

# Stochastic Modelling

isakhammer

2020

## Contents

<b>1</b>	<b>Lecture 1</b>	<b>3</b>
1.1	Practical Information . . . . .	3
1.2	Mathematical description . . . . .	3
1.3	Recall from Statistics Course . . . . .	3
1.3.1	Combining Event . . . . .	3
1.3.2	Probability . . . . .	4
1.3.3	Law of total probability . . . . .	4
1.3.4	Independence . . . . .	4
1.3.5	Random Variables . . . . .	5
1.3.6	Notation for random variables . . . . .	5
1.3.7	Discrete random variables . . . . .	5
1.3.8	CFD . . . . .	6
1.3.9	Continious random vairbales . . . . .	6
1.3.10	Important properties . . . . .	6
1.3.11	Expectation . . . . .	6
1.3.12	Variance . . . . .	7
1.3.13	Joint CDF . . . . .	7
1.3.14	Joint distrubution for discrete random variables . . . . .	7
1.3.15	Joint distrubution for continous random variables . . . . .	7
1.3.16	Independence . . . . .	8
<b>2</b>	<b>Lecture 3</b>	<b>9</b>
2.1	Randoms sum . . . . .	9
2.2	Self Study . . . . .	9
2.3	Stochastic process in descrete time . . . . .	9
2.4	Markov chain . . . . .	10
2.5	Doing n transitions. . . . .	12
<b>3</b>	<b>Lecture 4</b>	<b>14</b>
3.1	Introduction to first step analysis . . . . .	14
<b>4</b>	<b>Lecture 5</b>	<b>19</b>
4.1	Markov Chain in infinitive time . . . . .	21

<b>5</b>	<b>Lecture 2020-09-14</b>	<b>22</b>
	5.0.1 Section 4.5 . . . . .	22
	5.1 Why do we care so much about markov chains? . . . . .	23
	5.2 Continuous Time Markov Chain . . . . .	23
	5.2.1 Section 5.1.2 . . . . .	23
<b>6</b>	<b>Lecture 08/09/20</b>	<b>26</b>
<b>7</b>	<b>Lecture 2020-09-18</b>	<b>29</b>
	7.0.1 Little Oh notation . . . . .	29
<b>8</b>	<b>References</b>	<b>33</b>

# 1 Lecture 1

## 1.1 Practical Information

Two projects

- The projects count 20% and exam 80%.
- Must be done with two people.
- If you want to do statistics is it worth learning  $R$ .

### Course Overview

- Markov chains for discrete time and discrete outcome.
  - Set of states and discrete time points.
  - Transition between states
  - Future depends on the present, but not the past.
- Continuous time Markov chains. (continuous time and discrete outcome.)
- Brownian motion and Gaussian processes (continuous time and continuous outcome.)

## 1.2 Mathematical description

**Definition 1.1.** A *stochastic process*  $\{x(t), t \in T\}$  is a family of random variables, where  $T$  is a set of indices, and  $X(t)$  is a random variable for each value of  $t$ .

## 1.3 Recall from Statistics Course

A random experiment is performed the outcome of the experiment is random.

- The set of possible outcomes is the **sample space**  $\omega$ 
  - An **event**  $A \subset \omega$  if the outcome is contained in  $A$
  - The **complement** of an event  $A$  is  $A^c = \omega \setminus A$
  - The **null event**  $\emptyset$  is the empty set  $\emptyset = \omega \setminus \omega$

### 1.3.1 Combining Event

Let  $A$  and  $B$  be events

- The **union**  $A \cup B$  is the event that at least one of  $A$  and  $B$  occur.
- the **intersection**  $A \cap B$  is the event that both  $A$  and  $B$  occur.

The events  $A_1, A_2, \dots$  are called disjoint (or **mutually exclusive**) if  $A_i \cap A_j = \emptyset$  for  $i \neq j$

### 1.3.2 Probability

$Pr$  is called a probability on  $\omega$  if

- $Pr \{\omega\} = 1$
- $0 \leq P\{A\} \leq 1$  for all events  $A$
- For  $A_1, A_2, \dots$  that are mutually exclusive

$$P\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \sum_{i=1}^{\infty} P\{A_i\}$$

We call  $P\{A\}$  the probability of  $A$ .

### 1.3.3 Law of total probability

Let  $A_1, A_2, \dots$  be a partition of  $\omega$  ie

- $\omega = \bigcup_{i=1}^{\infty} A_i$
- $A_1, A_2, A_3, \dots$  are mutually exclusive.

Then for any event  $B$

$$P\{B\} = \sum_{i=1}^{\infty} P\{B \cap A_i\}$$

**This concept is very important.**

### 1.3.4 Independence

Event  $A$  and  $B$  are independent of

$$P\{A \cap B\} = P\{A\} P\{B\}$$

Events  $A_1, \dots, A_n$  are independent if for any subset

$$P\left\{\bigcap_{j=1}^k A_{i_j}\right\} = \prod_{j=1}^k P\{A_{i_j}\}$$

In this case  $P\{\bigcap_{i=1}^n A_i\} = \prod_{i=1}^n P\{A_i\}$

### 1.3.5 Random Variables

**Definition 1.2.** A *random variable* is a real-valued function on the sample space. Informally: A random variable is a real valued variable that takes on its value by chance.

**Example.**

- Throw two dice.  $X$  = sum of the two dice
- Throw a coin.  $X$  is 1 for heads and  $X$  is 0 for tails.

### 1.3.6 Notation for random variables

We use

- upper case letters such as  $X$ ,  $Y$  and  $Z$  to represent random variables.
- lower case letters as  $x$ ,  $y$ ,  $z$  to denote the real-valued realized value of a the random variable.

Expression such as  $\{X \leq x\}$  denators the event that  $X$  assumes a valye less than or earl to the real number  $x$ .

### 1.3.7 Discrete random variables

The random variable  $X$  is **discrete** if it has a finite or countablle number of possible outcomes  $x_1, x_2, \dots$

- The **probability mass function**  $p_x(x)$  is given by

$$p_x(x) = P\{X = x\}$$

and satisfies

$$\sum_{i=1}^{\infty} p_x(x_i) = 1 \quad \text{and} \quad 0 \leq p_x(x_i) \leq 1$$

- The **cumulative distribution function** (CDF) a of  $X$  can be written

$$F_x(x) = P\{X \leq x\} = \sum_{i: x_i \leq x} p_x(x_i)$$

### 1.3.8 CFD

The CDF of  $X$  may also be called the **distribution function** of  $X$

Let  $F_x(x)$  be the CDF of  $X$ , then

- $F_x(x)$  is monotonally increasing.
- $F_x$  is a stepfunction, which is a piece-wise constant with jumps at  $x_i$ .
- $\lim_{x \rightarrow \infty} F_x(x) = 1$
- $\lim_{x \rightarrow -\infty} F_x(x) = 0$

### 1.3.9 Continuous random variables

A **continuous** random variable takes values on a continuous scale.

- The CDF,  $F_x(x) = P(X \leq x)$  is continuous.
- The **probability density function** (PDF)  $f_x(x) = F'_x(x)$  can be used to calculate probabilities

$$\begin{aligned} Pr\{a < X < b\} &= Pr\{a \leq X < b\} = Pr\{a < X \leq b\} \\ &= Pr\{a \leq X \leq b\} = \int_a^b f_x(x) dx \end{aligned}$$

### 1.3.10 Important properties

- CDF:
  - Monotonically increasing
  - continuous
  - $\lim_{x \rightarrow \infty} F_x = 1$  and  $\lim_{x \rightarrow -\infty} F_x(x) = 0$
- PDF
  - $f_x(x) \geq 0$  for  $x \in \mathbb{R}$
  - $\int_{-\infty}^{\infty} f_x(x) dx = 1$

### 1.3.11 Expectation

Let  $g : \mathbb{R} \rightarrow \mathbb{R}$  be a function and  $X$  be a random variable.

- If  $X$  is discrete, the expected value of  $g(X)$  is

$$E[g(X)] = \sum_{x: p_x(x) > 0} g(x) p_x(x)$$

- If  $X$  is continuous, the expected value of  $g(X)$  is

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_x(x) dx$$

### 1.3.12 Variance

The variance of the random variable  $X$  is

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

Important properties of expectation and variance.

- Expectations is linear

$$E[aX + bY + c] = aE[X] + bE[Y] + c.$$

- Variance scales quadratically and is invariaient to the addition of constants

$$\text{Var}[aX + b] = a^2 \text{Var}[X]$$

- fir independent stochastic variables.

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$

### 1.3.13 Joint CDF

If  $(X, Y)$  is a pair for random variables, their **joint comulative distribution function** is given by

$$F_{X,Y} = F(x, y) = \text{Pr}\{X \leq x \cap Y \leq y\}$$

### 1.3.14 Joint distrubution for discrete random variables

If  $X$  and  $Y$  are discrete, the **joint probability mass function**  $p_{x,y} = \text{Pr}\{X = x, Y = y\}$ . can be used to compute probabilities

$$\text{Pr}\{a < X < b, c < Y \leq d\} = \sum_{a < x \leq b} \sum_{c < y \leq d} p_{X,Y}(x, y)$$

### 1.3.15 Joint distrubution for continous random variables

If  $X$  and  $Y$  are continous the **joint probability density function**

$$f_{X,Y}(x, y) = f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$$

can be used to compute probabilities

$$\text{Pr}\{a < X \leq b, c < Y \leq d\} = \int_a^b \int_c^d f(x, y) dx dy$$

### 1.3.16 Independence

The random variables  $X$  and  $Y$  are independent if

$$\Pr\{X \leq a, Y \leq b\} = \Pr\{X \leq a\} \cdot \Pr\{Y \leq b\}, \quad \forall a, b \in \mathbb{R}$$

In terms of CDFs:  $F_{X,Y}(a, b) = F_X(a) \cdot F_Y(b) \quad \forall a, b \in \mathbb{R}$

Thus we have

- $p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$  for discrete random variables
- $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$  for continuous random variables.



## 2 Lecture 3

### 2.1 Randoms sum

Building on the hunter example from last week. we can more generally consider random sums

$$X = \begin{cases} 0, & N = 0 \\ \zeta_1 + \zeta_2 + \dots + \zeta_N, & N > 0 \end{cases}$$

where

- $N$  is a discrete random variable with values  $0, 1, \dots$
- $\zeta_1, \zeta_2, \dots$  are independent random variables
- $N$  is independent of  $\zeta_1, \zeta_2 + \dots + \zeta_N$
- **Notation**  $X = \sum_{i=1}^N \zeta_i = \zeta_1 + \zeta_2 + \dots + \zeta_N$

#### Example.

1. Insurance company

$N$  : Number of claims.

$\zeta_1, \zeta_2, \dots$  : Sizes of the claims

Total liability:

$$X = \zeta_1 + \zeta_2 + \dots + \zeta_N$$

2. Be careful!

$$\begin{aligned} \overbrace{E \left[ \sum_{i=1}^N \zeta_i \right]}^{\neq \sum_{i=1}^N E[\zeta_i]} &= E \left[ E \left[ \sum_{i=1}^N \zeta_i \mid N \right] \right] \\ &= E \left[ \sum_{i=1}^N E[\zeta_i \mid N] \right] \end{aligned}$$

### 2.2 Self Study

Section 2.2, 2.3, 2.4

### 2.3 Stochastic process in discrete time

**Definition 2.1.** A **discrete-time stochastic process** is a family of random variables  $[X_t : t \in T]$  where  $T$  is discrete.

- We use  $T = \{0, 1, 2, \dots\}$  and write  $X_n$  instead of  $X_t$
- we call  $X_n$  the **state** at time  $n = 0, 1, 2, 3, \dots$
- We call the set of all possible states the **state space**

Table 1: Table for example

Day	$n = 0$	$n = 1$	$n = 2$	$\dots$
Random Variable	$X_0$	$X_1$	$X_2$	$\dots$
Realization 1	$x_0 = 0$	$x_1 = 1$	$x_2 = 1$	$\dots$
Realization 2	$x_0 = 1$	$x_1 = 1$	$x_2 = 1$	$\dots$

**Example.**

$$X_n = \begin{cases} 1, & \text{if it rains on day } n \\ 0, & \text{no rain on day } n \end{cases}$$

State space =  $\{0, 1\}$

**We have a problem.** Need

$$Pr \{X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_0 = x_0\}.$$

for all  $n = 0, 1, 2, \dots$

## 2.4 Markov chain

**Definition 2.2** (Discrete time Markov Chain). A **Discrete time markoc chain** is a discrete time stochastic process  $\{X_n : n = 0, 1, \dots\}$  that statisfied the **markov property** such that

$$\begin{aligned} Pr \{X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} \\ = Pr \{X_{n+1} = j \mid X_n = i\} \end{aligned}$$

for  $n = 0, 1, 2, 3, \dots$  and for all states  $i$  and  $j$

**Definition 2.3** (One-step transition probabilities). We can define it as

- For a discrete Markov chain  $\{X_n : n = 0, 1, 2, \dots\}$  we call  $P_{ij}^{n,n+1} = Pr \{X_{n+1} = j, X_n = i\}$  the **one step trainstition probabilities**.

- We will assume **stationary transition probabilities** , i.e that

$$P_{ij}^{n,n+1} = P_{ij}$$

for  $n = 0, 1, 2, \dots$  and all states  $i$  and  $j$  .

Some of the properties

1. "You will always go somewhere"

$$\sum_j P_{ij} = 1 \quad \forall i$$

2. The markov chain can be described as follows.

$$\begin{aligned} & Pr \{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \\ &= Pr \{X_0 = i_0\} Pr \{X_1 = i_1 \mid X_0 = i_0\} \dots \\ &\quad Pr \{X_n = i_n \mid X_{n-1} = i_{n-1} \dots X_0 = i_0\} \\ &\quad \vdots \quad \text{Markov step} \\ &= Pr \{X_0 = i_0\} \cdot Pr \{X_1 = i_1 \mid X_0 = i_0\} \dots \\ &\quad Pr \{X_n = i_n \mid X_{n-1} = i_{n-1}\} \\ &= Pr \{X_0 = i_0\} P_{i_0, i_1} \cdot P_{i_1, i_2} \dots P_{i_{n-1}, i_n} \end{aligned}$$

Which is a major simplification.

**Definition 2.4** (Transition Probability Matrix). For a discrete time markov-chain with state space  $\{0, 1, \dots, N\}$  we call

$$\mathbf{P} = \begin{bmatrix} P_{00} & \dots & P_{0N} \\ P_{10} & \dots & \\ \vdots & & \ddots \\ P_{N0} & \dots & P_{NN} \end{bmatrix}$$

Is the transition matrix. For statespace  $\{0, 1, 2, \dots\}$  we envision an infinitely sized matrix.

**Example.**

- Markoc chain :  $\{X_n : n = 0, 1, 2, \dots\}$
- State space =  $\{0, 1\}$
- Transition Matrix

$$\mathbf{P} = \begin{bmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{bmatrix}$$

We can compute

$$\begin{aligned} Pr \{X_3 = 1 \mid X_2 = 0\} &= p_{01} \\ &= 0.1 \end{aligned}$$

$$\begin{aligned} Pr \{X_{10} = 0 \mid X_9 = 1\} &= P_{10} \\ &= 0.6 \end{aligned}$$

**Definition 2.5** (Transition Diagram). *Let  $\{X_n : n = 0, 1, \dots\}$  be a discrete time Markov chain. A **state transtion diagram** visualizes the transition probabilities as a weighted directed graph where the nodes are the states and the edges are the possible transitions marked with the transistion probabilities.*

**Example.** State space =  $\{0, 1, 2\}$  and

$$P = \begin{bmatrix} 0.95 & 0.05 & 9 \\ 0 & 0.9 & 0.1 \\ 0.01 & 0 & 0.99 \end{bmatrix}$$

Transistion diagram

Nice figure of the diagram

## 2.5 Doing n transitions.

**Theorem 2.1.** *For a Markoc chain  $\{X_n : n = 0, 1, \dots\}$  and any  $m \geq 0$  we have*

$$Pr \{X_{m-n} = j \mid X_m = i\} = P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)}, \quad n > 0$$

where we define

$$P_{ij}^{(0)} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

*Proof.* Set  $m = 0$  then is

$$\begin{aligned} P_{ij}^{(n+1)} &= Pr \{X_{n+1} = j \mid X_0 = i\} \\ &= \sum_k Pr \{X_{n+1} = j, X_1 = k \mid X_0 = i\} \\ &= \sum_k Pr \{X_{n+1} = j \mid X_1 = k, X_0 = i\} \cdot Pr \{X_1 = k \mid X_0 = i\} \\ &= \sum_k P_{kj}^{(h)} \cdot P_{ik} = \sum_k P_{ik} P_{kj}^{(h)} \end{aligned}$$

□

**Example.**  $\{X_n : n = 0, 1, 2, \dots\}$  is a markoc chain and

$$P = \begin{bmatrix} 0.1 & 0.9 \\ 0.6 & 0.4 \end{bmatrix}$$

Find  $P_{01}^{(4)}$ . **Solution.**

$$P^2 = \begin{bmatrix} 0.55 & 0.45 \\ 0.30 & 0.70 \end{bmatrix}$$

So by doing matrix multiplication and we end up with

$$P^4 = P^2 \cdot P^2 = \begin{bmatrix} 0.4375 & 0.5625 \\ 0.3750 & 0.6250 \end{bmatrix}$$

Which therefore ends up with the answer

$$P_{01}^{(4)} = 0.5625$$

### 3 Lecture 4

#### 3.1 Introduction to first step analysis

##### Input

- $i_0$  : starting state
- $P$  : transition probability matrix
- $T$ : number of time steps

##### Algorithm

1. Set  $x_0 = i_0$
2. for  $n = 1 \dots T$
3.     Simulate  $x_n$  from  $X_n \mid X_{n-1} = x_{n-1}$
4. end

**output** : One realization  $x_0, x_1, \dots, x_T$

##### Example.

$$P = \begin{pmatrix} 0.95 & 0.05 & 0 \\ 0 & 0.90 & 0.10 \\ 0.01 & 0 & 0.99 \end{pmatrix}$$

Let  $x_0 = 0$

1.  $x_0 = 0$
- 2.

$$\begin{aligned} Pr \{X_1 = 0 \mid X_0 = 0\} &= P_{00} = 0.95 \\ Pr \{X_1 = 1 \mid X_0 = 0\} &= P_{01} = 0.05 \\ Pr \{X_1 = 2 \mid X_0 = 0\} &= P_{02} = 0 \\ &\vdots \end{aligned}$$

Assume we get  $x_1 = 1$

3. States

•

$$\begin{aligned} 0 : P_{10} &= 0 \\ 1 : P_{11} &= 0.90 \\ 2 : P_{12} &= 0.10 \\ &\vdots \end{aligned}$$

General notes on simulation

- $Pr\{A\} \approx \frac{\text{times A occur}}{\text{Simulations}}$
- $E[X] \approx \frac{1}{N} \sum_{i=1}^N x_i$

**Example.** We have  $N = 100$  divided into two containers labelled  $A$  and  $B$ . At each time  $n$ , one ball is selected at random and moved to the container. Let  $Y_n$  denote the number of balls in container  $A$  at time  $n$ , and define  $X_n = Y_n - 50$ . Find the transition probabilities and simulate and plot one realization of

$$\{X_n : n = 0, 1, \dots, 500\}$$

**Answer**

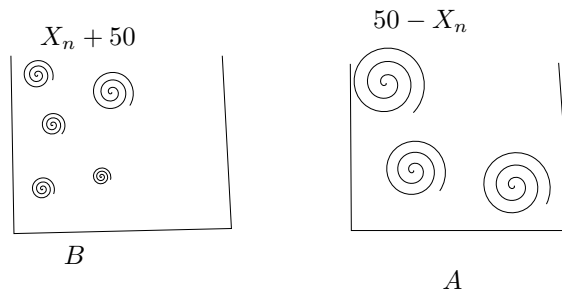


Figure 1: balls

- Only move One ball
- Can move only from  $i$  to  $j = i - 1$  or  $j = i + 1$

$$P_{ij} = \begin{cases} \frac{50-i}{100} & , j = i + 1 \\ \frac{50+i}{100} & , j = i - 1 \\ 0 & , \text{otherwise.} \end{cases}$$

**Motivation**

**Definition 3.1.** For a markov chain, a state  $i$  such that  $P_{ij} = 0 \forall j \neq i$  is

called **absorbing**.

**Example.** Let  $\{X_n\}$  be a Markov chain with transition probability matrix

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta & \gamma \\ 0 & 0 & 1 \end{pmatrix}$$

where  $\alpha, \beta, \gamma > 0$  and  $\beta = 1 - \alpha - \gamma$ . Assume  $x_0 = 1$

1. What is the expected time until absorption ?
2. What is the probability to be absorbed in state 0 ?

**Realization .**

4 steps to absorption  
 $\overbrace{1, 1, 1, 1, 1, 2}^{4 \text{ steps to absorption}}, 2, 2, \dots$

**Mathematically**

Let  $T = \min \{n \geq 0 : X_n = 0 \text{ or } X_n = 2\}$ . Then is

$$Q1 : E[T \mid X_0 = 1]$$

$$Q2 : Pr\{X_T = 0 \mid X_0 = 1\}$$

The idea of first step analysis is to define

- $T^{(n)} = \min \{n \geq 0 : X_{m \times n} = 0 \text{ or } X_{m+b} = 2\}$
- $T = T^{(0)}$
- $v_i^{(m)} = E[T^{(m)} \mid X_m = i]$
- $v_i = v_i^{(0)}$

Table 2: Let  $m$  be timesteps

$m$	0	2	3	4	5
$v_0^{(m)}$	0	0	0	0	0
$v_1^{(m)}$	$v_1$	$v_1$	$v_1$	$v_1$	$v_1$
$v_2^{(m)}$	0	0	0	0	0



### First step analysis for Q1

$$\begin{aligned}
 v_i &= \sum_{k=0}^2 Pr \{X_1 = k \mid X_0 = i\} (1 + v_k) \\
 &= \sum_{k=0}^2 P_{ik} (1 + v_k) = \sum_{k=0}^2 P_{ik} v_k + 1 \quad \text{which is true for } i = 0, 1, 2
 \end{aligned}$$

Which is reduced to linear algebra. Solving it by

$$\begin{aligned}
 v_0 &= v_2 = 0 \\
 \implies v_1 &= \alpha v_0 + \beta v_1 + \gamma v_2 + 1 \\
 \implies v_1 &= \frac{1}{1 - \beta} \quad [\text{Q1}]
 \end{aligned}$$

$$P_{ij} \implies i = \text{row}, \quad j = \text{column}$$

First step analysis and let

$$\begin{aligned}
 u_i &= Pr \{X_T = 0 \mid X_0 = i\} \\
 &\downarrow \\
 u_i &= \sum_{k=0}^2 P_{ik} u_k, \quad i = 0, 1, 2
 \end{aligned}$$

- Easy:  $u_0 = 1, u_2 = 0$
- Harder:  $u_1 = \alpha u_0 + \beta u_1 + \gamma u_2$  such that

$$u_1 = \alpha \frac{1}{1 - \beta} = \frac{\alpha}{\alpha - \beta} \quad [\text{Q2}]$$

**Example.** let  $[X_n]$  be a markov chain with transition matrix

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The starting state is  $x_0 = 1$ . Calculate the probability to be absorbed in the state  $D$ .

1. Define  $u_i = Pr \{ \text{absorbed in state } 0 \mid X_0 = i \}$  for  $i = 0, 1, 2, 3$
2. Get the easy ones out of the way. In this case  $u_0 = 1$  and  $u_3 = 0$
- 3.

$$\begin{aligned} u_1 &= P_{10}u_0 + P_{11}u_1 + P_{12}u_2 + P_{13}u_3 \\ &= 0.4 + 0.3u_1 + 0.2u_2 \\ u_2 &= P_{20}u_0 + P_{21}u_1 + P_{22}u_2 + P_{23}u_3 \\ &= 0.1 + 0.3u_1 + 0.3u_2 \end{aligned}$$

4. Solve for  $u_1$  and  $u_2$

## 4 Lecture 5

**Example.** Let  $P$  be the matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

With starting state  $x_0 = 1$

1. Define  $T = \min_{n \geq 0: X_n = 0} \quad X_n = 3$  and  $v_i = E[T \mid X_0 = i]$  for  $i = 0, 1, 2, 3$

2. Set  $v_0 = v_3 = 0$

3.

$$v_1 = P_{10}v_0 + P_{11}v_1 + P_{12}v_2 + P_{13}v_3 = 0.3v_1 + 0.2v_2 + 1$$

and

$$v_2 = P_{20}v_0 + P_{21}v_1 + P_{22}v_2 + P_{23}v_3 + 1 = 0.3v_1 + 0.3v_2 + 1$$

4. Solve the equations and end up with

$$v_1 = \frac{90}{43} \quad \text{and} \quad v_2 = \frac{100}{43}$$

**Theorem 4.1.** Let  $\{X_n\}$  be a discrete time Markov chain with state space  $S = \{0, 1, \dots, N\}$  and transition probability matrix  $\mathbf{P}$ . Let  $A \subset S$  be the set of absorbing state. Then

1. If  $v_i$  is the expected time to absorption conditional on  $X_0 = i$  then

$$v_i = 0, \quad i \in A$$

$$v_i = 1 + \sum_{i \in \mathbb{R}} P_{ik}v_k \quad i \in A^c$$

**Example.** A gambler has 10\$ and bets 1\$ If he wins the round, his fortune increases 1\$. The probability of winning each round is  $0 < p < 1$  and the probability of losing each round is  $q = 1 - p$ . The gambler will continue gambling until his fortune is \$  $N$  or 0\$ where  $N > 10$ . What is the probability the gambler will be ruined.

1. Extract the essential stuff.

$$X_n = \text{Fortune at time } n, \quad n = 0, 1, 2, \dots$$

$$\text{State space} = \{0, 1, \dots, N\}$$

$$\text{Target: } u_k = \Pr \{ \text{Absorption in state 0} \mid X_0 = k \}, \quad k = 0, 1, \dots, N$$

2. Visualize the transitions. Insert figure of transitions.
3. Make the eprobability matrix. The rows are "to" and the columns are "1"

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ q & 0 & p & 0 & \dots & 0 \\ 0 & q & 0 & p & \dots & \\ \vdots & & \ddots & & & \\ 0 & 0 & \dots & q & 0 & p \\ & & & & & 1 \end{bmatrix}$$

4. Set up the iteration

$$\begin{aligned} u_0 &= 1, \quad u_N = 0, \quad \text{Easy} \\ u_i &= P_{i,i,1}u_{i-1} + P_{i,i+1}u_{i+1} \\ &= qu_{i-1} + pu_{i+1}, \quad i = 1, 2, \dots, N-1 \end{aligned}$$

5. (a)

$$\begin{aligned} \overbrace{(p+q)}^{=1} u_i &= qu_{i-1} + pu_{i+1} \\ q[u_i - u_{i-1}] &= p[u_{i+1} - u_i] \\ \downarrow \quad \text{Trick} \quad \chi_i &= u_i - u_{i-1} \\ q\chi_1 &= p\chi_{i+1}, \quad \implies \chi_{i+1} = \frac{q}{p}\chi_i \quad i = 1, 2, \dots, N \end{aligned}$$

- (b)

$$\begin{aligned} \chi_1 + \chi_2 + \dots + \chi_k &= [u - u_0] + [u_2 - u_1] + \dots + [u_k - u_{k-1}] \\ \downarrow \quad \text{Telescoping sum} \\ \chi_1 \left[ 1 + \frac{q}{p} + \left(\frac{q}{p}\right)^2 + \dots + \left(\frac{q}{p}\right)^{k-1} \right] &= u_k - 1, \quad k = 1, 2, 3, \dots, N \end{aligned}$$

For  $k = N$  :

$$\begin{aligned} \chi_1 &= \frac{u_N - 1}{\sum_{k=0}^{N-1} \left(\frac{q}{p}\right)^k} = \frac{-1}{\sum_{k=0}^{N-1} \left(\frac{q}{p}\right)^k} \\ &= \begin{cases} -\frac{1}{N} & , q = p = \frac{1}{2} \\ -\frac{(1-\frac{q}{p})}{(1-(\frac{q}{p})^N)} & q \neq p \end{cases} \end{aligned}$$

(c) From the telescoping sum

$$u_k = 1 + \chi_1 \sum_{i=0}^{k-1} \left(\frac{q}{p}\right)^i$$

$$= \begin{cases} 1 - \frac{1}{N} \cdot k = \frac{N-k}{N}, & p = q = \frac{1}{2} \\ 1 - \frac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}, & p \neq q \end{cases}, \quad \text{where } k = 1, 2, \dots$$

6. The final step

$$u_{10} = \begin{cases} \frac{N-10}{N}, & p = q = \frac{1}{2} \\ \frac{\left(\frac{q}{p}\right)^{10} - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}, & q \neq p \end{cases}$$

*Remark.* • When  $N \rightarrow \infty$

$q \geq p \implies$  Almost certain you will loose.

$$q < p \implies P(\text{ruined}) = \left(\frac{q}{p}\right)^{10}$$

#### 4.1 Markov Chain in infinitive time

**Definition 4.1. Regular Markov Chain .** Consider a Markov chain  $\{X_n : n = 0, 1, \dots\}$  with finite state space  $\{0, 1, 2, \dots\}$  and transition matrix  $\mathbf{P}$ . IF there exists an integer  $k > 0$  so that all regular elements  $\mathbf{P}^k$  are strictly positive, we call  $\mathbf{P}$  and  $\{X_n\}$  regular.

*Remark.* 1.  $P$  is regular means that it exists an  $k > 0$  so that  $P_{ij}^{(k)} > 0 \quad \forall i, j$

2. If  $P_{ij}^{(k)} > 0 \quad \forall i, j$ , then is  $P_{ij}^{(K)} > 0 \quad \forall i, j$  and  $K \geq k$

## 5 Lecture 2020-09-14

Find Stationary distributions

(i)  $\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

- Positive recurrent, aperiodic and irreducible.
- $\implies$  Limiting distribution:

$$\pi = \left( \frac{1}{2}, \frac{1}{2} \right)$$

(ii)  $\mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

- Positive recurrent and irreducible.
- unique stationary distribution.
- $\pi = \left( \frac{1}{2}, \frac{1}{2} \right)$

(iii)  $\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

- Reducible!
- Part 1:

$$\pi_0 = 1\pi_0 + 0\pi_1 = \pi_0$$

$$\pi_1 = 0\pi_0 + 1\pi_1 = \pi_1$$

$$\implies \pi = 1 - \pi_1$$

$$\implies \pi = (\pi_1, 1 - \pi_1)$$

- Part 2:

Must have

$$\pi_0 \geq 0$$

$$\pi_1 \geq 0$$

$$\implies \pi = (\pi_0, 1 - \pi_0), \quad 0 \leq \pi_0 \leq 1$$

### 5.0.1 Section 4.5

Read it yourself .

## 5.1 Why do we care so much about markov chains?

- (i) Importance goes far beyond statistical modelling of physical phenomena.
- (ii) In the end of the 80s and start of 90s the computational power was growing stronger.
- (iii) We realized that we could sample from difficult distribution by constructing Markov chains whose stationary matched desired target distribution.
- (iv) The theory we have discussed of the theory developed to show that these methods worked.

## 5.2 Continuous Time Markov Chain

**Definition 5.1.** The stochastic variable  $X$  has a **Poisson distribution** with (mean) parameter  $\mu > 0$  if

$$p(x) = \frac{\mu^x}{x!} e^{-\mu}$$

We write  $X \sim \text{Poisson}(\mu)$

*Remark.*  $X \sim \text{Poisson}(10)$

- (i)  $E[X] = \mu$
- (ii)  $\text{Var}[X] = \mu$
- (iii)  $\text{SD}[X] = \sqrt{\mu}$

**Theorem 5.1.** If  $X \sim \text{Poisson}(\mu)$ ,  $Y \sim (\chi)$  and  $Y$  are independent.

**Theorem 5.2.** If  $N \sim \text{Poisson}(\mu)$  and  $M | N \sim \text{Binomial}(N, p)$  then

$$M \sim \text{Poisson}(\mu p)$$

*Remark.* (i)  $M = \sum_{k=1}^N I_k$ , where  $I_1, I_2, \dots \sim \text{Bernoulli}(p)$  and  $I_1, I_2, \dots$  and  $N$  are independent.

- (ii) This is called **thinning**.

### 5.2.1 Section 5.1.2

**Definition 5.2.** A *Possion process* with rate *intensity*  $\lambda > 0$  is an integer-valued stochastic process  $\{X(t) : t \geq 0\}$  for which.

- For any  $n > 0$  and any time point  $0 < t_0 < t_1 < \dots < t_n$  the increments

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$$

are independent

- For  $s \geq 0$  and  $t > 0$

$$X(s+t) - X(s) \sim \text{Poisson}(\lambda t)$$

- $X(0) = 0$

*Remark.* • 1. is called independent increments

- In 2, we have

$$X(s + \Delta t) - X(s) \sim \text{Poisson}(\lambda \Delta t)$$

- Illustration

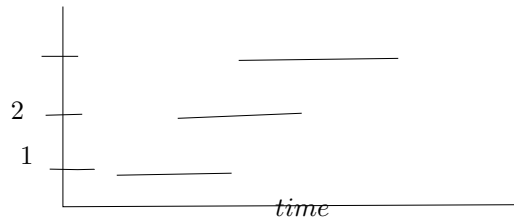


Figure 2: illustration

- $X(t) = X(t) - X(0) \sim \text{Poisson}(\lambda t)$



**Example.** We assume the arrival of customers to a store follows a Poisson process with rate  $\lambda = 4$  customers per hours. The store opens at 09:00. What is the probability that exactly one customer has arrived by 09:30 and exactly five customers have arrived by 11:30.

**Answer.** Let  $X(t)$  = arrivals by time  $t$  For  $t \geq 0$  (measured in hours). Then is the question

$$\begin{aligned}
 & Pr \left\{ X \left( \frac{1}{2} \right) = 1, X \left( \frac{5}{2} \right) = 5 \right\} \\
 & \quad \downarrow = \text{Rephrase as increments} \\
 & \quad = Pr \left\{ X \left( \frac{1}{2} \right) - X(0) = 1, X \left( \frac{5}{2} \right) - X \left( \frac{1}{2} \right) = 4 \right\} \\
 & \quad \downarrow \text{Independent increments} \\
 & \quad = Pr \left\{ X \left( \frac{1}{2} \right) - X(0) = 1 \right\} \cdot \overbrace{Pr \left\{ X \left( \frac{5}{2} \right) - X \left( \frac{1}{2} \right) = 4 \right\}}^{Poisson(2\lambda)} \\
 & \quad \quad \quad \underbrace{Poisson(\frac{1}{2}\lambda)} \\
 & \quad = \frac{2^1}{1!} e^{-2} \cdot \frac{8^4}{4!} e^{-8} \\
 & \quad = 0.0155
 \end{aligned}$$

**Example.** Assume the arrival of customers to follows an inhomogeneous Poisson process with rate  $\lambda(t) = t$ ,  $t \geq 0$ . Assume the store opens at 09:00. What is the probability that no-one has arrived at 10:00.

**Answer.**

$$X(1) - X(0) \sim Poisson \left( \overbrace{\int_0^1 t dt}^{=\frac{1}{2}} \right)$$

## 6 Lecture 08/09/20

Equivalent classes and classifications of states in Markov chains.

Things to check

- Understand why regularity fails.
- Extend regularity to infinite spaces.

**Example** Let  $\{X_n : 0, 1, \dots, N\}$  be a markov chain.

- (a) It can go from  $0 \rightarrow 0$  and  $1 \rightarrow$  with probabilities  $p_{00} = p_{11} = 1$ , two separate markov chains. Realizations :

0, 0, 0, 0, 0, 0, ...

1, 1, 1, 1, 1, 1, ...

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \implies P^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

**Definition 6.1.** Let  $\{X_n : 0, 1, \dots\}$  be a Markov chain with state space  $\{0, 1, \dots\}$  then is

- (i) State  $j$  is **accessible** from state  $i$  if  $\exists n \geq 0$  so that  $P^{(n)}_{ij} > 0$
- (ii) If states  $i$  and  $j$  are accessible from each other they are said to **communicate** we write  $i \sim j$ . If states  $i$  and  $j$  do not communicate we write  $i \not\sim j$

*Remark.* If  $i \not\sim j$ , then either (or both)

- (a) (i)  $P^{(n)}_{ij} = 0, \quad \forall n \geq 0$   
(ii)  $P_{ji} = 0, \quad \forall n \geq 0$

- (b) Only the graph matters, not the values of the edges.

(c)  $P^{(0)}_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

**Theorem 6.1.** Communication is an **equivalence relation**

- (i) **reflexive**,  $i \sim i$
- (ii) **symmetric**  $i \sim j \implies j \sim i$
- (iii) **Transitive**  $i \sim j$  and  $j \sim k$  implies  $i \sim k$

A equivalence relation induces **equivalence classes** of sets of states that communicate.

*Proof.* (i)  $P_{ii}^{(0)} = 1 \implies i \sim i$

(ii) By definition is this true.

(iii) (a)  $i \sim j \implies \exists n \geq 0 : P_{ij}^{(n)} > 0$

$$j \sim k \implies \exists m \geq 0 : P_{jk}^{(m)} > 0$$

(b) Chapman-kilogram

$$P_{ik}^{(n+m)} = \sum_{r=0}^{\infty} P_{ir}^{(n)} P_{rk}^{(m)} \geq P_{ij}^{(n)} P_{jk}^{(m)}$$

$\implies k$  is accessible from  $i$ .

(c) Show yourself

**$i$  is accessible from  $k$**

□

**Definition 6.2.** A Markov chain is **irreducible** if  $\sim$  (communication) induces exactly one equivalent class. If not, it is called reducible.

**Definition 6.3.** The **period** of state  $i$ , written as  $d(i)$  is

$$d(i) = \gcd \left\{ n \geq 1 : P_{ii}^{(n)} > 0 \right\}$$

If  $P_{ii}^{(n)} = 0$  for all  $n \geq 1$ , we define  $d(i) = 0$ . If  $d(i) = 1$ , we call the state  $i$  is **aperiodic**.

**Theorem 6.2.** if  $i \sim j$ , then  $d(i) = d(j)$

*Remark.* The period is a property of the equivalence class.

**Notation** The state space may be infinite:  $\{0, 1, \dots\}$ . We introduce

(i) The probability the first return happend after exactly  $n$  steps

$$f_{ii}^{(n)} = \Pr \{X_n = i, X_\mu \neq i, \mu = 1, 2, \dots, n-1 \mid X_0 = i\} \quad n > 0$$

We will define  $f_{ii}^{(0)} = 0$

(ii) The probability of returning at some time

$$f_{ii} = \sum_{k=0}^{\infty} f_{ii}^{(k)} = \lim_{n \rightarrow \infty} \sum_{k=0}^n f_{ii}^{(k)}.$$

*Remark.*  $f_{ii} < 1 \leftrightarrow$  Positive probability of never returning to  $i$

**Definition 6.4.** State  $i$  is **recurrent** if the probability of returning to state  $i$  in a finite number of timesteps is one  $f_{ii} = 1$ . A state that is not recurrent  $f_{ii} < 1$  is called **transient**.

**Theorem 6.3.** A state  $i$  is recurrent if and only if

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty$$

Equivalently, state  $i$  is transient if and only if

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$$

*Proof.* (i)

$$\begin{aligned} \sum_{n=1}^{\infty} P_{ii}^{(n)} &= \sum_{n=1}^{\infty} E[\mathbb{I}\{X_n = i\} \mid X_0 = i] \\ &= E\left[\sum_{n=1}^{\infty} \mathbb{I}\{X_n = i \mid X_0 = i\}\right] \\ &= E[M \mid X_0 = i] \\ M &\rightarrow \text{Returns to state.} \end{aligned}$$

$$(ii) \ E[M \mid X_0 = i] = \begin{cases} f_{ii} \frac{1}{1-f_{ii}}, & f_{ii} < 1 \\ \infty, & f_{ii} = 1 \end{cases}$$

□

## 7 Lecture 2020-09-18

Read Section 5.1.4 by yourself.

### Section 5.2 Motivation

- (a)  $\{X(t) : t \geq 0\}$  with rate  $\lambda_1 = 5$ ,  $0 \leq t \leq 10$

$$E[X(t)] = \lambda t = 5t,$$

- (b)  $\{Y(t) : t \geq 0\}$  with rate  $\lambda_2 = t$ ,  $0 \leq t \leq 10$

$$E[Y(t)] = \frac{t^2}{2}$$

Do scatterplot on the project when working on poisson distribution.

**Theorem 7.1.** Let  $p_1, p_2, \dots \in [0, 1]$  be a sequence such that  $\lim_{n \rightarrow \infty} np_n = \lambda < \infty$ , then

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \lambda^n \frac{1}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$

*Remark.* In TMA4295 Statistical Inference we will say that  $Binomial(n, p_n)$  converges in  $Poission(\lambda)$  as  $n \rightarrow \infty$ .

*Remark.* .

- (i)  $p_n \rightarrow 0$ , but  $n \rightarrow \infty$ .  $np_n \rightarrow \lambda$  when  $n \rightarrow \infty$
- (ii) Many trials ( $n \gg 1$ ) and success is rare ( $p \ll 1$ )  $\implies$  Nr of Successes  
Poisson distribution.

Typical examples

- Customers arrivals.
- Car accident.
- Telephone calls.

### 7.0.1 Little Oh notation

- (i) You may be familiar with the expressions such as

$$n = o(n^2), \quad \text{as } n \rightarrow \infty$$

May be thought as "n is much smaller than  $n^2$  as  $n \rightarrow \infty$ "

(ii) We are going to mostly work with expressions of the form

$$h^2 = o(h), \quad h \rightarrow 0^+$$

May be thought as " $h^2$  is much smaller than  $h$  as  $h \rightarrow 0^+$ "

**Definition 7.1.** Let  $f$  and  $g$  be real functions. We use *little-oh-notation* in the two following ways

$$(i) \quad f(n) = o(g(n)), \quad n \rightarrow \infty \implies \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$$

$$(ii) \quad f(h) = o(g(h)) \quad h \rightarrow 0^+ \implies \lim_{h \rightarrow 0^+} \frac{f(h)}{g(h)} = 0$$

**Example.** Are the following statements false or true?

$$(i) \quad h^2 = o(h) \quad h \rightarrow 0^+$$

$$\lim_{h \rightarrow 0^+} \frac{h^2}{h} = 0$$

True

$$(ii) \quad h^2 = o(h) \quad h \rightarrow \infty$$

$$\lim_{h \rightarrow \infty} \frac{h^2}{h} = \lim_{h \rightarrow \infty} h = \infty$$

False

$$(iii) \quad \sqrt{h} = o(h) \quad h \rightarrow 0^+$$

$$\lim_{h \rightarrow 0^+} \frac{\sqrt{h}}{h} = \infty$$

False

$$(iv) \quad h \rightarrow o(1) \quad h \rightarrow 0^+$$

$$\lim_{h \rightarrow 0^+} \frac{h}{1} = 0$$

True

*Remark.*

$$h^p = o(h) \quad h \rightarrow 0^+ \implies p > 1$$

**Definition 7.2.** A *Cprocess* is a stochastic process  $\{N(t) : t \geq 0\}$  so that

(i)  $N(t)$  is a integer for  $t \geq 0$

(ii)  $N(t) \geq 0$ , for  $t \geq 0$

(iii) If  $s \geq t$ , then  $N(s) \leq N(t)$

We sometimes write

$$N(a, b) = N(b) - N(a) = \text{Number of events in } (a, b], \quad 0 \leq a \leq b$$

However, the notation will not be used in the lecture.

**Definition 7.3.** Let  $\{N(t) : t \geq 0\}$  be a counting process. Then  $\{N(t) : t \geq 0\}$  is a **Poisson process** with **rate (intensity)**  $\lambda > 0$  if

(i) For every integer  $m > 1$  for any timepoints

$$0 = t_0 < t_1 < \dots < t_m$$

$$N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_m) - N(t_{m-1})$$

"independent increments"

(ii) For  $t \geq 0$  and  $h > 0$ , the distribution of  $N(t+h) - N(t)$  only depends on  $h$  and  $t$ . "Stationary Increments"

(iii)  $\Pr\{N(t+h) - N(t) = 1\} = \lambda h + o(h), \quad h \rightarrow 0^+ \quad \forall t \geq 0$

(iv)  $\Pr\{N(t+h) - N(t) = 0\} = 1 - \lambda h + o(h), \quad h \rightarrow 0^+ \quad \forall t \geq 0$

(v)  $N(0) = 0$

For def iii and iv can be described as

$$\begin{aligned} \Rightarrow \Pr\{N(t+h) - N(t) \geq 2\} &= 1 - \overbrace{[\lambda h + o(h)]}^{1 \text{ event}} - \overbrace{[1 - \lambda h + o(h)]}^{0 \text{ events}} \\ &= o(h) \\ &\Rightarrow \text{Events cannot occur at the same time} \\ &\Rightarrow \text{Jumps are of size 1} \end{aligned}$$

Recall

**Definition 7.4.** (Simplified version.) A **Poisson process** with rate **rate**  $\lambda > 0$  is an integer valued stochastic process  $\{N(t) : t \geq 0\}$  for which

(i) Increments are independent,

(ii) For  $s \geq 0$  and  $t > 0$

$$N(s+t) - N(s) \sim \text{Poisson}(\lambda t)$$

$$(iii) \ N(0) = 0$$

**Theorem 7.2.** *Definition of simplified and general of a Poisson process are equivalent.*

*Proof.* Let's call the simplified version P1 and the general version P2, then we need to prove

- Prove that  $P1 \implies P2$ : i), ii) and v) is proved by definition.

$$\begin{aligned} Pr \{N(t+h) - N(t) = 1\} &= \frac{(\lambda h)^1}{1!} e^{-\lambda h} \\ &= \lambda h (1 - \lambda h o(h)), \quad \text{as } h \rightarrow 0^+ \\ &= \lambda h - \lambda^2 h^2 + \lambda h o(h) \\ &= \lambda h + o(h) \end{aligned}$$

This type of manipulations are important on the exam.

For iv):

$$\begin{aligned} Pr \{N(t+h) - N(t) = 0\} &= \frac{(\lambda h)^0}{0!} e^{-\lambda h} \\ &= 1 \cdot (1 - \lambda h + o(h)) \\ &= 1 - \lambda h + o(h), \quad \forall t \geq 0 \end{aligned}$$

- Prove that  $P2 \implies P1$ : i) and iii) are proved by definition.

For ii): Set  $s = 0$  Need to show that

$$N(h) - N(0) \sim \text{Poisson}(\lambda h)$$

(i) Divide  $(0, h]$  into equal size sub-intervals.

$$\implies t_i = \frac{i}{m}, \quad i = 0, 1, \dots, m.$$

(ii) Let

$$\varepsilon = \begin{cases} 1, & \text{at least one event in } (t_{i-1}, t_i] \\ 0, & \text{Otherwise} \end{cases},$$

$$i = 1, 2, \dots, m \text{ Then we can let } S_m = \sum_{i=1}^m \varepsilon_i.$$



(iii)  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m \sim \text{Bernoulli}(p_m)$  where  $p_m = \frac{\lambda h}{m} + o\left(\frac{h}{m}\right)$  as  $m \rightarrow \infty$ .

Let  $S = \lim_{m \rightarrow \infty} S_m a$  we get

$$\lim_{m \rightarrow \infty} m o_m = \lim_{m \rightarrow \infty} (\lambda h + o(1)) = \lambda h$$

This is called the "Law of rare events"  $S \sim \text{Poisson}(\lambda h)$ .

(iv)

$$\Pr\{N(h) - N(0) \neq S_m\} \leq \sum_{i=1}^m \Pr\{N(t_i) - N(t_{i-1}) \geq 2\}$$

□

## 8 References