STAT333 Applied Probability
Notes Summary

Professor Pengfei Li
Summarized by Iris Jiang

Spring 2019

Contents

Chapter 1

Introduction

1.1 Distributions

Bernoulli trials
Definition

1. Each trial has 2 outcomes: "s" (success) or "f" (failure)
2. All trials are independent
3. Probability of "s" ($P(s)$) on each trail are the same

Notation: $p = P(success)$, $q = 1 - p = P(failure)$
Bernoulli Random Variables: $\sim Bernoulli(p)$, p is the probability of success

Let $I_i = \begin{cases} 1 & \text{if "s" appears on the } i^{th} \text{ trail} \\ 0 & \text{otherwise} \end{cases}$

Then, $P(I_i = 1) = p$ & $P(I_i = 0) = q$, where $I_1, I_2, ..., I_n$ are a sequence of i.i.d. (independent identically distributed) Bernoulli rvs
Formulas: $X \sim Bernoulli(p)$

- $P(X = 0) = 1 - p$ and $P(X = 1) = q$
- $E(X) = p$
- $Var(X) = p(1 - p)$
- $pgf = 1 - p + ps$

Binomial rvs
Notation: $\sim Bin(n, p)$
Range: $x = \{0, 1, 2, ..., n\}$ $\quad$ $x =$ number of "s" in n Bernoulli trials
Formulas: $X \sim Bin(n, p)$

- $P(X = k) = \binom{n}{k} P^k (1 - p)^n - k,\ k = 0, 1, ..., n$
- $E(X) = np$
- $Var(X) = np(1 - p)$

- $pgf = (1-p+ps)^n$

Results:

1. $x = \sum_{i=1}^{n} I_i$
2. If $x_1 \sim Bin(n_1, p)$ $x_2 \sim Bin(n_2, p)$, and x_1, x_2 are independent, then $x_1 + x_2 \sim Bin(n_1 + n_2, p)$

Geometric rvs

Geometric rv is a waiting time rv.

Definition

$x =$ number of trails to get first "s" in the sequence of Bernoulli trials

Range: $x = \{1, 2, ...\}$

Formulas: $X \sim Geo(p)$

- $P(X = k) = p(1-p)^{k-1}, k = 1, 2, ...$
- $E(X) = \frac{1}{p}$
- $Var(X) = \frac{1-p}{p^2}$
- $pgf = \frac{ps}{1-(1-p)s}$

Property: no-memory property

$$P(x > n + m | x > m) = P(x > n) = P(\underbrace{x - m}_{Remaining\ Time} > n \,|\, \underbrace{x > m}_{at\ time\ m,\ we\ do\ not\ observe\ "s"})$$

The property tells us given that we do not observe the event "s", the remaining time $\sim Geo(p)$

1.2 Indicator rv

Definition

For a given event A, we define $I_A = \begin{cases} 1 & \text{if A occurs} \\ 0 & \text{otherwise} \end{cases}$

Properties

1. $E(I_A) = P(I_A = 1) = P(A) = p$
2. $Var(I_A) = P(I_A = 1)P(I_A = 0) = pq$
3. $E(I^2) = E(I)$

1.3 Useful Relationships

- $P(E \cap F) = P(F|E)P(E)$
- $Var(X) = E(X^2) - [E(X)]^2$
  $E(X^2) = Var(X) + [E(X)]^2$

- $Cov(X,Y) = E(XY) - E(X)E(Y)$
  If X and Y are independent, then $Cov(X,Y) = 0$ or $E(XY) = E(X)E(Y)$

- $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
  If X and Y are independent, then $Var(X+Y) = Var(X) + Var(Y)$

Chapter 2

Waiting Time RVs

2.1 Classification of T_E

1. If $P(T_E < \infty) < 1 \Rightarrow T_E$ is improper
2. If $P(T_E < \infty) = 1 \Rightarrow T_E$ is proper
   (a) If $E(T_E) = \infty \Rightarrow T_E$ is null proper
   (b) If $E(T_E) < \infty \Rightarrow T_E$ is short proper

Comments

1. If T_E is improper $\Rightarrow E(T_E) = \infty$
2. If $E(T_E) < \infty \Rightarrow T_E$ is short proper.
   (We do not need to verify $P(T_E < \infty) = 1$)

2.2 Notes

Denote R as the remaining time for event X, then R and X follow the same distribution and we have $E(X) = E(R)$ and $E(X^2) = E(R^2)$

Chapter 3

Conditional Expectation

3.1 Joint RVs

pmf and pdf

- $$f_X(x) = \begin{cases} \sum_y f_{X|Y}(x|y) & \text{Discrete RV} \\ \int_{-\infty}^{\infty} f_{X|Y}(x|y)dy & \text{Continuous RV} \end{cases}$$

- $$f_Y(y) = \begin{cases} \sum_x f_{X|Y}(x|y) & \text{Discrete RV} \\ \int_{-\infty}^{\infty} f_{X|Y}(x|y)dx & \text{Continuous RV} \end{cases}$$

Property

1. If $X\&Y$ are independent, then $g(x)\&h(y)$ are independent
2. If $X\&Y$ are independent, then $E[g(x)h(y)] = E[g(x)]E[h(y)]$

3.2 Conditional Expectation

Conditional Distribution

For a given y, the conditional pmf/pdf for X given $Y = y$ is

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

Conditional Expectation

- The conditional expectation of x given $Y = y$ is

$$E(X|Y = y) = \begin{cases} \sum_x x \times f_{X|Y}(x|y) & \text{Discrete RV} \\ \int_{-\infty}^{\infty} x \times f_{X|Y}(x|y)dx & \text{Continuous RV} \end{cases}$$

- The conditional expectation of $g(x)$ given $Y = y$ is

$$E[g(x)|Y = y] = \begin{cases} \sum_x g(x) \times f_{X|Y}(x|y) & \text{Discrete RV} \\ \int_{-\infty}^{\infty} g(x) \times f_{X|Y}(x|y)dx & \text{Continuous RV} \end{cases}$$

Properties

1. Conditional expectation has all properties of normal expectations
2. Substitution Rule

$$E[x \times \underbrace{g(Y)}_{random\, variable} |Y = y] = E(x \times \underbrace{g(y)}_{constant} |Y = y) = g(y)E(X|Y = y)$$

   In general: $E[\underbrace{h(X,Y)}_{function\, of\, X\&Y} |Y = y] = E[\underbrace{h(X,y)}_{function\, of\, X only} |Y = y]$

3. Independence Property
   If $X\&Y$ are independent then

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = f_X(x)$$

$$\Rightarrow E(X|Y = y) = E(x)\, and\, E[g(x)|Y = y] = E[g(x)]$$

3.3 Expectation by Conditioning

Double Expectation Theorem

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

Step 1. What is $E(X|Y)$

(a) $E(X|Y)$ is a random variable; and depends on Y

(b) given $Y = y$, the function of $g(Y)$: $g(y) = E(X|Y = y)$

Step 2. How to get $E(X|Y = y)$

(a) Figure out $g(y) = E(X|Y = y)$, by definition or properties

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

Step 3. How to apply $E(X) = E[E(X|Y)]$

$$E(X) = E[E(X|Y)] = E[g(Y)]$$

$$= \begin{cases} \sum_y gy \times f_Y(y) \\ \int_{-\infty}^{\infty} gy \times f_Y(y)dy \end{cases} = \begin{cases} \sum_y E(X|Y = y) \times f_Y(y) & \text{discrete case} \\ \int_{-\infty}^{\infty} E(X|Y = y) \times f_Y(y)dy & \text{continuous case} \end{cases}$$

3.4 Probability by Conditioning

$$P(A) = \begin{cases} \sum_y P(A|Y = y)f_Y(y) & \text{discrete Y} \\ \int_{-\infty}^{\infty} P(A|Y = y)f_Y(y)dy & \text{continuous Y} \end{cases}$$

3.5 Variance by Conditioning

Method 1. By definition,

$$\begin{aligned} Var(X) &= E(X^2) - [E(X)]^2 \\ &= E(X^2|Y) - [E(X|Y)]^2 \end{aligned}$$

Method 2. Conditional Variance Formula

$$Var(X|Y = y) = E(X^2|Y = y) - [E(X|Y = y)]^2$$

3.5.1 Finding $Var(X|Y)$

Two Steps

Step 1. Find $h(y) = Var(X|Y = y)$

Step 2. Apply $h(y)$ to Y get $Var(X|Y) = h(Y)$

Note. *If X and Y are independent, $Var(X|Y = y) = Var(X)$*

Theorem.

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$

3.5.2 Finding Expectation and Variance for Compound Random Variable

Definition. *Suppose $X_1, ..., X_n$ are a sequence of iid rvs. N_i is a rv only takes non-negative integers. Further $N, X_1, ..., X_n$ are independent. Then*

$$W = \sum_{i=1}^{N} X_i$$

is called a compound rv.
(If $N = 0$, then $W = 0$)

Theorem.

$$E(W) = E(N) \times E(X_1)$$

$$Var(W) = E(N) * Var(X_1) + Var(N) \times [E(X_1)]^2$$

Chapter 4

Probability Generating Function

4.1 Generating Function

$$A(s) = a_0 + a_1 s + a_2 s^2 + ... = \sum_{n=0}^{\infty} a_n s^n$$

4.1.1 Properties of Generating Function

1. Summation

$$C(s) = A(s) \pm B(s) = \sum_{n=0}^{\infty} (a_n \pm b_n) s^n$$

$$c_n = a_n \pm b_n$$

$$R(C) = min(R(A), R(B))$$

2. Product

$$C(s) = A(s) \times B(s) = \sum_{n=0}^{\infty} c_n s^n$$

$$(\sum_{n=0}^{\infty} a_n s^n)(\sum_{n=0}^{\infty} b_n s^n) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} s^n$$

$$c_n = \underbrace{\sum_{k=0}^{n} a_k b_{n-k}}_{\text{n+1 terms}}$$

$$R(c) = min(R(A), R(b))$$

4.2 Probability Generating Function

$$G_X(s) = \sum_{n=0}^{\infty} p_n s^n = \sum_{n=0}^{\infty} P(X = n) s^n$$

If x is a proper random variable, then $\begin{cases} P(x = \infty) = 0 \\ P(x < \infty) = 1 \end{cases}$, then

$$G_X(s) = \sum_{n=0}^{\infty} P(X = n)s^n = E[s^X]$$

4.2.1 Properties of PGF

1. pgf helps to find $\{p_n = P(X = n)\}_{n=0}^{\infty}$
   (a)
   $$\begin{cases} p_0 = G_X(0) \\ p_n = \frac{G_X^{(n)}(0)}{n!} \end{cases}$$
   (b) Use properties of gfs to recover $\{p_n\}_{n=0}^{\infty}$
2. pgf helps up to classify random variable

   Note.
   $$P(X < \infty) = \sum_{n=0}^{\infty} P(X = n) = \sum_{n=0}^{\infty} p_n = G_X(1)$$
   $$\begin{cases} G_X(1) = 1 & \Rightarrow \text{ X is proper} \\ G_X(1) < 1 & \Rightarrow \text{ X is improper} \\ G_X(1) > 1 & \Rightarrow \text{ you did something wrong} \end{cases}$$
3. If X is proper, then
   $$E(X) = G'_X(1)$$
   $$Var(X) = \underbrace{G''_X(1) + G'_X(1)}_{E(X^2)} - [G'_X(1)]^2$$
4. Uniqueness Theorem
   Two random variables X and Y have the same distribution if and only if
   $$G_X(s) = G_Y(s)$$
5. Independence Property
   Suppose X_1 and X_2 are non-negative random variable with the ranges $\{0, 1, 2, ...\} \cup \{\infty\}$. Further, X_1 and X_2 are independent. Then
   $$G_{X_1+X_2} = G_{X_1}(s) \times G_{X_2}(s)$$
   Note that if X_1 and X_2 are proper, then $G_{X_1+X_2}(s) = E[s^{X_1+X_2}]$

4.2.2 pgf of distributions

1. pgd for indicate rv I_A

$$s_A^I = \begin{cases} s^1 & I_A = 1, P(I_A = 1) = p \\ s^0 & I_A = 0, P(I_A = 0) = p \end{cases}$$

$$E[s^{I_A}] = s^1 \times p + s^0 \times q = ps + q$$

$$\Rightarrow G_{I_A}(s) = ps + q$$

$$R_{I_A} = \infty$$

2. $X \sim Bin(n, p)$

$$G_X(s) = G_{\sum_{i=1}^n I_i}(s) = \prod_{i=1}^n \underbrace{G_{I_i}(s)}_{ps+q} = (ps + q)^n$$

$$R_X = \infty$$

3. $X \sim Geo(p)$

$$G_X(s) = E[s^X] = \sum_{n=1}^\infty P(X = n) \times s^n =$$

$$R_X = \frac{1}{1-p}$$

4. $X \sim NegBin(r, p)$

$$G_X(s) = G_{\sum_{i=1}^n X_i}(s) = \prod_{i=1}^r G_{X_i}(s) = (\frac{ps}{1-(1-p)s})^r$$

$$R_X = \frac{1}{1-p}$$

5. $X \sim Pois(\lambda)$

$$G_X(s) = E(s^X) = \sum_{n=0}^\infty P(X = n)s^n = \sum_{n=0}^\infty \frac{\lambda^n e^{-n}}{n!} s^n = \underbrace{\sum_{n=0}^\infty \frac{(\lambda s)^n}{n!}}_{e^{\lambda s}} e^{-\lambda} = e^{\lambda s - \lambda}$$

$$R_X = \infty$$

4.3 Simple Random Walk

Definition. *Let X_0 be the starting point of the process ($X_0 = 0$) and X_n be the position of the process after n steps. Then $\{X_n\}_{n=0}^\infty$ is called a simple(ordinary) random walk*

4.3.1 Notations

- $\lambda_{0,0}$: the event that returning to 0, given the process starting with 0
- $\lambda_{0,k}$: the event that visiting to k, given the process starting with 0
- $T_{0,k}$: waiting time for observing the first $\lambda_{0,k}$
  $= min\{n \geq 1, X_n = k | X_0 = 0\}$
- $G_{0,0}(s) = \sum_{n=0}^{\infty} P(T_{0,0} = n) s^n$
- $G_{0,0}(s) = \sum_{n=0}^{\infty} P(T_{0,k} = n) s^n$

4.3.2 Properties for $T_{0,k}$

1. For positive integer k

$$\begin{aligned} T_{0,k} &= T_{0,1} + T1,2 + \cdots + T_{k-1,k} \\ &= \sum_{i=1}^{k} T_{i-1,i} \\ T_{i,j} &= \text{waiting time for visiting } j\text{, starting from } i \\ &= min\{n \geq 1, X_n = j | X_0 = i\} \end{aligned}$$

$$G_{0,k}(s) = [G_{0,1}(s)]^k$$

for $k > 0$

4.3.3 Properties for $G_{0,k}(s)$

$$G_{0,k}(s) = \sum_{n=0}^{\infty} P(T_{0,k}) s^n$$

1.
$$G_{0,0}(s) = 1 - \sqrt{1 - 4pqs^2}$$

2.
$$G_{0,1}(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}$$

3. $k > 0$
$$G_{0,k}(s) = [G_{0,1}(s)]^k = [\frac{1 - \sqrt{1 - 4pqs^2}}{2qs}]^k$$

4.
$$G_{0,-1}(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps}$$

5. For $k < 0$
$$G_{0,k}(s) = [G_{0,-1}(s)]^{|k|} = [\frac{1 - \sqrt{1 - 4pqs^2}}{2ps}]^{|k|}$$

4.3.4 Classify $T_{0,k}$

For ordinary random walk:

$$P(T_{0,k} < \infty) = G_{0,k}(1)$$

$$E(T_{0,k}) = G'_{0,k}(1)$$

For $\lambda_{0,k}, k > 0$

1. $P(T_{0,k} < \infty) = 1$ and $E(T_{0,k}) = \frac{k}{p-q}$ if $p > q$ (short proper)
2. $P(T_{0,k} < \infty) = 1$ and $E(T_{0,k}) = \infty$ if $p = q = \frac{1}{2}$ (null proper)
3. $P(T_{0,k} < \infty) = (\frac{p}{q})^k < 1$ and $E(T_{0,k}) = \infty$ if $p < q$ (improper)

For $\lambda_{0,k}, k < 0$

1. $P(T_{0,k} < \infty) = 1$ and $E(T_{0,k}) = \frac{|k|}{p-q}$ if $p > q$ (short proper)
2. $P(T_{0,k} < \infty) = 1$ and $E(T_{0,k}) = \infty$ if $p = q = \frac{1}{2}$ (null proper)
3. $P(T_{0,k} < \infty) = (\frac{q}{p})^{|}k| < 1$ and $E(T_{0,k}) = \infty$ if $p < q$ (improper)

Chapter 5

Discrete Markov Process

5.1 Definition and Notations

Suppose we have a sequence of rvs $\{X_n\}_{n=0}^{\infty}$

- n is called **time have**
- $\{X_n\}_{n=0}^{\infty}$ is called **stochastic process**
- **state space** is all possible values of $\{X_n\}_{n=0}^{\infty}$; notation: S
- For $i \in S$, we call i **state** i

Definition. *$\{X_n\}_{n=0}^{\infty}$ is called a **discrete Markov process** (Markov chain) if*

1. *state space S is discrete or countable*
2. *$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, ..., X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j|X_n = i) = P(X_1 = j|X_0 = i)$, denoted by p_{ij}*

5.2 Property

1. **Markov Property**
   Given the current information (X_n), the future (X_{n+1}) does not depend on the history ($X_{n-1}, X_{n-2}, ..., X_0$)
2. **Time Homogeneous Property**
   Conditional probs so not depend on starting time, they only depend on step size

5.3 One-Step Transition Probs and Matrix

Definition. *Let*

$$p_{\underbrace{i}_{\text{current state}}\underbrace{j}_{\text{future state}}} = P(X_1 = j|X_0 = i) = P(X_{n+1} = j|X_n = i)$$

*called **one-step transition prob from state** i **to state** j*

$$
\begin{array}{cc}
 & \begin{array}{ccc} \text{column:} & \text{future} & \text{state} \end{array} \\
\underline{\underline{P}} = \begin{array}{c} \text{row:} \\ \text{current} \\ \text{state} \end{array} & \begin{pmatrix} \ddots & \cdots & \ddots \\ \vdots & p_{ij} & \vdots \\ \cdots & \cdots & \cdots \end{pmatrix}
\end{array}
$$

$i \in S, j \in S$

Property

1. $p_{ij} \geq 0$
2. $\sum_{j \in S} p_{ij} = 1$, sum of each row = 1

5.4 Chapman-Kolmogorov Equations

5.4.1 n-step Transition Probs and Matrix

Notations.

$$P_{ij}^{(n)} = P(X_n = j | X_0 = i) = P(X_{n+m} = j | X_m = i)$$

$$\underline{\underline{P}}^{(n)} = (p_{ij}^{(n)})_{i \in S, j \in S} \text{ and } \underline{\underline{P}}^{(1)} = \underline{\underline{P}}$$

$$
\begin{aligned}
\pi_i^{(0)} &= P(X_0 = i), i \in S \\
\pi_j^{(n)} &= P(X_n = j), j \in S \\
\underline{\underline{\pi}}^{(0)} &= (\pi_i^{(0)}) = (P(X_0 = i))_{i \in S} \\
&\Rightarrow \text{ row vector/distribution of } X_0 \\
\underline{\underline{\pi}}^{(n)} &= (\pi_j^{(0)}) = (P(X_n = j))_{j \in S} \\
&\Rightarrow \text{ row vector/distribution of } X_n
\end{aligned}
$$

CK-Equations

Theorem 1. CK-equation 1 - n-step transition matrix

Pointwise Form:

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{(ik)}^{(n)} p_{kj}^{(m)}$$

Matrix Form:

$$\underline{\underline{P}}^{(n)} = \underline{\underline{P}}^n$$

Theorem 2. CK-equation 2 - n-step transition probs

Pointwise Form:

$$\pi_j^{(n)} = P(X_n = j) = \sum_{i \in S} \pi_i^{(0)} p_{ij}^{(n)}$$

Matrix Form:

$$\underline{\underline{\pi}}^{(n)} = \underline{\underline{\pi}}^{(0)} \ \underline{\underline{P}}^n$$

5.5 Classification of States

Method:

1. $\lambda_{i,i}$ = returning state i, given $X_0 = i$
2. $T_{i,i}$ = waiting time to observe first $\lambda_{i,i} = min\{n \geq 1, X_n = i | X_0 = i\}$

5.5.1 Classification by $T_{i,i}$

1. we call state i **transient** if $T_{i,i}$ is improper, i.e., $P(T_{i,i} < \infty) < 1$
2. we call state i **null recurrent** if $T_{i,i}$ is null proper, i.e., $P(T_{i,i} < \infty) = 1$ and $E(T_{i,i}) = \infty$
3. we call state i **positive recurrent** if $T_{i,i}$ is short proper, i.e., $P(T_{i,i} < \infty) = 1$ and $E(T_{i,i}) < \infty$

5.5.2 Classification by $\lambda_{i,i}$

recurrent: on average, we can observe $\lambda_{i,i}$ infinite number of times
transient: on average, we can observe $\lambda_{i,i}$ finite number of times

5.5.3 Period of State

Definition. *Period of state i is defined as $d = gcd\{n | p_{ii}^{(n)} > 0 \;\&\; n \geq 1\}$*

1. If $d = 1$, state i is called aperiodic
   $p_{ii}^{(n)} > 0$ for all $n \geq 1$ or
   $\exists N$ such that $p_{ii}^{(n)} > 0$ when $n \geq N$
2. If $d > 1$, state i is called periodic we only have $p_{ii}^{(nd)} > 0$; for other steps $p_{ii}^{(n)} = 0$

5.5.4 Methods to Classify State i Based on $\underline{\underline{P}}$

Definition. *If a Markov process only has one class, then it is called **irreducible***

5.5.5 Theorems about Class

Theorem 1. Let C be a class in Markov process, then

(a) All states in C have same period
(b) All states in C have same classification

Theorem 2. Period of a special class
Let C be a class, If $\exists i \in C$, such that $p_{ii} > 0$, then all states have period 1

Definition. *A class C is said to be **closed** if it is impossible to leave the class.
i.e., $\forall i \in C$, $j \notin C$, then $p_{ij} = 0$*

Definition. *A class C is said to be **open** if it is possible to leave the class.
i.e., $\forall i \in C$, $j \notin C$, then $p_{ij} > 0$*

Theorem 3. (a) All states in open class are transient
(b) If C is **closed** and **has finite number of states**, then all states are **positive recurrent**

5.5.6 Stationary Distribution

$$\underline{\underline{\pi}}\,\underline{\underline{P}} = \underline{\underline{\pi}} \text{ and } \pi_i \geq 0, \sum_{i \in S} \pi_i = 1$$

1. if $\underline{\underline{\pi}}^{(n)} = \underline{\underline{\pi}}$, then $\underline{\underline{\pi}}^{(n+1)} = \underline{\underline{\pi}}^{(n+2)} = \cdots = \underline{\underline{\pi}}$

Theorem. *An **irreducible** Markov process has stationary distribution if and only if all states are **positive recurrent**, and then in this case*

$$E(T_{i,i}) = \frac{1}{\pi_i}$$

5.5.7 Positive Recurrent States

Two ways to say if the states in a class are positive recurrent:

1. irreducible and finite number of states $\Rightarrow$ positive recurrent
2. $\underline{\underline{\pi}}$ exists and is unique $\Rightarrow$ positive recurrent
(this is the only way to find $E(T_{i,i})$)

Chapter 6

Poisson Process

6.1 Exponential Process

1. probability density function

$$f(x) = \begin{cases} \lambda e^{-\lambda} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

2. tail probability

$$P(X > t) = e^{-\lambda t}$$

$t > 0$

3.

$$E(x) = \frac{1}{\lambda} \text{ and } Var(x) = \frac{1}{\lambda^2}$$

4. No-memory Property

$$P(X > t + s | x > s) = P(X > t)$$

5. Alarm clock lemma
If $X_i \sim exp(\lambda_i),\ i = 1, ..., n\ \ X_1, ..., X_n$ are independent, then

   (a)

$$min(X_1, ..., X_n) \sim exp(\sum_{i=1}^{n} \lambda_i)$$

   (b)

$$P(X_i = min(X_1, ..., X_n)) = \frac{\lambda_i}{\sum_{k=1}^{n} \lambda_k}$$

6. If $X_1, X_2, ... \sim exp(\lambda)$, $N \sim Geo(p),\ 0 < p < 1$, X_i's and N are independent, then $\sum_{i=1}^{N} X_i \sim exp(\lambda p)$

6.2 Poisson Process

Definition. *A counting process $X(t), t \geq 0$ is a Possion Process with rate λ if*

1. $x(0) = 0$

2. *If* $0 \leq s_1 < s_2 \leq t_1 < t_2$*, then* $\underbrace{X(t_2) - X(t_1)}_{\text{number of events in } (t_1, t_2]}$ *and* $\underbrace{X(s_2) - X(s_1)}_{\text{number of events in } (s_1, s_2]}$ *are independent*

3. $X(t+s) - X(s) \sim Pois(\lambda t)$

Property. 1. *In a small interval, we can only observe 0 or 1 event*

2. $T_1, ..., T_i \sim exp(\lambda)$*,* T_i *is waiting time for the i-th event*

3. *Suppose* $0 < s < t$*, then* $X(s)|X(t) = n \sim Bin(n, \frac{s}{t})$

4. *Suppose* $X(t)$ = *number of events in* $(0, t]$ *and follows Poisson Process with rate* λ*, further events can be classified into 2 types*

$$\begin{aligned} \text{type I} &: \text{prob} = p \\ \text{type II} &: \text{prob} = q = 1 - p \end{aligned}$$

and all events are independent. Let

$$\begin{aligned} X_i(t) &= \text{number of type I events in } (0, t] \\ X_i(t) &= \text{number of type II events in } (0, t] \end{aligned}$$

Then

(a) $X_1(t) \sim Pois(\lambda pt)$

(b) $X)_2(t) \sim Pois(\lambda qt)$

(c) $X_1(t)$ *and* $X_2(t)$ *are independent*

Chapter 7

Continuous Markov Process

7.1 Generator Matrix

$$\underline{\underline{R}} = \begin{array}{c c} & \begin{array}{ccc} \text{column:} & \text{future} & \text{state} \end{array} \\ \begin{array}{c} \text{row:} \\ \text{current} \\ \text{state} \end{array} & \begin{pmatrix} \ddots & \cdots & \ddots \\ \vdots & r_{ij} & \vdots \\ \cdots & \cdots & \cdots \end{pmatrix} \end{array} = \underline{\underline{P}}'(0)$$

7.2 Stationary Distribution

$$\left.\begin{array}{l} \underline{\underline{\pi}}\,\underline{\underline{R}} = \underline{\underline{0}} \\ \sum_{i \in S} \pi_i = 1 \end{array}\right\} \Rightarrow \underline{\underline{\pi}}$$

7.3 Calculating L_i and E_{ij}

$$\begin{aligned} L_i &= \text{waiting time for process to stay in state } i \text{ before jumping to other states} \\ &= min\{t \geq 0, X(t) \neq i | X(0) = i\} \\ E_{ij} &= P(\text{jump to state } i \mid \text{process leaves state } i) \end{aligned}$$

Note $E_{ii} = 0$

Property. *1.* $L_i \sim exp(-r_{ii})$

2. $E_{ij} = \frac{r_{ij}}{-r_{ii}} = \frac{r_{ij}}{|r_{ii}|}$ *for* $i \neq j$

7.4 Birth-Death Process

Definition. *The continuous time Markov Process* $\{X(t), t \geq 0\}$ *is a birth-death process if*

1. *At each jump, the process can either*

$$i \rightarrow \begin{cases} i+1 \text{ jump up by 1 unit [birth]} \\ i-1 \text{ jump down by 1 unit [death]} \end{cases}$$

2. when $X(t) = i$

$$L_{Bi} = \text{ waiting time for a birth} \sim exp(\overbrace{\lambda_i}^{\text{birth rate}})$$
$$L_{Di} = \text{ waiting time for a death} \sim exp(\underbrace{\mu_i}_{\text{death rate}})$$

Also, $\mu_0 = 0$ *and* $L_i = min(L_{Bi}, L_{Di}) \sim exp(\lambda_i + \mu_i)$

7.4.1 Generator Matrix in Birth-Death Process

$$r_{ii} = -(\lambda_i + \mu_i)$$
$$E_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}$$
$$E_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}$$
$$E_{i,j} = 0 \text{ for } |i-j| \geq 2$$

$$\underline{\underline{R}} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & \cdots \end{array} \\ \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & -\lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & -\lambda_2 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}$$

7.4.2 Stationary Distribution

$$\pi_j = \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j}\pi_0,\ j \geq 1$$
$$\pi_0 = \frac{1}{1 + \sum_{j=1}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j}}$$

Check:

1. If $\sum_{j=1}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j} = \infty$, then
$$\pi_0 = 0 \text{ and } \pi_{j,j\geq 1} = 0$$
$$\Rightarrow \underline{\underline{\pi}} \text{ does not exist}$$

2. If $\sum_{j=1}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j} < \infty$, then
$$\pi_0 = \frac{1}{1 + \sum_{j=1}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j}} \text{ and } \pi_j = \frac{\lambda_0\lambda_1\cdots\lambda_{j-1}}{\mu_1\mu_2\cdots\mu_j}\pi_0$$