assume that each caller speaks either English or French, but not both.
  - Suppose that the arrival for each type of calls follows a Poisson process.
  - The arrival rates for English and French calls are 2 and 1 calls per minute, respectively.
  - Assume that call arrivals for different types are independent.

*(a)* What is the probability that the 2nd English call will arrive after minute 5?

*(c)* What is the expected time for the 1st call that can be answered by a bilingual operator (that speaks both English and French) to arrive?

*(d)* Suppose that the call center is staffed by bilingual operators that speak both English and French.

- Let $N(t)$ be the total number of calls of both types that arrive during $(0, t]$. What kind of a process is $N(t)$?
- What is the mean and variance of $N(t)$?

*(g)* What is the probability that two calls arrive in the interval 0 to 10 minutes and three calls arrive in the interval from 5 to 15 minutes?

## blank

# *blank*

## *Previous Exercise*

- Calls to a center follow a Poisson process with rate 120 calls per hour.
  - Each all has probability 1/4 from a male customer.
  - The call center opens at 9am each morning.

*(d)* What is the expected elapse of time from 9am until the second female customer arrives?

## *blank*

# *Previous Exercise*

Assume that call arrival to a call center follows a NHPP. The call center opens from 9am to 5pm. During the first hour, the arrival rate increases linearly from 0 at 9am to 60 calls per hour at 10am. After 10am, the arrival rate is constant at 60 calls per hour.

(a) Plot the arrival rate function $\lambda(t)$ as a function of time $t$; indicate clearly the time unit used.

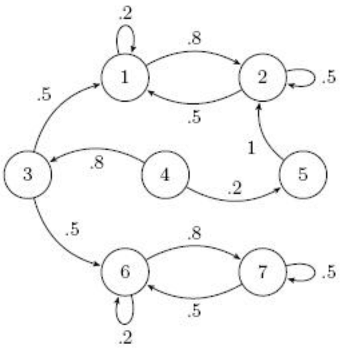
(b) Find the probability that exactly 5 calls have arrived by 9:10am.

# *blank*

# *blank*

## *Exercise 1*

- Calculate $\mathbf{P}^{100}$ for the following DTMC. Make sure you can retrieve every element of the $\mathbf{P}^{100}$ matrix.



(Solution)

$$\mathbf{P} = \begin{pmatrix} .2 & .8 & 0 & 0 & 0 & 0 & 0 \\ .5 & .5 & 0 & 0 & 0 & 0 & 0 \\ .5 & 0 & 0 & 0 & 0 & .5 & 0 \\ 0 & 0 & .8 & 0 & .2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .2 & .8 \\ 0 & 0 & 0 & 0 & 0 & .5 & .5 \end{pmatrix}.$$

Step 1) Identifying classes.
Recurrent classes: $\{1, 2\}$, $\{6, 7\}$. Transient classes: $\{3\}$, $\{4\}$, $\{5\}$
(So far, how many classes do you see in this example? There are two recurrent classes and three transient classes. Therefore, there are total five classes.)

Step 2) Set up $p^{100}$ and identify as many zeros as possible.

$$\mathbf{P}^{100} \approx \lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} x & x & 0 & 0 & 0 & 0 & 0 \\ x & x & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & & \\ & & 0 & 0 & 0 & & \\ & & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{pmatrix}.$$

Step 3) Identify the "$x$" in above matrix. The "$x$" are limiting probability from a recurrent class to the recurrent class itself. Thus, we can find limiting probability by getting stationary distribution of the small Markov chain of recurrent class. \
For the first recurrent class of $\{1, 2\}$,

$$\begin{aligned} \pi_1 &= 0.2\pi_1 + 0.5\pi_2 \\ \pi_2 &= 0.8\pi_1 + 0.5\pi_2 \\ \pi_1 + \pi_2 &= 1 \end{aligned}$$

Thus, $(\pi_1, \pi_2) = (5/13, 8/13)$. Similarly, we have $(\pi_6, \pi_7) = (5/13, 8/13)$. Now you have identified the "$x$"s in the previous matrix arriving to

$$\mathbf{P}^{100} \approx \lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & & \\ & & 0 & 0 & 0 & & \\ & & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & 0 & 5/13 & 8/13 \\ 0 & 0 & 0 & 0 & 0 & 5/13 & 8/13 \end{pmatrix}.$$

Step 4) We shall find the limiting probability going from transient state to recurrent class, i.e. $f_{3,\{1,2\}}, f_{4,\{1,2\}}, f_{5,\{1,2\}}$. Once you find these three numbers out, you will automatically get $f_{3,\{6,7\}}, f_{4,\{6,7\}}, f_{5,\{6,7\}}$.

$$\mathbf{P}^{100} \approx \lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ 5/13 f_{3,\{1,2\}} & 8/13 f_{3,\{1,2\}} & 0 & 0 & 0 & 5/13 f_{3,\{6,7\}} & 8/13 f_{3,\{6,7\}} \\ 5/13 f_{4,\{1,2\}} & 8/13 f_{4,\{1,2\}} & 0 & 0 & 0 & 5/13 f_{4,\{6,7\}} & 8/13 f_{4,\{6,7\}} \\ 5/13 f_{5,\{1,2\}} & 8/13 f_{5,\{1,2\}} & 0 & 0 & 0 & 5/13 f_{5,\{6,7\}} & 8/13 f_{5,\{6,7\}} \\ 0 & 0 & 0 & 0 & 0 & 5/13 & 8/13 \\ 0 & 0 & 0 & 0 & 0 & 5/13 & 8/13 \end{pmatrix}$$

Step 5) It remains us to find $f_{3,\{1,2\}}, f_{4,\{1,2\}}, f_{5,\{1,2\}}$ as explained above. So let's begin filling this out by calculating the absorption probabilities into class $\{1,2\}$

$$
\begin{aligned}
f_{3,\{1,2\}} &= 0.5 \\
f_{4,\{1,2\}} &= 0.8 f_{3,\{1,2\}} + 0.2 f_{5,\{1,2\}} \\
f_{5,\{1,2\}} &= 1
\end{aligned}
$$

Solving by substituting the first and third equations in the second, we get

$$
\begin{aligned}
f_{3,\{1,2\}} &= 1/2 \\
f_{4,\{1,2\}} &= 3/5 \\
f_{5,\{1,2\}} &= 1
\end{aligned}
$$

Obviously, the absorption probabilities into class $\{6,7\}$, can be calculated directly since for the transient states the sum of all absorption probabilities is 1.

$$
\begin{aligned}
f_{3,\{6,7\}} &= 1 - f_{3,\{1,2\}} = 1/2 \\
f_{4,\{6,7\}} &= 1 - f_{4,\{1,2\}} = 2/5 \\
f_{5,\{6,7\}} &= 1 - f_{5,\{1,2\}} = 0
\end{aligned}
$$

Step 6) Deliver solution.

$$\mathbf{P}^{100} \approx \lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ 5/26 & 4/13 & 0 & 0 & 0 & 5/26 & 4/13 \\ 3/13 & 24/65 & 0 & 0 & 0 & 2/13 & 16/65 \\ 5/13 & 8/13 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5/13 & 8/13 \\ 0 & 0 & 0 & 0 & 0 & 5/13 & 8/13 \end{pmatrix}$$

# Exercise 2

- Carefully state the definitions for the terms below.
  - Reach
  - Communicate
  - Class
  - Irreducible
  - Reducible
  - Period
  - Periodic
  - Aperiodic
  - Recurrent
  - Absorbing state
  - Transient

(Solution)

- A state $i$ can ***reach*** a state $j$ and write $i \to j$ if $\exists n$ s.t. $P_{ij}^n > 0$.
- State $i$ and $j$ are said to ***communicate*** and write $i \leftrightarrow j$ if $i \to j$ and $j \to i$.
- A group of states that communicate is said to be a ***class***.
- MC $X_n$ is said to be ***irreducible*** if all states communicate. (Or, equivalently, MC $X_n$ is said to be ***irreducible*** if $\exists$ only one class.)
- MC $X_n$ is said to be ***reducible*** if there are more than one class in MC.
- For a state $i \in S$, ***period*** $d(i) := gcd\{n, P_{ii}^n > 0\}$
- MC $X_n$ is said to be ***periodic*** if $\exists\, i$ with $d(i) > 1$
- MC $X_n$ is said to be ***aperiodic*** if not ***periodic***. (Or, equivalently, MC $X_n$ is said to be ***aperiodic*** if all states have period of 1.)
- A state $i$ is said to be ***recurrent*** if, starting from $i$, the probability of getting back to $i$ in some future is 1.
- A state $i$ is said to be ***absorbing state***, as a special case of recurrent state, if $p_{ii} = 1$.
- A state $i$ is said to be ***transient*** if, starting from $i$, the probability of getting back to $i$ is less than 1. (Remark. The chance of staying in transient class after infinite amount of time is zero.)

## *Exercise 3*

(Solutions to questions below can be all found in the lecture note.)

- What is pmf and cdf for the Poisson distribution?
- For $X \sim Poi(\lambda)$, show that $\mathbb{E}X = \lambda$.
- What is the definition of the stochastic matrix? What is the definition of the doubly stochastic matrix and what is an important theorem regarding its stationary distribution?
- Suppose $X_1 \sim exp(\lambda_1)$, $X_2 \sim exp(\lambda_2)$ and they are independent, then what is the distribution of the random variable $min(X_1, X_2)$? Prove your answer.
- Prove that the time to first arrival in $PP(\lambda)$ follows $exp(\lambda)$.

```
"Quality is never an accident. It is always the result of intelligent effort.
- John Ruskin."
```