随机过程

教授：吴明燕

笔记由 Dafu Zhu 编写

基于 2025 春季厦大数院《随机过程》

最后修改：2025/05/21

目录

1 概率论准备知识

成绩：平时（作业 + 考勤）+ 期中论文 + 期末

概率论准备知识

概率论中, 随机变量的本质是可测函数.

$$X : \Omega \to S$$

S 的 σ-代数记为 $\mathcal{S}$, 是个 Borel σ-代数（由开集/闭集生成）

Q: 为什么要给 Ω 一个 σ-代数？

A: 样本空间是抽象的, 给它 σ-代数赋予它结构, 相当于对信息进行重整/提取

概率测度的本质是集函数,

$$\text{集合} \to \text{函数}$$

将信息具象化,

$$\begin{aligned} \mathbb{P} :& \mathcal{F} \to [0,1] \\ & A \to \mathbb{P}(A) \end{aligned}$$

随机过程：一族随机变量 $\{X_t\}_{t\in\mathbb{T}}$

其中 $\mathbb{T}$ 为指标集, $X_t : \Omega \to S$

Example 1.1

$\mathbb{T} = \mathbb{N}_0$: 时间离散；$\mathbb{T} = [0, T]$: 时间连续

$$X : (\Omega, \mathcal{F}, \mathbb{P}) \to (S, \mathcal{S}, \mu_X)$$

思考：什么是随机过程的分布 $\{\mu_t\}_{t\in\mathbb{T}}$？

1.1 事件概率

1.1.1 事件域

Definition 1.1 (样本空间、事件)
样本点、样本空间、事件和事件的运算：

- 样本点 ω：一次试验的结果
- 样本空间 Ω：全体样本点
- 事件：Ω 的子集
- 事件的运算：集合的运算，即交并补 $(A \cap B, A \cup B, A^c)$

Definition 1.2
若 $A \cap B = \varnothing$，则称 A 与 B 不相交，更一般地，若 $A_i \cap A_j = \varnothing (i \neq j)$，则称 $\{A_i\}_{i \geqslant 1}$ 互不相交

Definition 1.3 (σ-代数)
称 $\mathcal{F} \subset 2^{\Omega} = \{A | A \subset \Omega\}$ 是一个 σ-代数/事件域（其中 2^{Ω} 表示所有 Ω 的子集构成的集合，是一个集类）若

1. $\Omega \subset \mathcal{F}$
2. (对补封闭) $A \in \mathcal{F} \to A^c \in \mathcal{F}$
3. (对可列并封闭) $A_n \in \mathcal{F}, n \geqslant 1 \Rightarrow \cup_{n \geqslant 1} A_n = \cup_{n=1}^{\infty} A_n \in \mathcal{F}$

σ 代数是满足以上特定条件的集类，是由 Ω 的子集构成的集合
注：σ 代数对有限交/有限并/可列交封闭

现在给出了一个定义，我们会想“为什么定义会这样给呢”，现在要举一些例子说明“定义有意义”

Example 1.2
最小的 σ 代数：$\{\varnothing, \Omega\}$
最大的 σ 代数：2^{Ω}

以上这两个例子一个太小、一个太大，似乎没意义，所以叫它们“平凡的”

Example 1.3
$A \subset \Omega, \sigma(\{A\}) = \sigma(A) = \{A, A^c, \Omega, \varnothing\} = \sigma(A^c)$
这是由 A 生成的 σ 代数

Definition 1.4 (划分/分割)
称 $\Pi_{\Omega} := \{\Lambda_n, n \geqslant 1\}$ 是 Ω 的一个分划，若 $\Omega = \sum_{n \geqslant 1} \Lambda_n$

1. $\Omega = \cup_{n \geqslant 1} \Lambda_n$
2. $\{\Lambda_n\}_{n \geqslant 1}$ 互不相交

Example 1.4

$\Omega = \sum_{n\geqslant 1}\Lambda_n, \Pi_\Omega := \{\Lambda_n\}_{n\geqslant 1}$

$$\sigma(\Pi_\Omega) = \left\{\sum_{k\in J}\Lambda_k, J\subset \mathbb{N}\right\}$$

Problem 1 (作业 1-1)

证明：

1. $\sigma(\Pi_\Omega)$ 是一个 σ 代数
2. $\sigma(\Pi_\Omega)$ 是包含集类 Π_Ω 的最小 σ 代数

$(S,\mathcal{S}) = (S, 2^S)$: S 可列时, 取 2^S 为 σ 代数

$(S,\mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$: S 为实数集时, 取博雷尔集 $\mathcal{B}(\mathbb{R})$ 为 σ 代数

1.1.2 概率测度

Definition 1.5 (概率测度)

$(\Omega,\mathcal{F})$ 称 $\mathbb{P}:\mathcal{F}\to[0,1]$ 是概率测度

1. 非负性
2. 归一性
3. 可列可加性 *

其中, 可列可加性的表述为：设 $\{E_n, n\geqslant 1\}$ 是 $\mathcal{F}$ 中互不相交的集合序列 $(E_i\cap E_j=\varnothing, i\neq j)$, 则

$$\mathbb{P}(\cup_{n=1}^\infty E_n) = \sum_{n=1}^\infty \mathbb{P}(E_n)$$

Property 1.1

$\mathbb{P}$ 满足有限可加性（可列可加一定有限可加, 如果既不是可列可加、也不是有限可加, 则不可测）

Corollary 1.1

1. $\mathbb{P}(A) = 1-\mathbb{P}(A^c)$
2. 若 $A\subset B$, 则 $\mathbb{B} = \mathbb{A} + \mathbb{P}(BA^c) \geqslant \mathbb{P}(A)$
3. $\mathbb{P}(A\cup B) = \mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B)$

Remark 1. 引用知乎上三维之外的大白话解释可列可加性：

首先, 在我们总是习惯于处理有限相加, 而很少遇到无限相加的情况. 从测度论内容理解, 有限相加与事实（数学的）不符, 比如 $(0,1)$ 区间有不可数个点, 每个点的测度（理解为直径吧）是 0, 按照习惯想法（有限相加）, 直径的加和（总宽度）应该为 0, 显然, $(0,1)$ 区间的宽度不可能是 0;

如果规定为“只要是无穷多个点相加，其宽度就不再是 0”的话，还是存在矛盾，我们知道，区间 $(0,1)$ 上的有理数是是无穷多个的（而且是可列的），那么其宽度就应该为 1，可是无理数还是不可数的呢——理解为无理数是有理数的无穷大量或有理数是无理数的无穷小量，那么无理数的宽度是多少呢？即使还是 1，显然 $(0,1)$ 区间的宽度不可能是 2 吧!?

于是，勒贝格说道：在测量长度、面积、体积时，我们采用可列可加性，即可列个点相加，规定其宽度（测度）为 0，如果点的个数超过了可列个（这时必是连续统的），那么，就不满足了——即这些点的总宽度就不是 0 了，而是具有了非 0 的宽度（正测度），当然，具有测度的这些点是紧挨在一起的，否则不一定有测度，比如康托大师制造的三分集就很诡异.

到这里，可列可加性事实上讲完了，再啰嗦一下次可列可加性. 这是因为不论作为集合，还是概率上的事件（也是集合），一般是存在公共元素的，因此，一般情形下，当然满足次可列可加性的性质了，可列可加性只有在集合之间的距离大于 0 或事件之间完全独立的情形下，才会满足.

Property 1.2 (次可列可加性)

$A_n \subset \mathcal{F}, n \geqslant 1$

$$\mathbb{P}(\bigcup_{n\geqslant 1} A_n) \leqslant \sum_{n\geqslant 1} \mathbb{P}(A_n)$$

证明：$\cup_{n\geqslant 1} A_n = \sum_{n\geqslant 1} B_n$，其中 $B_1 = A_1, B_2 = A_2 \cap (A_1)^c, \cdots, B_n = A_n \cap A_1^c \cap A_2^c \cap \cdots \cap A_{n-1}^c$

$B_n \subset A_n$，由可列可加性和 Corollary 1.1(2)

Problem 2 (作业 1-2)

证明 $\cup_{n\geqslant 1} A_n = \sum_{n\geqslant 1} B_n$

证明：

1. 先证 $\bigcup_{n\geqslant 1} A_n \subseteq \sum_{n\geqslant 1} B_n$.
   假设 $x \in \bigcup_{n\geqslant 1} A_n$，
   若 $x \in A_1$，则 $x \in B_1$，
   若 $x \in A_2$ 且 $x \notin A_1$，则 $x \in B_2$
   $\cdots$
   若 $x \in A_n$ 且 $x \notin A_1$, $x \notin A_2$, $\ldots$, $x \notin A_{n-1}$，则 $x \in B_n$
   $\forall x \in \bigcup_{n\geqslant 1} A_n$，都有 $x \in \bigcup_{n\geqslant 1} B_n$
   $\because B_i \cap B_j = \varnothing, i \neq j, \therefore \bigcup_{n\geqslant 1} B_n = \sum_{n\geqslant 1} B_n$, $x \in \sum_{n\geqslant 1} B_n$.
2. 再证 $\sum_{n\geqslant 1} B_n \subseteq \bigcup_{n\geqslant 1} A_n$
   假设 $x \in \sum_{n\geqslant 1} B_n$，则 $\exists n_0 \in \mathbb{N}^+$，使得 $x \in B_{n_0}$，
   由 B 的定义

$$B_{n_0} = A_{n_0} \cap \left(\bigcap_{k=1}^{n_0-1} A_k^c \right)$$

   $\therefore x \in A_{n_0} \subseteq \bigcup_{n\geqslant 1} A_n$
   $\therefore \bigcup_{n\geqslant 1} A_n = \sum_{n\geqslant 1} B_n$ □

Property 1.3 (连续性)
(1) $A_n \uparrow$ 单调上升, 即 $A_n \subset A_{n+1}$, $\lim_{n\to\infty} A_n = \cup_{n\geqslant 1} A_n$, 则 $\mathbb{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty}\mathbb{P}(A_n)$
(2) $B_n \downarrow$ 单调下降, 即 $B_n \supset B_{n+1}$, $\lim_{n\to\infty} B_n = \cap_{n\geqslant 1} B_n$, 则 $\mathbb{P}(\lim_{n\to\infty} B_n) = \lim_{n\to\infty}\mathbb{P}(B_n)$

证明: (1) $\cup_{n\geqslant 1} A_n = A_1 + A_2 \setminus A_1 + A_3 \setminus A_2 + \cdots$

$$\begin{aligned}
\mathbb{P}(\bigcup_{n\geqslant 1} A_n) &= \mathbb{P}(A_1) + \sum_{n\geqslant 1}\mathbb{P}(A_{n+1}\setminus A_n) \\
&= \mathbb{P}(A_1) + \lim_{m\to\infty}\sum_{n=1}^{m}\mathbb{P}(A_{n+1}\setminus A_n) \\
&= \mathbb{P}(A_1) + \lim_{m\to\infty}\sum_{n=1}^{m}[\mathbb{P}(A_{n+1}) - \mathbb{P}(A_n)] \\
&= \mathbb{P}(A_1) + \lim_{m\to\infty}[\mathbb{P}(A_{m+1}) - \mathbb{P}(A_1)] \\
&= \lim_{m\to\infty}\mathbb{P}(A_{m+1}) \\
&= \lim_{n\to\infty}\mathbb{P}(A_n) \quad \square
\end{aligned}$$

(2) $B_n \downarrow B \Rightarrow \forall n, B_{n+1} \subseteq B_n \Rightarrow \forall B_n^c \subseteq B_{n+1}^c$

$$\begin{aligned}
\mathbb{P}(B) = \mathbb{P}(\cap_{n\geqslant 1} B_n) &= 1 - \mathbb{P}((\cap_{n\geqslant 1} B_n)^c) \\
&= 1 - \mathbb{P}(\cup_{n\geqslant 1} B_n^c) \\
&= 1 - \mathbb{P}(B_1^c \cup (\cup_{n\geqslant 2}(B_n^c \setminus B_{n-1}^c))) \\
&= 1 - \mathbb{P}(B_1^c) - \sum_{n\geqslant 2}(\mathbb{P}(B_n^c) - \mathbb{P}(B_{n-1}^c)) \\
&= 1 - \mathbb{P}(B_1^c) - \lim_{m\to\infty}\sum_{n=2}^{m}(\mathbb{P}(B_n^c) - \mathbb{P}(B_{n-1}^c)) \\
&= 1 - \mathbb{P}(B_1^c) - \lim_{m\to\infty}(\mathbb{P}(B_m^c) - \mathbb{P}(B_1^c)) \\
&= 1 - \mathbb{P}(B_1^c) - \lim_{n\to\infty}\mathbb{P}(B_n^c) + \mathbb{P}(B_1^c) \\
&= 1 - \lim_{n\to\infty}\mathbb{P}(B_n^c) \\
&= \lim_{n\to\infty}\mathbb{P}(B_n) \quad \square
\end{aligned}$$

第二个等式用到 De Morgan's Law

1.2 独立性

Definition 1.6 (事件间的独立性)
$(\Omega,\mathcal{F},\mathbb{P}), A,B\in\mathcal{F}$, 称 A 与 B 独立, 若 $\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)$, 记为 $A\perp\!\!\!\perp B$

Definition 1.7 (事件间的相互独立)
$\{A_n\}_{n\geqslant 1}\subset\mathcal{F}$, 称其相互独立, 若 $\forall J\subset\mathbb{N},\#J\geqslant 2$

$$\mathbb{P}(\bigcap_{k\in J}A_k)=\prod_{k\in J}\mathbb{P}(A_k)$$

Property 1.4
$A\perp\!\!\!\perp B\Rightarrow A\perp\!\!\!\perp B^c, A^c\perp\!\!\!\perp B, A^c\perp\!\!\!\perp B^c$

Definition 1.8 (σ 代数间的独立性)
$(\Omega,\mathcal{F}_1,\mathbb{P}),(\Omega,\mathcal{F}_2,\mathbb{P})$ 称 $\mathcal{F}_1$ 与 $\mathcal{F}_2$ 独立, 若 $\forall A_1\in\mathcal{F}_1, A_2\in\mathcal{F}_2$, 有 $A_1\perp\!\!\!\perp A_2$, 记为 $\mathcal{F}_1\perp\!\!\!\perp\mathcal{F}_2$

Definition 1.9 (σ 代数间相互独立)
$(\Omega,\mathcal{F}_k,\mathbb{P})(k\geqslant 1)$ 称 $\{\mathcal{F}_k\}_{k\geqslant 1}$ 相互独立, 若 $\forall J\subset\mathbb{N},\#J\geqslant 2,\forall A_k\in\mathcal{F}_k(k\in J)$, 有

$$\mathbb{P}(\bigcap_{k\in J}A_k)=\prod_{k\in J}P(A_k)$$

Property 1.5
$\{\mathcal{F}_k\}_{k\geqslant 1}$ 相互独立 $\Leftrightarrow\forall A_k\in\mathcal{F}_k,\mathbb{P}(\cap_{k\geqslant 1}A_k)=\prod_{k=1}^{\infty}\mathbb{P}(A_k)$

证明: $\Rightarrow$ 显然, J 取 $\mathbb{N}$ 即可, $\mathbb{N}\subset\mathbb{N}$
$\Leftarrow$ 注意到右侧 $\forall A_k\in\mathcal{F}$ 对于左侧条件 $\forall A_k\in\mathcal{F}(k\in J)$ 更加一般, 所以证 $\Leftarrow$ 的过程也是从一般到特殊. 从 $\cap_{k\geqslant 1}A_k\to\cap_{k\in J}A_k$ 即从 $k\in\mathbb{N}\to k\in J$. 思路是把 $k\in\mathbb{N}$ 分成 $k\in J$ 和 $k\in J^c$, 在 $k\in J^c$ 上取 $A_k=\Omega$, 再利用性质 $\Omega\perp\!\!\!\perp A$.
对于 $\forall J\subseteq\mathbb{N}$

$$\bigcap_{k\geqslant 1}A_k=\left(\bigcap_{k\in J}A_k\right)\cap\left(\bigcap_{k\in J^c}\Omega\right)$$

$$\begin{aligned}\mathbb{P}(\bigcap_{k\geqslant 1}A_k)&=\mathbb{P}\left((\bigcap_{k\in J}A_k)\cap(\bigcap_{k\in J^c}\Omega)\right)\\&=\mathbb{P}(\bigcap_{k\in J}A_k)\mathbb{P}(\bigcap_{k\in J^c}\Omega)\qquad[\Omega\perp\!\!\!\perp A_k]\\&=\mathbb{P}(\bigcap_{k\in J}A_k)\end{aligned}$$

$$\prod_{k\geqslant 1}\mathbb{P}(A_k)=\prod_{k\in J}\mathbb{P}(A_k)\cdot\prod_{k\in J^c}\mathbb{P}(\Omega)=\Pi_{k\in J}\mathbb{P}(A_k)$$

又因为 $\mathbb{P}(\cap_{k\geqslant 1}A_k)=\prod_{k=1}^{\infty}\mathbb{P}(A_k)$

$$\mathbb{P}(\bigcap_{k\in J}A_k)=\prod_{k\in J}\mathbb{P}(A_k)\quad\square$$

Definition 1.10 (离散随机变量)

令取值空间 $S=\{x_k\}_{k\geqslant 1}$ (x_k 互不相同), $\Omega=\sum_{k\geqslant 1}\Lambda_k$ (划分), 则称

$$X(\omega)=\sum_{k\geqslant 1}x_k\mathbb{I}_{\Lambda_k}(\omega),\omega\in\Omega \tag{1.1}$$

为离散随机变量. 其中

$$\mathbb{I}_{\Lambda_k}(\omega)=\begin{cases}1 & \text{if } \omega\in\Lambda_k\\ 0 & \text{if } \omega\notin\Lambda_k\end{cases}$$

这个定义的核心思想是:

- 对于每个样本点 $\omega\in\Omega$, $X(\omega)$ 的取值是 x_k, 当且仅当 $\omega\in\Lambda_k$
- 因此, X 的取值由样本点 ω 所在的划分 Λ_k 决定

由于随机变量是个可测函数

$$X:(\Omega,?)\to(S,2^S)$$

那么 X 生成的 σ 代数表示为 $\sigma(X):=X^{-1}(2^S)=\{X^{-1}(A)|A\in 2^S\}$

Property 1.6

$\sigma(X):=X^{-1}(2^S)$, 则

1. $\sigma(X)=\sigma(\Pi_\Omega)$ 故称 $\sigma(X)$ 为由 X 生成的 σ 代数. 其中 $\Pi_\Omega=\{\Lambda_k,k\geqslant 1\},\Lambda_k=\{X=x_k\}$
2. $X:(\Omega,\sigma(X))\to(S,2^S)$. 这个记号的解释是 $\forall A\in 2^S,X^{-1}(A)=\{\omega\in\Omega|X(\omega)\in A\}\in\sigma(X)$

证明: 要证 $\sigma(X)=\sigma(\Pi_\Omega)$, 即证两个集合互相包含

$\sigma(\Pi_X)=\{\sum_{k\in J}\Lambda_k|J\subseteq\mathbb{N}\}$ 由划分生成, $\sigma(X)=X^{-1}(2^S)$ 由 X 生成

下证 $\sigma(X)\subseteq\sigma(\Pi_X)$

$$\begin{aligned}\forall A\in 2^S,X^{-1}(A)&=\{\omega|X(\omega)\in A\}\\&=\sum_{x_k\in A}\{\omega\in\Omega|X(\omega)=x_k\}\\&=\sum_{x_k\in A}\{X=x_k\}\\&=\sum_{x_k\in A}\Lambda_k\in\sigma(\Pi_X)\end{aligned}$$

第二个等式用到离散 r.v. 定义1.10

下证 $\sigma(\Pi_X) \subseteq \sigma(X)$

$$
\begin{aligned}
J \subseteq \mathbb{N}, \quad \sum_{k\in J} \Lambda_k &= \sum_{k\in J} \{\omega | X(\omega) = x_k\} \\
&= \{\omega | X(\omega) \in \{x_k, k \in J\}\} \\
&= X^{-1}(\{x_k, k\in J\}) \in \sigma(X)
\end{aligned}
$$

最后一个等式中 $\{x_k, k \in J\} \in 2^S$ □

Example 1.5

$X = \mathbb{I}_A$ 由划分的定义 $\Pi_X = \{\Lambda_k\}_{k\geqslant 1}, \Lambda_k = \{X = x_k\}$, 知道划分将全集分成两部分

$$
\begin{aligned}
\Pi_X &= \{\{X = 1\}, \{X = 0\}\} \\
&= \{\{\omega \in \Omega | X(\omega) = 1\}, \{\omega \in \Omega | X(\omega) = 0\}\} \\
&= \{A, A^c\}
\end{aligned}
$$

$\sigma(\Pi_A) = \{\varnothing, A, A^c, \Omega\} = \sigma(A) = \sigma(A^c)$

其中 $\sigma(\Pi_A)$ 由划分生成, $\sigma(A)$ 由 A 生成, 两者相等

另外, $\sigma(X) = \sigma(\mathbb{I}_A) = \sigma(\Pi_X) = \{\varnothing, A, A^c, \Omega\} = \sigma(A) \Rightarrow \sigma(\mathbb{I}_A) = \sigma(A)$

Definition 1.11 (离散随机变量间的独立性)

$X : \Omega \to S_1, Y : \Omega \to S_2$ 为两离散随机变量, 称 $X \perp\!\!\!\perp Y$, 若 $\sigma(X) \perp\!\!\!\perp \sigma(Y)$[定义1.8], 即 $X^{-1}(2^{S_1}) \perp\!\!\!\perp X^{-1}(2^{S_2})$

即 $\forall E_1 \subseteq S_1, E_2 \subseteq S_2$, 有 $\mathbb{P}(X \in E_1, Y \in E_2) = \mathbb{P}(X \in E_1)\mathbb{P}(Y \in E_2)$

S_1, S_2 分别为 X, Y 的取值空间, $E_1 \subseteq S_1$ 为 X 的一个取值, $X \in E_1 := \{\omega \in \Omega | X(\omega) \in E_1\}$, E_2 同理

Theorem 1.1

$X \perp\!\!\!\perp Y \Leftrightarrow \forall x \in S_X, y \in S_Y$ 有$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$

证明: $\Rightarrow$ 一般到特殊, 取 $E_1 = \{x\}, E_2 = \{y\}$, 由 $\{x\} \in S_X, \{y\} \in S_Y$ 易证

$\Leftarrow$

$$
\begin{aligned}
\mathbb{P}(X \in E_1, Y \in E_2) &= \mathbb{P}(\bigcup_{x\in E_1} \{X = x\} \cap \{Y \in E_2\}) \\
&= \sum_{x\in E_1} \mathbb{P}(\{X = x\} \cap \sum_{y \in E_2} \{Y = y\}) \\
&= \sum_{x\in E_1} \sum_{y\in E_2} \mathbb{P}(X = x, Y = y) \\
&= \sum_{x\in E_1} (\sum_{y\in E_2} \mathbb{P}(X = x)\mathbb{P}(Y = y)) \\
&= \sum_{x\in E_1} \mathbb{P}(X = x)\mathbb{P}(Y \in E_2) \\
&= \mathbb{P}(X \in E_1)\mathbb{P}(Y \in E_2)
\end{aligned}
$$

第一个等式中, $\{X=x\}\cap\{Y\in E_2\}$ 看作一整个集合 $\subseteq\{X=x\}$, 因为离散、每个 x 不相交, 所以这是个不交并, 由练习2, 可以改写成加法形式.
第四个等式由条件 $\mathbb{P}(X=x,Y=y)=\mathbb{P}(X=x)\mathbb{P}(Y=y)$ 成立. □

Theorem 1.2
$X\perp\!\!\!\perp Y\Leftrightarrow\forall x\in S_X,y\in S_Y,\mathbb{P}(X\leqslant x,Y\leqslant y)=\mathbb{P}(X\leqslant x)\mathbb{P}(Y\leqslant y)$

用 Theorem 1.1证明
$\Rightarrow$ 已知 $X\perp\!\!\!\perp Y$, 由定义1.11, $\forall E_1\subseteq S_1,E_2\subseteq S_2$, 有 $\mathbb{P}(X\in E_1,Y\in E_2)=\mathbb{P}(X\in E_1)\mathbb{P}(Y\in E_2)$. 取 $E_1=\{\omega|X(\omega)\leqslant x\},E_2=\{\omega|Y(\omega)\leqslant y\}$
$\Leftarrow$

$$\begin{aligned}\mathbb{P}(X=x,Y=y)&=\mathbb{P}(X\leqslant x,Y\leqslant y)-\mathbb{P}(X\leqslant x^-,Y\leqslant y)-\mathbb{P}(X\leqslant x,Y\leqslant y^-)+\mathbb{P}(X\leqslant x^-,Y\leqslant y^-)\\&=\mathbb{P}(X\leqslant x)\mathbb{P}(Y\leqslant y)-\mathbb{P}(X\leqslant x^-)\mathbb{P}(Y\leqslant y)-\mathbb{P}(X\leqslant x)\mathbb{P}(Y\leqslant y^-)+\mathbb{P}(X\leqslant x^-)\mathbb{P}(Y\leqslant y^-)\\&=[\mathbb{P}(X\leqslant x)-\mathbb{P}(X\leqslant x^-)][\mathbb{P}(Y\leqslant y)-\mathbb{P}(Y\leqslant y^-)]\\&=\mathbb{P}(X=x)\mathbb{P}(Y=y)\end{aligned}$$

其中 x^-,y^- 为小于 x,y 的最大值, 由于离散, $\{X\leqslant x\}-\{X\leqslant x^-\}=\{X=x\},\{Y\leqslant y\}-\{Y\leqslant y^-\}=\{Y=y\}$

Definition 1.12
称一列离散随机变量 $\{X_n\}_{n\geqslant 1}$ 相互独立, 若 $\sigma(X_n),n\geqslant 1$ 相互独立

Theorem 1.3
$\{A_n\}_{n\geqslant 1}$ 事件列下列等价

1. $\{A_n\}_{n\geqslant 1}$ 相互独立
2. $\sigma(A_n),n\geqslant 1$ 相互独立
3. $\mathbb{I}_{A_n},n\geqslant 1$ 相互独立

证明. 1. 由例题1.5, $\sigma(\mathbb{I}_{A_n})=\sigma(A_n)$, 所以 $(2)\Leftrightarrow(3)$
2. 下证 $(2)\rightarrow(1)$, 一般到特殊, $A_n\subseteq\sigma(A_n)$
3. 下证 $(1)\rightarrow(2)$, $\sigma(A_n)=\{A_n,A_n^c,\varnothing,\Omega\}$, $\varnothing\perp\!\!\!\perp A_n,\Omega\perp\!\!\!\perp A_n$, 由性质1.4, $\varnothing\perp\!\!\!\perp A_n^c,\Omega\perp\!\!\!\perp A_n^c$
由 Property 1.5, $\forall A_k\in\sigma(A_n),\mathbb{P}(\cap_{k\geqslant 1}A_k)=\prod_{k=1}^{\infty}\mathbb{P}(A_k)$
由于条件 (1), 上面等式成立 $\Rightarrow$ 满足 σ 代数相互独立的定义. □

1.3 条件概率与条件独立

Definition 1.13 (条件概率)

$B \in \mathcal{F}, \mathbb{P}(B) > 0$ 定义

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)} =: \mathbb{P}_B(A) \quad \forall A \in \mathcal{F}$$

Theorem 1.4 (乘法公式)

$\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B)$,

$$\mathbb{P}(\bigcap_{k=1}^{n} A_k) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1A_2)\cdots\mathbb{P}(A_n|\bigcap_{k=1}^{n-1} A_k) \tag{1.2}$$

Theorem 1.5 (全概公式)

(1) $\Omega = \sum_{k\geqslant 1} \Lambda_k$ 划分 $\mathbb{P}(\Lambda_k) > 0$, 则 $\forall A \in \mathcal{F}$,

$$\mathbb{P}(A) = \sum_{k\geqslant 1} \mathbb{P}(A|\Lambda_k)\mathbb{P}(\Lambda_k)$$

(2) * 一般地, $\{B_n\}_{n\geqslant 1}$ 互不相交, $\mathbb{P}(B) > 0, \mathbb{P}(\sum_{n\geqslant 1} B_n) = 1$, 则 $\forall A \in \mathcal{F}$

$$\mathbb{P}(A) = \sum_{n\geqslant 1} \mathbb{P}(A|B_n)\mathbb{P}(B_n)$$

注：$\mathbb{P}(\cdot) = 1$ 不一定是全集, 但概率测度是 1. 同样, $\mathbb{P}(\cdot) = 0$ 不一定是 $\varnothing$, 而是叫零测集

证明：

(1) 由 $A = A \cap \Omega = A \cap (\sum_{k\geqslant 1} \Lambda_k) = \sum_{k\geqslant 1}(A \cap \Lambda_k)$, A 被划分成若干不相交的集合 $A \cap \Lambda_k$, 根据可列可加性, 得到

$$\mathbb{P}(A) = \sum_{k\geqslant 1} \mathbb{P}(A \cap \Lambda_k) = \sum_{k\geqslant 1} \mathbb{P}(A|\Lambda_k)\mathbb{P}(\Lambda_k)$$

(2) $\Omega = (\sum_{n\geqslant 1} B_n) + (\sum_{n\geqslant 1} B_n)^c = \sum_{n\geqslant 0} B_n$, 其中 $B_0 = (\sum_{n\geqslant 1} B_n)^c$

$\mathbb{P}(B_0) = 1 - \mathbb{P}(\sum_{n\geqslant 1} B_n) = 0 \to 0 \leqslant \mathbb{P}(AB_0) \leqslant \mathbb{P}(B_0) = 0$

左边不等号成立是因为概率测度非负, 右边不等号成立是因为 $AB_0 \subseteq B_0$, 所以 $\mathbb{P}(AB_0) = 0$

$$\begin{aligned}\mathbb{P}(A) &= \sum_{n\geqslant 0} \mathbb{P}(AB_n) \quad [\text{可列可加性}] \\ &= \sum_{n\geqslant 1} \mathbb{P}(AB_n) \quad [\mathbb{P}(AB_0) = 0] \\ &= \sum_{n\geqslant 1} \mathbb{P}(A|B_n)\mathbb{P}(B_n) \quad [\text{全概公式}] \quad \square\end{aligned}$$

Theorem 1.6

$\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$

$$A \perp\!\!\!\perp B \Leftrightarrow \mathbb{P}(A|B) = \mathbb{P}(A) \Leftrightarrow \mathbb{P}(B|A) = \mathbb{P}(B)$$

$\mathbb{P}(A|B)$ 见定义1.13

Theorem 1.7

$\mathbb{P}_B : \mathcal{F} \to [0,1]$ 也是 $(\Omega, \mathcal{F})$ 上的概率测度 [定义1.5]

Property 1.7

$\mathbb{P}(C) > 0, \mathbb{P}(B) > 0$, 则

$$\mathbb{P}_B(\cdot|C) = \mathbb{P}(\cdot|BC) = \mathbb{P}_{BC}(\cdot)$$

$\mathbb{P}_B(\cdot|C)$ 见定义1.13

Definition 1.14

称 C 条件发生下, A 与 B 独立, 若

$$\mathbb{P}_C(AB) = \mathbb{P}_C(A)\mathbb{P}_C(B) \tag{1.3}$$

记为 $A \perp\!\!\!\perp_C B$（条件独立）

Theorem 1.8

$\mathbb{P}(C) > 0, \mathbb{P}(BC) > 0$ 则 $A \perp\!\!\!\perp_C B \Leftrightarrow \mathbb{P}_C(A|B) = \mathbb{P}_C(A)$

证明. 由 $A \perp\!\!\!\perp_C B, \mathbb{P}_C(AB) = \mathbb{P}_C(A)\mathbb{P}_C(B)$

$$\mathbb{P}_C(A|B) = \frac{\mathbb{P}_C(AB)}{\mathbb{P}_C(B)} = \mathbb{P}_C(A)$$

□

1.4 期望与条件期望

1.4.1 离散随机变量的期望

Definition 1.15 (X 的期望)

$X:\Omega\to S$

$$\mathbb{E}(X)=\sum_{x\in S}x\mathbb{P}(X=x)=\mathbb{E}^{\mathbb{P}}(X)$$

当此求和绝对收敛

注：$\mathbb{E}^{\mathbb{P}}(X)$ 强调这是在概率测度 $\mathbb{P}$ 下的期望

Definition 1.16 ($g(X)$ 的期望)

$g:\mathbb{R}\to\mathbb{R}$

$$\mathbb{E}g(X)=\sum_{x\in S}g(x)\mathbb{P}(X=x)$$

当此求和绝对收敛

关于“求和绝对收敛”的讨论：

Example 1.6

$\mathbb{E}(\mathbb{I}_A)=\mathbb{P}(A),A\in\mathcal{F}$

Example 1.7

X 是离散随机变量，由定义1.10, $X=\sum_{x\in S}x\mathbb{I}_{A_x}$，其中 $A_x:=\{X=x\}$. B 是任意的，求 $\mathbb{E}(\mathbb{I}_BX)$

Remark 2. 对于 $A_x:=\{X=x\}$ 应这样理解，A_x 是样本空间 Ω 的一个子集，包含了所有使得 $X(\omega)=x$ 的样本点 ω.

根据离散随机变量的定义，$X(\omega)=x_k$ 当且仅当 $\omega\in\Lambda_k$. 因此对于每个 $x_k\in S$，有

$$A_{x_k}=\{X=x_k\}=\{\omega\in\Omega|X(\omega)=x_k\}=\Lambda_k$$

所以 $A_x=\{A_{x_k}\}_{k\geqslant1}$ 就是离散随机变量的划分

对于 $X=\sum_{x\in S}x\mathbb{I}_{A_x}$ 可以这样理解. 对于每个 $x\in S$, $\mathbb{I}_{A_x}(\omega)$ 是事件 $A_x=\{X=x\}$ 的指示函数

$$\mathbb{I}_{A_x}(\omega)=\begin{cases}1 & \textit{if } X(\omega)=x\\ 0 & \textit{if } X(\omega)\neq x\end{cases}$$

Solution. 要先求 $\mathbb{E}(|\mathbb{I}_BX|)<\infty$ 说明期望存在

对 $\forall\omega\in B$

$$\begin{aligned}\mathbb{I}_BX(\omega)&=\mathbb{I}_B(\omega)\sum_{x\in S}(x\cdot\mathbb{I}_{A_x}(\omega))\\&=\sum_{x\in S}x\mathbb{I}_{A_x\cap B}(\omega)\end{aligned}$$

其中 $\mathbb{I}_{A_x\cap B}$ 也可记为 $\mathbb{I}_{A_xB}$
$\{A_xB, x\in S\}\cup\{B^c\}$ 构成了样本空间 Ω 的一个划分. 因为 A_x 本身是对 Ω 的一个划分, 其与 B 的交是对 B 的划分. 并上 B^c, 则满足划分的定义*1.4*
对于 $\omega\in\Omega$, 由划分

$$\mathbb{I}_BX(\omega)=0\cdot\mathbb{I}_{B^c}(\omega)+\sum_{x\in S}x\mathbb{I}_{A_x\cap B}$$

$$\therefore \mathbb{E}|\mathbb{I}_BX|=\sum_{x\in S}|x|\mathbb{P}(A_xB)\leqslant\sum_{x\in S}|x|\mathbb{P}(A_x)=\mathbb{E}|X|<\infty$$

最后一个等号参考期望的定义*1.15*

$$\mathbb{E}(\mathbb{I}_BX)=\sum_{x\in S}x\mathbb{P}(A_xB)=\sum_{x\in S}x\mathbb{P}(\{X=x\}\cap B)$$

Theorem 1.9
$\mathbb{E}(aX+bY)=a\mathbb{E}X+b\mathbb{E}Y$

离散随机变量有两种表达形式, 如定义1.10和练习1.7所示

$$X=\sum_{x\in S}x\mathbb{I}_{\{X=x\}}=\sum_{k\geqslant1}x_k\mathbb{I}_{\Lambda_k}$$

$$\sum_{x\in S}x\mathbb{P}(X=x)=\sum_{k\geqslant1}\mathbb{P}(X=x_k)$$

只有在“求和绝对收敛”(见定义1.15)的条件下, 等式才成立

Remark 3.

1. $\sum_{x\in S}$ *(1)* 级数的重排 *(2)* 可和族
2. X 是离散随机变量, $g:\mathbb{R}\to\mathbb{R}$, 则

$$g(X)=\sum_{x\in S}g(x)\mathbb{I}_{\{X=x\}}$$

   是一个离散随机变量, 且 $\sigma(g(X))\subseteq\sigma(X)$. 下面说明这个结论
   当 $x_1\neq x_2$ 时可能 $g(x_1)=g(x_2)$, 因此

$$\Pi_X=\{\{X=x\}|x\in S\}\neq\Pi_{g(X)}$$

   其实 $\Pi_{g(X)}\subseteq\sigma(\Pi_X)$, 因为对于 $x_1\neq x_2$ 但 $g(x_1)=g(x_2)$ 的情况, 比如在 Π_X 上 x_1,x_2 对应的样本空间是 Ω_1,Ω_2, 但在 $\Pi_{g(X)}$ 上是 $\Omega_1\cup\Omega_2$. 这一项在 Π_X 里有, 因为 σ 代数对可列并封闭. 但 Ω_1,Ω_2 分别在 $\Pi_{g(X)}$ 上没有. 把 σ 代数理解成信息, 则 $g(X)=y$ 提供的信息是比直接提供 x 的值要少的(在 $g(\cdot)$ 已知的情况下).
3. $X\perp\!\!\!\perp Y$, $g,h:\mathbb{R}\to\mathbb{R}$, 则 $g(X)\perp\!\!\!\perp h(Y)$. 因为 $\sigma(X)\perp\!\!\!\perp\sigma(Y)$, 而 $\sigma(g(X))\subseteq\sigma(X),\sigma(h(Y))\subseteq\sigma(Y)$
   如果 X,Y 是连续随机变量, 则对 g,h 有其他要求. 特殊地, 结论 *3* 对 g,h 连续时成立.

Theorem 1.10

(1) $X \perp\!\!\!\perp Y, \mathbb{E}|X| < \infty, \mathbb{E}|Y| < \infty$, 则 $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

(2) $X_1, X_2, \cdots, X_n$ 相互独立, 则 $\mathbb{E}(X_1 \cdots X_n) = \mathbb{E}X_1 \cdots \mathbb{E}X_n$

(3) $X \perp\!\!\!\perp Y, g, h : \mathbb{R} \to \mathbb{R}, \mathbb{E}|g(X)| < \infty, \mathbb{E}|h(Y)| < \infty$

$$\Rightarrow g(X) \perp\!\!\!\perp h(Y), \mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$$

Theorem 1.11

若 $X \geqslant 0$ 取整数值, 则 $\mathbb{E}(X) = \sum_{k \geqslant 1} \mathbb{P}(X \geqslant k)$

证明:

1.4.2 条件期望

1° 关于"给定集合"的条件期望

Definition 1.17

$(\Omega, \mathcal{F}, \mathbb{P}), X : \Omega \to S, A \in \mathcal{F}, \mathbb{P}(A) > 0, \mathbb{E}|X| < \infty$, 定义 X 关于 A 的条件期望

$$\begin{aligned} \mathbb{E}(X|A) &:= \sum_{x \in S} \mathbb{P}(X = x|A) \\ &= \sum_{x \in S} x\mathbb{P}_A(X = x) \\ &= E^{\mathbb{P}_A}(X) \end{aligned}$$

Property 1.8 (线性性)

$\mathbb{E}(aX + bY|A) = a\mathbb{E}(X|A) + b\mathbb{E}(Y|A)$

证明:(用期望的性质)

Example 1.8

$\mathbb{E}(\mathbb{I}_B|A) = 1 \cdot \mathbb{P}(B|A) + 0 \cdot \mathbb{P}(B^c|A) = \mathbb{P}(B|A)$

Example 1.9

$B \perp\!\!\!\perp A \Rightarrow \mathbb{E}(\mathbb{I}_B|A) = \mathbb{E}(\mathbb{I}_B)$

Property 1.9

$\mathbb{E}|X| < \infty, \mathbb{P}(A) > 0, X \perp\!\!\!\perp \mathbb{I}_A \Rightarrow \mathbb{E}(X|A) = \mathbb{E}(X)$

证明：

$\because X \perp\!\!\!\perp \mathbb{I}_A, \therefore \{X=x\} \perp\!\!\!\perp A$

$$\mathbb{E}(X|A) = \sum_{x\in S} x\mathbb{P}(X=x|A) = \sum_{x\in S} x\mathbb{P}(X=x) = \mathbb{E}(X)$$

其中

$$\sum_{x\in S} x\mathbb{P}(X=x|A) = \sum_{x\in S} x\frac{\mathbb{P}(\{X=x\}\cap A)}{P(A)} = \mathbb{E}(X\mathbb{I}_A)/\mathbb{P}(A)$$

最后一个等号由例题1.7

至此没有用到独立性, 可以得到以下推论

Corollary 1.2

$\mathbb{E}(X|A) = \mathbb{E}(X\mathbb{I}_A)/\mathbb{P}(A)$

Problem 3 (作业 2-1)

Y 在 A 上取常数 c, 证明：$\mathbb{E}(XY|A) = c\mathbb{E}(X|A)$

$2°$ 关于“给定划分生成的 σ 代数”的条件期望

Definition 1.18

设 $\Pi = \{\Lambda_k\}_{k\geqslant 1}$ 是 Ω 的划分, X 为离散随机变量, $\mathbb{E}|X| < \infty$, 定义

$$\mathbb{E}(X|\sigma(\Pi))(\omega) := \mathbb{E}(X|\Lambda_k)$$

当 $\omega \in \Lambda_k$, 即

$$\mathbb{E}(X|\sigma(\Pi)) = \sum_{k\geqslant 1} \mathbb{I}_{\Lambda_k}\mathbb{E}(X|\Lambda_k)$$

期望的本质是积分, 现在因为数分里的积分不够用了, 我们要定义新积分, 希望它也能保留原先的好性质

Property 1.10 (线性性)

$\mathbb{E}(aX+bY|\sigma(\Pi)) = a\mathbb{E}(X|\sigma(\Pi)) + b\mathbb{E}(Y|\sigma(\Pi))$

证明：$\omega \in \Lambda_k$, $LHS = \mathbb{E}(aX+bY|\Lambda_k) = a\mathbb{E}(X|\Lambda_k) + b\mathbb{E}(Y|\Lambda_k)$

第二个等号由性质1.8成立.

Example 1.10

$$\begin{aligned}
\mathbb{E}(X|\{\varnothing,\Omega\}) &= \mathbb{E}(X|\sigma(\Omega)) \\
&= \mathbb{I}_\Omega \mathbb{E}(X|\Omega) \qquad [\text{定义}(1.18)] \\
&= \sum_{x\in S} x\mathbb{P}(X=x|\Omega) \qquad [\text{定义}(1.17), \Omega \perp\!\!\!\perp X] \\
&= \sum_{x\in S} x\mathbb{P}(X=x) \\
&= \mathbb{E}(X)
\end{aligned}$$

独立可以理解为：什么信息也没提供

Example 1.11

$$\begin{aligned}
\mathbb{E}(\mathbb{I}_B|\sigma(A)) &= \mathbb{E}(\mathbb{I}_B|\{A, A^c, \Omega, \varnothing\}) \\
&= \mathbb{E}(\mathbb{I}_B|\sigma(A, A^c)) \\
&= \mathbb{I}_A \mathbb{E}(\mathbb{I}_B|A) + \mathbb{I}_{A^c}\mathbb{E}(\mathbb{I}_B|A^c)
\end{aligned}$$

更进一步，若 $A \perp\!\!\!\perp B$，由 $\sigma(B) \perp\!\!\!\perp \sigma(A) \to \sigma(\mathbb{I}_B) \perp\!\!\!\perp \sigma(\mathbb{I}_A) \Rightarrow \mathbb{E}(\mathbb{I}_B|\sigma(A)) = \mathbb{E}(\mathbb{I}_B)$

可以把这个结果推广：

Property 1.11

$\sigma(X) \perp\!\!\!\perp \sigma(\Pi)$，则 $\mathbb{E}(X|\sigma(\Pi)) = \mathbb{E}(X)$

证明：$\Pi_X = \{\{X=x\}|x\in S\}$，默认 x 不相同

$\sigma(X) = \sigma(\Pi_X) = \{\{X=x\}|x\in S\}$

不妨设 $\Pi = \{\Lambda_k, k \geqslant 1\}$

则 $\sigma(X) \perp\!\!\!\perp \sigma(\Pi) \Rightarrow \forall x \in S, k \geqslant 1, \{X=x\} \perp\!\!\!\perp \Lambda_k$

$$\begin{aligned}
\mathbb{E}(X|\sigma(\Pi)) &= \sum_{k\geqslant 1} \mathbb{I}_{\Lambda_k}\mathbb{E}(X|\Lambda_k) \\
&= \sum_{k\geqslant 1} \mathbb{I}_{\Lambda_k} \sum_{x\in S} x\mathbb{P}(X=x|\Lambda_k) \\
&= \sum_{k\geqslant 1} \mathbb{I}_{\Lambda_k} \sum_{x\in S} x\mathbb{P}(X=x) \\
&= \sum_{k\geqslant 1} \mathbb{I}_{\Lambda_k}\mathbb{E}(X) \\
&= \mathbb{I}_\Omega \mathbb{E}(X) \\
&= \mathbb{E}(X)
\end{aligned}$$

Example 1.12

$\mathbb{E}(X|\sigma(X)) = X$

$\sigma(X)$ 作为条件相当于知道了与 X 相关的所有信息, 即提取已知量

证明: $\sigma(X) = \sigma(\Pi_X)$, 其中 $\mathbb{I}_X = \{\{X = x\}|x \in S\}$

$$\begin{aligned}
\mathbb{E}(X|\sigma(X)) &= \sum_{x\in S} \mathbb{I}_{\{X=x\}}\mathbb{E}(X|X = x) \\
&= \sum_{x\in S} \mathbb{I}_{\{X=x\}}\mathbb{E}(X\mathbb{I}_{\{X=x\}})/\mathbb{P}(X = x) \quad [\text{Cor } (1.2)] \\
&= \sum_{x\in S} \mathbb{I}_{\{X=x\}} \cdot \frac{x \cdot \mathbb{P}(X = x) + 0 \cdot \mathbb{P}(X \neq x)}{\mathbb{P}(X = x)} \\
&= \sum_{x\in S} x\mathbb{I}_{\{X=x\}} = X \quad \square
\end{aligned}$$

Property 1.12 (提取已知量)

设 $\Pi = \{\Lambda_k, k \geqslant 1\}$ 为 Ω 的划分, $\mathbb{E}|X| < \infty, \mathbb{E}|XY| < \infty$, 则当 $\sigma(X) \subseteq \sigma(\Pi)$ 时, 有

1. $\mathbb{E}(X|\sigma(\Pi)) = X$
2. $\mathbb{E}(XY|\sigma(\Pi)) = X\mathbb{E}(Y|\sigma(\Pi))$

特别地, 取 $X = \mathbb{I}_A, A \in \sigma(\Pi)$, 则

1. $\mathbb{E}(\mathbb{I}_A|\sigma(\Pi)) = \mathbb{I}_A$
2. $\mathbb{E}(\mathbb{I}_AY|\sigma(\Pi)) = \mathbb{I}_A\mathbb{E}(Y|\sigma(\Pi))$

证明: 只需证 (2), 因为从 (2)→(1) 即 $Y = \mathbb{I}_\Omega$

$X = \sum_{x\in S} x\mathbb{I}_{A_x}$, 其中 $A_x := \{X = x\}$

(Step 1) $\sigma(X) = \{\sum_{x\in S'_X} A_x|S'_X \subseteq S_X\}$

$\sigma(X) = \{\sum_{k\in J} \Lambda_k|J \subseteq \mathbb{N}\}$

已知: $\sigma(X) \subseteq \sigma(\Pi) \Rightarrow \exists$ 一族 $\{x_k\}_{k\geqslant 1}$（可能有相同元素）, 使得 $X = \sum_{k\geqslant 1} x_k\mathbb{I}_{\Lambda_k}$, 其中 $\cup_{k\geqslant 1}\{x_k\} = S_x$ (S_x 为取值空间)

注: Π 是 $\Pi_X = \{A_x|x \in S\}$ 的加细划分

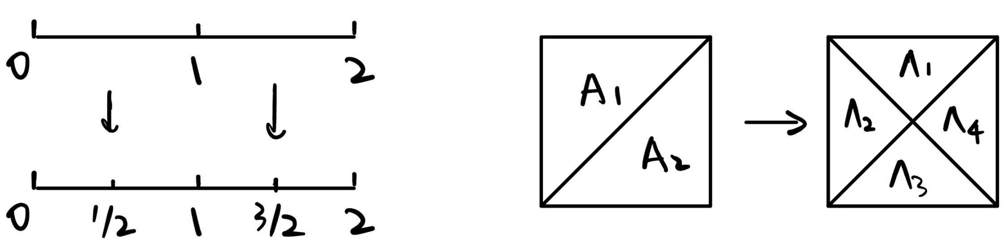


图 1: 加细划分

(Step 2) 对于 $\omega \in \Lambda_j, \forall j \geqslant 1$

$$\begin{aligned}
\mathbb{E}(XY|\sigma(\Pi))(\omega) &= \mathbb{E}(\sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k} Y|\sigma(\Pi))(\omega) \qquad [X = \sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k}] \\
&= \mathbb{E}(\sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k} Y|\Lambda_j) \qquad [\sigma(\Pi)\text{定义}] \\
&= \mathbb{E}(\sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k} Y \mathbb{I}_{\Lambda_j})/\mathbb{P}(\Lambda_j) \qquad [\text{Cor } (1.2)] \\
&= \mathbb{E}(Y x_j \mathbb{I}_{\Lambda_j})/\mathbb{P}(\Lambda_j) \qquad [\mathbb{I}_{\Lambda_k}\mathbb{I}_{\Lambda_j}\text{当}\Lambda_k \neq \Lambda_j\text{时} = 0] \\
&= x_j \mathbb{E}(Y\mathbb{I}_{\Lambda_j})/\mathbb{P}(\Lambda_j) \\
&= x_j \mathbb{E}(Y|\mathbb{I}_{\Lambda_j}) \\
&= X(\omega)\mathbb{E}(Y|\mathbb{I}_{\Lambda_j})
\end{aligned}$$

$$\Rightarrow \mathbb{E}(XY|\sigma(\Pi)) = X \sum_{j\geqslant 1} \mathbb{I}_{\Lambda_j} \mathbb{E}(Y|\Lambda_j) = X\mathbb{E}(Y|\sigma(\Pi))$$

数学上有种现象叫“法国人的伎俩”，即把定理当定义用．严格地讲，这么做有时会出现存在性和唯一性不满足的问题．下面介绍一个常被当做定义用的定理：

Theorem 1.12

$\Pi = \{\Lambda_k, k \geqslant 1\}$ 为 Ω 的划分, $\mathbb{E}|X| < \infty$. 记 $Y := \mathbb{E}(X|\sigma(\Pi)) = \sum_{k\geqslant 1} \mathbb{I}_{\Lambda_k} \mathbb{E}(X|\Lambda_k)$, 则

1. Y 仍是一个离散随机变量, 且 $\mathbb{E}|Y| \geqslant \mathbb{E}|X| < \infty$
2. $\sigma(Y) \subseteq \sigma(\Pi)$ (记作 $Y \in \sigma(\Pi)$, 即 Y 的所有信息都在 $\sigma(\Pi)$ 里)
3. $\forall A \in \sigma(\Pi)$, 有 $\mathbb{E}(Y\mathbb{I}_A) = \mathbb{E}(X\mathbb{I}_A)$

证明：(1)$E|X| = \sum_{x\in S_x} |x|\mathbb{P}(X = x) < \infty$

$$\mathbb{E}|Y| = \sum_{k\geqslant 1} |\mathbb{E}(X|\Lambda_k)|\mathbb{P}(\Lambda_k) \geqslant \sum_{k\geqslant 1}\sum_{x\in S} |x|\mathbb{P}(\{X = x\} \cap \Lambda_k)$$

逻辑上, 现在第一个等号不成立, 但之后 $< \infty$ 一写出来, 之前的所有等号立刻成立, 此处只为书写简便

$$\mathbb{E}|X| = \sum_{x\in S_x} |x|\mathbb{P}(X = x) = \sum_{x\in S} |x| \sum_{k\geqslant 1} \mathbb{P}(\Lambda_k \cap \{X = x\})$$

我们知道 $\sum_{x\in S} |x| \sum_{k\geqslant 1} \mathbb{P}(\Lambda_k \cap \{X = x\})$ 绝对收敛, 若求和次序交换后的 $\sum_{k\geqslant 1}\sum_{x\in S} |x|\mathbb{P}(\{X = x\} \cap \Lambda_k)$ 也绝对收敛, 则 $\mathbb{E}|Y| < \infty$ 得证. 有一个引理可以保证绝对收敛：

Lemma 1.1 ([7].P280. 推论)

从 273-280

Corollary 1.3

来自 Thm 1.12(1).

1. (重期望公式)
$$\mathbb{E}|\mathbb{E}(X|\sigma(\Pi))| = \mathbb{E}|X|, \mathbb{E}(\mathbb{E}(X|\sigma(\Pi))) = \mathbb{E}(X) \tag{1.4}$$
2. $|\mathbb{E}(X|\Lambda_k)| \leqslant \mathbb{E}(|X| \mid \Lambda_k), |\mathbb{E}(X|\sigma(\Pi))| \leqslant \mathbb{E}(|X| \mid \sigma(\Pi))$

(2) 由定义, $Y = \sum_{k\geqslant 1} y_k \mathbb{I}_{\Lambda_k}$, 其中 $y_k := \mathbb{E}(X|\Lambda_k)$
记 $S_Y = \cup_{k\geqslant 1}\{y_k\}$, 注意到, 可能 $\exists i \neq j$, 但 $y_i = y_j$
故 $J_y = \{k|y_k = y\}(y \in S_Y)$ 中个数可能大于 1

$$Y = \sum_{y\in S_Y} y\mathbb{I}_{\sum_{k\in J_y}\Lambda_k}$$

$$\{Y = y\} = \sum_{k\in J_y} \Lambda_k \in \sigma(\Pi)$$

$$\sigma(Y) \subseteq \sigma(\Pi) \quad \square$$

(3) $\mathbb{E}(Y\mathbb{I}_A) = \mathbb{E}(\mathbb{I}_A \mathbb{E}(X|\sigma(\Pi)))$

$$\begin{aligned}\mathbb{E}(Y\mathbb{I}_A) &= \mathbb{E}(\mathbb{I}_A \mathbb{E}(X|\sigma(\Pi))) \\ &= \mathbb{E}(\mathbb{E}(X\mathbb{I}_A|\sigma(\Pi))) \qquad [A \in \sigma(\Pi), \text{性质}(1.12)] \\ &= \mathbb{E}(X\mathbb{I}_A) \qquad [\text{Cor } (1.3)]\end{aligned}$$

3° 关于离散随机变量的条件期望

Definition 1.19

概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$, X, Y 为离散随机变量, $\mathbb{E}|X| < \infty$. 定义 $\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma(Y)) = \mathbb{E}(X|\sigma(\Pi_Y))$, 称为 X 关于 Y 的条件期望

注: $\omega = \{Y = y\} \in \Pi_Y$ 或 $Y(\omega) = y$, $\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X|Y = y)$

Example 1.13

$\mathbb{E}(X|\Pi_\Omega) = \mathbb{E}(X|\sigma(\Omega)) = \mathbb{E}(X)$

Example 1.14

$\mathbb{I}_A \perp\!\!\!\perp \mathbb{I}_B \Rightarrow \mathbb{E}(\mathbb{I}_A|\mathbb{I}_B) = [\text{Exa}(1.11)]\mathbb{E}(\mathbb{I}_A)$

Example 1.15

$\mathbb{E}(X|X) = \mathbb{E}(X|\sigma(X)) = X[\text{Exa } 1.12]$

Property 1.13

假设以下期望、条件期望都有意义

1. $\mathbb{E}(aX+bY|Z)=a\mathbb{E}(X|Z)+b\mathbb{E}(Y|Z)$
2. $X \perp\!\!\!\perp Y \Rightarrow \mathbb{E}(X|Y)=\mathbb{E}(X)$
3. $\sigma(X)\subseteq\sigma(Z)\Rightarrow\mathbb{E}(XY|Z)=X\mathbb{E}(Y|Z)$
4. $\mathbb{E}(\mathbb{E}(X|Z))=\mathbb{E}(X)$
5. $|\mathbb{E}(X|Z)|\leqslant\mathbb{E}(|X|\mid Z)$

4° 关于多个离散随机变量的条件期望

$\mathbb{E}(Y|X_1,\cdots,X_n)$

1. 由 $X_1,\cdots,X_n$ 生成的 σ 代数 $\sigma(X_1,\cdots,X_n)$
2. $:=\mathbb{E}(Y|\sigma(X_1,\cdots,X_n))$

怎样生成 σ 代数可以包含 $X_1,\cdots,X_n$ 尽可能多的信息?
直觉是 $\bigcup_{k=1}^{\infty}\sigma(X_k)$, 然而它不一定是 σ 代数, 因为它对可列并不封闭.
每个 $\sigma(X_k)$ 是一个 σ 代数, 因此它对可列并封闭.
然而, $\bigcup_{k=1}^{\infty}\sigma(X_k)$ 只是将每个 $\sigma(X_k)$ 中的集合简单地并在一起, 并没有保证这些集合的可列并仍然在 $\bigcup_{k=1}^{\infty}\sigma(X_k)$ 中.
例如, 假设 $X_k\in\sigma(X_k)$, 那么 X_k 在 $\bigcup_{k=1}^{\infty}\sigma(X_k)$ 中, 但 $\bigcup_{k=1}^{\infty}X_k$ 可能不在 $\bigcup_{k=1}^{\infty}\sigma(X_k)$ 中, 因为它可能不属于任何一个单独的 $\sigma(X_k)$. 问题出在 $\bigcup_{k=1}^{\infty}\sigma(X_k)$ 缺少 $\{\sigma(X_k)\}_{k\geqslant1}$ 交互的部分
怎样把 $\bigcup_{k=1}^{\infty}\sigma(X_k)$ 变成 σ 代数?

Definition 1.20 (多个离散随机变量的条件期望)

定义由离散随机变量 $X_1,\cdots,X_n$ 生成的 σ 代数

$$\begin{aligned}\sigma(X_1,\cdots,X_n) &:= (X_1,\cdots,X_n)^{-1}(2^{S_1}\times\cdots\times2^{S_n})\\ &:= \{\underbrace{(X_1,\cdots,X_n)^{-1}(A_1\times\cdots\times A_n)}_{\text{柱集}}|A_1\times\cdots\times A_n\subseteq\underbrace{S_1\times\cdots\times S_n}_{\text{乘积空间}}\}\\ &= \{\bigcap_{k=1}^{\infty}X_k^{-1}(A_k)|A_k\in2^{S_k},1\leqslant k\leqslant n\}\end{aligned}$$

Theorem 1.13

令 $x_k = \sum_{i\geqslant 1} x_{k,i}\mathbb{I}_{\Lambda_{k,i}}, 1 \leqslant k \leqslant n$, 为离散随机变量, 对每一个 k, $\Pi_k := \{\Lambda_{k,i}|i \geqslant 1\}$ 为 Ω 的划分, 定义

$$\Pi_{(X_1,\cdots,X_n)} := \{\Lambda_{1,i_1} \cap \cdots \cap \Lambda_{n,i_n}|i_k \geqslant 1, 1 \leqslant k \leqslant n\}$$

则

1. $\Pi_{(X_1,\cdots,X_n)}$ 是 Ω 的划分, 且

$$\sigma(\Pi_{(X_1,cdots,X_n)}) = \left\{ \sum_{\substack{(i_1,\cdots,i_n)\\ \in J_1\times\cdots\times J_n}} (\Lambda_{1,i_1} \cap \cdots \cap \Lambda_{1,i_n})|J_k \subseteq \mathbb{N}, 1 \leqslant k \leqslant n \right\}$$

2. $\sigma(X_1, \cdots, X_n) = \sigma(\Pi_{(X_1,\cdots,X_n)})$（即定义1.20是有意义的, well-defined, make sense, 良定义）

Problem 4 (作业 2-2)

证明 Theorem 1.13在 $n = 2$ 时成立

Definition 1.21

$\mathbb{E}|Z| < \infty$ 定义

$$\mathbb{E}(Z|X_1, \cdots, X_n) = \mathbb{E}(Z|\sigma(X_1, \cdots, X_n)) := \mathbb{E}(Z|\sigma(\Pi_{(X_1,\cdots,X_n)}))$$

Definition 1.22

$(\Omega, \mathcal{F}, \mathbb{P}), Y : \Omega \to S_Y, X_1 : \Omega \to S_1, X_2 : \Omega \to S_2$ 为离散随机变量, 称 Y 和 (X_1, X_2) 独立, 若 $\sigma(Y) \perp\!\!\!\perp \sigma(X_1, X_2)$. $[\sigma(Y) = Y^{-1}(2^{S_Y}), \sigma(X_1, X_2) = (X_1, X_2)^{-1}(2^{S_1} \times 2^{S_2})]$
即 $\forall A \subseteq S_Y, B \subseteq 2^{S_1} \times 2^{S_2}, B = B_1 \times B_2$, 有

$$\mathbb{P}(Y \in A, (X_1, X_2) \in B) = \mathbb{P}(Y \in A)\mathbb{P}((X_1, X_2) \in B)$$

其中 $\mathbb{P}((X_1, X_2) \in B) = \mathbb{P}(X_1 \in B_1, X_2 \in B_2)$

Problem 5 (作业 2-3)

证明:

$$\begin{aligned} Y \perp\!\!\!\perp (X_1, X_2) \Leftrightarrow & \forall y \in S_Y, x_1 \in S_1, x_2 \in S_2 \\ & 有\mathbb{P}(Y = y, (X_1, X_2) = (x_1, x_2)) \\ & = \mathbb{P}(Y = y)\mathbb{P}((X_1, X_2) = (x_1, x_2)) \end{aligned}$$

有了上述定义, 可以推广:

1. $(Y_1, \cdots, Y_n) \perp\!\!\!\perp (X_1, \cdots, X_n)$
2. $Y \perp\!\!\!\perp_A (X_1, \cdots, X_n)(A \in \mathcal{F}, \mathbb{P}(A) > 0)$

Property 1.14

$Y \perp\!\!\!\perp (X_1, X_2) \Rightarrow Y \perp\!\!\!\perp X_1, Y \perp\!\!\!\perp X_2$

证明：在定义1.22中取 $B_2 = \Omega$

$$\begin{aligned}
\mathbb{P}(Y \in A, X_1 \in B_1) &= \mathbb{P}(Y \in A, X_1 \in B_1, X_2 \in S_2) \\
&= \mathbb{P}(Y \in A)\mathbb{P}(X_1 \in B_1, X_2 \in S_2) \qquad [Y \perp\!\!\!\perp (X_1, X_2)] \\
&= \mathbb{P}(Y \in A)\mathbb{P}(X_1 \in B_1)
\end{aligned}$$

注：看到 $\Rightarrow$ 要自然地问, 反过来 $\Leftarrow$ 成立吗？做数学要多问自己一些问题, 即便没有答案

Corollary 1.4

$(Y_1, \cdots, Y_n) \perp\!\!\!\perp (X_1, \cdots, X_n) \Rightarrow Y_k \perp\!\!\!\perp X_j, 1 \leqslant k \leqslant m, 1 \leqslant j \leqslant n$

1.5 随机过程

1.5.1 什么是随机过程

Definition 1.23 (随机过程)
设 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间, $(S, \mathcal{S})$ 为可测空间, $\mathbb{T}$ 为指标集/参数集, 称随机变量族

$$\{X_t : (\Omega, \mathcal{F}, \mathbb{P}) \to (S, \mathcal{S}) | t \in \mathbb{T}\}$$

为 (S 值) 随机过程 X. 其中 $(S, \mathcal{S})$ 称为 X 的状态空间
注:

1. $forall t \in \mathbb{T}$, X_t 为随机变量
2. $\mathbb{T}$ 为时间集, X_t 为过程 X 在时刻 t 的状态

$\mathbb{T} \backslash S \subseteq \mathbb{R}$	离散$(e.g.\ \mathbb{N})$　连续$(e.g.\ \mathbb{R}, \mathbb{R}^+)$
可数集$(e.g.\ \mathbb{N}, \mathbb{Z})$	离散时间/参数的随机过程
连续统$(e.g.\ [0, T], \mathbb{R}^+)$	连续时间/参数的随机过程

1.5.2 随机过程的分布

1. $\forall t \in \mathbb{T}, X_t : \Omega \to S$ 为随机变量/可测映射
2. $X : \mathbb{T} \times \Omega \to S$ 二元映射
3. $X : \Omega \to S^{\mathbb{T}}$ 其中 $S^{\mathbb{T}} = \{f | f : \mathbb{T} \to \S\}$, $X : \omega \to X(\omega) = X(\cdot, \omega)$

分布可用有限维分布族刻画

Definition 1.24
固定样本点 ω, 则 $X.(\omega)$ 为 $\mathbb{T} \to S$ 的映射, 即 $X.(\omega) \in S^{\mathbb{T}}$, 称 $X.(\omega)$ 是过程 X 的一个实现/样本路径/样本函数

Definition 1.25
$\forall n \geqslant 1, t_1, t_2, \cdots, t_n$ 称

$$(x_1, x_2, \cdots, x_n) \mapsto F_{t_1, t_2, \cdots, t_n}(x_1, x_2, \cdots, x_n) = \mathbb{P}(X_{t_1} \leqslant x_1, \cdots, X_{t_n} \leqslant x_n)$$

为 X 的 n 维分布

Definition 1.26 (过程的有限维分布族)
定义

$$\{F_{t_1, t_2, \cdots, t_n} | n \geqslant 1, t_1, \cdots, t_n \in \mathbb{T}\}$$

1.5.3 随机过程的存在性

1. (抽象的) 从概率论/测度论出发去证明随机过程存在性, 不写出具体形式, 满足随机过程符合给定的有限维分布族即可
2. (具体的) 构造性证明

Property 1.15

随机过程的有限维分布族具有以下两个性质

1. (对称性) 重排, 设 $\sigma:\{1,\cdots,n\}\to\{1,\cdots,n\}$ 为双射, 则

$$F_{t_{\sigma(1)},\cdots,t_{\sigma(n)}}(x_{\sigma(1)},\cdots,x_{\sigma(n)})=F_{t_1,\cdots,t_n}(x_1,\cdots,x_n)$$

2. (相容性) $m\geqslant n$

$$F_{t_1,\cdots,t_n,t_{n+1},\cdots,t_m}(x_1,\cdots,x_n,+\infty,\cdots,+\infty)=F_{t_1,\cdots,t_n}(x_1,\cdots,x_n)$$

   注：相容性类比从高维向低维的投影, $\mathbb{P}(X\leqslant+\infty)=F_X(+\infty)=1$

这两个性质是随机过程存在的必要条件

Theorem 1.14 (Kolmogorov 定理)

设分布函数族

$$\{F_{t_1,\cdots,t_n}|t_1,\cdots,t_n\in\mathbb{T},n\geqslant1\}$$

满足对称性, 相容性, 则必存在一个随机过程 $\{X_t,t\in\mathbb{T}\}$ 使得上述分布函数族 F 是 X 的有限维分布族

1.5.4 随机过程的基本类型

1. 离散时间马氏链（由条件概率定义）
2. Poisson 过程
3. 更新过程
4. 连续时间马氏链
5. 离散时间 Martingale （由条件期望定义）
6. 布朗运动

Definition 1.27

对连续时间的随机过程 $\{X_t,t\in\mathbb{T}\}$

1. 若对一切的 $t_0<t_1<\cdots<t_n$ 有 $X_{t_1}-X_{t_0},\cdots,X_{t_n}-X_{t_{n-1}}$ 相互独立, 则过程 X 是独立增量过程（e.g. 布朗运动）
2. 若对每一个 $S\in\mathbb{T},X_{t+s}-X_t$ 对一切的 t 都有相同分布, 称 X 为平稳增量过程

2 马氏链

2.1 离散时间马氏链

马尔可夫性 $\leftrightarrow$ 已知现在, 过去与未来不相干/独立

Definition 2.1 ((离散时间) 马氏链)
称 S 值随机过程 $\{X_n, n \geqslant 0\}$ 为马氏链, 若 X 满足以下马氏性: $\forall n \geqslant 0, x_0, x_1, \cdots, x_n, y \in S$,

$$\mathbb{P}(\underbrace{X_{n+1} = y}_{\text{未来}} \,|\, \underbrace{X_0 = x_0, \cdots, X_{n-1} = x_{n-1}}_{\text{过去}}, \underbrace{X_n = x_n}_{\text{现在}}) = \mathbb{P}(X_{n+1} = y | X_n = x_n) \tag{M_1}$$

其中 X_0 的分布称为 X 的初始分布

Definition 2.2
当 S 为有限集, 称链为有限链, 当 S 为无限集, 称链为无限链

注: 改写 (M_1)

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

$$RHS = \mathbb{P}_{X_n = x_n}(X_{n+1} = y)$$

$$\begin{aligned} M_1 &\Leftrightarrow \{X_{n+1} = y\} \perp\!\!\!\perp_{\{X_n = x_n\}} \{X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\} \\ &\Leftrightarrow X_{n+1} \perp\!\!\!\perp_{\{X_n = x_n\}} (X_0, \cdots, X_{n-1}) \end{aligned}$$

(M_1) 未来 $\perp\!\!\!\perp_{\text{现在}}$ 过去

$$\mathbb{P}_{\text{现在}}(\text{未来}|\text{过去}) = \mathbb{P}_{\text{现在}}(\text{未来})$$

Lemma 2.1 (马氏性的等价表示)
[Grimmett [3]] 下面三个命题等价

1. (M_1) 马氏性
2. (M_2) $\forall k \geqslant 0, 0 \leqslant n_1 < \cdots < n_k \leqslant n$, 对于 $y, x_{n_1}, \cdots, x_{n_k} \in S$,

$$\mathbb{P}(X_{n+1} = y | X_{n_1} = x_{n_1}, \cdots, X_{n_k} = x_{n_k}) = \mathbb{P}(X_{n+1} = y | X_{n_k} = x_{n_k}) \tag{2.1}$$

   即

$$\{X_{n+1} = y\} \perp\!\!\!\perp_{\{X_{n_k} = x_{n_k}\}} \{X_{n_1} = x_{n_1}, \cdots, X_{n_{k-1}} = x_{n_{k-1}}\}$$

3. (M_3) 对 $\forall m \geqslant 1, n \geqslant 0, \{y, x_i, 0 \leqslant i \leqslant n\} \subseteq S$, 有

$$\mathbb{P}(X_{n+m} = y | X_0 = x_0, \cdots, X_n = x_n) = \mathbb{P}(X_{n+m} = y | X_n = x_n) \tag{2.2}$$

   即

$$\{X_{n+m} = y\} \perp\!\!\!\perp_{\{X_n = x_n\}} \{X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\}$$

证明：思路 $1 \leftrightarrows 3 \leftrightarrows 2$

$(2) \to (3)$, 先处理一些记号的问题. 记 (2) 中的 n 为 $n^{(2)}$, (3) 中的 n 为 $n^{(3)}$. 则取 $n_k = n^{(3)} = n^{(2)} + 1 - m \leqslant n^{(2)}$, 所以 $n^{(3)} + m = (n^{(2)} + 1 - m) + m = n^{(2)} + 1$, 即已知 (2) 可推 (3)

$(3) \to (1)$, 取 $m = 1$, 显然

只需证 $(3) \to (2), (1) \to (3)$

这里回顾独立的三种写法

1. $A \perp\!\!\!\perp_B C$ 记号
2. $\mathbb{P}_B(A, C) = \mathbb{P}_B(A)\mathbb{P}_B(C)$ 定义
3. $\mathbb{P}_B(A|C) = \mathbb{P}_B(A)$ 定理

(Step 1) 证明 $(3) \to (2)$

思路：(2)(3) 条件不同, 想要由 (3) 推 (2), 则切换到 (2) 的条件概率测度, 展开, 再用 (3) 的条件瘦身

对 $\forall k \geqslant 2, 0 \leqslant n_1 < n_2 < \cdots < n_k = n$

令 $J = \{0, 1, \cdots, n_k - 1, n_k\} \backslash \{n_1, \cdots, n_k\}$, $\tilde{\mathbb{P}}(\cdot) = \mathbb{P}(\cdot | X_{n_1} = x_{n_1}, \cdots, X_{n_k} = x_{n_k})$

$$\begin{aligned}
\tilde{\mathbb{P}}(X_{n+1} = y) &= \sum_{x_j \in S, j \in J} \tilde{\mathbb{P}}(X_{n+1} = y | X_j = x_j, j \in J) \cdot \tilde{\mathbb{P}}(X_j = x_j, j \in J) \qquad [\text{全概公式}] \\
&= \mathbb{P}(X_{n+1} = y | X_{n_k} = x_{n_k}) \sum_{x_j \in S, j \in J} \tilde{\mathbb{P}}(X_j = x_j, j \in J) \qquad [(3), \mathbb{P}_C(\cdot|A) = \mathbb{P}_C(\cdot)] \\
&= \mathbb{P}(X_{n+1} = y | X_{n_k} = x_{n_k})
\end{aligned}$$

其中, 记号 $\sum_{x_j \in S, j \in J}$ 中的下标意为：假设 J 中元素个数为 $\#J = u$, 则 $(x^{(1)}, \cdots, x^{(u)}) \in S^u$. 从简单的开始, $\sum_{x \in S} \mathbb{P}(X = x) = \mathbb{P}(\Omega), \sum_{(x,y) \in S^2} \mathbb{P}(X = x, Y = y) = \mathbb{P}(\Omega), \cdots, \sum_{(x^{(1)}, \cdots, x^{(u)}) \in S^u} \mathbb{P}(X^{(1)} = x^{(1)}, \cdots, X^{(u)} = x^{(u)}) = \mathbb{P}(\Omega) = 1$

(Step 2) 下证 $(1) \to (3)$

1. $m = 1$ 时, 即 (1)

2. 假设 $m = k$ 时 (3) 成立, 即 $\forall n \geqslant 1, \{y, x_i, n \geqslant i \geqslant 0\} \subseteq S$,

$$\{X_{n+k} = y\} \perp\!\!\!\perp_{\{X_n = x_n\}} \{X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\} \xrightarrow{\text{性质}(1.14)} \{X_{n+k} = y\} \perp\!\!\!\perp_{\{X_n = x_n\}} \{X_{n-1} = x_{n-1}\}$$

$$\begin{aligned}
\mathbb{P}(X_{n+k} = y | X_0 = x_0, \cdots, X_n = x_n) &= \mathbb{P}(X_{n+k} = y | X_n = x_n) \\
&= \mathbb{P}(X_{n+k} = y | X_n = x_n, X_{n-1} = x_{n-1})
\end{aligned} \tag{*}$$

当 $m = k + 1$ 时, 对 $\forall \{y, x_i, n \geqslant i \geqslant 0\} \subseteq S$

令 $\tilde{\mathbb{P}}_n(\cdot) := \mathbb{P}(\cdot | X_0 = x_0, \cdots, X_n = x_n)$

$$\begin{aligned}
\tilde{\mathbb{P}}_n(X_{n+k+1} = y) &\overset{\text{Thm}(1.5)}{=} \sum_{x_{n+1} \in S} \tilde{\mathbb{P}}_n(X_{n+k+1} = y | X_{n+1} = x_{n+1}) \cdot \tilde{\mathbb{P}}_n(X_{n+1} = x_{n+1}) \\
&= \sum_{x_{n+1} \in S} \mathbb{P}(X_{n+k+1} = y | X_{n+1} = x_{n+1}, X_n = x_n) \cdot \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \quad [(*), \text{归纳法假设}] \\
&\overset{(1.2)}{=} \sum_{x_{n+1} \in S} \mathbb{P}(X_{n+k+1} = y, X_{n+1} = x_{n+1}, X_n = x_n) / \mathbb{P}(X_n = x_n) \\
&= \mathbb{P}(X_{n+k+1} = y, X_n = x_n) / \mathbb{P}(X_n = x_n) \\
&= \mathbb{P}(X_{n+k+1} = y | X_n = x_n)
\end{aligned}$$

即 $m=k+1$ 得证 $\square$

证明 (Step 2) 时如果在 x_{n+k} 处展开而不是在 x_{n+1}, 也是可以的. 实际上在 $x_{n+j}, \forall j, 1 \leqslant j \leqslant k$ 展开都可以, 关键在于用性质1.14和全概公式1.5凑出乘法公式(1.2), 消元即可.

Remark 4. 三种写法的直觉

1. M_1: 未来"下一步"跟过去"每一步"都无关
2. M_2: 未来"下一步"跟过去的"任意若干步"都无关
3. M_3: 未来"下 m 步"跟过去"每一步"都无关

可以推出, 由 (2)(3), 下式也成立:

对 $\forall m \geqslant 1, n \geqslant 0, \{y, x_i, 0 \leqslant i \leqslant n\} \subseteq S$

$$\mathbb{P}(X_{n+m}=y|X_{n_1}=x_{n_1}, \cdots, X_{n_k}=x_{n_k}) = \mathbb{P}(X_{n+m}=y|X_{n_k}=x_{n_k})$$

Corollary 2.1

若 X 是马氏链, 则 $\forall n \geqslant 1, \{x_i, n \geqslant i \geqslant 0, y\} \subseteq S$, 有

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

补充记号:

- 乘积空间
$$S^n := \underbrace{S \times \cdots \times S}_{\text{n 个}}$$
- 乘积 σ 代数
$$\bigotimes_n 2^S := \underbrace{2^S \times \cdots \times 2^S}_{\text{n 个}}$$

Property 2.1 (马氏性的等价条件)

下列三个命题等价

1. 马氏性 (M_1)
2. 对 $\forall n \geqslant 1, m \geqslant 1, A \in \otimes_n 2^S, B \in \otimes_m 2^S$, 即 $(A \subseteq S^n, B \subseteq S^m)$, 有
$$\begin{aligned}&\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_{n-1}) \in A, (X_{n+1}, \cdots, X_{n+m}) \in B) \\ =&\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_{n-1}) \in A) \cdot \mathbb{P}_{\{X_n=x_n\}}((X_{n+1}, \cdots, X_{n+m}) \in B)\end{aligned} \tag{2.3}$$
即 $(X_0, \cdots, X_{n-1}) \perp\!\!\!\perp_{\{X_n=x_n\}} (X_{n+1}, \cdots, X_{n+m})$ 的定义
3. $\mathbb{P}_{\{X_n=x_n\}}((X_{n+1}, \cdots, X_{n+m}) \in B|(X_0, \cdots, X_{n-1}) \in A) = \mathbb{P}_{\{X_n=x_n\}}((X_{n+1}, \cdots, X_{n+m}) \in B)$

证明: (2) $\Leftrightarrow$ (3), 独立的定义和定理, 显然

(3) $\rightarrow$ (1), 取 $k=0$ 显然

只需证 (1) $\rightarrow$ (3)

只需证 (3) 对简单事件 A, B (单点集合) 成立, 即 $\forall n \geqslant 1, m \geqslant 1, \{x_0, x_1, \cdots, x_{n+m} \subseteq S\}$, 有

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

其中 $x_{n+1}^{n+m} = (x_{n+1}, \cdots, x_{n+m}), x_0^{n-1} = (x_0, \cdots, x_{n-1})$
* 只要对单点集合成立, 对一般情况也成立, 证明见 Theorem 1.1
令

$$\tilde{\mathbb{P}}_n(\cdot) := \mathbb{P}_{\{X_n=x_n\}}(\cdot | (X_0, \cdots, X_{n-1}) = x_0^{n-1}) = \mathbb{P}(\cdot | (X_0, \cdots, X_n) = x_0^n)$$

只证 $m = 2$, 即由

$$\tilde{\mathbb{P}}_n(X_{n+1} = x_{n+1}) = \mathbb{P}_{\{X_n=x_n\}}(X_{n+1} = x_{n+1})$$

证得

$$\tilde{\mathbb{P}}_n(X_{n+1} = x_{n+1}, X_{n+2} = x_{n+2}) = \mathbb{P}_{\{X_n=x_n\}}(X_{n+1} = x_{n+1}, X_{n+2} = x_{n+2})$$

$$\begin{aligned}
\tilde{\mathbb{P}}_n((X_{n+1}, X_{n+2}) = (x_{n+1}, x_{n+2})) &= \tilde{\mathbb{P}}_n(X_{n+1} = x_{n+1}) \cdot \tilde{\mathbb{P}}_n(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}) \\
&\overset{(M_1)}{=} \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \cdot \mathbb{P}(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}) \\
&\overset{\text{Cor}(2.1)}{=} \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \cdot \mathbb{P}(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}, X_n = x_n) \\
&= \mathbb{P}_{\{X_n=x_n\}}(X_{n+1} = x_{n+1}) \cdot \mathbb{P}_{\{X_n=x_n\}}(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}) \\
&\overset{(1.2)}{=} \mathbb{P}_{\{X_n=x_n\}}((X_{n+1}, X_{n+2}) = (x_{n+1}, x_{n+2}))
\end{aligned}$$

Corollary 2.2
设 X 为马氏链, 则对每一个 $n \geqslant 1, m \geqslant 1, u_k < u_{k+1}, 0 \leqslant k \leqslant n + m - 1$, 有

$$(X_{u_0}, \cdots, X_{u_{n-1}}) \perp\!\!\!\perp_{\{X_{u_n}=x_{u_n}\}} (X_{u_{n+1}}, \cdots, X_{u_{n+m}})$$

2.2 时齐马氏链与转移概率

Definition 2.3 (时间齐次马氏链)
称马氏链 $X:\{X_n, n\geqslant 0\}$ 为时齐的或时间齐次马氏链，若对 $\forall n\geqslant 0, i,j\in S$

$$\mathbb{P}(X_{n+1}=j|X_n=i)=\mathbb{P}(X_1=j|X_0=i)$$

Definition 2.4
X 是时齐马氏链，称

$$p_{ij}:=p_{i,j}=\mathbb{P}(X_1=j|X_0=i) \qquad i,j\in S$$

为 X 从状态 i 到 j 的（一步）转移概率，并称矩阵

$$P=\begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

为（一步）转移（概率）矩阵

若不加说明，则默认讨论的马氏链都是时齐的

注：

$$\begin{aligned}\mathbb{P}(x_{n+1}=y)&=\sum_{x\in S}\mathbb{P}(X_{n+1}=y|X_n=x)\cdot\mathbb{P}(X_n=x)\\&=\sum_{x\in S}p_{xy}\cdot\mathbb{P}(X_n=x)\end{aligned}$$

Theorem 2.1 (转移矩阵的刻画)
转移矩阵是一个随机矩阵，即

1. $\forall i,j\in S, p_{ij}\geqslant 0$
2. $\forall i\in S, \sum_{j\in S}p_{ij}=1$

即转移矩阵的每一行 $(p_{ij})_{j\in S}$ 为 S 上的一个概率分布

注：另一种随机矩阵是指元素为随机变量的矩阵，和这里讲的没有关系

证明：

$$\sum_{j\in S}\mathbb{P}(X_1=j|X_0=i)=\mathbb{P}(X_1\in S|X_0=i)=\mathbb{P}(\Omega|X_0=i)=1$$

Definition 2.5 (时齐马氏链)
设 $X = \{X_n, n \geqslant 0\}$ 为一随机过程, 若

1. 初值 X_0 满足分布 $\mu = (\mu_i)_{i\in S}$, 即 $\mathbb{P}(X_0 = i) = \mu_i, i \in S$
2. 存在一个随机矩阵 $P = (p_{ij})_{i,j\in S}$ 使得 $\forall n \geqslant 1, i_0, \cdots, i_{n-1}, i, j \in S$

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \cdots, X_{n-1} = i_{n-1}, X_n = i) = p_{ij}$$

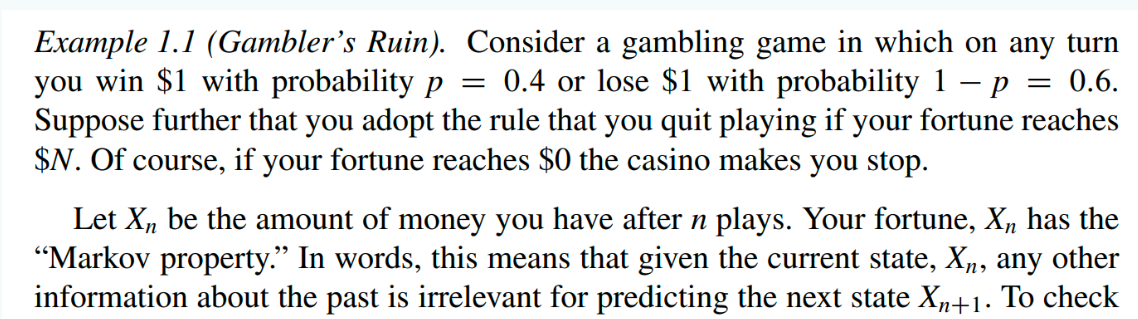
则称 X 具有初始分布 μ 和转移矩阵 P 的（时齐）马氏链, 记作 $X \sim \mathrm{Markov}(\mu, P)$

上述定义与 (M_1) 马氏链定义2.1等价
证明：$(2) \to (M_1)$

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j | X_n = i) &= \sum_{(i_0,\cdots,i_{n-1})\in S^n} \mathbb{P}(X_{n+1} = j | X_n = i, X_0 = i_0, \cdots, X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) \\
&= \sum_{(i_0,\cdots,i_{n-1})\in S^n} p_{ij} \cdot \mathbb{P}(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) = p_{ij}
\end{aligned}$$

所以 $\mathbb{P}(X_{n+1} = j | X_0 = i_0, \cdots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i)$
即然有 (M_1), 为什么还要定义2.5？因为该定义决定了马氏链的有限维分布

Example 2.1 (Gambler's Ruin)
[Durrett [2],P1]



Example 1.1 (Gambler's Ruin). Consider a gambling game in which on any turn you win \$1 with probability $p = 0.4$ or lose \$1 with probability $1 - p = 0.6$. Suppose further that you adopt the rule that you quit playing if your fortune reaches \$$N$. Of course, if your fortune reaches \$0 the casino makes you stop.

Let X_n be the amount of money you have after n plays. Your fortune, X_n has the "Markov property." In words, this means that given the current state, X_n, any other information about the past is irrelevant for predicting the next state X_{n+1}. To check

图 2: Gambler's Ruin

Claim 2.1. $\{X_n, n \geqslant 0\}$ 为（时齐）马氏链

1. 对于 $0 < i_0, \cdots, i_{n-1} < N, n \geqslant 0$ 有

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = i + 1 | X_n = i, X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) \\
=&\mathbb{P}(X_{n+1} = i + 1 | X_n = i) = 0.4 = \mathbb{P}(\text{第}n+1\text{次赌局赢一元})
\end{aligned}$$

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = i - 1 | X_n = i, X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) \\
=&\mathbb{P}(X_{n+1} = i - 1 | X_n = i) = 0.6 = \mathbb{P}(\text{第}n+1\text{次赌局输一元})
\end{aligned}$$

2. $\mathbb{P}(X_{n+1} = 0 | X_n = 0, X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) = 1 = \mathbb{P}(X_{n+1} = 0 | X_n = 0)$
$\mathbb{P}(X_{n+1} = N | X_n = N, X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) = 1 = \mathbb{P}(X_{n+1} = N | X_n = N)$

最后一个等号是由题目设定得到，从 $0 \to 0$ 或 $N \to N$ 的概率都为 1，因为游戏结束
综上，$p(i, i+1) = 0.4, 0 < i < N, p(i, i-1) = 0.6, 0 < i < N, p(0,0) = p(N,N) = 1$
e.g.

When $N = 5$ the matrix is

	0	**1**	**2**	**3**	**4**	**5**
0	1.0	0	0	0	0	0
1	0.6	0	0.4	0	0	0
2	0	0.6	0	0.4	0	0
3	0	0	0.6	0	0.4	0
4	0	0	0	0.6	0	0.4
5	0	0	0	0	0	1.0

图 3: N=5

Example 2.2 (Two-Stage Markov Chains)
[Durrett [2], P7]

Example 1.10 (Two-Stage Markov Chains). In a Markov chain the distribution of X_{n+1} only depends on X_n. This can easily be generalized to case in which the distribution of X_{n+1} only depends on (X_n, X_{n-1}). For a concrete example consider a basketball player who makes a shot with the following probabilities:

1/2 if he has missed the last two times
2/3 if he has hit one of his last two shots
3/4 if he has hit both of his last two shots

图 4: Two-Stage Markov Chains

1. $\mathbb{P}(X_{n+1} = H | X_n = M, X_{n-1} = M) = 1/2$
2. $\mathbb{P}(X_{n+1} = H | X_n = M, X_{n-1} = H) = \mathbb{P}(X_{n+1} = H | X_n = H, X_{n-1} = M) = 2/3$
3. $\mathbb{P}(X_{n+1} = H | X_n = H, X_{n-1} = H) = 3/4$

Claim 2.2. $Y_n = (X_n, X_{n-1}), n \geqslant 1$ 则 $\{Y_n, n \geqslant 1\}$ 是（时齐）马氏链，$Y_n : \Omega \to \{HH, HM, MH, MM\}$

证明：

$$
\begin{aligned}
&\mathbb{P}(Y_{n+1} = HH | Y_n = HH, Y_j = (x_j, x_{j-1}), 1 \leqslant j \leqslant n-1) \\
=&\mathbb{P}(X_{n+1} = H, X_n = H | X_n = H, X_{n-1} = H, X_j = x_j, X_{j-1} = x_{j-1}, 0 \leqslant j \leqslant n-1) \\
=&\mathbb{P}(X_{n+1} = H | X_n = H, X_{n-1} = H) \\
=&3/4 \qquad [3.]
\end{aligned}
$$

对 1.2. 同理 □

Proposition 2.1 (初见马氏链的有限维分布)

设 $P=(p_{ij})_{i,j\in S}$ 为随机矩阵, $\mu=(\mu_i)_{i\in S}$ 为概率分布, $X=\{X_n,n\geqslant 0\}$ 为 S 值离散时间随机过程. 则过程 $X\sim \text{Markov}(\mu,P)$ 当且仅当对任意的 $n\geqslant 0, i_0,i_1,\cdots,i_n\in S$, X 有有限维分布:

$$\mathbb{P}(X_0=i_0,X_1=i_1,\cdots,X_n=i_n)=\mu_{i_0}\prod_{k=0}^{n-1}p_{i_k i_{k+1}} \tag{2.4}$$

证明: $\Rightarrow$

$$\begin{aligned}
&\mathbb{P}(X_0=i_0,X_1=i_1,\cdots,X_n=i_n)\\
=&\mathbb{P}(X_0=i_0)\mathbb{P}(X_1=i_1|X_0=i_0)\cdots\mathbb{P}(X_n=i_n|X_0=i_0,\cdots X_{n-1}=i_{n-1}) \quad [\text{乘法公式}(1.2)]\\
=&\mathbb{P}(X_0=i_0)\mathbb{P}(X_1=i_1|X_0=i_0)\cdots\mathbb{P}(X_n=i_n|X_{n-1}=i_{n-1}) \quad [\text{Markov}]\\
=&\mu_{i_0}p_{i_0,i_1}\cdots p_{i_{n-1},i_n}
\end{aligned}$$

严格地讲, $\mathbb{P}(\cdot|A)$ 需保证 $\mathbb{P}(A)>0$. 对 $\mathbb{P}(A)=0$ 情况的分类讨论, 见 Resnick [4], prop 2.1.1

$\Leftarrow$

1. $n=0, \mathbb{P}(X_0=i_0)=\mu_{i_0}\Rightarrow X_0\sim(\mu_i)_{i\in S}$
2. $n\geqslant 1$

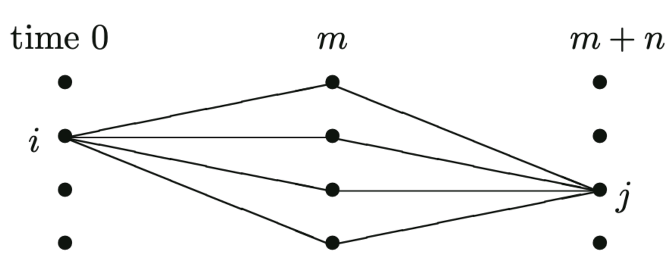
$$\mathbb{P}(X_{n+1}=i_{n+1}|X_0=i_0,\cdots,X_n=x_n)=\frac{\mathbb{P}(X_0=i_0,\cdots,X_{n+1}=i_{n+1})}{\mathbb{P}(X_0=i_0,\cdots,X_n=i_n)}=p_{i_n,i_{n+1}}$$

由时齐马氏链定义, 初始分布和转移矩阵都符合定义2.5

$$\therefore \quad X\sim \text{Markov}(\mu,P) \quad \square$$

对于 $\mathbb{P}(X_0=i_0,X_1=i_1,\cdots,X_n=i_n)$, 如果我们想把 X_1 挖掉, 即

$$\begin{aligned}
\mathbb{P}(X_0=i_0,X_2=i_2,\cdots,X_n=i_n)&=\sum_{i_1\in S}\mathbb{P}(X_0=i_0,X_1=i_1,\cdots,X_n=i_n)\\
&=\mu_{i_0}\sum_{i_1\in S}(P_{i_0,i_1}P_{i_1,i_2})\cdots P_{i_{n-1},i_n}
\end{aligned}$$



2.3 多步转移概率与矩阵乘法

Definition 2.6
设 $X=\{X_n,n\geqslant 0\}$ 为马氏链, 称

$$p_{ij}(m,m+n):=\mathbb{P}(X_{n+m}=j|X_m=i)\quad (i,j\in S,m,n\geqslant 0)$$

为 X 的 n 步转移概率, 并称 $P(m,m+n)=(p_{ij}(m,m+n))_{i,j\in S}$ 为 X 的 n 步转移（概率）矩阵, 其中

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

当 X 时齐, $P(m,m+1)=(p_{ij}(m,m+1))_{i,j\in S}=(p_{ij}(0,1))_{i,j\in S}=(p_{ij})_{i,j\in S}$
可见 $n=1$ 时, $P(m,m+1)$ 与 m 无关. 那 $n>1$ 时呢?

2.3.1 Chapman-Kolmogorov 方程

Theorem 2.2 (C-K 方程)
设 $\{X_n,x\geqslant 0\}$ 为马氏链

$$p_{ij}(m,m+n+r)=\sum_{k\in S}p_{ik}(m,m+n)p_{kj}(m+n,m+n+r)\tag{2.5}$$

其中 $i,j\in S,m,n,r\geqslant 0$, 即

$$P(m,m+n+r)=P(m,m+n)P(m+n,m+n+r)$$

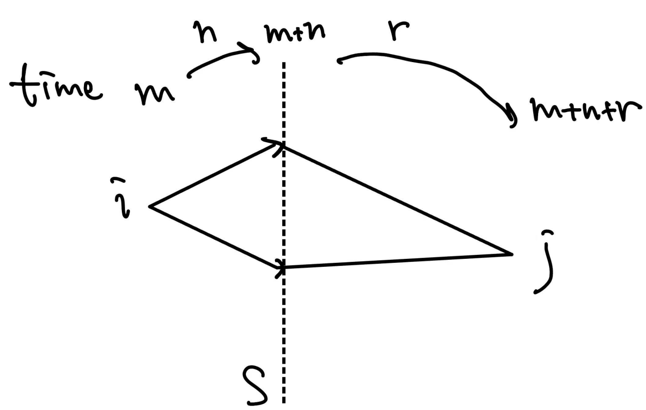


图 5: Multi-steps

证明:

$$\begin{aligned}
p_{ij}(m,m+n+r) &= P(X_{m+n+r}=j|X_m=i) \\
&= \sum_{k\in S} P(X_{m+n+r}=j, X_{m+n}=k|X_m=i) \\
&= \sum_{k\in S} \mathbb{P}_{\{X_m=i\}}(X_{m+n+r}=j|X_{m+n}=k)\mathbb{P}_{\{X_m=i\}}(X_{m+n}=k) \quad [\text{乘法公式}(1.2)] \\
&= \sum_{k\in S} p_{ik}(m,m+n)p_{kj}(m+n,m+n+r) \quad [\text{Markov}]
\end{aligned}$$

Corollary 2.3

设 X 为具有（一步）转移矩阵 P 的时齐马氏链, 则

1. $\forall m,n \geqslant 0$, 有 $P(m,m+n)=P(0,n)=P^n$. 其中, 约定 $P^0=I$（单位矩阵）
   从而, 可记 X 的 n 步转移概率为 $p_{ij}(n)$ 或 $p_{ij}^{(n)}$, n 步转移概率矩阵为 $P(n)$, 且有
   $$P(n)=P^n=(p_{ij}^{(n)})_{i,j\in S}$$
2. C-K 方程可改写为
   $$p_{ij}(m+n)=\sum_{k\in S} p_{ik}^{(m)}p_{kj}^{(n)}$$
   $P(m+n)=P(m)P(n)$, 即 $P^{m+n}=P^mP^n$

证明:

$$\begin{aligned}
P(m,m+n) &= P(m,m+1)\cdot P(m+1,m+n) \quad [\text{C-K}] \\
&= P\cdot P(m+1,m+n) \quad [\text{时齐}] \\
&= P^n \quad \square
\end{aligned}$$

Proposition 2.2

$\forall n\geqslant 0, P(n)=P^n$ 仍是一个随机矩阵 (Theorem 2.1)

证明: $n=2$ 时, $P^2=(p_{ij}(2))_{i,j\in S}$

$\Rightarrow$

$$\begin{aligned}
\sum_{j\in S} p_{ij}(2) &= \sum_{j\in S}\sum_{k\in S} p_{ik}p_{kj} \quad [\text{C-K, 默认}p_{ik}(1)=p_{ik}] \\
&= \sum_{k\in S}\sum_{j\in S} p_{ik}p_{kj} \\
&= \sum_{k\in S} p_{ik}\cdot(\sum_{j\in S} p_{kj}) \\
&= \sum_{k\in S} p_{ik}=1 \quad \square
\end{aligned}$$

第二个等号, 级数可交换是因为非负, 要么有限（收敛）、要么 $+\infty$（发散）

2.3.2 马氏链的任意有限维分布

Proposition 2.3

$X \sim \text{Markov}(\mu, P)$, 其中 $\mu = (\mu_i)_{i\in S}, P = (p_{ij})_{i,j\in S}$, 则

$$\mathbb{P}(X_{u_1} = i_1, \cdots, X_{u_n} = i_n) = \mu_{i_1}^{(u_1)} \prod_{k=1}^{n-1} p_{i_k, i_{k+1}}^{(u_{k+1}-u_k)}$$

其中, $0 < u_1 < u_2 < \cdots < u_n$, $i_1, i_2, \cdots, i_n \in S$, $\mu^{(u_1)} = (\mu_i^{(u_1)})_{i\in S}$ 为 X_{u_1} 的有限维分布

证明:

$$\begin{aligned}\mathbb{P}(X_{u_1} = i_1, \cdots, X_{u_n} = i_n) &= \mathbb{P}(X_{u_1} = i_1) \cdot \mathbb{P}(X_{u_2} = i_2 | X_{u_1} = i_1) \cdots \mathbb{P}(X_{u_n} = i_n | X_{u_1} = i_1, \cdots, X_{u_{n-1}} = i_{n-1}) \\ &= (\mu_{i_1}^{(u_1)}) p_{i_1,i_2}^{(u_2-u_1)} \cdots p_{i_{n-1},i_n}^{(u_n-u_{n-1})} \quad [\text{Markov}] \\ &= \mu_{i_1}^{(u_1)} \prod_{k=1}^{n-1} p_{i_k,i_{k+1}}^{(u_{k+1}-u_k)}\end{aligned}$$

用概率表示不够直观, 尝试用转移矩阵来表示

Lemma 2.2

$\mu^{(m+n)} = \mu^{(n)} P^m (\forall m, n \geqslant 0)$, 即

$$\mu_j^{(m+n)} = (\mu^{(n)} P^m)_j = \sum_{i\in S} \mu_i^{(n)} p_{ij}^{(m)}$$

特别地, 取 $n = 0$, 则 $\mu^{(m)} = \mu \cdot P^m$ (μ 看成行向量), 即 $\mu_j^{(m)} = (\mu P^m)_j = \sum_{i\in S} \mu_i \cdot p_{ij}^{(m)}$

证明:

$$\begin{aligned}\mu_j^{(n+m)} = \mathbb{P}(X_{n+m} = j) &= \sum_{i\in S} \mathbb{P}(X_{n+m} = j | X_n = i) \mathbb{P}(X_n = i) \\ &= \sum_{i\in S} p_{ij}(m) \mu_i^{(n)} \\ &= (\mu^{(n)} P^m)_j \quad \square\end{aligned}$$

$\Rightarrow \mu^{(m+n)} = \mu^{(n)} P^m$

Theorem 2.3 (任意有限维分布 II)

$\forall 0 \leqslant u_1 < u_2 < \cdots < u_n, i_1, \cdots, i_n \in S$

$$\mathbb{P}(X_{u_1} = i_1, \cdots, X_{u_n} = i_n) = (\mu P_{i_1}^{u_1}) \prod_{k=1}^{n-1} P_{i_k,i_{k+1}}^{u_{k+1}-u_k}$$

其中, $P_{i,j}^m =: (P^m)_{i,j} =: p_{i,j}^{(m)}$

讨论随机过程地存在性:

抽象地, $\mu, P \xrightarrow{\text{Thm (1.14)}}$ 有限维分布族 $\to X \sim \text{Markov}(\mu, P)$, μ, P 可以刻画具备对称性、相容性的有限维分布

具体地, 参考 Resnick [4], P62, Section 2.1

2.4 （从固定点出发的）马氏链

固定 $i \in S$, 定义 $\mathbb{P}_i(\cdot) = \mathbb{P}(\cdot|X_0 = i)$, $\mathbb{E}_i(X) = \mathbb{E}(X|X_0 = i) = \sum_{x \in S} x\mathbb{P}_i(X = x)$

2.4.1 链的状态：常返和暂留

Definition 2.7

称状态 i 为常返的, 若

$$\mathbb{P}_i(X_n = i \text{对某个} n \geqslant 1) = 1$$

如果上面的概率 < 1, 则称为暂留的/非常返的

注：i 常返 $\Leftrightarrow \mathbb{P}_i(\cup_{n \geqslant 1}\{X_n = i\}) = 1$

思考：i 常返 $\Leftrightarrow$ “不停地/无数次回到 i”

$\Leftrightarrow \mathbb{P}_i(\omega|\omega \in$ 无数多个$\{X_n = i\})$

$\Leftrightarrow \mathbb{P}_i(\omega|\omega \in \cap_{k \geqslant 1} \cup_{n \geqslant k} \{X_n = i\}, \forall k)$

$\Leftrightarrow \mathbb{P}_i(X_n = i, i.o.)$ (infinitely often)

无数多次返回 i 可严格定义为：

$$\bigcap_{k \geqslant 1} \bigcup_{n \geqslant k} \{X_n = i\}$$

集合的语言中, $\cup$ 即 $\exists$, $\cap$ 即 $\forall$, 因此

- $\bigcup_{n \geqslant k}\{X_n = i\}$ 表示 $\exists n_0 \geqslant k$ 使得 $X_{n_0} = i$
- 对 $\forall k$ 取交集 $\bigcap_{k \geqslant 1}$, 即无论 k 多大, 总存在更大的 n 满足 $X_n = i$, 从而保证无限次返回

即 $\forall k, \exists n_k, \text{s.t. } \{X_{n_k} = i\}$ 发生

$$k = 1, n_1 \geqslant k$$

$$k = n_1 + 1, n_2 \geqslant n_1 + 1 > n_1$$

$$\cdots$$

Remark 5 (如何进一步理解). 无界和 ∞ 的区别是什么？

无界：$\forall M > 0, \exists k, s.t. \ |x_k| > M$

Example 2.3

$1, 2, 3, 4, \cdots$ 为 ∞/无界

$1, 0, 2, 0, 3, 0, 4, \cdots$ 并非 ∞, 但是无界的

迁移到 $\bigcap_{k \geqslant 1} \bigcup_{n \geqslant k}$ 的例子

Example 2.4

$A_1=\{0,1\}, A_2=\{2\}, A_3=\{0,3\}, \cdots$, 则

$$\bigcap_{k=1}^{\infty}\bigcup_{n\geqslant k}A_n=\{0\}, \qquad \bigcap_{k=1}^{\infty}A_k=\varnothing$$

其中 $\bigcap_{k=1}^{\infty}\bigcup_{n\geqslant k}$ 也即 $\limsup$

但我们推理得到的“常返”和定义里的并不等价

$$\bigcap_{k\geqslant 1}\bigcup_{n\geqslant k}\{X_n=i\}\not\Leftrightarrow\bigcup_{n\geqslant 1}\{X_n=i\}$$

且 LHS 是 RHS 的子集, 因此由定义的 $\mathbb{P}(RHS)=1$ 不能推出 $\mathbb{P}(LHS)=1$. 于是我们疑惑为什么会叫它常返. 这里要用到高阶知识“停时”, 我们最后会回到这个问题.
下面给出几种判断常返/暂留的方法.

2.4.2 从数学角度：并改写成不交并

i 常返 $\Leftrightarrow \mathbb{P}_i(\cup_{n\geqslant 1}\{X_n=i\})=1$
$\quad\Leftrightarrow \mathbb{P}_i(\text{有限步到达}i)=1$
$\quad\Leftrightarrow \mathbb{P}(\text{从}i\text{出发条件下, 有限时间内回到}i)=1$
$B_1(i)=\{X_1=i\}, B_2(i)=\{X_2=i\}\backslash\{X_1=i\}=\{X_2=i, X_1\neq i\}, \cdots, B_n(i)=\{X_n=i, X_{n-1}\neq i\cdots, X_1\neq i\}$
由练习2,

$$\sum_{n\geqslant 1}B_n(i)=\bigcup_{n\geqslant 1}\{X_n=i\} \tag{2.6}$$

i 常返 $\Leftrightarrow 1=\mathbb{P}_i(\sum_{n\geqslant 1}B_n(i))=\sum_{n\geqslant 1}\mathbb{P}_i(B_n(i))$, 第二个等号由可列可加性得到（定义1.5）

$$\begin{aligned}\mathbb{P}_i(B_n(j))=\mathbb{P}_i(X_n=j, X_{n-1}\neq j, \cdots, X_1\neq j)&=\mathbb{P}_i(\text{首次访问}j\text{的时刻为}n)\\&=\mathbb{P}_i(\text{走}n\text{步首次到达}j)\end{aligned}$$

故

$$\mathbb{P}_i\left(\sum_{n\geqslant 1}B_n(j)\right)=\mathbb{P}_i(\text{首次访问}j\text{的时刻为有限时间})=\mathbb{P}_i(\text{有限时间内首次访问}j)$$

记号

$$\begin{cases}f_{ij}:=\mathbb{P}_i\left(\sum_{n\geqslant 1}B_n(j)\right)=\sum_{n\geqslant 1}\mathbb{P}_i(B_n(j))=\mathbb{P}_i(\text{首次访问}j\text{的时刻为有限时间})\\ f_{ij}(n):=\mathbb{P}_i(B_n(j))=\mathbb{P}_i(\text{首次访问}j\text{的时刻为}n)\end{cases} \tag{2.7}$$

Proposition 2.4

（不交并视角下）常返和暂留的等价命题

1. i 常返 $\iff$
$$1 = f_{ii} = \sum_{n\geqslant 1} f_{ii}(n) \tag{2.8}$$

2. i 暂留 $\iff$
$$1 > f_{ii} = \sum_{n\geqslant 1} f_{ii}(n) \tag{2.9}$$

证明：由可列可加性，$f_{ii} = \sum_{n\geqslant 1} f_{ii}(n)$ 总是成立. 而 $f_{ii} = \mathbb{P}_i(\sum_{n\geqslant 1} B_n(i)) = \mathbb{P}_i(\cup_{n\geqslant 1}\{X_n = i\})$，即 i 常返的定义，因此 $f_{ii} = 1$. 若 i 暂留，则 $f_{ii} \neq 1$，由概率测度的定义，f_{ii} 不能大于 1，所以 $f_{ii} < 1$. □

2.4.3 从“多步转移概率”角度判别

定义新记号（P 不是转移矩阵）

$$P_{ij}(s) := \sum_{n\geqslant 0} s^n p_{ij}(n) \qquad F_{ij}(s) := \sum_{n\geqslant 0} s^n f_{ij}(n)$$

其中，$p_{ij}(0) = \delta_{ij}, f_{ij}(0) = 0$

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

注：当 $|s| < 1$ 时，$P_{ij}(s), F_{ij}(s)$ 绝对收敛

由 Abel 连续性定理，

$$\lim_{s\uparrow 1} F_{ij}(s) = \sum_{n\geqslant 1} f_{ij}(n) = f_{ij} \in [0,1]$$

$$\lim_{s\uparrow 1} P_{ij}(s) = \sum_{n\geqslant 0} p_{ij}(0) = \text{finite}/+\infty$$

Lemma 2.3 (Grimmett [3], Thm 6.3.3)

设 $|s| < 1$，则

$$P_{ij}(s) = \delta_{ij} + P_{jj}(s)F_{ij}(s) \tag{2.10}$$

其中

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

证明：构造不交并，$B_m(i) = \{X_m = i, X_{m-1} \neq i, \cdots, X_1 \neq i\}, m \geqslant 1$

$\sum_{m\geqslant 1} B_m(i) = \cup_{n\geqslant 1}\{X_n = i\}, B_m(i) \subseteq \{X_n \neq i\}, m \geqslant n+1$

$$\begin{aligned} p_{ij}(n) = \mathbb{P}_i(X_n = j) &= \mathbb{P}_i(\{X_n = j\} \cap \sum_{m\geqslant 1} B_m(j)) \\ &= \sum_{m\geqslant 1} \mathbb{P}_i(\{X_n = j\} \cap B_m(j)) \quad = \sum_{m=1}^{n} \mathbb{P}_i(\{X_n = j\} \cap B_m(j)) \end{aligned} \tag{2.11}$$

最后一个等号成立是因为 $m \geqslant n+1$ 时 $\{X_n = j\} \cap B_m(j)$ 为空集

$$\sum_{m=1}^{n} \mathbb{P}_i(\{X_n = j\} \cap B_m(j)) = \sum_{m=1}^{n} \mathbb{P}_i(X_n = j | B_m(j)) \mathbb{P}_i(B_m(j)) \tag{2.12}$$

其中 $X_m = j, X_{n-1} \neq j, \cdots, X_1 \neq j,\ X_{n-1} \in S\backslash\{j\}$

用一般而非单点的马氏性(2.2)

$$\begin{aligned} \sum_{m=1}^{n} \mathbb{P}_i(X_n = j | B_m(j)) \mathbb{P}_i(B_m(j)) &= \sum_{m=1}^{n} \mathbb{P}(X_n = j | X_m = j) \cdot f_{ij}(m) \\ &= \sum_{m=1}^{n} p_{jj}(n-m) \cdot f_{ij}(m) \end{aligned} \tag{2.13}$$

整合(2.11), (2.12), (2.13),

$$p_{ij}(n) = \sum_{m=1}^{n} p_{jj}(n-m) \cdot f_{ij}(m) \tag{2.14}$$

当 $n \geqslant 1$ 时,

$$\begin{aligned} P_{ij}(s) &= s^0 p_{ij}(0) + \sum_{n\geqslant 1} s^n \cdot p_{ij}(n) \\ &= \delta_{ij} + \sum_{n\geqslant 1} s^n \sum_{m=1}^{n} p_{jj}(n-m) f_{ij}(m) \\ &= \delta_{ij} + \sum_{n\geqslant 1} \sum_{m=1}^{n} s^n p_{jj}(n-m) f_{ij}(m) \\ &= \delta_{ij} + \sum_{n\geqslant 1} \sum_{m=1}^{n} (s^{n-m} p_{jj}(n-m))(s^m f_{ij}(m)) \\ &= \delta_{ij} + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mathbb{I}_{\{1\leqslant m\leqslant n\}} (s^{n-m} p_{jj}(n-m))(s^m f_{ij}(m)) \end{aligned}$$

【重要技巧】把 $\mathbb{I}_{\{1\leqslant m\leqslant n\}}(s^{n-m} p_{jj}(n-m))(s^m f_{ij}(m))$ 看作 $a_{n,m}$, 由 Lemma 1.1考察绝对收敛

$0 \leqslant s < 1, |s| = s$

正向级数一定有意义, 就看是有限/∞

$$\begin{aligned}
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\mathbb{I}_{\{1\leqslant m\leqslant n\}}s^{n-m}p_{jj}(n-m)s^m f_{ij}(m)\\
=&\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\mathbb{I}_{\{1\leqslant m\leqslant n\}}s^{n-m}p_{jj}(n-m)s^m f_{ij}(m)\\
=&\sum_{m=1}^{\infty}(\sum_{n=m}^{\infty}s^{n-m}p_{jj}(n-m))s^m f_{ij}(m)\\
=&(\sum_{k=0}^{\infty}s^k p_{jj}(k))(\sum_{m=1}^{\infty}s^m f_{ij}(m))<\infty
\end{aligned}$$

其中 $k=n-m$. 因为 $s^0 f_{ij}(0)=0$, 则 $\sum_{m=0}^{\infty}s^m f_{ij}(m)=\sum_{m=1}^{\infty}s^m f_{ij}(m)$. 代回原式

$$P_{ij}(s)=\delta_{ij}+P_{jj}(s)\cdot F_{ij}(s)$$

Proposition 2.5

(多步转移概率视角下)常返和暂留的等价命题

1. j 常返 $\iff$
$$1=f_{jj}\Leftrightarrow\sum_{n\geqslant 0}p_{jj}(n)=\infty \tag{2.15}$$
2. j 暂留 $\iff$
$$\begin{aligned}
1>f_{jj}\Leftrightarrow&\sum_{n\geqslant 0}p_{jj}(n)<\infty\\
\Rightarrow&\sum_{n\geqslant 0}p_{ij}(n)<\infty,\forall i\in S\\
\Rightarrow&\lim_{n\to\infty}p_{ij}(n)=0,\forall i\in S
\end{aligned}\tag{2.16}$$

证明: 只证 (1). $|s|<1$ 时, $P_{ij}(s)=\delta_{ij}+P_{jj}(s)F_{ij}(s)$
令 $i=j$, $P_{jj}(s)=1+P_{jj}(s)F_{jj}(s)$

$$P_{jj}(s)=\frac{1}{1-F_{jj}(s)}\tag{2.17}$$

由(2.8), j 常返 $\iff 1=f_{jj}=F_{jj}(1)\overset{\text{Abel}}{=}\lim_{s\uparrow 1}F_{jj}(s)$
对(2.17), 令 $s\to 1$, 有 $\sum_{n\geqslant 0}p_{jj}(n)=+\infty$ □

Problem 6 (作业 5-1)

证明: j 暂留 $\Rightarrow\sum_{n\geqslant 0}p_{ij}(n)<\infty,\forall i\in S$

证明. 由(2.10), 若 $s\uparrow 1$,

$$\sum_{n\geqslant 0}p_{ij}(n)=\delta_{ij}+\left(\sum_{n\geqslant 0}p_{jj}(n)\right)\left(\sum_{n\geqslant 0}f_{ij}(n)\right)\leqslant\delta_{ij}+\sum_{n\geqslant 0}p_{jj}(n)<\infty$$

其中 $\sum_{n\geqslant 0} f_{ij}(n) = f_{ij} \in [0,1]$. □

2.4.4 从"首次回访时间"角度判别

$$
\begin{aligned}
j\text{常返} &\iff 1 = \mathbb{P}_j(X_n = j \text{ 对某个} n \geqslant 1) = \mathbb{P}_j(\text{有限时间内回访}j) \\
&\iff 1 = f_{jj} = \mathbb{P}_j(\underbrace{\text{首次回访}j\text{的时刻}}_{T_j<\infty}\text{有限}) \\
& \quad 1 = \sum_{n\geqslant 1} f_{jj}(n) = \sum_{n\geqslant 1} \mathbb{P}_j(\underbrace{\text{首次回访}j\text{的时刻}}_{T_j=n}\text{是}n)
\end{aligned}
$$

Definition 2.8 (首次回访时间)

首次回访 j 的时刻

$$T_j = \min\{n \geqslant 1 | X_n = j\} \tag{2.18}$$

约定 $\min \varnothing = +\infty$

注: $\{T_j = \infty\} \iff \{\omega | \{n \geqslant 1 | X_n(\omega) = j\} = \varnothing\} \iff \{\omega | X_n(\omega) \neq j, \forall n \geqslant 1\} = \cap_{n\geqslant 1}\{X_n \neq j\}$
$\{T_j < \infty\} = (\{T_j = \infty\})^c = (\cap_{n\geqslant 1}\{X_n \neq j\})^c = \cup_{n\geqslant 1}(\{X_n \neq j\})^c = \cup_{n\geqslant 1}\{X_n = j\}$
这个式子串联起常返的定义和 T_j 的关系, 于是有以下性质.

Property 2.2

$f_{jj} = \mathbb{P}_j(T_j < \infty), f_{jj}(n) = \mathbb{P}_j(T_j = n)$

由(2.7)知, $f_{ij}, f_{ij}(n)$ 是由不交并定义的, 对于"首次回访时间"这一角度, 定义新的符号

$$\rho_{ij} := \mathbb{P}_i(T_j < \infty) \tag{2.19}$$

Proposition 2.6

(首次回访时间视角下) 常返和暂留的等价命题

1. j 常返 $\iff$
$$1 = \rho_{jj} = \mathbb{P}_j(T_j < \infty) \iff 0 = \mathbb{P}_j(T_j = \infty) \tag{2.20}$$
2. j 暂留 $\iff$
$$1 > \rho_{jj} = \mathbb{P}_j(T_j < \infty) \iff 0 < \mathbb{P}_j(T_j = \infty) \tag{2.21}$$

证明: 由(2.15), (2.16), j常返 $\iff 1 = f_{jj} = \rho_{jj} \iff 0 = 1 - \rho_{jj}$, 其余同理. □

Definition 2.9 (平均回访时间)

j 的平均回访时间

$$m_j := \mathbb{E}_j T_j = \mathbb{E}(T_j | X_0 = j) \tag{2.22}$$

Theorem 2.4

用不交并表示平均回访时间.

$$m_j = \mathbb{E}_j T_j = \begin{cases} \sum_{n\geqslant 1} n f_{jj}(n) & j\text{常返} \\ \infty & j\text{暂留} \end{cases} \tag{2.23}$$

证明:

(1) j 暂留 $\Rightarrow \mathbb{P}_j(T_j = \infty) > 0$

$T_j = T_j \mathbb{I}_{\{T_j=\infty\}} + T_j \mathbb{I}_{\{T_j<\infty\}}$

$\mathbb{E}_j T_j = \mathbb{E}_j T_j \mathbb{I}_{\{T_j=\infty\}} + \mathbb{E}_j T_j \mathbb{I}_{\{T_j<\infty\}} \geqslant \mathbb{E}_j T_j \mathbb{I}_{\{T_j=\infty\}} = \infty \cdot \mathbb{P}_j(T_j = \infty) = \infty$

(2) j 常返 $\Rightarrow \mathbb{P}_j(T_j = \infty) = 0$

取期望时不起作用, 因为 $0 \cdot \infty$ 是不定形

$$\mathbb{E}_j T_j = \mathbb{E}_j T_j \mathbb{I}_{\{T_j<\infty\}} = \mathbb{E}_j T_j \mathbb{I}_{\sum_{n\geqslant 1}\{T_j=n\}} \overset{(1.1)}{=} \mathbb{E}_j \sum_{n\geqslant 1} T_j \mathbb{I}_{\{T_j=n\}} = \sum_{n\geqslant 1} n \mathbb{P}_j(T_j = n) = \sum_{n\geqslant 1} n f_{jj}(n)$$

Definition 2.10 (正常返/零常返)

j 常返时

1. $\mathbb{E}_j T_j < \infty$ 称 j 是正常返
2. $\mathbb{E}_j T_j = \infty$ 称 j 是零常返（平均意义上再也不回来）

j 常返 $\iff 1 = f_{jj} \iff \sum_{n\geqslant 0} p_{jj}(n) = \infty$

$\iff 1 = \rho_{jj} = \mathbb{P}_j(T_j < \infty) \iff 0 = \mathbb{P}_j(T_j = \infty)$

$$\begin{aligned} \mathbb{P}_j(T_j < \infty) &= \mathbb{P}(\text{从}j\text{出发条件下, 首次回到}j\text{的时刻有限}) \\ &= \mathbb{P}(\text{从}j\text{出发条件下, 有限时间内至少访问}j\text{有 1 次}) \\ &= \mathbb{P}(\text{从}j\text{出发条件下, 有限时间内回访}j\text{的次数} \geqslant 1) \end{aligned} \tag{2.24}$$

Definition 2.11 (访问次数)

链在时刻 0 之后, 访问 j 的次数

$$N(j) := \#\{n \geqslant 1 | X_n = j\} = \sum_{n\geqslant 1} \mathbb{I}_{\{X_n=j\}} \tag{2.25}$$

注: $N(j) : \Omega \to \{0, 1, 2, \cdots\} \cup \{+\infty\}$

至此, 我们已经从四个角度表示了常返

1. 常返的定义
2. 不交并
3. 多步转移概率
4. 首次访问时间

做个阶段性小结, 回顾 i 常返的几种等价表示

$$
\begin{aligned}
i\text{常返} &\overset{\text{Def}}{\iff} 1=\mathbb{P}_i(\bigcup_{n\geqslant 1}\{X_n=i\}) \\
&\iff 1=f_{ii}:=\sum_{n\geqslant 1}f_{ii}(n)=\sum_{n\geqslant 1}\mathbb{P}_i(X_1\neq i,\cdots,X_{n-1}\neq i,X_n=i) \\
&\iff \sum_{n\geqslant 1}p_{ii}(n)=\infty \\
&\iff 1=\rho_{ii}=\mathbb{P}_i(T_i^{(1)}:=T_i<\infty)
\end{aligned}
$$

由(2.24), $\mathbb{P}_i(T_i<\infty)=\mathbb{P}_i(N(i)\geqslant 1)$, 因此可得“回访次数”视角下 i 常返的条件

Proposition 2.7

(回访次数视角下) 常返和暂留的等价命题

1. i 常返 $\iff$
$$\mathbb{P}_i\{N(i)\geqslant 1\}=1 \tag{2.26}$$
2. i 暂留 $\iff$
$$\mathbb{P}_i\{N(i)\geqslant 1\}<1 \tag{2.27}$$

另一方面, 我们从“常返”的文字含义推理.

无数次地回访 $\leftrightarrow$ 访问次数 $=\infty$ $\leftrightarrow$

$$\mathbb{P}_i(N(i)=\infty)=1$$

两种表述的等价条件互相等价吗? 即

$$\mathbb{P}_i(N(i)\geqslant 1)\overset{?}{\iff}\mathbb{P}_i(N(i)=\infty)$$

需要 Strong Markov Property (SMP) 使上面 $\Leftrightarrow$ 成立. 这里先补充一些关于 $N(j)$ 的内容, 然后再回到证明.

考察 $\{N(y)=\infty\}=\cap_{k\geqslant 1}\{N(y)\geqslant k\}$

由概率测度的连续性 (性质1.3)

$$\mathbb{P}_x(N(y)=\infty)=\lim_{k\to\infty}\mathbb{P}_x(N(y)\geqslant k)$$

其中,

$$
\begin{aligned}
\mathbb{P}_x(N(y)\geqslant k) &= \mathbb{P}(\text{从}x\text{出发条件下, 访问}y\text{的次数}\geqslant k) \\
&= \mathbb{P}(\text{从}x\text{出发条件下, 至少访问}y\text{有}k\text{次}) \\
&= \mathbb{P}(\text{从}x\text{出发条件下, 第}k\text{次访问}y\text{的时刻有限})
\end{aligned}
$$

Definition 2.12 (第 k 次访问时间)

由 $T_y^{(1)}:=T_y:=\min\{n\geqslant 1|X_n=y\},T_y^{(2)}:=\min\{n>T_y^{(1)}|X_n=y\},\cdots$, 得到

$$T_y^{(k)}:=\min\{n>T_y^{(k-1)}|X_n=y\},\quad \forall k\geqslant 2 \tag{2.28}$$

Claim 2.3. $N(y) \geqslant k$ 与“第 k 次访问 y 的时刻有限”等价，

$$\mathbb{P}_x(N(y) \geqslant k) = \mathbb{P}_x(T_y^{(k)} < \infty) \tag{2.29}$$

Definition 2.13
(1) 第 k 次访问概率

$$\rho_{xy}^{(k)} := \mathbb{P}_x(T_y^{(k)} < \infty) \tag{2.30}$$

其中，$\rho_{xy}^{(1)} = \rho_{xy}$，rf.(2.19)
(2) 第 k 次回访概率

$$\rho_{yy}^{(k)} := \mathbb{P}_y(T_y^{(k)} < \infty) \tag{2.31}$$

注：$\rho_{yy}^{(2)} \overset{?}{=} \rho_{yy} \cdot \rho_{yy}$
直观上是这样，但严格证明要求 SMP
这是因为不同时间对应的是不同的随机过程，如

- $t = 0$ 时，过程是 $\{X_n, n \geqslant 0\}$
- $t = T_j$ 时，过程是 $\{X_{T_j+n}, n \geqslant 0\}$

SMP 是一个使得 $X_{T_j+n} = X_n, \forall T_j$ 的性质，之后会详细说. 以上结论可总结成下面引理.

Lemma 2.4
(由 SMP 知) $\rho_{xy}^{(k)} = \rho_{xy}\rho_{yy}^{(k-1)}$. 特别地，$\rho_{yy}^{(k)} = \rho_{yy}^k$

接着我们回到证明

$$\mathbb{P}_i(N(i) \geqslant 1) = 1 \iff \mathbb{P}_i(N(i) = \infty) = 1 \tag{2.32}$$

证明：$\Leftarrow$ 显然，因为 $\{N(i) = \infty\}$ 相对 $N(i) \geqslant 1$ 是小集合
$\Rightarrow$ 由于 $\{N(i) = \infty\} = \cap_{k\geqslant 1}\{N(i) \geqslant k\}$，$\{N(i) \geqslant k\}$ 单调下降，所以 $\cap_{k\geqslant 1}\{N(i) \geqslant k\} = \lim_{k\to\infty}\{N(i) \geqslant k\}$.

$$\mathbb{P}_i(N(i) = \infty) = \lim_{k\to\infty} \mathbb{P}_i(N(i) \geqslant k) \overset{\text{SMP}}{=} \lim_{k\to\infty} \rho_{ii}^k = 1$$

暂留的证明同理：

$$\begin{aligned} i\text{暂留} &\iff 1 > \rho_{ii} \\ &\iff \mathbb{P}_i(N(i) = \infty) = \lim_{k\to\infty} \rho_{ii}^k = 0 \end{aligned}$$

2.4.5 从“平均回访次数”角度判别

回顾(2.25)，

$$N(y) = \sum_{n\geqslant 1} \mathbb{I}_{\{X_n = y\}} = \sum_{k\geqslant 1} \mathbb{I}_{\{N(y)\geqslant k\}}$$

Lemma 2.5 (Durrett [2], lem 1.11)

$$\mathbb{E}_y N(y) = \begin{cases} \infty & y\text{常返} \\ \dfrac{\rho_{yy}}{1-\rho_{yy}} & y\text{暂留} \end{cases} \tag{2.33}$$

证明：

$$\mathbb{E}_y N(y) = \mathbb{E}_y \sum_{k\geqslant 1} \mathbb{I}_{\{N(y)\geqslant k\}} \overset{\text{Exa(1.6)}}{=} \sum_{k\geqslant 1} \mathbb{P}_y(N(y)\geqslant k) = \sum_{k\geqslant 1} \mathbb{P}_y(T_y^{(k)}<\infty) = \sum_{k\geqslant 1} \rho_{yy}^{(k)} \overset{\text{SMP}}{=} \sum_{k\geqslant 1} \rho_{yy}^k$$

$\rho_{yy}=1 \Rightarrow \mathbb{E}_y N(y)=\infty$

$\rho_{yy}<1 \Rightarrow \mathbb{E}_y N(y) = \rho_{yy}/(1-\rho_{yy})$ □

Proposition 2.8

（平均回访次数视角下）常返和暂留的等价命题

1. i 常返 $\iff$

$$\mathbb{E}_i N(i) = \infty \tag{2.34}$$

2. i 暂留 $\iff$

$$\mathbb{E}_i N(i) < \infty \tag{2.35}$$

证明 (1)：$\Rightarrow$ 由(2.33), 显然

$\Leftarrow$ $N(y)$ 为非负 r.v., 有当 $k\to\infty$ 时, $\forall\omega, \sum_{n=1}^k \mathbb{I}_{\{X_n=y\}}(\omega) \uparrow \sum_{n=1}^\infty \mathbb{I}_{\{X_n=y\}}(\omega)$

$$\mathbb{E}_y N(y) := \lim_{k\to\infty} \mathbb{E}_y \sum_{n=1}^k \mathbb{I}_{\{X_n=y\}} = \lim_{k\to\infty} \sum_{n=1}^k \mathbb{E}_y \mathbb{I}_{\{X_n=y\}} = \lim_{k\to\infty} \sum_{n=1}^k \mathbb{P}_y(X_n=y) = \sum_{n=1}^\infty p_{yy}(n) = \infty$$

由(2.15), y 常返 □

将上面几个角度总结成下面定理

Theorem 2.5 (链的状态：等价表述)

$$\begin{aligned}
i\text{常返} &\overset{\text{Def}}{\iff} 1 = \mathbb{P}_i(\bigcup_{n\geqslant 1}\{X_n=i\}) \quad [\text{回访发生的概率}] \\
&\iff 1 = f_{ii} := \sum_{n\geqslant 1} f_{ii}(n) = \sum_{n\geqslant 1} \mathbb{P}_i(X_1\neq i,\cdots,X_{n-1}\neq i, X_n=i) \\
&\iff \sum_{n\geqslant 1} p_{ii}(n) = \infty \\
&\iff 1 = \rho_{ii} = \mathbb{P}_i(T_i^{(1)} := T_i < \infty) = \mathbb{P}_i\{N(i)\geqslant 1\} = 1 \\
&\overset{\text{why the name}}{\iff} \mathbb{P}_i(N(i)=\infty) = \lim_{k\to\infty} \mathbb{P}_i(N(i)\geqslant k) \overset{\text{SMP}}{=} \lim_{k\to\infty} \rho_{ii}^k = 1 \\
&\iff \mathbb{E}_i N(i) = \infty
\end{aligned}$$

2.4.6 停时与强马氏性

Definition 2.14 (停时/Stopping time)
随机变量 $\tau:\Omega\to\{0,1,2,\cdots\}\cup\{+\infty\}$, 满足 $\forall\infty>n\geqslant 0,\{\tau=n\}\in\sigma(X_0,\cdots,X_n)$, 称 τ 是关于 $(X_n)_{n\geqslant 0}$ 的停时

Example 2.5
首次回访时刻是一个停时 $T_y^{(1)}:=T_y:=\min\{n\geqslant 1|X_n=y\}$.

$$\begin{aligned}\{T_y^{(1)}=n\}&=\{X_1\neq y,\cdots,X_{n-1}\neq y,X_n=y\}\quad n\geqslant 1\\&=\{(X_0,X_1,\cdots,X_n)\in S\times(S\backslash\{y\})\times\cdots\times(S\backslash\{y\})\times\{y\}\}\\&\in\sigma(X_0,\cdots,X_n)=(X_0,\cdots,X_n)^{-1}(\bigotimes_{n+1}2^S)\end{aligned}$$

Definition 2.15 (停止 σ 代数)
τ 是关于 $(X_n)_{n\geqslant 0}$ 的停时, 定义

$$\mathcal{F}_\tau:=\{A|A\cap\{\tau=n\}\in\sigma(X_0,\cdots,X_n),\forall n\}$$

注：$B\in\mathcal{F}_\tau\Leftrightarrow B$ 是由 $X_0,\cdots,X_\tau$ 决定的事件（这是直观上的解释, 因为 τ 是随机的, 我们不知道“$\cdots$”是什么）$\Leftrightarrow B\cap\{T=n\}\in\sigma(X_0,\cdots,X_n),\forall n$
停止 σ-代数的定义是为了形式化“到随机时间 τ 为止的信息”. 因为 τ 本身是随机的, 我们不能直接写 $\sigma(X_0,\cdots,X_\tau)$（因为 τ 不确定）, 所以需要通过对所有可能的 $\tau=n$ 进行分解.

Problem 7 (作业 5-2)
设 τ 为关于 $(X_n)_{n\geqslant 0}$ 的停时, 即对任意的 $\infty>n\geqslant 0$, 有

$$\{\tau=n\}\in\sigma(X_0,\cdots,X_n)$$

证明:

1. (停止 σ 代数的定义) $\mathcal{F}_\tau:=\{A|A\cap\{\tau=n\}\in\sigma(X_0,\cdots,X_n),\forall n\geqslant 1\}$ 是一个 σ-代数
2. $\sigma(X_\tau)\in\mathcal{F}_\tau$

以下内容来自强马氏性讲义 [1].

Lemma 2.6 (马氏性的小推广)
若 X 为马氏链, 则对任意 $n,m\geqslant 0,x_n\in S,A\in\otimes_{n+1}2^S,B\in\otimes_{m+1}2^S$（即 $A\subset S^{n+1},B\subset S^{m+1}$）, 有

$$\begin{aligned}&\mathbb{P}_{\{X_n=x_n\}}((X_0,\cdots,X_n)\in A,(X_n,\cdots,X_{n+m})\in B)\\=&\mathbb{P}_{\{X_n=x_n\}}((X_0,\cdots,X_n)\in A)\times\mathbb{P}_{\{X_n=x_n\}}((X_n,\cdots,X_{n+m})\in B)\end{aligned}\tag{2.36}$$

即 $(X_0,\cdots,X_n)\perp\!\!\!\perp_{\{X_n=x_n\}}(X_n,\cdots,X_{n+m})$ 的定义, rf.(1.3)

证明：回顾马氏性, rf.(2.3)

不妨设 $A = A_0 \times \cdots \times A_n, B = B_n \times \cdots \times B_{n+m}$

(Case 1) 若 $x_n \notin A_n$ 或 $x_n \notin B_n$, 则

$$\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_n) \in A) = 0 \text{ 或} \mathbb{P}_{\{X_n=x_n\}}((X_n, \cdots, X_{n+m}) \in B) = 0$$

且

$$\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_n) \in A, (X_n, \cdots, X_{n+m}) \in B) = 0$$

从而, (2.36) 得证

(Case 2) 设 $x_n \in A_n$, 且 $x_n \in B_n$. 若 $n = 0, m = 0$, 则显然有

$$\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_n) \in A) = \mathbb{P}_{\{X_n=x_n\}}((X_n, \cdots, X_{n+m}) \in B) = 1$$

$$\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_n) \in A, (X_n, \cdots, X_{n+m}) \in B) = 1$$

此时, 显然有 (2.36) 成立. 若 $n \geqslant 1, m \geqslant 1$, 则

$$\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_n) \in A) = \mathbb{P}_{\{X_n=x_n\}}\{X_j \in A_j, 0 \leqslant j \leqslant n-1\}$$

$$\mathbb{P}_{\{X_n=x_n\}}((X_n, \cdots, X_{n+m}) \in B) = \mathbb{P}_{\{X_n=x_n\}}\{X_j \in B_j, n+1 \leqslant j \leqslant n+m\}$$

且

$$\begin{aligned}
&\mathbb{P}_{\{X_n=x_n\}}((X_0, \cdots, X_n) \in A, (X_n, \cdots, X_{n+m}) \in B) \\
=&\mathbb{P}_{\{X_n=x_n\}}\{X_j \in A_j, 0 \leqslant j \leqslant n-1, X_k \in B_k, n+1 \leqslant k \leqslant n+m\} \\
\overset{(2.3)}{=}&\mathbb{P}_{\{X_n=x_n\}}\{X_j \in A_j, 0 \leqslant j \leqslant n-1\}\mathbb{P}_{\{X_n=x_n\}}\{X_k \in B_k, n+1 \leqslant k \leqslant n+m\}
\end{aligned}$$

故而 (2.36) 得证. 对于其他情形 $n \geqslant 1, m = 0$ 或 $n = 0, m \geqslant 1$, 可类似证明. $\square$

Proposition 2.9 (强马氏性)

$X := \{X_n, n \geqslant 0\} \sim \text{Markov}(\mu, P)$, τ 是关于 $(X_n)_{n\geqslant 0}$ 的停时, 则

1. 在 $\{\tau < \infty\}$ 和 $\{X_\tau = x\}$ 条件下

$$(X_{\tau+n})_{n\geqslant 0} \sim \text{Markov}(\delta_x, P)$$

   其中 $\delta_x = (\delta_{xy})_{y\in S}$, 记号

$$\delta_{xy} = \begin{cases} 1 & y = x \\ 0 & y \neq x \end{cases}$$

   注: $(X_{\tau+n})_{n\geqslant 0} \sim \text{Markov}(\delta_x, P)$ under $\mathbb{P}(\cdot|\tau < \infty, X_\tau = x)$. 在原先的概率测度 $\mathbb{P}$ 下, $(X_{\tau+n})_{n\geqslant 0}$ 不是马氏链

2. $\forall J \subseteq \mathbb{N}_0, \#J < \infty$, 有

$$\sigma(X_{\tau+n}, n \in J) \perp\!\!\!\perp_{\{\tau<\infty, X_\tau = x\}} \mathcal{F}_\tau$$

   注:

   (a) $(X_{\tau+n})_{n\geqslant 0}$ 与 $X_0, \cdots, X_\tau$ 独立

   (b) $(X_{\tau+n})_{n\geqslant 0} \perp\!\!\!\perp \mathcal{F}_\tau$ under $\mathbb{P}(\cdot|\tau < \infty, X_\tau = x)$

证明: (1) 回顾马氏链的有限维分布, rf.(2.4),
根据此结论, 我们只需考察链 $\{X_{\tau+n}, n \geqslant 0\}$ 的有限维分布.
(Step 1) 设 $j_0 \neq x$. 则对任意的 $n \geqslant 1, m \geqslant 0$, 有

$$\mathbb{P}(X_{\tau+0} = j_0, \cdots, X_{\tau+n} = j_n, \tau = m, X_\tau = x) = 0$$

关于 $m \geqslant 0$ 求和, 并注意到 $\{\tau < \infty\} = \sum_{m=0}^{\infty}\{\tau = m\}$, 得:

$$\mathbb{P}(X_{\tau+0} = j_0, \cdots, X_{\tau+n} = j_n, \tau < \infty, X_\tau = x) = 0$$

两边同除 $P(\tau < \infty, X_\tau = x)$, 有

$$\mathbb{P}(X_{\tau+0} = j_0, \cdots, X_{\tau+n} = j_n|\tau < \infty, X_\tau = x) = 0 = \delta_{xj_0} \prod_{k=0}^{n-1} p_{j_k j_{k+1}}$$

(Step 2) 设 $j_0 = x$. 注意到, 对任意的 $m \geqslant 0$, 有

$$\{\tau = m\} \in \sigma(X_0, \cdots, X_m)$$

故而, 由(2.36): 对任意的 $n \geqslant 1, m \geqslant 0$, 有

$$\{\tau = m\} \perp\!\!\!\perp_{\{X_m = x\}} \{X_m = j_0, X_{m+1} = j_1, \cdots, X_{m+n} = j_n\}$$

从而, 对任意的 $m \geqslant 0$, 有

$$\begin{aligned}
\mathbb{P}(X_{\tau+0}=j_0,\tau=m,X_\tau=x) &= \mathbb{P}(X_{m+0}=j_0,\tau=m,X_m=x) \\
&\overset{(1.2)}{=} \mathbb{P}(X_{m+0}=j_0,\tau=m|X_m=x)\times\mathbb{P}(X_m=x) \\
&\overset{(M_1)}{=} \mathbb{P}(X_{m+0}=j_0|X_m=x)\times\mathbb{P}(\tau=m|X_m=x)\times\mathbb{P}(X_m=x) \\
&\overset{(1.2)}{=} 1\times\mathbb{P}(\tau=m,X_m=x) \\
&= \delta_{xx}\mathbb{P}(\tau=m,X_m=x)
\end{aligned}$$

以及, 对任意的 $n \geqslant 1, m \geqslant 0$, 有

$$\begin{aligned}
&\mathbb{P}(X_{\tau+0}=j_0,X_{\tau+1}=j_1,\cdots,X_{\tau+n}=j_n,\tau=m,X_\tau=x) \\
=&\mathbb{P}(X_{m+1}=j_1,\cdots,X_{m+n}=j_n,\tau=m,X_m=x) \\
\overset{(1.2)}{=}&\mathbb{P}(X_{m+1}=j_1,\cdots,X_{m+n}=j_n,\tau=m|X_m=x)\times\mathbb{P}(X_m=x) \\
\overset{(M_1)}{=}&\mathbb{P}(X_{m+1}=j_1,\cdots,X_{m+n}=j_n|X_m=x)\times\mathbb{P}(\tau=m|X_m=x)\times\mathbb{P}(X_m=x) \\
=&\frac{\mathbb{P}(X_{m+1}=j_1,\cdots,X_{m+n}=j_n,X_m=x)}{\mathbb{P}(X_m=x)}\times\mathbb{P}(\tau=m|X_m=x)\times\mathbb{P}(X_m=x) \\
\overset{(2.4)}{=}&\frac{\mu_x^{(m)}p_{x,j_1}p_{j_2,j_3}\cdots p_{j_{n-1},j_n}}{\mu_x^{(m)}}\times\mathbb{P}(\tau=m,X_m=x) \\
=&\delta_{xx}p_{x,j_1}p_{j_2,j_3}\cdots p_{j_{n-1},j_n}\times\mathbb{P}(\tau=m,X_m=x)
\end{aligned}$$

其中, $\mu_x^{(m)} := P(X_m=x)$. 综上, 关于 $m \geqslant 0$ 求和（注意到 $\{\tau<\infty\}=\sum_{m=0}^{\infty}\{\tau=m\}$）, 再两边同除以 $P(\tau<\infty,X_\tau=x)$, 得：当 $j_0=x$, 对任意的 $n \geqslant 0$, 有

$$\mathbb{P}(X_{\tau+0}=j_0,\cdots,X_{\tau+n}=j_n|\tau<\infty,X_\tau=x)=\delta_{xj_0}\prod_{k=0}^{n-1}p_{j_kj_{k+1}}.$$

(Step 3) 综上, 链的 $(X_{\tau+n})_{n\geqslant 0}$ 的有限维分布为

$$\mathbb{P}(X_{\tau+0}=j_0,\cdots,X_{\tau+n}=j_n|\tau<\infty,X_\tau=x)=\delta_{xj_0}\prod_{k=0}^{n-1}p_{j_kj_{k+1}}.$$

即有 $(X_{\tau+n})_{n\geqslant 0}\sim \mathrm{Markov}(\delta_x,P)$ under $\mathbb{P}(\cdot|\tau<\infty,X_\tau=x)$.
(2) 作业.

Problem 8 (作业 5-3)
$\forall J\subseteq\mathbb{N}_0, \#J<\infty$, 有

$$\sigma(X_{\tau+n},n\in J)\perp\!\!\!\perp_{\{\tau<\infty,X_\tau=x\}}\mathcal{F}_\tau$$

注：

1. $(X_{\tau+n})_{n\geqslant 0}$ 与 $X_0,\cdots,X_\tau$ 独立
2. $(X_{\tau+n})_{n\geqslant 0}\perp\!\!\!\perp\mathcal{F}_\tau$ under $\mathbb{P}(\cdot|\tau<\infty,X_\tau=x)$

用下面方法表述两次返回之间的等待时间 $S_y^{(k)}$

$T_y^{(0)} := 0, T_y^{(1)} := T_y$, 对于 $k \geqslant 2, T_y^{(k)} := \min\{n \geqslant T_{y+1}^{(k-1)} | X_n = y\}$

$$S_y^{(k)} = \begin{cases} T_y^{(k)} - T_y^{(k-1)} & \text{若} T_y^{(k-1)} < \infty \\ 0 & \text{否则} \end{cases}$$

注意到 $T_y^{(k)} = T_y^{(k-1)} + \min\{n \geqslant 1 | X_{T_y^{(k-1)}+n} = y\}$

$$\Rightarrow S_y^{(k)} = \min\{n \geqslant 1 | X_{T_y^{(k-1)}+n} = y\}, \text{if } T_y^{(k-1)} < \infty$$

即 $(X_{T_y^{(k-1)}+n})_{n \geqslant 0}$ 的首次回访时刻 $S_y^{(k)}$

Lemma 2.7

对 $k \geqslant 2$, 有 $\sigma(S_y^{(k)}) \perp\!\!\!\perp_{\{T_y^{(k-1)} < \infty\}} \mathcal{F}_{T_y^{(k-1)}}$, 且

$$\mathbb{P}(S_y^{(k)} < \infty | T_y^{(k-1)} < \infty) = \mathbb{P}(T_y^{(1)} < \infty | X_0 = y) =: \rho_{yy} \tag{2.37}$$

Corollary 2.4

对 $k \geqslant 0$ 有 $\rho_{yy}^{(k)} = \rho_{yy}^k$, 即

$$\mathbb{P}_y(N(y) \geqslant k) = \mathbb{P}_y(T_y^{(k)} < \infty) = \rho_{yy}^k$$

证明：第 k 次访问 y 的时刻有限，即第 $k-1$ 次访问 y 的时刻有限且时间间隔有限，即

$$\begin{aligned} \mathbb{P}_y(T_y^{(k)} < \infty) &= \mathbb{P}_y(S_y^{(k)} < \infty, T_y^{(k-1)} < \infty) \\ &= \mathbb{P}_y(S_y^{(k)} < \infty | T_y^{(k-1)} < \infty) \mathbb{P}_y(T_y^{(k-1)} < \infty) \end{aligned}$$

递归

$$\begin{aligned} \mathbb{P}_y(T_y^{(k)} < \infty) &= \mathbb{P}_y(S_y^{(k)} < \infty | T_y^{(k-1)} < \infty) \cdots \mathbb{P}_y(S_y^{(2)} < \infty | T_y^{(1)} < \infty) \mathbb{P}_y(T_y^{(1)} < \infty) \\ &\overset{(2.37)}{=} \rho_{yy}^k \quad \square \end{aligned}$$

2.5 类结构

2.5.1 状态 i 间的关系：可达与互通

Definition 2.16 (可达)

$i,j \in S$, 若 $\exists n \geqslant 0, \text{s.t. } p_{ij}(n) > 0$, 则称 i 可达 j, 记作 $i \to j$

注：$i \to i, p_{ii}(0) = 1 > 0$ 包括在内

Definition 2.17 (互通)

若 $i \to j, j \to i$ 称 i 与 j 互通, 记作 $i \leftrightarrow j$

Theorem 2.6

对不同的 i 与 j, 下面命题等价

1. $i \to j$
2. $0 < f_{ij} = \rho_{ij} = \mathbb{P}_i(T_j < \infty)$ [Durrett [2], Def 1.1]
3. $\exists$ 某些状态, $i_0 = i, i_1, \cdots, i_{n-1}, i_n = j, \text{s.t. } p_{i_0,i_1} \cdots p_{i_{n-1},i_n} > 0$
4. $\mathbb{P}_i(\exists n \geqslant 0, X_n = j) > 0$

Problem 9 (作业 6-1)

证明：Theorem 2.6命题的等价性, 即 $1 \Leftrightarrow 2, 1 \Leftrightarrow 3, 1 \Leftrightarrow 4$

Problem 10 (作业 6-2)

定义 first hitting time

$$H^j := \min\{n \geqslant 0 | X_n = j\}$$

1. 证明：H^j 是一个关于 $(X_n)_{n\geqslant 0}$ 的停时
2. 利用 H^j 定义“可达”, 并且证明该新定义与原定义等价

Property 2.3 (Durrett [2], Lem1.4)

若 $i \to j, j \to k$, 则 $i \to k$

证明. $i \to j \iff \exists n \geqslant 0, \text{s.t. } p_{ij}(n) > 0$

$j \to k \iff \exists m \geqslant 0, \text{s.t. } p_{jk}(m) > 0$

$$p_{ik}(n+m) \overset{\text{C-K}}{=} \sum_{r \in S} p_{ir}^{(n)} p_{rk}^{(m)} \geqslant p_{ij}^{(n)} p_{jk}^{(m)} > 0, \quad \therefore i \to k$$

□

Property 2.4

互通关系 ($\leftrightarrow$) 在 S 上是等价关系, 即

1. (自反的) $i \leftrightarrow i$
2. (对称的) $i \leftrightarrow j$, 则 $j \leftrightarrow i$
3. (传递的) $i \leftrightarrow j, j \leftrightarrow k$, 则 $i \leftrightarrow k$

2.5.2 常返与暂留是类性质

Lemma 2.8 (Durrett [2], Thm1.5&Lem1.6)

设 $i \to j, \rho_{ij} > 0$, 则

1. i 常返的 $\Rightarrow \rho_{ji} = 1(> 0 \Rightarrow j \to i)$
2. $\rho_{ji} < 1 \Rightarrow i$ 非常返/暂留的

注: 直观上 (2)$i \to j \overset{\text{prob}>0}{\nrightarrow} i$, 则 i 暂留

证明. (1), (2) 是逆否命题等价, 因此只需证一个即可. 下面证明 (2).
$i \neq j, \rho_{ji} < 1 \Rightarrow 0 < 1 - \rho_{ji} = 1 - \mathbb{P}_j(T_i < \infty) = \mathbb{P}_j(T_i = \infty)$
为了证 i 暂留, 即证 $\rho_{ii} < 1$, 即 $\mathbb{P}_i(T_i = \infty) > 0$
(Step 1) $\rho_{ij} > 0, i \to j \Rightarrow \exists k \geqslant 1, \text{s.t. } p_{ij}(k) > 0$

$$K := \min\{k \geqslant 1 | p_{ij}(k) > 0\}$$

由 C-K 方程(2.5), $\exists$ 与 i, j 不同的状态 $i_1, \cdots, i_{K-1}$, s.t.

$$p_{i,i_1} p_{i_1,i_2} \cdots p_{i_{K-1},j} > 0$$

注: 已知 $\mathbb{P}_j(T_i = \infty) > 0$, 要证 $\mathbb{P}_i(T_i = \infty) > 0$, 思路是将起始点从 i 挪到 j.

(Step 2)

$$
\begin{aligned}
\mathbb{P}_i(T_i=\infty) &= \mathbb{P}_i(\bigcap_{n\geqslant 1}\{X_n\neq i\}) \\
&\geqslant \mathbb{P}_i(\bigcap_{n=0}^{K-1}\{X_n=i_n\}, X_K=j, \bigcap_{n\geqslant K+1}\{X_n\neq i\}) \\
&= \mathbb{P}(\bigcap_{n=0}^{K-1}\{X_n=i_n\}, X_K=j, \bigcap_{n\geqslant K+1}\{X_n\neq i\})/\mathbb{P}(X_0=i) \\
&= \mathbb{P}(\bigcap_{n=0}^{K-1}\{X_n=i_n\}|X_K=j)\cdot\mathbb{P}(\bigcap_{n\geqslant K+1}\{X_n\neq i\}|X_K=j,\bigcap_{k=0}^{K-1}\{X_k=i_k\})\cdot\mathbb{P}(X_K=j)/\mathbb{P}(X_0=i) \\
&\xlongequal{\text{Markov}} \mathbb{P}(\bigcap_{n=0}^{K-1}\{X_n=i_n\}|X_K=j)\cdot\mathbb{P}(\bigcap_{n\geqslant K+1}\{X_n\neq i\}|X_K=j)\cdot\mathbb{P}(X_K=j)/\mathbb{P}(X_0=i) \\
&= \underbrace{\mathbb{P}(\bigcap_{n=0}^{K-1}\{X_n=i_n\}, X_K=j)}_{\text{有限维分布}}\cdot\mathbb{P}(\bigcap_{n\geqslant K+1}\{X_n\neq i\}|X_K=j)/\mathbb{P}(X_0=i) \\
&= \frac{\mu_i^{(0)}\cdot p_{i,i_1}\cdot p_{i_1,i_2}\cdots p_{i_{K-1},j}}{\mu_i^{(0)}}\cdot\lim_{m\to\infty}\mathbb{P}(\bigcap_{n=K+1}^{K+m}\{X_n\neq i\}|X_K=j)
\end{aligned}
$$

由 (Step 1), $p_{i,i_1}\cdot p_{i_1,i_2}\cdots p_{i_{K-1},j}>0$, 因此只需证明后面概率的极限也为正，即可证明 $\mathbb{P}_i(T_i=\infty)>0$.
假设 $(X_n)_{n\geqslant 0}\sim \text{Markov}(\mu,P)$
$\tau_1=0,\tau_2=K$ 为关于 $(X_n)_{n\geqslant 0}$ 的停时, 则由 SMP 知

1. 在 $\tilde{\mathbb{P}}(\cdot):=\mathbb{P}(\cdot|X_K=j)$ 下, $(X_{K+n})_{n\geqslant 0}\sim\text{Markov}(\delta_j,P)$
2. 在 $\mathbb{P}_j(\cdot):=\mathbb{P}(\cdot|X_0=j)$ 下, $(X_n)_{n\geqslant 0}\sim\text{Markov}(\delta_j,P)$

发现在测度 $\tilde{\mathbb{P}}(\cdot)$ 下的 $(X_{K+n})_{n\geqslant 0}$, 与测度 $\mathbb{P}_j(\cdot)$ 下的 $(X_n)_{n\geqslant 0}$ 遵循同样的有限维分布

$$
\begin{aligned}
&\mathbb{P}(\bigcap_{n=K+1}^{K+m}\{X_n\neq i\}|X_K=j) \\
=&\tilde{\mathbb{P}}(X_{K+n}\neq i, 1\leqslant n\leqslant m) \\
\overset{\text{SMP}}{=}&\mathbb{P}_j(X_n\neq i,1\leqslant n\leqslant m)\xrightarrow{m\to\infty}\mathbb{P}_j\left(\bigcap_{n\geqslant 1}\{X_n\neq i\}\right)=\mathbb{P}_j(T_i=\infty)=1-\rho_{ji}>0
\end{aligned}
$$

因此 $\mathbb{P}_i(T_i=\infty)>0$ □

Corollary 2.5
$i\to j$, i 常返 $\Rightarrow j$ 常返

证明. $i\neq j, i\to j$, i 常返, 由 Lemma 2.8, 知 $\rho_{ji}=1>0$, 所以 $j\to i, i\leftrightarrow j$
$\exists m,n\geqslant 0,\text{s.t. } p_{ij}(m)>0, p_{ji}(n)>0$
$\forall r\geqslant 0, p_{jj}(n+r+m)\overset{\text{C-K}}{\geqslant} p_{ji}(n)p_{ii}(r)p_{ij}(m)$

两边同时求和

$$\sum_{r\geqslant 0} p_{jj}(n+r+m) \geqslant p_{ji}(n)\left(\sum_{r\geqslant 0} p_{ii}(r)\right)p_{ij}(m) = \infty$$

其中 $p_{ji}(n) > 0, p_{ij}(m) > 0, \sum_{r\geqslant 0} p_{ii}(r) = \infty$($i$ 常返)

$$\therefore \sum_{r\geqslant 0} p_{jj}(r) = \infty \quad \Rightarrow j\text{常返}$$

□

Corollary 2.6

若 $i \leftrightarrow j$, 则

$$i\text{常返} \iff j\text{常返}$$

Definition 2.18 (集合的不可约)

$C \subseteq S, \forall i,j \in C$, 有 $i \leftrightarrow j$, 则称 C 不可约

Definition 2.19 (链的不可约)

若 S 不可约, 则称链不可约

Theorem 2.7

若 $C \subseteq S$ 不可约, 则 C 中状态要么全是常返的, 要么全是暂留的

2.5.3 状态空间分解

Definition 2.20 (闭集)

$C \subseteq S$, 若 $i \in C, j \notin C \Rightarrow p_{ij} = 0$, 则称 C 为闭集

Problem 11 (作业 6-3)

证明 $C \subseteq S$ 闭集等价于

$$i \in C, i \to j \Rightarrow j \in C \quad (j \notin C \Rightarrow i \nrightarrow j)$$

Example 2.6

若 $\{i\}$ 闭, 则 $\forall j \neq i, p_{ij} = 0 \Leftrightarrow p_{ii} = 1$, 称 i 为吸收态

Theorem 2.8

每一个有限的不可约闭集都是常返的

证明之前先介绍一个 Lemma

Lemma 2.9

每一个有限闭集中都至少有一个常返态

证明. (反证法) 设 C 为有限闭集, 非常返的

$\forall i \in C \Rightarrow i$ 暂留 $\Rightarrow \sum_{n\geqslant 1} p_{ji}(n) < \infty, \forall j \in S$. (由(2.16))

$$\infty \overset{C\text{有限}}{>} \sum_{i\in C}\sum_{n\geqslant 1} p_{ji}^{(n)} = \sum_{n\geqslant 1}\sum_{i\in C} p_{ji}^{(n)}$$

这里不是 $i \in S$ 而是 $i \in C$, 所以还要考虑 $i \in C^c$

$\forall i \in C^c, j \in C \overset{C\text{闭}}{\Rightarrow} j \nrightarrow i \Rightarrow \forall n \geqslant 0, p_{ji}(n) = 0$

$$\infty > \sum_{n\geqslant 1}\sum_{i\in C} p_{ji}^{(n)} = \sum_{n\geqslant 1}\left(\sum_{i\in S} p_{ji}^{(n)}\right) = \sum_{n\geqslant 1} 1 = \infty$$

矛盾 $\square$

在一个不可约闭集 C 中, 至少有一个常返态 $i \in C$, 由不可约定义和 Lemma 2.6, $\forall j \in C, j \leftrightarrow i, j$ 常返

Corollary 2.7

状态空间 S 有限, 则 S 中必存在一个常返态

Theorem 2.9 (分解定理)

状态空间 S 可分解为

$$S = T + R_1 + R_2 + \cdots$$

其中 T 中所有状态非常返, R_r 为常返不可约闭集

证明. (Step 1) 首先把所有暂留态拿出来

$$T := \{j \in S | j\text{暂留}\}$$

(Step 2) $i_1 \in S\backslash T \neq \varnothing$(若 $S\backslash T = \varnothing$, 则在 Step 1 结束)

i_1 常返, 定义 $R_1 = \{j \in S | j \leftrightarrow i_1\}$

$R_1 \subseteq S\backslash T$, R_1 常返互通类

(Step 3) $i_2 \in S\backslash(T\cup R_1), R_2 = \{j \in S | i_2 \leftrightarrow j\} \Rightarrow R_2 \subseteq S\backslash(T\cup R_1)$

若 $j \in R_2, j \in R_1 \Rightarrow j \leftrightarrow i_2, j \leftrightarrow i_1 \Rightarrow i_1 \leftrightarrow i_2 \Rightarrow i_2 \in R_1$, 矛盾

(Step 4) 迭代 $\square$

2.6 平稳分布与特殊例子

$$(X_n)_{n\geqslant 0} \sim \text{Markov}(\mu^{(0)}, P)$$

初始分布 $\mu^{(0)} = (\mu_i^{(0)})_{i\in S}$, 其中 $\mu_i^{(0)} = \mathbb{P}(X_0 = i)$
在 n 时刻的分布, $\mu^{(n)} = (\mu_i^{(n)})_{i\in S}$, 其中 $\mu_i^{(n)} = \mathbb{P}(X_n = i)$

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

Definition 2.21 (平稳分布)
称概率分布 $\pi = (\pi_i)_{i\in S}$ 是转移矩阵 P 的平稳分布, 若

$$\pi P = \pi \tag{2.38}$$

注: $\mu^{(n+1)} = \pi P^{n+1} = \pi P \cdot P^n = \pi P^n = \pi P = \pi = \mu^{(0)}$

Problem 12 (作业 6-4)
设 $(X_n) \sim \text{Markov}(\pi, P)$, π 是 P 的平稳分布, 证明: 对固定 $m \geqslant 0$, 有 $(X_{m+n})_{n\geqslant 0} \sim \text{Markov}(\pi, P)$

2.6.1 双随机链 (Doubly Stochastic Chain)

回顾随机矩阵定义2.1, 现在由行和为 1, 拓展到列和也为 1.

Definition 2.22
称转移矩阵 $(p_{xy})_{x,y\in S}$ 是双随机的, 若 $\sum_{x\in S} p_{xy} = 1$

Theorem 2.10
设 $P = (p_{xy})_{x,y\in S}$ 为具有 $N < \infty$ 个状态的马氏链的转移概率矩阵, 且有均匀分布 $\pi_x = \frac{1}{N}, x \in S$ 则下面两个命题等价:

1. π_x 是 P 的平稳分布
2. P 双随机

证明. $\forall y \in S, \pi_y = \frac{1}{N}$

$$(\pi P)_y = \sum_{x\in S} \pi_x p_{xy} = \frac{1}{N} \sum_{x\in S} p_{xy}$$

1. P 双随机, $(\pi P)_y = 1/N = \pi_y, (\forall y \in S) \Rightarrow \pi P = \pi$
2. $(\pi P)_y = \pi_y, \forall y \in S, \frac{1}{N} \sum_{x\in S} p_{xy} = 1/N \Rightarrow \sum_{x\in S} p_{xy} = 1 \Rightarrow P$ 双随机

□

2.6.2 细致平衡条件 (Detailed Balance Condition)

Definition 2.23

称概率分布 π 满足 DBC, 若

$$\pi_x p_{xy} = \pi_y p_{yx} \quad (\forall x, y \in S) \tag{2.39}$$

注：DBC 是 $\pi P = \pi$ 的充分不必要条件.

证明. $\pi P = \pi \iff (\pi P)_y = \pi_y, \forall y \iff \sum_{x\in S} \pi_x p_{xy} = \pi_y, \forall y$. 由于 $\pi_y = \pi_y(\sum_{x\in S} p_{yx}) = \sum_{x\in S} \pi_y p_{yx}$,

$$\pi P = \pi \iff \sum_{x\in S} \pi_x p_{xy} = \sum_{x\in S} \pi_y p_{yx} \tag{2.40}$$

由 (2.39) 可以推出 (2.40) 右边等式, 但反之不然. □

Example 2.7 (DBC 反例)

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

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.1 & 0.6 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}$$

P 双随机, $\pi = (1/3, 1/3, 1/3)$ 是 P 的平稳分布

Claim 2.4. π 不满足 *DBC*

反证：π 满足 DBC$\Rightarrow \pi_x p_{xy} = \pi_y p_{yx}$ 与 $p_{12} = 0.5 \neq p_{21} = 0.3$ 矛盾 □

Example 2.8 (生灭链)

状态空间 $S = \{l, l+1, \cdots, r-1, r\} \subseteq N_0$, 设 P 满足

1. 一步转移不超过 1, 当 $|x-y| \geqslant 2$ 时, $p_{xy} = 0$
2. $p_{x,x+1} = p_x (\forall x < r)$
3. $p_{x,x-1} = q_x (\forall x > l)$
4. $p_{x,x} = 1 - p_x - q_x (\forall x \in S)$

求出 P 的满足 DBC 条件的平稳分布 π, rf.(2.39).

1. $|x-y| \geqslant 2$ 时, $p_{xy} = p_{yx} = 0$
2. $x = y$ 时, $p_{xy} = p_{yx}, \pi_x = \pi_y$
3. $y = x+1$ 时, $(x < r), \pi_x p_{x,x+1} = \pi_{x+1} p_{x+1,x}$

$$\pi_{x+1} = \frac{\pi_x p_{x,x+1}}{p_{x+1,x}} = \pi_x \frac{p_x}{q_{x+1}}$$

$$\pi_{l+n} = \overbrace{\underbrace{\pi_l \frac{p_l}{q_{l+1}}}_{\pi_{l+1}} \frac{p_{l+1}}{q_{l+2}}}^{\pi_{l+2}} \cdots \frac{p_{l+n-1}}{q_{l+n}}$$

令

$$a_0=1, a_1=\frac{p_l}{q_{l+1}}, a_2=\frac{p_lp_{l+1}}{q_{l+1}q_{l+2}}, \cdots, a_{r-l}=\frac{p_lp_{l+1}\cdots p_{r-1}}{q_{l+1}q_{l+2}\cdots q_r}$$

则 $\pi=(\pi_l a_0, \pi_l a_1, \cdots, \pi_l a_{r-l})$. 又 $\sum_{x\in S}\pi_x=1$, 则

$$\pi_l\sum_{0\leqslant n\leqslant r-l}a_n=1\Rightarrow\pi_l=\frac{1}{\sum_{0\leqslant n\leqslant r-l}a_n}$$

记 $a:=\sum_{0\leqslant n\leqslant r-l}a_n$, 则 $\pi=(a_0/a, a_1/a, \cdots, a_{r-l}/a)$ □

2.6.3 可逆性

Theorem 2.11

设 $(X_n)_{n\geqslant 0}\sim\text{Markov}(\pi,P)$, 其中 π 是 P 的平稳分布. 固定 n, 令 $Y_m:=X_{n-m}(0\leqslant m\leqslant n)$, 则

$$(Y_m)_{0\leqslant m\leqslant n}\sim\text{Markov}(\pi,\hat{P})$$

其中 $\hat{P}=(\hat{p}_{ij})_{i,j\in S}$

$$\hat{p}_{ij}=\frac{\pi_j p_{ji}}{\pi_i}$$

这里 $\hat{p}_{ij}$ 称为对偶 (dual) 转移概率

证明. 先验证 $(Y_m)_{0\leqslant m\leqslant n}\sim\text{Markov}(\pi,\hat{P})$, 用定义/有限维分布. 这里用定义验证.
(Step 1) 验证 $\hat{P}$ 是随机矩阵

1. 元素非负 $\hat{p}_{ij}\geqslant 0$
2. $\sum_{j\in S}\hat{p}_{ij}=\sum_{j\in S}(\pi_j p_{ji}/\pi_i)=\pi_i/\pi_i=1$, rf.(2.38)

(Step 2) 验证初始分布. 由 (2.38) , 初始分布 $Y_0=X_n\sim\pi$
(Step 3) 验证马氏性

$$\begin{aligned}\mathbb{P}(Y_{m+1}=i_{m+1}|Y_m=i_m,\cdots,Y_1=i_1,Y_0=i_0)&=\frac{\mathbb{P}(Y_{m+1}=i_{m+1},\cdots,Y_0=i_0)}{\mathbb{P}(Y_m=i_m,\cdots,Y_0=i_0)}\\&=\frac{\mathbb{P}(X_{n-m-1}=i_{m+1},\cdots,X_{n-1}=i_1,X_n=i_0)}{\mathbb{P}(X_{n-m}=i_m,\cdots,X_n=i_0)}\\&=\frac{\pi_{i_{m+1}}P_{i_{m+1},i_m}\cdots P_{i_1,i_0}}{\pi_{i_m}P_{i_m,i_{m-1}}\cdots P_{i_1,i_0}}\\&=\frac{\pi_{i_{m+1}}P_{i_{m+1},i_m}}{\pi_{i_m}}=\hat{p}_{i_m,i_{m+1}}\end{aligned}$$

□

Corollary 2.8 (可逆性)

若 P 的平稳分布为 π, 满足 DBC 条件(2.39), 则 $\hat{P}=P$. 即原来的链 $\overset{(d)}{=}$ 逆向链（记号 $\overset{(d)}{=}$ 表示同分布）

证明. $\hat{p}_{ij}=(\pi_j p_{ji})/\pi_i\overset{\text{DBC}}{=}(\pi_i p_{ij})/\pi_i=p_{ij}$ □

2.6.4 求 P 的平稳分布（若唯一）

$$\begin{cases} \pi P = \pi \\ \sum_x \pi_x = 1 (\pi_x \geqslant 0, \forall x) \end{cases} \Rightarrow \begin{cases} \pi (P - \mathbb{I}) = 0 \\ \sum_x \pi_x = 1 \end{cases}$$

Example 2.9

例 [Durrett, 1.19]

Example 1.19 (Brand Preference (Continuation of 1.5)).

	1	**2**	**3**
1	.8	.1	.1
2	.2	.6	.2
3	.3	.3	.4

$$P - \mathbb{I} = \begin{bmatrix} -0.2 & 1 & 1 \\ 0.2 & -0.4 & 0.2 \\ 0.3 & 0.3 & -0.6 \end{bmatrix}$$

$$\begin{cases} \pi P = \pi \\ \sum_x \pi_x = 1 (\pi_x \geqslant 0, \forall x) \end{cases} \Rightarrow \begin{cases} -0.2\pi_1 + 0.2\pi_2 + 0.3\pi_3 = 0 \\ \pi_1 - 0.4\pi_2 + 0.3\pi_3 = 0 \\ \pi_1 + 0.2\pi_2 - 0.6\pi_3 = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

前三个等式是线性相关的，删去一个等式

$$A = \begin{bmatrix} -0.2 & 1 & 1 \\ 0.2 & -0.4 & 1 \\ 0.3 & 0.3 & 1 \end{bmatrix}, \quad b = [0, 0, 1]$$

$\pi A = b \Rightarrow \pi = bA^{-1}$，即 A^{-1} 的最后一行

2.7 极限行为与平稳分布的存在唯一性

研究 $\lim\limits_{n\to\infty} p_{ij}^{(n)}$

1. 由 (2.16), j 暂留 $\Rightarrow \sum\limits_{n\geqslant 0} p_{ij}^{(n)} < \infty, \forall i \in S \Rightarrow \lim\limits_{n\to\infty} p_{ij}^{(n)} = 0, \forall i \in S$. 下面可以把注意力放在常返上
2. $\lim\limits_{n\to\infty} p_{ij}^{(n)}$ 不存在的反例

$$S = \{1,2\}, P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, P^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, P^{2n+1} = P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$p_{ij}^{(2n)} \neq p_{ij}^{(2n+1)}, \forall i,j \in S$, 所以 $p_{ij}^{(n)}$ 不收敛

Definition 2.24 (周期)

令 $I_x := \{n \geqslant 1 | P_{xx}^{(n)} > 0\}$, 定义 x 的周期 $d(x) = \gcd(I_x)$

1. $d(x) > 1$, 称 x 周期的
2. $d(x) = 1$, 称 x 非周期的
3. $I_x = \varnothing$, 称 x 周期为 ∞

注: gcd 为 greatest common divisor 最大公因数.

Definition 2.25

称链是周期的, 若所有状态是周期的

Theorem 2.12 (收敛定理)

马氏链不可约, 非周期, 且存在平稳分布 π, 则

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \quad (\forall i, j \in S)$$

注：找到周期不是件容易的事, 我们通常讨论非周期的链

Problem 13 (作用 7-1)

设 S 有限, $\exists i \in S$, s.t. $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j (\forall j \in S)$. 证明：$\pi = (\pi_j)_{j\in S}$ 是 $P = (p_{ij})_{i,j\in S}$ 的平稳分布

Theorem 2.13 (渐进频率)
马氏链不可约, 常返, 则
$$\lim_{n\to\infty}\frac{N_n(y)}{n}=\frac{1}{\mathbb{E}_y T_y} \tag{2.41}$$
注:

1. $N_n(y)=\sum_{k=1}^{n}\mathbb{I}_{\{X_k=y\}}$($n$ 时刻前, 访问 y 的总次数)
2. 考虑 $\frac{N_n(y)}{n}$, 表示 n 时刻前访问 y 的频率/时间比例, 因此 $\lim_{n\to\infty}\frac{N_n(y)}{n}$ 为在状态 y 上花费的时间比例的极限
3. $\mathbb{E}_y T_y=\begin{cases}<\infty & \text{y正常返}\\ \infty & \text{y暂留/零常返}\end{cases}$, rf. (Def 2.10).

证明. Durrett (3ed), Thm 1.20, p47. □

Theorem 2.14
马氏链不可约

1. (平稳分布唯一性, Durrett, Thm 1.21) 若平稳分布存在, 则
$$\pi_y=\frac{1}{\mathbb{E}_y T_y} \tag{2.42}$$
则 π 唯一
2. (平稳测度存在性) 若马氏链常返, 则 $\exists$ 平稳测度, $\mu=(\mu_x)_{x\in S}$, 且 $\mu_x>0,\forall x$. 令 $T_x=\min\{n\geqslant 1|X_n=x\}$.
$$\mu_x(y)=\sum_{n=0}^{\infty}\mathbb{P}_x(X_n=y,T_x>n) \tag{2.43}$$

注: $\mu=(\mu_x)_{x\in S}$ 是一个平稳测度, 若

1. (测度) $\mu_x\geqslant 0,\forall x\in S$
2. $\mu P=\mu$

证明. 1. (2.42): Durrett (3ed), Thm 1.21, p47.
2. (2.43): Durrett (3ed), Thm 1.24, p48.

□

相对于上面的大定理, 下面的推论对我们更有用

Corollary 2.9
马氏链具有有限状态, 不可约, 则

1. 存在唯一平稳分布 $\pi=(\pi_x)_{x\in S}$, 且 $\pi_x=\frac{1}{\mathbb{E}_x T_x}>0,\forall x\in S$
2. $\lim_{n\to\infty}\frac{N_n(y)}{n}=\frac{1}{\mathbb{E}_x T_x}=\pi_x$

证明. 1. S 有限不可约, 闭集 $\Rightarrow$ 不可约, 常返, rf. (Thm 2.8).

   (a) 由 Thm 2.14 (2) 知, 存在 $\mu = (\mu_x)_{x\in S}, \mu_x \geqslant 0, \mu P = \mu$. 令 $\pi_x = \dfrac{\mu_x}{\sum_{x\in S}\mu_x}$(正则化 μ), $\pi_x > 0$, 且

   $$\pi P = \frac{1}{\sum_{x\in S}\mu_x}\mu P = \frac{1}{\sum_{x\in S}\mu_x}\mu = \pi$$

   (b) 由 Thm 2.14 (1) 知, π 唯一且 $\pi_x = \dfrac{1}{\mathbb{E}_x T_x}$

2. 由 Thm 2.13

□

2.8 首达时及其应用

Definition 2.26 (首达时)

首达时 (first hitting time) 定义为

$$V_A := \min\{n \geqslant 0 | X_n \in A\} \tag{2.44}$$

注：前面提到的首次回访时间 (first passage time) 是要求 $n \geqslant 1$, rf. (2.18).

2.8.1 击中概率 (hitting time) 与离出分布

Definition 2.27 (击中概率)

击中概率定义为

$$h_x^A := \mathbb{P}_x(V_A < \infty) \tag{2.45}$$

特别地，A 为闭集，称 h_x^A 为吸收概率

下面介绍 h_x^A 的一个性质

Lemma 2.10

$h^A := (h_x^A)_{x\in S}$ 满足下列方程

$$\begin{cases} h_x^A = 1 & x \in A \\ h_x^A = \sum_y p_{xy} h_y^A & x \notin A \end{cases} \tag{2.46}$$

其中 $x \notin A$ 的情况对应卷积 $f(x) = (f * g)(x) = \sum_{y\in s} f(y)g(x-y)$.

击中概率是上述方程的一个解，之后我们将验证其唯一性.

证明. $x \in A \Rightarrow V_A = 0$, 所以 $h_x^A = \mathbb{P}_x(V_A < \infty) = 1$.
$x \notin A \Rightarrow V_A \geqslant 1$, 考虑一步转移情况 (one step reasoning) $\leftarrow$ 证明思想

$$h_x^A = \sum_{y\in S} \mathbb{P}_x(V_A < \infty, X_1 = y) = \sum_{y\in S} \mathbb{P}(V_A < \infty | X_1 = y, X_0 = x)\mathbb{P}(X_1 = y | X_0 = x)$$

Claim 2.5. $\mathbb{P}(V_A < \infty | X_1 = y, X_0 = x) = h_y^A, \forall y \in S, x \notin A$

利用马氏性,

$$\begin{aligned} \mathbb{P}(V_A < \infty | X_1 = y, X_0 = x) &\xlongequal{x\notin A} \mathbb{P}(\bigcup_{n\geqslant 1}\{X_n \in A\} | X_1 = y, X_0 = x) \\ &\xlongequal{\text{Markov}} \mathbb{P}(\bigcup_{n\geqslant 1}\{X_n \in A\} | X_1 = y) \\ &\xlongequal{\text{SMP}} \mathbb{P}_y(\bigcup_{n\geqslant 0}\{X_n \in A\}) = \mathbb{P}_y(V_A < \infty) = h_y^A \end{aligned}$$

□

Example 2.10

$a,b\in S, V_a:=V_{\{a\}}, V_b:=V_{\{b\}}$, 考虑 $h(x)=\mathbb{P}_x(V_a<V_b)$, 则 $h=(h(x))_{x\in S}$ 满足下列方程

$$\begin{cases} h(a)=1, h(b)=0 \\ h(x)=\sum_y p_{xy}h(y) \quad x\neq a,b \end{cases}$$

证明. (和上述引理证明过程一样) 使用一步展开方法,

$$h(x)=\mathbb{P}_x(V_a<V_b)=\sum_{y\in S}\mathbb{P}_x(V_a<V_b|X_1=y)\mathbb{P}_x(X_1=y)$$

只需证 $\mathbb{P}_x(V_a<V_b|X_1=y)=h(y), \forall x\neq a,b,y\in S, \to V_a\geqslant 1$, 就满足 $h(x)=\sum_{y\in S}p_{xy}h(y)$.

$$\begin{aligned}
\text{LHS} &= \mathbb{P}_x(1\leqslant V_a<\infty, V_a<V_b|X_1=y) \\
&\xlongequal{x\neq a,b} \mathbb{P}_x\left(\bigcup_{m\geqslant 1}\left\{\{X_m=a\}\cap\bigcap_{1\leqslant k\leqslant m}\{X_k\neq a,b\}\right\}\middle|X_1=y\right) \\
&= \mathbb{P}_x\left(\sum_{m\geqslant 1}\left\{\{X_m=a\}\cap\bigcap_{1\leqslant k\leqslant m}\{X_k\neq a,b\}\right\}\middle|X_1=y\right) \\
&= \sum_{m\geqslant 1}\mathbb{P}_x\left(\{X_m=a\}\cap\bigcap_{1\leqslant k\leqslant m}\{X_k\neq a,b\}\middle|X_1=y\right) \\
&\xlongequal{\text{Markov}} \sum_{m\geqslant 1}\mathbb{P}(V_a=m, V_a<V_b|X_1=y) \\
&\xlongequal{\text{SMP}} \sum_{m\geqslant 1}\mathbb{P}_y(V_a=m, V_a<V_b) \\
&= \mathbb{P}_y(V_a<V_b)=h(y)
\end{aligned}$$

□

Theorem 2.15

$A,B\subseteq S, A\cap B=\varnothing$, 令 $C=S-A\cup B$. 若 C 有限, $\mathbb{P}_x(V_A\wedge V_B<\infty)>0, \forall x\in C$, 则方程

$$\begin{cases} h(x)=1 & x\in A \\ h(x)=\sum_{y\in S}p_{xy}h(y) & x\in C \\ h(x)=0 & x\in B \end{cases} \tag{2.47}$$

存在唯一非负解 $h(x)=\mathbb{P}_x(V_A<V_B), \forall x\in S$ (不证明)

注:

1. $\wedge$ 是取小符号, $V_A\wedge V_B:=\min\{V_A,V_B\}$
2. $\mathbb{P}_x(V_a\wedge V_b<\infty)>0 \iff x\to a$ 或$x\to b$
3. $A\cap B=\varnothing$ 时, $V_A\wedge V_B=V_{A\cup B}$

Problem 14 (作业 8-1)

证明：

1. $\mathbb{P}_x(V_a \wedge V_b < \infty) > 0 \iff x \to a$ 或$x \to b$
2. $A \cap B = \varnothing$ 时, $V_A \wedge V_B = V_{A\cup B}$

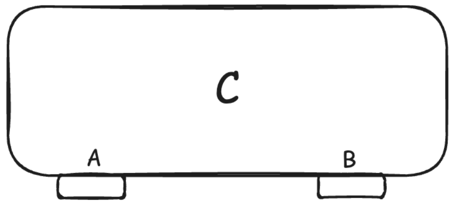


图 6: An example

$\tau_C = \min\{n \geqslant 0 | X_n \notin C\}$ 为首次离出时刻/逃逸时刻.
$\tau = \min\{n \geqslant 0 | X_n \in A \cup B\}, A \cap B = \varnothing$.
$\mathbb{P}_x(X_{\tau_C} \in A) = \mathbb{P}_x(V_A < V_B)$ 为逃逸概率/离出分布.
特别的, $A = \{a\}, B = \{b\}$, a, b 为吸收态, $x \to a (x \neq a)$, $a \nrightarrow x, \rho_{ax} = 0 < 1$, 由 Lem 2.8, x 暂留.
$\mathbb{P}_x(V_a < V_b) = \mathbb{P}_x(V_a < \infty)$ 为吸收概率.
$\tau_C = V_{A\cup B} = V_A \wedge V_B$
$\therefore \mathbb{P}_x(V_A \wedge V_B < \infty) = \mathbb{P}_x(\tau_C < \infty) > 0$.

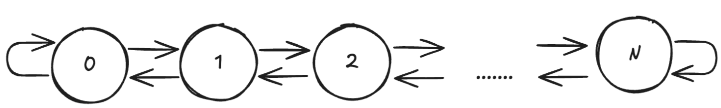
Example 2.11

X_n: 财富, $X_n = 0$ 或 N 时游戏结束, 问: 赌徒破产概率.

解.

$$\begin{cases} p(x, x+1) = p & 0 < x < N \\ p(x, x-1) = q = 1-p & 0 < x < N \\ p(0,0) = 1, p(N,N) = 1 & x = 0, x = N \end{cases}$$

$0, N$ 为吸收态. 令 $h(x) := \mathbb{P}_x(V_0 < \infty) = \mathbb{P}_x(V_0 < V_N)$, $x = 1, \cdots, N-1$. S 有限, 不可约, 则 $\forall 0 < x < N, x \to 0, x \to N$.



由 Thm 2.15, $h(x) = (h(x))_{x\in S}$ 是下列方程的唯一非负解.

$$\begin{cases} h(0) = 1, h(N) = 0 \\ h(x) = \sum_y p_{xy} h(y) = p(x, x+1)h(x+1) + p(x, x-1)h(x-1) & 0 < x < N \end{cases}$$

$h(x) = ph(x+1) + qh(x-1), 0 < x < N$. $p(h(x+1) - h(x)) = q(h(x) - h(x-1))$.

$$h(x+1) - h(x) = \frac{q}{p}(h(x) - h(x-1)) = \left(\frac{q}{p}\right)^x (h(1) - h(0)), \quad \forall 0 \leqslant x \leqslant N$$

$$\begin{aligned} h(x) &= h(0) + \sum_{k=0}^{x-1}(h(k+1) - h(k)) \\ &= h(0) + (h(1) - h(0)) \sum_{k=0}^{x-1} \left(\frac{q}{p}\right)^x, \quad \forall 0 \leqslant x \leqslant N \end{aligned}$$

令 $\theta = q/p$.

1. $\theta = 1$ 时, $h(x) = h(0) + x(h(1) - h(0))$, $0 = h(N) = 1 + N(h(1) - h(0))$, $h(1) - h(0) = -1/N$, $h(x) = 1 + (-1/N)x = (N-x)/N$
2. $\theta \neq 1$ 时, 同理.

$\square$

还有一种用线性代数方法求

$h[A] := (h(x))_{x \in A}$ (列), $P[C, C] = (p_{ij})_{i,j \in C}$, $h(x) = \mathbb{P}_x(V_A < V_B)$.

$$\begin{cases} h(x) = 1 & x \in A \\ h(x) = 0 & x \in B \\ h(x) = \sum_{y \in S} p_{xy} h(y) & x \in C \end{cases} \Rightarrow \begin{cases} h[A] = \mathbf{1} \\ h[B] = \mathbf{0} \\ h[C] = P[C, C]h[C] + P[C, A]\mathbf{1} \end{cases}$$

因为 A, B, C 互不相交,

$$\begin{aligned} h(x) &= \sum_{y \in S} p_{xy} h(y) \\ &= \sum_{y \in A} p_{xy} h(y) + \sum_{y \in B} p_{xy} h(y) + \sum_{y \in C} p_{xy} h(y) \\ &= \sum_{y \in A} p_{xy} + \sum_{y \in C} p_{xy} h(y) \\ &= (P[x, A])_{x \in C}\mathbf{1} + (P[x, C])_{x \in C}(h(y))_{y \in C} \\ &= P[C, A]\mathbf{1} + P[C, C]h[C], \forall x \in C \end{aligned}$$

$\Rightarrow h[C] = (I_{|C|} - P[C, C])^{-1} P[C, A]\mathbf{1}$

2.8.2 平均首达时与离出时刻

Definition 2.28

定义平均首达时.

$$k_x^A := \mathbb{E}[V_A | X_0 = x] = \begin{cases} \sum_{n \geqslant 0} n\mathbb{P}_x(V_A = n) & \mathbb{P}_x(V_A < \infty) = 1 \\ \infty & \mathbb{P}_x(V_A = \infty) > 0 \end{cases}$$

思考: $\mathbb{P}_x(V_A < \infty) = 1$ 与常返的区别是?

常返 $\mathbb{P}_x(T_x<\infty)=1$ 是 $\mathbb{P}_x(V_A<\infty)=1$ 的充分不必要条件.

Lemma 2.11

$k^A=(k_x^A)_{x\in S}$ 满足

$$\begin{cases} k_x^A=0 & x\in A \\ k_x^A=1+\sum_{y\in S}p_{xy}k_y^A & x\notin A \end{cases}$$

证明. 当 $x\in A$ 时, $V_A=0, k_x^A=0$. 当 $x\notin A$ 时,

$$\mathbb{E}_x V_A=\sum_{y\in S}\mathbb{E}_x[V_A\mathbb{I}_{\{X_1=y\}}]\xlongequal{\text{Cor }(1.2)}\sum_{y\in S}\mathbb{E}_x[V_A|X_1=y]\mathbb{P}_x(X_1=y)$$

Claim 2.6. $\mathbb{E}_x[V_A|X_1=y]=\mathbb{E}[V_A+1|X_0=y], \forall y\in S, x\notin A$.

证明. 见 HW Week8 作业 2. □

将 Claim 2.6 代回 $\mathbb{E}_x V_A$.

$$\begin{aligned}\sum_{y\in S}\mathbb{E}_x[V_A|X_1=y]\mathbb{P}_x(X_1=y)&=\sum_{y\in S}p_{xy}(1+\mathbb{E}[V_A|X_0=y])\\&=\sum_{y\in S}p_{xy}+\sum_{y\in S}p_{xy}k_y^A\\&=1+\sum_{y\in S}p_{xy}k_y^A\end{aligned}$$

□

Theorem 2.16

令 $C=S-A, A\subseteq S$. 若 C 有限, 且 $\forall x\in C, \mathbb{P}_x(V_A<\infty)$=1. 则

$$\begin{cases} g(x)=0 & x\in A \\ g(x)=1+\sum_{y\in C}p_{xy}g(y) & x\in C \end{cases} \tag{2.48}$$

存在唯一非负解 $g(x)=\mathbb{E}_x V_A$.

$g[C]=\mathbf{1}+P[C,C]g[C]$. $g[C]=(I_{|C|}-P[C,C])^{-1}\mathbf{1}\overset{\text{书上}}{=}(\mathbf{I}-\gamma)^{-1}\mathbf{1}$

继续 Gambler's Ruin 例题

离出时刻 $\tau_C:=\min\{n\geqslant 0|X_n\neq C\}=V_{C^c}$

离出分布 X_{τ_C} 的分布

1. $X_{\tau_C}\in C^c$, 故 $\{X_{\tau_C}=C^c\}=\Omega$
2. 令 $x\in C, A\subseteq C^c, \mathbb{P}_x(X_{\tau_C}\in A)=\mathbb{P}_x(V_A<V_{C^c})$

Example 2.12 (等待 HT 出现的时间)

X_n: $n+1$ 时刻硬币朝上的图案, $n \geqslant 0$, $S_x = \{H,T\}$. 令 $Y_n = (X_n, X_{n+1}), n \geqslant 0$, 考虑其为马氏链且 $S = \{HH, HT, TH, TT\}$.

$$\mathbf{P} = \begin{array}{c} \\ \mathbf{HH} \\ \mathbf{HT} \\ \mathbf{TH} \\ \mathbf{TT} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{HH} & \mathbf{HT} & \mathbf{TH} & \mathbf{TT} \end{array} \\ \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix} \end{array}$$

令 T_{HT} 为出现 HT 所需的硬币投掷数. 求 $\mathbb{E}T_{HT}$.

解. $A = \{HT\}, V_A := \min\{n \geqslant 0 | Y_n = (X_n, X_{n+1}) \in A\}$, $T_{HT} = V_A + 2$

(Step 1) 求 $\mathbb{E}_x V_A$. S 有限, 不可约. 由 Thm 2.16, $g(x) := \mathbb{E}_x V_A$ 是下列方程的唯一非负解.

$$\begin{cases} g[A] = 0 \\ g[C] = \mathbf{1}_{|C|} + P[C,C]g[C] \end{cases}$$

$$g[C] = (I_{|C|} - P[C,C])^{-1}\mathbf{1}_{|C|} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} \begin{array}{c} \mathbf{HH} \\ \mathbf{TH} \\ \mathbf{TT} \end{array}$$

(Step 2)

$$\begin{aligned} \mathbb{E}T_{HT} &= 2 + \mathbb{E}V_A \\ &= 2 + \sum_{x \in S} \mathbb{E}V_A \mathbb{I}_{\{Y_0 = x\}} \\ &\xlongequal{\text{Cor (1.2)}} 2 + \sum_{x \in S} \mathbb{E}[V_A | Y_0 = x]\mathbb{P}(Y_0 = x) \\ &= 2 + \sum_{x \in S} g(x)\mathbb{P}(X_0 = H, X_1 = T) \\ &= 2 + \frac{1}{4}(0 + 2 + 2 + 4) = 4 \end{aligned}$$

□

2.9 具有无限状态的马氏链

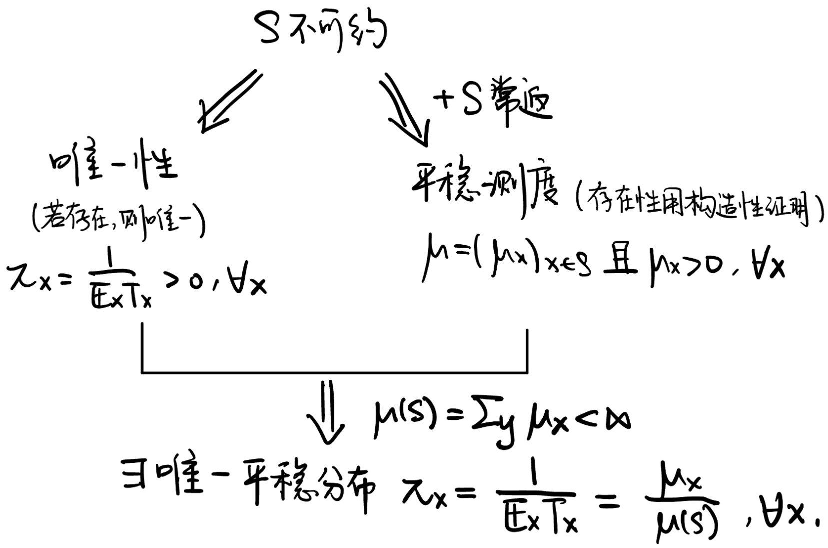


图 7: Summary

“构造性证明” 见 (2.43).

Claim 2.7. *S 不可约, 若存在平稳分布, 则 $\pi_x > 0, \forall x \in S$.*

证明. 设 $\exists x$, s.t. $\pi_x = 0$. $\pi = \pi P^n, \forall n \geqslant 0$. $\pi_x = \sum_y \pi_y p_{yx}^{(n)}, \forall n \geqslant 0$. $\Rightarrow \pi_y p_{yx}^{(n)} = 0, \forall y \in S, n \geqslant 0$. 又因 S 不可约, $\forall y \in S, y \to x$. 所以 $\exists n_y \geqslant 0$, s.t. $p_{yx}^{(n_y)} > 0$. $\Rightarrow \forall y \in S, \pi_y = 0$. 这与 $\sum_y \pi_y = 1$ 矛盾 □

Lemma 2.12

S 不可约, $\boxed{\text{若存在平稳测度 } \mu = (\mu_x)_{x\in S}, \mu_x > 0, \forall x}$ (即 “若 S 常返成立”), $\boxed{\mu(S) = \sum_{x\in S} \mu_x < \infty}$, 则存在唯一平稳分布

$$\pi_x = \frac{1}{\mathbb{E}_x T_x} = \frac{\mu_x}{\mu(S)} > 0, \forall x \tag{2.49}$$

注: $\pi_x = 1/(\mathbb{E}_x T_x) > 0, \forall x$ 为必要条件

$\Rightarrow \mathbb{E}_x T_x < \infty, \forall x \Rightarrow \forall x$, x 正常返 $\Rightarrow S$ 正常返. (这是必要条件, 那反过来是否充分? 能否推出框内条件?)

$i \leftrightarrow j$, 则由 Cor 2.6, i 正常返 $\iff$ j 正常返, i 零常返 $\iff$ j 零常返, i 非周期 $\iff$ j 非周期. $p_{jj} > 0 \Rightarrow d(j) = 1$.

Theorem 2.17

S 不可约, 则下列结论等价

1. 某个状态正常返
2. 存在平稳分布 π
3. 所有状态正常返

证明. (3) $\Rightarrow$ (1) 显然, (2) $\Rightarrow$ (3) 在上面注记中已证. 要证 (1) $\Rightarrow$ (2).
S 不可约, 存在 x 正常返 $\Rightarrow$ S 不可约, 常返. 由 Thm 2.14 及 (2.43), 存在平稳测度 $\mu_y = \sum_{n\geqslant 0}\mathbb{P}_x(T_x > n, X_n = y), \forall y \in S$. 由 Lem 2.12 框内条件知, 平稳测度要求 $\mu(S) < \infty$.

$$\begin{aligned}\mu(S) = \sum_{y\in S}\mu_y &= \sum_{n\geqslant 0}\sum_{y\in S}\mathbb{P}_x(T_x > n, X_n = y)\\ &= \sum_{n\geqslant 0}\mathbb{P}_x(T_x > n) = \mathbb{E}_x T_x < \infty\end{aligned}$$

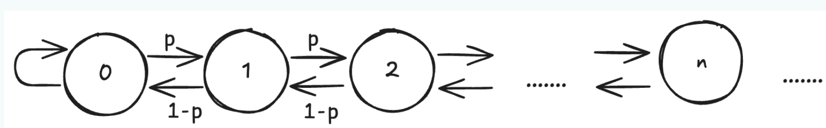
注: $T_x = \sum_{n\geqslant 0}\mathbb{I}_{\{T_x > n\}} = \sum_{n\geqslant 1}\mathbb{I}_{\{T_x \geqslant n\}}$.
存在平稳分布 $\pi_x = \mu_x/\mu(S) > 0, \forall x$. □

Corollary 2.10
对不可约链, 下列情况之一必发生

1. S 非常返, 则不存在平稳分布
2. S 常返, 则对 $\forall x \in S$, 存在平稳测度 $\mu^{(x)} = (\mu_y^{(x)})_{y\in S}$. (由 Thm 2.14 定义).
   (a) $\forall x, \mu^{(x)}(S) = \mathbb{E}_x T_x < \infty$, 则 S 正常返, 则存在唯一平稳分布
   (b) $\forall x, \mu^{(x)}(S) = \mathbb{E}_x T_x = \infty$, 则 S 零常返, 则不存在平稳分布

2.9.1 广义生灭链

Example 2.13 (带反射壁的随机游动 (Durrett 1.54))
质点在 $S = \{0, 1, 2, \cdots\}$ 移动, 规定



$p \in (0,1)$, X_n: 表示 n 时刻质点所在位置. 则

1. $p \in (0, 1/2)$ 时, 存在唯一平稳分布, S 正常返
2. $p > 1/2$ 时, S 非常返, 不存在平稳分布
3. $p = 1/2$ 时, S 零常返, 不存在平稳分布

证明. 1. (Step 1) S 不可约, 则平稳分布若存在则唯一
(Step 2) (存在性) 回顾 Exa 2.8, 与当前例子的区别是状态空间 S 是否有限. 设 π 满足 DBC, $\pi_x p_{xy} = \pi_y p_{yx}, \forall x, y$. $\pi_i p_{i,i+1} = \pi_{i+1}p_{i+1,i}, \forall i \geqslant 0$. $\pi_i p = \pi_{i+1}(1-p), \forall i \geqslant 0$.

$$\pi_{i+1} = \frac{p}{1-p}\pi_i = \left(\frac{p}{1-p}\right)^{i+1}\pi_0, \forall i \geqslant 0$$

$$\pi_i = \left(\frac{p}{1-p}\right)^{i}\pi_0, \forall i \geqslant 0 \tag{*}$$

又由 $\sum_i \pi_i = 1$, 有

$$1 = \pi_0 \sum_{i \geqslant 0} \left(\frac{p}{1-p}\right)^i = \pi_0 \frac{1}{1-\frac{p}{1-p}} \Rightarrow \pi_0 = \frac{1-2p}{1-p} \in (0,1)$$

代回 $(*)$, 得 $\pi_i > 0, \forall i$. 其中无穷级数 $\sum_{i \geqslant 0} \left(\frac{p}{1-p}\right)^i$ 当 $|p/(1-p)| < 1$ 时收敛, 即 $p < 1/2$ 时. $(\sum_{i=0}^{\infty} r^i = 1/(1-r), \forall |r| < 1)$

2. 只需证状态 0 暂留, 又因为 $0 \to 1$, Lem 2.8, 故只需证 $1 > \rho_{1,0} = \mathbb{P}_1(T_0 < \infty) = \mathbb{P}_1(V_0 < \infty)$. 考察 $\mathbb{P}_x(V_0 < \infty)$. 注意到

$$\{V_0 < \infty, X_0 = x \neq 0\} = \bigcup_{M \geqslant 0} \{V_0 = M, X_0 = x \neq 0\} = \bigcup_{M \geqslant 0} \bigcup_{N \geqslant x+M} \{V_0 < V_N, X_0 = x \neq 0\}$$

$X_0 = x \neq 0, \forall N \geqslant x + M + 10000000$, $M = V_0 < V_N < V_{N+1}$.
令 $h(x) := \mathbb{P}_x(V_0 < V_N)$, 则

$$\begin{cases} h(0) = 1, h(N) = 0 \\ h(x) = ph(x+1) + (1-p)h(x-1) \quad \forall x \neq 0, N \end{cases}$$

由 Exa 2.11, $\theta = (1-p)/p < 1$ 时, $h(x) = (\theta^x - \theta^N)/(1-\theta^N), \forall x \in S$.

3. 考虑一步转移情况.

$$\begin{aligned} \mathbb{E}_0 T_0 &= \mathbb{E}_0[T_0 \mathbb{I}_{\{X_1=0\}}] + \mathbb{E}_0[T_0 \mathbb{I}_{\{X_1=1\}}] \\ &= \mathbb{E}_0[T_0 | X_1 = 0] p_{0,0} + \mathbb{E}_0[T_0 | X_1 = 1] p_{0,1} \\ &\overset{\text{Markov}}{=\!=\!=} \mathbb{E}[V_0 + 1 | X_0 = 0] \frac{1}{2} + \mathbb{E}[V_0 + 1 | X_0 = 1] \frac{1}{2} \\ &= \frac{1}{2}(\mathbb{E}_0(1) + \mathbb{E}_1(1)) + \frac{1}{2} \mathbb{E}_1 V_0 \\ &= 1 + \frac{1}{2} \mathbb{E}_1 V_0 \end{aligned} \tag{*}$$

考察 $g(x) = \mathbb{E}_x V_0$.
(Step 1) 考察 $\mathbb{E}_x(V_0 \wedge V_N)$, 同 Exa 2.12 类似计算得到 $\mathbb{E}_x(V_0 \wedge V_N) = x(N-x)$.
(Step 2) $\mathbb{E}_1 V_0 \geqslant \mathbb{E}_1(V_0 \wedge V_N) = N - 1 \to +\infty, (N \to +\infty)$. 代回 $(*)$ 得 $\mathbb{E}_0 T_0 = +\infty$

□

3 泊松过程

3.1 指数分布，泊松分布

3.1.1 指数分布

Definition 3.1 (指数分布)
称随机变量 T 服从 "参数/速率 $\lambda(\lambda>0)$ 的指数分布", 记作 $T\sim \mathrm{EXP}(\lambda)$, 若分布

$$F_T(t)=\begin{cases}1-e^{-\lambda t} & t\geqslant 0\\ 0 & t<0\end{cases} \tag{3.1}$$

$$f_T(t)=\begin{cases}\lambda e^{-\lambda t} & t\geqslant 0\\ 0 & t<0\end{cases} \tag{3.2}$$

Property 3.1 (矩)
$T\sim \mathrm{EXP}(\lambda)$, 则

1. $\mathbb{E}T=1/\lambda$
2. $\mathbb{E}T^2=2/\lambda$
3. $\mathrm{Var}(T)=\mathbb{E}T^2-(\mathbb{E}T)^2=1/\lambda^2$

Property 3.2 (Scaling)
$T\sim \mathrm{EXP}(\lambda), S\sim \mathrm{EXP}(1)$, 则 $S/\lambda \overset{(d)}{=} T, S \overset{(d)}{=} \lambda T$.

Property 3.3 (无记忆性)
$T\sim \mathrm{EXP}(\lambda)$, 则 $\mathbb{P}(T>t+s|T>t)=\mathbb{P}(T>s)$
注: 等价于 $\bar{F}(t+s)=\bar{F}(t)\bar{F}(s)$, 其中 $\bar{F}(t):=1-F(t)$.

Property 3.4 (指数分布的排序)
$T_i\sim \mathrm{EXP}(\lambda_i)$, 独立. 令 $V=\min\{T_1,T_2,\cdots,T_n\}$, $I:=\min\{i|T_i=V\}$ (随机下标)

1. $V\sim \mathrm{EXP}\left(\sum_{k=1}^n\lambda_k\right)$, 即
$$\mathbb{P}(V>t)=\exp\left(-\sum_{k=1}^n\lambda_k\cdot t\right) \tag{3.3}$$
2. $\mathbb{P}(T_i=V)=\lambda_i/(\sum_{k=1}^n\lambda_k)$
3. $I \perp\!\!\!\perp V$

证明. (1) $\mathbb{P}(V>t)=\mathbb{P}(T_k>t,1\leqslant k\leqslant n)=\prod_{k=1}^n\mathbb{P}(T_k>t)=\prod_{k=1}^n e^{-\lambda_k t}=\exp(-\sum_{k=1}^n\lambda_k\cdot t),t\geqslant 0.$

(2) 设 $S\sim\mathrm{EXP}(\lambda),U\sim\mathrm{EXP}(\mu),S\perp\!\!\!\perp U$

$$
\begin{aligned}
\mathbb{P}(S=\min\{S,U\})&=\mathbb{P}(S\leqslant U)\\
&\xlongequal{S\perp\!\!\!\perp U}\int_0^\infty\mathbb{P}(S\leqslant t)f_U(t)dt\\
&=\int_0^\infty(1-e^{-\lambda t})f_U(t)dt\\
&=1-\int_0^\infty e^{-\lambda t}\mu e^{-\mu t}dt\\
&=1-\frac{\mu}{\lambda+\mu}\int_0^\infty(\lambda+\mu)e^{-(\lambda+\mu)t}dt=\frac{\lambda}{\lambda+\mu}
\end{aligned}\tag{*}
$$

令 $S=T_i,U=\min\{T_1,\cdots,T_{i-1},T_{i+1},\cdots,T_n\}$, 由 (1) 结论, 则 $\lambda=\lambda_i,\mu=(\sum_{k=1}^n\lambda_k)-\lambda_i$. 代回 (*), 得 $\lambda_i/(\sum_{k=1}^n\lambda_k)$.

(3) 令 $\tilde{V}_i:=\min\{T_1,\cdots,T_{i-1},T_{i+1},\cdots,T_n\}$.

$\{I=i\}=\{T_i<\tilde{V}_i\}+\cup_{j=i+1}^n\{T_i\leqslant\tilde{V}_i,T_i=T_j\}$. 其中 $\{T_i<\tilde{V}_i\}$ 为唯一最小, $T_i\leqslant\tilde{V}_i$ 中至少一个与之相等.

$$
\begin{aligned}
\mathbb{P}\left(\bigcup_{j=i+1}^n\{T_i\leqslant\tilde{V}_i,T_i=T_j\}\right)&\leqslant\sum_{j=i+1}^n\mathbb{P}(T_i=T_j)\\
&\xlongequal{T_i\perp\!\!\!\perp T_j}\sum_{j=i+1}^n\int_0^\infty\mathbb{P}(T_i=t)f_{T_j}(t)dt=0
\end{aligned}
$$

其中 $\mathbb{P}(T_i=t)=0$. 由此得 $\mathbb{P}(I=i)\stackrel{(d)}{=}\mathbb{P}(T_i<\tilde{V}_i)$.

$$
\begin{aligned}
\mathbb{P}(I=i,V>t)&=\mathbb{P}(T_i<\tilde{V}_i,V>t)\\
&=\mathbb{P}(\tilde{V}_i>T_i>t)\\
&\xlongequal{T_i\perp\!\!\!\perp\tilde{V}_i}\int_0^\infty\mathbb{P}(\tilde{V}_i>s>t)f_{T_i}(s)ds\\
&=\int_0^\infty\mathbb{I}_{s>t}\mathbb{P}(\tilde{V}_i>s)f_{T_i}(s)ds\\
&=\int_t^\infty\exp\left(-\sum_{k\neq i}\lambda_k\cdot s\right)\cdot\lambda_i e^{-\lambda_i s}ds\\
&\stackrel{(1)}{=}\frac{\lambda_i}{\sum_{k=1}^n\lambda_k}\cdot\int_t^\infty\left(\sum_{k=1}^n\lambda_k\right)\exp\left(-\sum_{k=i}^n\lambda_k\cdot s\right)ds\\
&\xlongequal{(1,2)}\mathbb{P}(V>t)\mathbb{P}(I=i)
\end{aligned}
$$

由 (2), $\mathbb{P}(T_i=V)=\mathbb{P}(T_i\leqslant\tilde{V}_i)=\lambda_i/(\sum_{k=1}^n\lambda_k)$. 所以 $\mathbb{P}(I=i)=\mathbb{P}(T_i<\tilde{V}_i)=\lambda_i/(\sum_{k=1}^n\lambda_k)$, 在测度意义上相等. □

Theorem 3.1 (指数分布随机变量之和)
设 $\tau_1, \tau_2, \cdots$ 独立同分布, $\tau_1 \sim \mathrm{EXP}(\lambda)$, 则对 $n \geqslant 1$, 有

$$T_n = \sum_{k=1}^{n} \tau_k \sim \Gamma(n, \lambda) \tag{3.4}$$

即

$$f_{T_n}(t) = \begin{cases} \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} & t \geqslant 0 \\ 0 & t < 0 \end{cases} \tag{3.5}$$

其中约定 $0^0 = 1, 0! = 1$.

证明. 1. $n = 1$ 显然.

2. 假设 $n = k$ 成立. 下证 $n = k + 1$ 也成立.

$$T_{k+1} = T_k + \tau_{k+1}, T_k \perp\!\!\!\perp T_{k+1}$$

$$\begin{aligned} f_{T_{k+1}}(t) &= (f_{T_k} * f_{\tau_{k+1}})(t) \\ &= \int_0^t f_{T_k}(s) f_{\tau_{k+1}}(t-s) ds \\ &= \int_0^t \lambda e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!} \lambda e^{-\lambda(t-s)} ds \\ &= \frac{\lambda^{k+1}}{(k-1)!} e^{-\lambda t} \int_0^t s^{k-1} ds \\ &= \frac{\lambda^{k+1} e^{-\lambda t}}{(k-1)!} \left[\frac{s^k}{k} \bigg|_0^t \right] = \lambda e^{-\lambda t} \frac{(\lambda t)^k}{k!} \end{aligned}$$

□

注: 概率密度函数之和的分布等于这两个密度函数的卷积. 从分布函数出发推导

$$\begin{aligned} F_Z(z) = \mathbb{P}(X + Y \leqslant z) &\xlongequal{X \perp\!\!\!\perp Y} \iint_{x+y \leqslant z} f_X(x) f_Y(y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} f_X(x) \left(\int_{-\infty}^{z-x} f_Y(y)\, dy \right) dx \\ &= \int_{-\infty}^{\infty} f_X(x) F_Y(z-x)\, dx \end{aligned}$$

对分布函数求导得到密度函数:

$$\begin{aligned} f_Z(z) = \frac{d}{dz} F_Z(z) &= \frac{d}{dz} \int_{-\infty}^{\infty} f_X(x) F_Y(z-x)\, dx \\ &= \int_{-\infty}^{\infty} f_X(x) \frac{d}{dz} F_Y(z-x)\, dx \\ &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx \\ &= (f_X * f_Y)(z) \end{aligned}$$

3.1.2 泊松分布

Definition 3.2

称 X 服从 "均值/参数为 $\lambda(\lambda>0)$ 的泊松分布" (记作 $X\sim\text{Poisson}(\lambda)$), 若

$$\mathbb{P}(X=n)=e^{-\lambda}\frac{\lambda^n}{n!} \tag{3.6}$$

Property 3.5 (矩)

$X\sim\text{Poisson}(\lambda)$, 对 $\forall k\geqslant 1$,

$$\mathbb{E}X(X-1)\cdots(X-k+1)=\lambda^k \tag{3.7}$$

特别地,

1. $k=1$ 时, $\mathbb{E}X=\lambda$
2. $\text{Var}(X)=\mathbb{E}X^2-(\mathbb{E}X)^2=\mathbb{E}X(X-1)+\mathbb{E}X-(\mathbb{E}X)^2\overset{k=2}{=}\lambda^2+(\lambda-\lambda^2)=\lambda$

证明.

$$\begin{aligned}
\text{LHS}&=\sum_{n=0}^{\infty}n(n-1)\cdots(n-k+1)e^{-\lambda}\frac{\lambda^n}{n!}\\
&=\sum_{n=k}^{\infty}\frac{n!}{(n-k)!}e^{-\lambda}\frac{\lambda^n}{n!}\\
&=\lambda^k e^{-\lambda}\sum_{n=k}^{\infty}\frac{\lambda^{n-k}}{(n-k)!}\\
&\xlongequal{m=n-k}\lambda^k e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}\\
&=\lambda^k e^{-\lambda}e^{\lambda}=\lambda^k
\end{aligned}$$

□

Theorem 3.2 (Durrett Thm 2.4, 泊松随机变量之和)

$X_k\sim\text{Poisson}(\lambda_k), k\geqslant 1$ 独立, 则

$$\sum_{k=1}^{N}X_k\sim\text{Poisson}\left(\sum_{k=1}^{N}\lambda_k\right) \tag{3.8}$$

证明. $N=2$.

$$\mathbb{P}(X_1+X_2=n)\xlongequal{X_1\perp\!\!\!\perp X_2}\sum_{m=0}^{\infty}\mathbb{P}(X_1+m=n)\mathbb{P}(X_2=m)$$

公式: $\mathbb{E}|g(X,Y)|<\infty, X \perp\!\!\!\perp Y$, 则 $\mathbb{E}g(X,Y)=\mathbb{E}(\mathbb{E}g(x,Y)|_{x=X})$.

$$\begin{aligned}\mathbb{P}(X_1+X_2=n) &= \mathbb{E}\mathbb{I}_{\{X_1+X_2=n\}} \\ &\xlongequal{X_1 \perp\!\!\!\perp X_2} \mathbb{E}_{X_2}(\mathbb{E}_{X_1}\mathbb{I}_{\{X_1+m=n\}}|_{m=X_2}) \\ &= \mathbb{E}_{X_2}(\mathbb{P}(X_1+m=n)|_{m=X_2}) \\ &= \sum_{m=0}^{\infty}\mathbb{P}(X_1+m=n)\mathbb{P}(X_2=m)\end{aligned}$$

$\mathbb{E}_{X_1}, \mathbb{E}_{X_2}$ 表示关于 X_1, X_2 求期望. 利用独立改写成卷积.

$$\begin{aligned}\mathbb{P}(X_1+X_2=n) &\xlongequal{X_1 \perp\!\!\!\perp X_2} \sum_{m=0}^{n}\mathbb{P}(X_1=n-m)\mathbb{P}(X_2=m) \\ &= \sum_{m=0}^{\infty} e^{-\lambda_1}\frac{\lambda_1^{n-m}}{(n-m)!}e^{-\lambda_2}\frac{\lambda_2^m}{m!} \\ &= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}\cdot\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^m\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-m} \\ &= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}\end{aligned}$$

其中, 由二项式定理 $(a+b)^n=\sum_{k=0}^n\binom{n}{k}a^kb^{n-k}$ 知,

$$\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^m\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-m}=\left(\frac{\lambda_1}{\lambda_1+\lambda_2}+\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^n=1$$

$\square$

Theorem 3.3 (Durrett(3ed), Thm 2.5, 二项分布的泊松逼近)
$X_n \sim \text{Binomial}(n, \frac{\lambda}{n}) \xrightarrow{n\to\infty} \text{Poisson}(\lambda), \lambda>0$. 其中 $\xrightarrow{n\to\infty}$ 表示依分布收敛.

证明.

$$\begin{aligned}\mathbb{P}(X_n=k) &= \frac{n!}{k!(n-k)!}(\frac{\lambda}{n})^k(1-\frac{\lambda}{n})^{n-k} \\ &= \frac{\lambda^k}{k!}\cdot\left[\frac{n!}{(n-k)!\cdot n^k}\cdot(1-\frac{\lambda}{n})^n\cdot(1-\frac{\lambda}{n})^{-k}\right] \\ &=: \frac{\lambda^k}{k!}\cdot[\mathcal{J}_{n,1}\cdot\mathcal{J}_{n,2}\cdot\mathcal{J}_{n,3}]\end{aligned}$$

$$\lim_{n\to\infty}\mathcal{J}_{n,1}=\lim_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{n\cdot n\cdots n}=1$$

$$\lim_{n\to\infty}\mathcal{J}_{n,2}=\lim_{n\to\infty}\left[(1+\frac{-\lambda}{n})^{n/(-\lambda)}\right]^{-\lambda}=e^{-\lambda}$$

$\lim_{n\to\infty}\mathcal{J}_{n,3}=1^{-k}=1$. 代回得

$$\lim_{n\to\infty}\mathbb{P}(X_n=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$$

$\square$

3.2 泊松过程的定义

Definition 3.3 (计数过程)
若随机变量 $N(t)$ 表示时间段 $[0,t]$ 内某事件发生的次数, 则称 $\{N(t), t \geqslant 0\}$ 是一个计数过程, 若

1. $N(t) \geqslant 0$ 取整数值
2. 若 $s < t$, 则 $N(s) \leqslant N(t)$, 且 $N(t) - N(s)$ 表示时间段 $(s,t]$ 内事件的发生次数

Definition 3.4 (泊松过程 I)
称一个计数过程 $\{N(t), t \geqslant 0\}$ 是一个带有参数/速率 $\lambda(\lambda > 0)$ 的泊松过程, 若

1. $N(0) = 0$, 即 $\mathbb{P}(N(0) = 0) = 1$
2. (独立增量) $\forall 0 \leqslant t_1 < t_2 < \cdots < t_n$, 有
$$N(t_2) - N(t_1), \cdots, N(t_n) - N(t_{n-1})$$
相互独立
3. $\forall t > 0, s \geqslant 0, N(t+s) - N(s) \sim \text{Poisson}(\lambda t)$

Property 3.6
由 Def 3.4 (3) 易知, $(N(t))_{t \geqslant 0}$ 具有平稳增量.
$$N(t+s) - N(s) \xlongequal{(d)} N(t) - N(0) \xlongequal{(d)} N(t) \sim \text{Poisson}(\lambda t)$$
注: $\mathbb{E}N(t) = \lambda t \Rightarrow \lambda = \mathbb{E}N(t)/t$ 过程的速率

Property 3.7 (Durrett, Lem 2.5)
对固定 $s \geqslant 0, \{N(t+s) - N(s), t \geqslant 0\}$ 仍是一个带有速率 λ 的 Poisson 过程, 且 $N(t+s) - N(s) \perp\!\!\!\perp N(r), \forall 0 \leqslant r \leqslant s, t \geqslant 0$

Definition 3.5 (泊松过程 II: 到达时间间隔)
令 $\tau_1, \tau_2, \cdots$ 为一列独立同分布的随机变量, $\tau_1 \sim \text{EXP}(\lambda)$,
$$T_n = \begin{cases} \sum_{k=1}^n \tau_k & n \geqslant 1 \\ 0 & n = 0 \end{cases}$$
$N(t) = \max\{n | T_n \leqslant t\}$, 则称 $\{N(t), t \geqslant 0\}$ 为带有速率 λ 的 Poisson 过程
注: “无穷小定义”, 见 Sheldon Ross 随机过程 [5].

注:

1. $T_n (n \geqslant 1)$: 第 n 个顾客的到店时刻

2. $\tau_n(n \geqslant 1)$: 第 n 个和第 $n-1$ 个顾客的到店时间间隔
3. $N(t)$: t 时刻之前到达的顾客总数

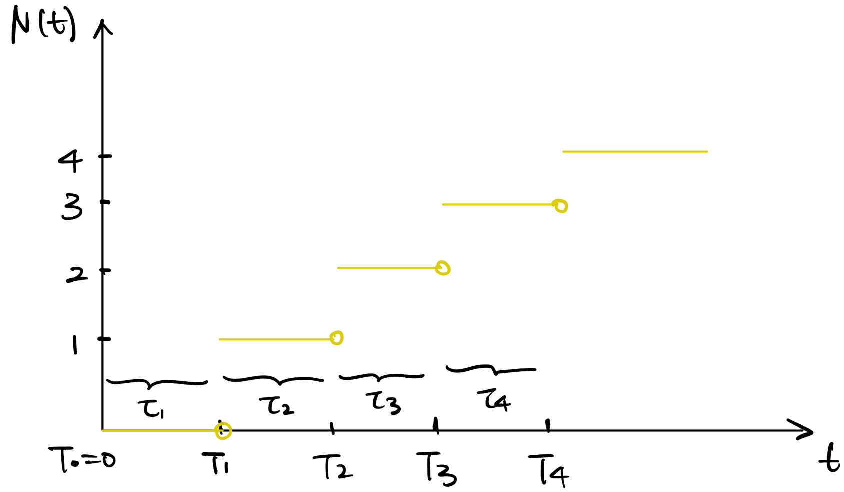


图 8: Arrival

注意 $T_1, T_2, \cdots$ 是随机的

1. $n \geqslant 1, \tau_n = T_n - T_{n-1} \sim \text{EXP}(\lambda), T_n \sim \Gamma(n, \lambda)$
2. $\{N(t) = n\} = \{T_n \leqslant t < T_{n+1}\}$
3. $\{N(t) \geqslant n\} = \{t \geqslant T_n\}$
4. $\{N(t) < n\} = \{t < T_n\}$

Theorem 3.4

两种定义是等价的, 即 Def 3.4 $\iff$ Def 3.5

Proposition 3.1

Def 3.5 $\Rightarrow$ Def 3.4

证明. 1. $\{N(0) = 0\} = \{T_1 > 0\} = \{\tau_1 > 0\} \Rightarrow \mathbb{P}(N(0) = 0) = \mathbb{P}(\tau_1 > 0) = 1$

先引入下面引理.

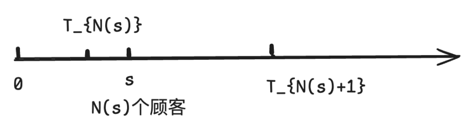


图 9: 固定起始时间 s

Lemma 3.1 (Lem 2.5')

对固定 $s \geqslant 0$ (图 9), 令 $\tau_1^s := T_{N(s)+1} - s$. $\tau_n^s := \tau_{N(s)+n}, n \geqslant 2$.

$$T_n^s := \begin{cases} \sum_{k=1}^n \tau_k^s & n \geqslant 1 \\ 0 & n = 0 \end{cases}$$

$N^s(t) := \max\{n | T_n^s \leqslant t\}$. 则

(a) $N^s(t) = N(t+s) - N(s)$

(b) $\forall k \geqslant 1, (\tau_1^s, \cdots, \tau_k^s) \perp\!\!\!\perp N(s)$, 即 $\tau_k^s = \tau_k \sim \text{EXP}(\lambda)$, $\tau_1^s, \tau_2^s, \cdots$ 相互独立

(c) $\{N^s(t) = N(t+s) - N(s), t \geqslant 0\}$ 为带有速率 λ 的泊松过程, 且 $N(t+s) - N(s) \perp\!\!\!\perp N(r), \forall 0 \leqslant r \leqslant s, t \geqslant 0$

2. (独立增量) 由 Lem 3.1 (c) 及数学归纳法
   $n = 2$, 对 $0 = t_0 \leqslant t_1 < t_2$, 有 $N(t_2) - N(t_1) \perp\!\!\!\perp N(t_1) = N(t_1) - N(t_0)$.
   $n = k$, 假设对 $0 = t_0 \leqslant t_1 < t_2 < \cdots < t_n$, 有 $N(t_k) - N(t_{k-1}), \cdots, N(t_1) - N(t_0)$ 相互独立. 考虑 $n = k+1$, 对于第 $k+1$ 个增量 $N(t_{k+1}) - N(t_k)$, 由 Lem 3.1 (c), 令 $s = t_k, N^s(t_{k+1} - t_k) \perp\!\!\!\perp N(r), \forall r \leqslant t_k$, 因此由 Thm 1.3, $N^s(t_{k+1} - t_k) \perp\!\!\!\perp \sigma(\bigcup_{r \leqslant t_k} N(r))$. $\forall 1 \leqslant j \leqslant k, N^s(t_{k+1} - t_k) \perp\!\!\!\perp (N(t_j) - N(t_{j-1}))$. 又因 $\forall 1 \leqslant j \leqslant k, N(t_j) - N(t_{j-1})$ 相互独立, 则 $\forall 1 \leqslant j \leqslant k+1, N(t_j) - N(t_{j-1})$ 相互独立.
3. 由 Lem 3.1 (a)(c) 知, 只需证 $N(t) \sim \text{Poisson}(\lambda t)$, 则有 $N(t+s) - N(s) \sim \text{Poisson}(\lambda t)$. 这是因为 Lem 3.1 (a)(c) 已经把 $N(t+s) - N(s)$ 表达为一个"从 s 开始重新计时"的新泊松过程 $N^s(t)$, 而我们知道 $N^s(t)$ 的构造方式与原始的 $N(t)$ 完全一致, 只不过起点平移到了 s.

   (a) $\mathbb{P}(N(t) = 0) = \mathbb{P}(T_1 > t) = \mathbb{P}(\tau_1 > t) = e^{-\lambda t}(\lambda t)^0/(0!)$

   (b) $n > 0$ 时,

$$\begin{aligned}
\mathbb{P}(N(t) = n) &= \mathbb{P}(T_n \leqslant t < T_{n+1}) \\
&= \mathbb{P}(T_n \leqslant t < T_n + \tau_{n+1}) \\
&\xlongequal{T_n \perp\!\!\!\perp \tau_{n+1}} \int_0^{+\infty} \mathbb{P}(u \leqslant t < u + \tau_{n+1}) f_{T_n}(u) du \\
&= \int_0^{+\infty} \mathbb{I}_{\{u \leqslant t\}} \mathbb{P}(\tau_{n+1} > t - u) f_{T_n}(u) du \\
&= \int_0^t e^{-\lambda(t-u)} \lambda e^{-\lambda u} \frac{(\lambda u)^{n-1}}{(n-1)!} du \\
&= \frac{\lambda^n e^{-\lambda t}}{(n-1)!} \int_0^t e^{\lambda u} \cdot e^{-\lambda u} u^{n-1} du \\
&= \frac{\lambda^n e^{-\lambda t}}{(n-1)!} \cdot \frac{1}{n} \cdot u^n |_0^t = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \sim \text{Poisson}(\lambda t)
\end{aligned}$$

□

下面证明 Lem 3.1.

证明. (a) 要证 $N^s(t) = N(t+s) - N(s)$.

$n \geqslant 1, T_n^s = T_{N(s)+1} - s + \tau_{N(s)+2} + \cdots + \tau_{N(s)+n} = T_{N(s)+n} - s, T_{N(s)} \leqslant s$

$$\begin{aligned}
N^s(t) &= \max\{n \geqslant 0 | T_n^s \leqslant t\} \\
&= \max\{n \geqslant 0 | T_{N(s)+n} \leqslant t + s\} \\
&\xlongequal{m=N(s)+n} \max\{m - N(s) \geqslant 0 | T_m \leqslant t + s\} \\
&= \max\{m \geqslant 0 | T_m \leqslant t + s\} - N(s) \\
&= N(t+s) - N(s)
\end{aligned}$$

(b) 要证明

$\forall k \geqslant 1$, $\boxed{(\tau_1^s, \cdots, \tau_k^s) \perp\!\!\!\perp N(s)}$, 即 $\tau_k^s = \tau_k \sim \mathrm{EXP}(\lambda)$, $\tau_1^s, \tau_2^s, \cdots$ 相互独立

实际上方框内的陈述更强, 目前无法证明, 所以只证后面的部分.

(1) $k = 1$,

$$\begin{aligned}
\mathbb{P}(\tau_1^s > t_1, N(s) = n) &= \mathbb{P}(T_{N(s)+1} - s > t_1, N(s) = n) \\
&= \mathbb{P}(T_{n+1} > t_1 + s, s \geqslant T_n) \\
&= \mathbb{P}(\tau_{n+1} > t_1 + s - T_n, T_n \leqslant s) \\
&\xlongequal{T_n \perp\!\!\!\perp \tau_{n+1}} \int_0^{+\infty} \mathbb{P}(\tau_{n+1} > t_1 + s - u, u \leqslant s) f_{T_n}(u) du \\
&= \int_0^{+\infty} \mathbb{I}_{\{u \leqslant s\}} \mathbb{P}(\tau_{n+1} > t_1 + s - u) f_{T_n}(u) du \\
&\xlongequal{(*)} \int_0^{+\infty} \mathbb{I}_{\{u \leqslant s\}} \mathbb{P}(\tau_{n+1} > s - u) f_{T_n}(u) du \cdot \mathbb{P}(\tau_{n+1} > t_1) \\
&= \mathbb{P}(\tau_{n+1} > s - T_n, T_n \leqslant s) \mathbb{P}(\tau_{n+1} > t_1) \\
&= \mathbb{P}(T_{n+1} > s \geqslant T_n) \mathbb{P}(\tau_{n+1} > t_1) \\
&= \mathbb{P}(N(s) = n) \mathbb{P}(\tau_{n+1} > t_1)
\end{aligned}$$

$(*)$ 处用了指数分布的无记忆性,

$$\begin{aligned}
\mathbb{P}(\tau_{n+1} > t_1 + s - u) &= \mathbb{P}(\tau_{n+1} > t_1 + s - u | \tau_{n+1} > t_1) \cdot \mathbb{P}(\tau_{n+1} > t) \\
&= \mathbb{P}(\tau_{n+1} > s - u) \cdot \mathbb{P}(\tau_{n+1} > t)
\end{aligned}$$

最后关于 n 求和,

$$\sum_{n \geqslant 0} \mathbb{P}(\tau_1^s > t_1, N(s) = n) = \sum_{n \geqslant 0} \mathbb{P}(N(s) = n) \mathbb{P}(\tau_{n+1} > t_1)$$

其中 $\mathrm{LHS} = \mathbb{P}(\tau_1^s > t_1)$. 因为 $\mathbb{P}(\tau_{n+1} > t_1) \xlongequal{iid} \mathbb{P}(\tau_1 > t_1)$, $\Rightarrow \mathrm{RHS} = \mathbb{P}(\tau_1 > t_1) \Rightarrow \tau_1^s \sim \mathrm{EXP}(\lambda)$. 代回, 得 $\mathbb{P}(\tau_1 > t_1) = \mathbb{P}(\tau_1^s > t_1)$, 即 $\tau_1^s \perp\!\!\!\perp N(s)$.

注: $\mathbb{P}(T_{n+1} - s > t | N(s) = n) = \mathbb{P}(\tau_1^s > t | N(s) = n) = \mathbb{P}(\tau_1 > t)$. $\Rightarrow T_{n+1} - s \sim \mathrm{EXP}(\lambda)$ under $\mathbb{P}(\cdot | N(t) = n)$.

(2) $k \geqslant 2$ 时,

$$\begin{aligned}
A_n &:= \{N(s)=n, \tau_1^s > t_1, \tau_2^s > t_2, \cdots, \tau_k^s > t_k\} \\
&= \{T_n \leqslant s, T_{n+1} > t_1 + s\} \cap \left\{\bigcap_{i=2}^{k} \{\tau_{n+2} > t_i\}\right\} \\
\mathbb{P}(A_n) &= \mathbb{P}(T_n \leqslant s, T_{n+1} > t_1 + s) \prod_{i=2}^{k} \mathbb{P}(\tau_{n+i} > t_i) \\
&\xlongequal{k=1} \mathbb{P}(N(s)=n)\mathbb{P}(\tau_1 > t_1) \prod_{i=2}^{k} \mathbb{P}(\tau_{n+i} > t_i)
\end{aligned}$$

在概率测度意义下 $\tau_{n+i} \overset{(d)}{=} \tau_i$, 关于 n 求和, $(\tau_1^s, \cdots, \tau_k^s) \overset{(d)}{=} (\tau_1, \cdots, \tau_k)$, 即 $\tau_1^s, \cdots, \tau_k^s$ 相互独立, 代回前式得 $(\tau_1^s, \cdots, \tau_k^s) \perp\!\!\!\perp N(s)$.

(3) 类似前面两步的技巧.

(c) 不妨设 $0 \leqslant t_1 < \cdots < t_k$

$$\begin{aligned}
\{N^s(t_1)=m_1, \cdots, N^s(t_k)=m_k\} &= \{T_{m_1}^s \leqslant t_1 < T_{m_1+1}^s, \cdots, T_{m_k}^s \leqslant t_k < T_{m_k+1}^s\} \\
&= \left\{\sum_{i=1}^{m_1} \tau_i^s \leqslant t_1 < \sum_{i=1}^{m_1+1} \tau_i^s, \cdots, \sum_{i=1}^{m_k} \tau_i^s \leqslant t_k < \sum_{i=1}^{m_k+1} \tau_i^s\right\}
\end{aligned}$$

由 (b), $\tau_i^s \sim \mathrm{EXP}(\lambda), \forall i \geqslant 0$. 由 Def 3.5, $N^s(t_i), \forall i \geqslant 0$ 为速率 λ 的泊松过程.

□

Proposition 3.2

Def 3.4 ⇒ Def 3.5

证明. $\{N(t), t \geqslant 0\}$ 是计数过程, $T_n = \min\{t | N(t) \geqslant n\}, \tau_n = T_n - T_{n-1}$.

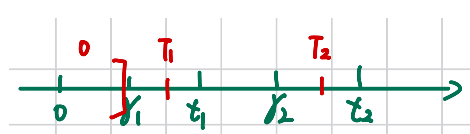
1. $t \neq 0$,

$$N(t) = \sum_{n=1}^{\infty} \mathbb{I}_{\{N(t) \geqslant n\}} = \sum_{n=1}^{\infty} \mathbb{I}_{\{T_n \leqslant t\}} = \max\{n | T_n \leqslant t\}$$

2. 下面只需证 $(\tau_k)_{k \geqslant 1} \overset{\text{iid}}{\sim} \mathrm{EXP}(\lambda)$

   (a) $k=1, \mathbb{P}(\tau_1 > t) = \mathbb{P}(T_1 > t) = \mathbb{P}(N(t)=0) = e^{-\lambda t}(\lambda t)^0/(0!) = e^{-\lambda t}$, 故 $\tau_1 \sim \mathrm{EXP}(\lambda)$.

   (b) 考察 $(T_1, T_2), S = \{(x_1, x_2) \in \mathbb{R}^2 | 0 < x_1 < x_2\}$. 设 $0 < r_1 < t_1 < r_2 < t_2$



$$
\begin{aligned}
&\mathbb{P}(r_1 \leqslant T_1 < t_1, r_2 \leqslant T_2 < t_2) \\
=&\mathbb{P}(N(r_1)=0, N(r_2)-N(t_1)=0, N(t_1)-N(r_1)=1, N(t_2)-N(r_2) \geqslant 1) \\
=&e^{-\lambda r_1} \cdot e^{-\lambda(r_2-t_1)} \cdot e^{-\lambda(t_1-r_1)} \cdot \frac{\lambda(t_1-r_1)}{1!} \cdot [1-e^{-\lambda(t_2-r_2)}] \\
=&\lambda(t_1-r_1)[e^{-\lambda r_2}-e^{-\lambda t_2}] = \lambda \int_{r_1}^{t_1} dx_1 \int_{r_2}^{t_2} \lambda e^{-\lambda x_2} dx_2
\end{aligned}
$$

$N(t_2)-N(r_2) \geqslant 1$ 是因为要求第二个事件 T_2 必须落在 $[r_2, t_2)$ 内，但并不限制该区间内后续事件的数量. 联合密度函数 $\lambda^2 e^{-\lambda x_2}$ 在区域 S (即 $0 < x_1 < x_2$) 上，所以 $f_{(T_1,T_2)}(x_1,x_2) = \lambda^2 e^{-\lambda x_2}\mathbb{I}_{\{(x_1,x_2)\in S\}}$.

$$
\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2) = \mathbb{P}(\tau_1 > t_1, T_2 - T_1 > t_2) = \int_{x_1 > t_1} \int_{x_2 - x_1 > t_2} \lambda^2 e^{-\lambda x_2} dx_2 dx_1 = e^{-\lambda(t_1+t_2)}
$$

又 $\mathbb{P}(\tau_1 > t_1) = e^{-\lambda t_1}$，所以 $\tau_2 \sim \mathrm{EXP}(\lambda)$ 且 $\tau_1 \perp\!\!\!\perp \tau_2$. 对 $k \geqslant 3$ 同理.

□

Definition 3.6 (Def 3.4 推广，非齐次的泊松过程)

称计数过程 $\{N(t), t \geqslant 0\}$ 为一个速率为 $\lambda(r)$ 的非齐次泊松过程，若

1. $N(0) = 0$
2. 独立增量性
3. 增量的分布

$$
N(t+s) - N(s) \sim \mathrm{Poisson}(\int_s^{t+s} \lambda(r) dr)
$$

Proposition 3.3

非齐次泊松过程的到达间隔时间列 $\tau_1, \tau_2, \cdots$ 不再服从指数分布，不再相互独立.

证明. 同上计算联合分布即可. □

3.3 复合泊松过程

Definition 3.7

$\{N(t), t \geqslant 0\}$ 是一个泊松过程, $\{Y_i\}_{i\geqslant 1}$ 独立同分布且 $\{Y_i\}_{i\geqslant 1} \perp\!\!\!\perp (N(t))_{t\geqslant 0}$. 令

$$S(t) := \begin{cases} \sum_{i=1}^{N(t)} Y_i & N(t) \neq 0 \\ 0 & N(t) = 0 \end{cases}$$

则 $\{S(t)\}_{t\geqslant 0}$ 称为一个复合泊松过程.

Theorem 3.5

$\{Y_i\}_{i\geqslant 1}$ 为独立同分布 r.v. 列, N 为取非负整数值的 r.v., $\boxed{\{Y_i\}_{i\geqslant 1} \perp\!\!\!\perp N}$. 令

$$S = \begin{cases} \sum_{i=1}^{N} Y_i & N \neq 0 \\ 0 & N = 0 \end{cases}$$

则

1. $\mathbb{E}S = \mathbb{E}N \cdot \mathbb{E}Y_i$
2. $\operatorname{Var}(S) = \mathbb{E}N \operatorname{Var}(Y_i) + \operatorname{Var}(N)(\mathbb{E}Y_i)^2$
3. 特别地, 若 $N \sim \text{Poisson}(\lambda)$, 则 $\mathbb{E}S = \lambda\mathbb{E}Y_i, \operatorname{Var}(S) = \lambda[\operatorname{Var}(Y_i) + (\mathbb{E}Y_i)^2] = \lambda\mathbb{E}Y_i^2$.

注:

1. 若 $N = n$ 非随机, 则 $\mathbb{E}S = n\mathbb{E}Y_i$
2. 若 $N = n$ 非随机, $\operatorname{Var}(S) = n\operatorname{Var}(Y_i)$. 若 Y_i 非随机, $\operatorname{Var}(S) = \operatorname{Var}(Y \cdot N) = Y^2 \operatorname{Var}(N)$.

证明. 1. 由重期望公式(1.4),

$$\mathbb{E}S = \mathbb{E}(\mathbb{E}(S|N)) = \mathbb{E}\left[\sum_{n\geqslant 0} \mathbb{E}(S|N=n)\mathbb{I}_{\{N=n\}}\right] = \sum_{n\geqslant 0} \mathbb{E}(S|N=n)\mathbb{P}(N=n)$$

其中 $\sum_{n\geqslant 0} \mathbb{E}(S|N=n)\mathbb{I}_{\{N=n\}}$ 为一个离散随机变量. $\mathbb{E}(S|N=n)$ 是一个数, 不具有随机性.

$$\mathbb{E}(S|N=n) = \mathbb{E}(\sum_{i=1}^{n} Y_i|N=n) \xlongequal{\{Y_i\} \perp\!\!\!\perp N} \mathbb{E}(\sum_{i=1}^{n} Y_i) = n\mathbb{E}Y_i$$

所以 $\mathbb{E}S = \mathbb{E}N\mathbb{E}Y_i$.

2. $\mathbb{E}(S^2|N=n) = \mathbb{E}[\sum_{i=1}^n Y_i]^2 = \operatorname{Var}(\sum_{i=1}^n Y_i) + [\mathbb{E}\sum_{i=1}^n Y_i]^2 = n\operatorname{Var}(Y_i) + n^2[\mathbb{E}Y_i]^2$. $\mathbb{E}S^2 = \mathbb{E}N \cdot \operatorname{Var}(Y_i) + \mathbb{E}N^2 \cdot (\mathbb{E}Y_i)^2$. $\operatorname{Var}(S) = \mathbb{E}S^2 - (\mathbb{E}S)^2 = \mathbb{E}N\operatorname{Var}(Y_i) + \operatorname{Var}(N)(\mathbb{E}Y_i)^2$

□

3.4 泊松过程的变换

1. 稀释 (thining): 把一个泊松过程拆分成若干个独立的泊松过程
2. 叠加 (superposition): 把若干个独立的泊松过程合并成一个泊松过程

3.4.1 稀释/可分解性

Theorem 3.6

$(N(t))_{t\geqslant 0}\sim \mathrm{PP}(\lambda)$, $(Y_i)_{i\geqslant 1}$ 为 iid 的离散随机变量列, $(Y_i)_{i\geqslant 1}\perp\!\!\!\perp(N(t))_{t\geqslant 0}$. $\mathbb{P}(Y_i=j)=p_j, 1\leqslant j\leqslant m$. 令

$$N_j(t):=\sum_{i=1}^{N(t)}\mathbb{I}_{\{Y_i=j\}}, 1\leqslant j\leqslant m$$

约定 $\sum_a^b(\cdot)=0, b<a$. 则

1. $N(t)=\sum_j N_j(t)$
2. $\{N_j(t), t\geqslant 0\}, 1\leqslant j\leqslant m$ 为相互独立的泊松过程, 且速率分别为 λp_j

注: $m=2$ 时, $\mathbb{P}(Y_i=1)=p, \mathbb{P}(Y_i=2)=1-p$. e.g. $N_1(t)$: $[0,t]$ 之间性别 1 的顾客数; $N_2(t)$: $[0,t]$ 之间性别 2 的顾客数.

证明. 下面证明 $N_1(t)$ 服从泊松过程, 且速率为 λp_1.

1. $N(0)=N_1(0)+N_2(0)=0$
2. (Step 1) Claim: $N_1(t)\sim \mathrm{Poisson}(\lambda pt), p_1=p$

$$\mathbb{P}(N_1(t)=n)=\sum_{m\geqslant 0}\mathbb{P}(N_1(t)=n|N(t)=n+m)\mathbb{P}(N(t)=n+m) \tag{3.9}$$

$(Y_i)_{i\geqslant 1}\overset{\text{iid}}{\sim}(N(r))_{r\geqslant 0}$, $\Rightarrow (\mathbb{I}_{\{Y_i=1\}})_{i\geqslant 1}\perp\!\!\!\perp(N(r))_{r\geqslant 0}$. 因为

$$\left\{\sum_{i=1}^m Y_i=n\right\}=\left\{(Y_1,\cdots,Y_m)\in\{(X_1,\cdots,X_m)\in S^m|\sum_{i=1}^m X_i=n\}\right\} \tag{3.10}$$

所以

$$\sum_{i=1}^{n+m}\mathbb{I}_{\{Y_i=1\}}=n\perp\!\!\!\perp(N(r))_{r\geqslant 0} \tag{3.11}$$

由 (3.11),

$$\mathbb{P}(\sum_{i=1}^{N(t)}\mathbb{I}_{\{Y_i=1\}}=n|N(t)=n+m)\xlongequal{\perp\!\!\!\perp}\mathbb{P}(\sum_{i=1}^{n+m}\mathbb{I}_{\{Y_i=1\}}=n)\xlongequal{\text{iid}}\binom{n+m}{n}p^n(1-p)^m \tag{3.12}$$

其中 $\mathbb{I}_{\{Y_i=1\}} \sim \text{Bernoulli}(p)$. 将 (3.12) 代回 (3.9), 得

$$\begin{aligned}
\mathbb{P}(N_1(t)=n) &= \sum_{m\geqslant 0}\frac{(n+m)!}{n!m!}p^n(1-p)^m e^{-\lambda t}\frac{(\lambda t)^{n+m}}{(n+m)!} \\
&= \frac{(\lambda pt)^n}{n!}e^{-\lambda t}\sum_{m\geqslant 0}\frac{(\lambda t(1-p))^m}{m!} \\
&= \frac{(\lambda pt)^n}{n!}e^{-\lambda t}e^{\lambda t(1-p)} \\
&= e^{-\lambda pt}\frac{(\lambda pt)^n}{n!}
\end{aligned}$$

(Step 2)

$$N_1(t+s)-N_1(s) = \sum_{i=N(s)+1}^{N(t+s)}\mathbb{I}_{\{Y_i=1\}} = \sum_{\tilde{i}=1}^{N(t+s)-N(s)}\mathbb{I}_{\{Y_{N(s)+\tilde{i}}=1\}}$$

由 (Step 1) 的证明步骤知, 需要条件 $(Y_{N(s)+i})_{i\geqslant 1} \perp\!\!\!\perp (N(t+s)-N(s))_{t\geqslant 0}$ 使得 $N(t+s)-N(s) \sim \text{Poisson}(\lambda pt)$

Lemma 3.2

令 $Y_i^s := Y_{N(s)+i}, N_1^s(t) := \sum_{i=1}^{N^s(t)}\mathbb{I}_{\{Y_i^s=1\}}$, 则

(a) $N_1^s(t) = N_1(t+s)-N_1(s)$

(b) $1^\circ (Y_i^s)_{i\geqslant 1}$, iid, $Y_i^s \overset{(d)}{=} Y_i$

$2^\circ (Y_i^s)_{i\geqslant 1} \perp\!\!\!\perp (N^s(t))_{t\geqslant 0}$

(c) $N_1^s(t) \sim \text{Poisson}(\lambda pt)$

证明. (a) 因为 $N^s(t) = N(t+s)-N(s)$

$$\begin{aligned}
N_1^s(t) &= \sum_{i=1}^{N^s(t)}\mathbb{I}_{\{Y_i^s=1\}} = \sum_{i=1}^{N(t+s)-N(s)}\mathbb{I}_{\{Y_{N(s)+i}=1\}} \\
&\xlongequal{\tilde{i}=N(s)+i} \sum_{\tilde{i}=N(s)+1}^{N(t+s)}\mathbb{I}_{\{Y_{\tilde{i}}=1\}} = N_1(t+s)-N_1(s)
\end{aligned}$$

(b) 1° 牢记 $(Y_i)_{i\geqslant 1}$ 是 iid 的

$$\begin{aligned}
\mathbb{P}(Y_{N(s)+i_1}=j_1, Y_{N(s)+i_2}=j_2) &= \sum_{n\geqslant 0}\mathbb{P}(Y_{N(s)+i_1}=j_1, Y_{N(s)+i_2}=j_2|N(s)=n)\mathbb{P}(N(s)=n) \\
&\xlongequal{\text{iid}} \sum_{n\geqslant 0}\mathbb{P}(Y_{i_1}=j_1)\mathbb{P}(Y_{i_2}=j_2)\mathbb{P}(N(s)=n) \\
&= \mathbb{P}(Y_{i_1}=j_1)\mathbb{P}(Y_{i_2}=j_2)
\end{aligned}$$

故 $(Y_{i_1}^s, Y_{i_2}^s) \overset{(d)}{=} (Y_{i_1}, Y_{i_2})$, $(Y_i^s)_{i\geqslant 1}$ 相互独立

2° Y_i^s 中存在 $N(s)$, 先将其固定.

$$\begin{aligned}&\sum_{m\geqslant 0}\mathbb{P}(Y_{N(s)+i}=j|N^s(t_1)=m_1,\cdots,N^s(t_k)=m_k,N(s)=m)\mathbb{P}(N(s)=m)\\=&\mathbb{P}(Y_{m+i}=j)\overset{1^\circ}{=}\mathbb{P}(Y_i^s=j)\end{aligned}$$

(c) 由 (Step 1) 步骤, (b) 独立性满足时, $N_1(t+s)-N_1(s)\sim \text{Poisson}(\lambda pt)$. 由 (a), $N_1(t+s)-N_1(s)$ 就是 $N_1^s(t)$, 所以 $N_1^s(t)\sim\text{Poisson}(\lambda pt)$.

□

由 Lem 3.2 (c) 知, $N_1(t+s)-N_1(s)\sim\text{Poisson}(\lambda pt)$

3. (独立增量性) $0=t_0<t_1<t_2, 0=n_0<n_1<n_2$.

$$\begin{aligned}A&:=\mathbb{P}(N_1(t_j)-N_1(t_{j-1})=m_j,1\leqslant j\leqslant 2|(N(t_1),N(t_2))=(n_1,n_2))\\&=\mathbb{P}\left(\sum_{i=N(t_{j-1})+1}^{N(t_j)}\mathbb{I}_{\{Y_i=1\}}=m_j,1\leqslant j\leqslant 2\middle|(N(t_1),N(t_2))=(n_1,n_2)\right)\\&\xlongequal{\mathbb{I}_{\{Y_i=1\}}\perp\!\!\!\perp(N(t))_{t\geqslant 0}}\mathbb{P}\left(\sum_{i=n_{j-1}+1}^{n_j}\mathbb{I}_{\{Y_i=1\}}=m_j,1\leqslant j\leqslant 2\right)\\&=\prod_{j=1}^{2}\mathbb{P}\left(\sum_{i=n_{j-1}+1}^{n_j}\mathbb{I}_{\{Y_i=1\}}=m_j\right)\end{aligned}$$

最后一个等式成立是因为, $j=1,2$ 时, 求和的区间分别为 $[n_0+1,n_1],[n_1+1,n_2]$, 因此相互独立.

$$\begin{aligned}&\sum_{n_2\geqslant n_1\geqslant 0}A\cdot\mathbb{P}(N(t_1)=n_1,N(t_2)=n_2)\\=&\sum_{n_2\geqslant n_1\geqslant 0}\prod_{j=1}^{2}\mathbb{P}(\sum_{i=n_{j-1}+1}^{n_j}\mathbb{I}_{\{Y_i=1\}}=m_j)\cdot\mathbb{P}(N(t_1)-N(t_0)=n_1)\cdot\mathbb{P}(N(t_2)-N(t_1)=n_2-n_1)\\=&\sum_{n_2\geqslant n_1\geqslant 0}\prod_{j=1}^{2}\mathbb{P}(\sum_{i=1}^{n_j-n_{j-1}}\mathbb{I}_{\{Y_i^{t_{j-1}}=1\}}=m_j)\mathbb{P}(\underbrace{N(t_j)-N(t_{j-1})}_{N^{t_{j-1}}(t_j-t_{j-1})}=n_j-n_{j-1})\\=&\sum_{n_2\geqslant n_1\geqslant 0}\prod_{j=1}^{2}\mathbb{P}(\sum_{i=1}^{N^{t_{j-1}}(t_j-t_{j-1})}\mathbb{I}_{\{Y_i^{t_{j-1}}=1\}}=m_j,N(t_j)-N(t_{j-1})=n_j-n_{j-1})\\=&\sum_{n_2\geqslant n_1\geqslant 0}\prod_{j=1}^{2}\mathbb{P}(N_1^{t_{j-1}}(t_j-t_{j-1})=m_j,N(t_j)-N(t_{j-1})=n_j-n_{j-1})\\=&\sum_{n_1\geqslant 0}\sum_{n_2\geqslant 0}\mathbb{I}_{\{n_2\geqslant n_1\}}\cdot\mathbb{P}(N_1(t_1)=m_1,N(t_1)=n_1)\cdot\mathbb{P}(N_1(t_2)-N_1(t_1)=m_2,N(t_2)-N(t_1)=n_2-n_1)\\=&\sum_{n_1\geqslant 0}\mathbb{P}(N_1(t_1)=m_1,N(t_1)=n_1)\mathbb{P}(N_1(t_2)-N_1(t_1)=m_2)\\=&\mathbb{P}(N_1(t_1)=m_1)\cdot\mathbb{P}(N_1(t_2)-N_1(t_1)=m_2)\end{aligned}$$

也就是

$$\mathbb{P}(N_1(t_j)-N_1(t_{j-1})=m_j,1\leqslant j\leqslant 2)=\mathbb{P}(N_1(t_1)=m_1)\cdot\mathbb{P}(N_1(t_2)-N_1(t_1)=m_2)$$

所以 $N_1(t_j)-N_1(t_{j-1}), j=1,2$ 相互独立.

$\square$

3.4.2 叠加

Theorem 3.7

$\{N_j(t), t\geqslant 0\}\sim \mathrm{PP}(\lambda_j), 1\leqslant j\leqslant k$, 相互独立.

$$\left(\sum_{j=1}^{k} N_j(t)\right)_{t\geqslant 0} \sim \text{Poisson}\left(\sum_{j=1}^{k}\lambda_j\right) \tag{3.13}$$

证明. $k=2$, 令 $N(t)=N_1(t)+N_2(t)$.

1. $N(0)=0$
2. $(N_1(t))_{t\geqslant 0} \perp\!\!\!\perp (N_2(t))_{t\geqslant 0}$. 故 $\forall 0=t_0<t_1<\cdots<t_n$, 有 $(N_1(t_1),\cdots,N_1(t_n))\perp\!\!\!\perp(N_2(t_1),\cdots,N_2(t_n))$.

$$\begin{aligned}&(N_1(t_1), N_1(t_2)-N_1(t_1),\cdots,N_1(t_n)-N_1(t_{n-1}))\\ &\perp\!\!\!\perp(N_2(t_1), N_2(t_2)-N_2(t_1),\cdots,N_2(t_n)-N_2(t_{n-1}))\end{aligned}$$

   $\Rightarrow N_1(t_j)-N_1(t_{j-1}), N_2(t_j)-N_2(t_{j-1}), 1\leqslant j\leqslant n$ 相互独立.
   $\Rightarrow (N_1(t_j)-N_1(t_{j-1}), N_2(t_j)-N_2(t_{j-1})), 1\leqslant j\leqslant n$ 相互独立
   $\Rightarrow N(t_j)-N(t_{j-1})=(N_1(t_j)-N_1(t_{j-1}))+(N_2(t_j)-N_2(t_{j-1})), 1\leqslant j\leqslant n$ 相互独立
3. $N(t+s)-N(s)=(N_1(t+s)-N_1(s))+(N_2(t+s)-N_2(s))\sim \text{Poisson}(\lambda_1 t+\lambda_2 t)$

$\square$

Example 3.1 (Durrett Exa 2.12, A Poisson Race)

Given a Poisson process of red arrivals with rate λ and an independent Poisson process of green arrivals with rate μ, what is the probability that we will get 6 red arrivals before a total of 4 green ones?

解. $(N_r(t))_{t\geqslant 0}\sim \mathrm{PP}(\lambda), (N_g(t))_{t\geqslant 0}\sim\mathrm{PP}(\mu)$ 相互独立, 问: $\mathbb{P}(T_4^g>T_6^r)$?
$N(t)=N_r(t)+N_g(t)\sim \mathrm{PP}(\lambda+\mu)$.
(Step 1) $A=\{T_4^g>T_6^r\}, B=\{[0,T_9]$之间至少有6个红队队员$\}$. Claim: $A=B$.

1. $(B\subseteq A)$ $[0,T_9]$ 之间红 $\geqslant 6$ 个, 绿 $\leqslant 3$ 个. $T_6^r\leqslant T_9\leqslant T_4^g$.
2. $(A\subseteq B, B^c\subseteq A^c)$ $[0,T_9]$ 之间红 $\leqslant 5$ 个, 绿 $\geqslant 4$ 个. $T_6^r\geqslant T_9\geqslant T_4^g$.

以上用叠加. 便于下面使用稀释的定理.
(Step 2) 将 $(N(t))_{t\geqslant 0}$ 按照 $(Y_i)_{i\geqslant 1}$ 稀释. 其中 $(Y_i)_{i\geqslant 1}$ iid 且 $(Y_i)_{i\geqslant 1}\perp\!\!\!\perp(N(t))_{t\geqslant 0}$, $\mathbb{P}(Y_i=r)=\lambda/(\lambda+\mu)$. 得到

1. $\tilde{N}_r(t)=\sum_{i=1}^{N(t)}\mathbb{I}_{\{Y_i=r\}}\sim\mathrm{PP}(\lambda)$. $\tilde{N}_g(t)=\sum_{i=1}^{N(t)}\mathbb{I}_{\{Y_i=g\}}\sim\mathrm{PP}(\mu)$. 两者相互独立.
2. $N(t)=\tilde{N}_r(t)+\tilde{N}_g(t)$

于是 $(T_4^g, T_6^r) \overset{(d)}{=} (\tilde{T}_4^g, \tilde{T}_6^r)$. 由 $\{T_n \leqslant t\} = \{N(t) \geqslant n\}$, 将 T 改写成 N, 则 $T, \tilde{T}$ 同分布, $N, \tilde{N}$ 同分布.

$$
\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(\tilde{T}_4^g > \tilde{T}_6^r) \\
&= \mathbb{P}(B) = \mathbb{P}(\tilde{N}_r(T_9) \geqslant 6) \\
&= \mathbb{P}(\sum_{i=1}^{N(T_9)=9} \mathbb{I}_{\{Y_i = r\}} \geqslant 6) \\
&= \sum_{k=6}^{9} \binom{9}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k}
\end{aligned}
$$

□

3.4.3 条件分布

Theorem 3.8 (到达时刻的条件分布)
$\{N(t), t \geqslant 0\} \sim \text{Poisson}(\lambda), \{T_k, k \geqslant 1\}$ 为其的到达时刻序列, 则 $\forall n \geqslant 1$, 有

$$
(T_1, \cdots, T_n | N(t) = n) \overset{(d)}{=} (V_1, \cdots, V_n) \tag{3.14}
$$

其中 $V_1 \leqslant V_2 \leqslant \cdots \leqslant V_n$ 是 $\{U_k, 1 \leqslant k \leqslant n\} \overset{\text{iid}}{\sim} \text{Uniform}[0, t]$ 的重排.

证明. Claim:

$$
f_{(T_1, \cdots, T_n | N(t) = n)}(x_1, \cdots, x_n) = \frac{n!}{t^n} \cdot \mathbb{I}_{\{0 < x_1 < x_2 < \cdots < x_n\}} \tag{3.15}
$$

为了让 $T_1 = t_1, T_2 = t_2, \cdots, T_n = t_n, N(t) = n$, 则 $\tau_1 = t_1, \tau_2 = t_2 - t_1, \cdots, \tau_n = t_n - t_{n-1}, \tau > t - t_n$.

$$
f_{T_1, \ldots, T_n}(x_1, \ldots, x_n) = \prod_{k=1}^{n} \lambda e^{-\lambda \tau_k} \cdot e^{-\lambda(t - t_n)} = \lambda^n e^{-\lambda t}
$$

$$
\mathbb{P}(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}
$$

相除得到 (3.15).

$$
f_{(V_1, \cdots, V_n)}(y_1, \cdots, y_n) = \frac{n!}{t^n} \cdot \mathbb{I}_{\{0 < y_1 < y_2 < \cdots < y_n\}} \tag{3.16}
$$

$S_n := \{\sigma : \{1, 2, \cdots, n\} \to \{1, 2, \cdots, n\} | \sigma \text{双射}\}$, $A_k := (x_k, y_k], k = 1, 2, \cdots, n$.
其中 $0 < x_1 < y_1 < x_2 < y_2 < \cdots < x_n < y_n \leqslant t$.

$$
\begin{aligned}
\mathbb{P}(V_1 \in A_1, \cdots, V_n \in A_n) &= \sum_{\sigma \in S_n} \mathbb{P}(V_1 \in A_1, \cdots, V_n \in A_n | U_1 \in A_{\sigma(1)}, \cdots, U_n \in A_{\sigma(n)}) \\
&\quad \times \mathbb{P}(U_1 \in A_{\sigma(1)}, \cdots, U_n \in A_{\sigma(n)}) \\
&\overset{\text{iid}}{=\!=} \sum_{\sigma \in S_n} \mathbb{P}(U_1 \in A_{\sigma(1)}) \cdots \mathbb{P}(U_n \in A_{\sigma(n)}) \\
&= n! \prod_{k=1}^{n} \left(\frac{y_k - x_k}{t}\right) \overset{?}{\to} f_{(V_1, \cdots, V_n)}(y_1, \cdots, y_n)
\end{aligned}
$$

得到 (3.16). □

Theorem 3.9

若 $0 < s < t, 0 \leqslant m \leqslant n$, 则

$$\mathbb{P}(N(s) = m | N(t) = n) = \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m} \tag{3.17}$$

即 $(N(s)|N(t) = n) \sim \text{Binomial}(n, \frac{s}{t})$

证明. $N(s) = \max\{n | T_n \leqslant s\} = \sum_{k=1}^{N(s)} \mathbb{I}_{\{T_k \leqslant s\}}$.

$$\begin{aligned}\mathbb{P}(N(s) = m | N(t) = n) &= \mathbb{P}(\sum_{k=1}^{N(s)} \mathbb{I}_{\{T_k \leqslant s\}} = m | N(t) = n) \\ &\overset{\text{(i)}}{=} \mathbb{P}(\sum_{k=1}^{n} \mathbb{I}_{\{V_k \leqslant s\}} = m) \\ &\overset{\text{(ii)}}{=} \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}\end{aligned}$$

其中

(i) 由 $(\mathbb{I}_{\{T_1 \leqslant s\}}, \cdots, \mathbb{I}_{\{T_n \leqslant s\}} | N(t) = n) \xlongequal[\text{Thm (3.8)}]{(d)} (\mathbb{I}_{\{V_1 \leqslant s\}}, \cdots, \mathbb{I}_{\{V_n \leqslant s\}})$.

(ii) 由 $\mathbb{I}_{\{V_k \leqslant s\}} \sim \text{Bernoulli}(s/t)$ 相互独立. □

4 更新过程

4.1 定义

Definition 4.1

设 $\{\tau_k, k \geqslant 1\} \overset{\text{iid}}{\sim} F(\cdot)$ 为非负随机变量列, 即 $\mathbb{P}(\tau_1 \leqslant t) = F(t)$, 其中 $F(0) = \mathbb{P}(\tau_1 \leqslant 0) = 0$ 则 $0 < \mathbb{E}\tau_1 < \infty$. 令 $T_n = \sum_{k=1}^{n} \tau_k, n \geqslant 1, T_0 = 0$, 则称由 $N(t) = \max\{n | T_n \leqslant t\}, t > 0$ 定义的计数过程为更新过程.

注:

1. τ_k: 第 k 个灯泡的寿命/更新时间间隔序列
2. T_n: 第 n 个灯泡损坏的时刻/第 n 次更新的时刻
3. $N(t)$: $[0,t]$ 中灯泡的损坏个数/更新的次数

$$N(t) = \sum_{n=1}^{\infty} n\mathbb{I}_{\{N(t)=n\}} = \sum_{n=1}^{\infty} n\mathbb{I}_{[T_n, T_{n+1})}(t)$$

Lemma 4.1

$\forall t \geqslant 0$, 有 $N(t) < \infty$ a.s. (almost surely/几乎必然/几乎处处), 即存在零测集 $\tilde{\Omega}^c$, s.t. $\forall \omega \in \tilde{\Omega}$, 有 $N(t)(\omega) < +\infty$, 即 $\boxed{\mathbb{P}(N(t) < +\infty) = 1}$.
注: 这样写的前提是 $N(t) < \infty$ 是可测集. 因为 $N(t)$ 是随机变量, 该前提成立.

注:

$$\mathbb{P}(N(t) = +\infty) = \lim_{n\to+\infty} \mathbb{P}(\underbrace{N(t) \geqslant n}_{\iff T_n \leqslant t}) = \lim_{n\to+\infty} F^{*n}(t)$$

其中 F^{*n} 为 F 的 n 重卷积.

Theorem 4.1 (强大数定律)

$\{X_k, k \geqslant 1\}$ iid, $\mathbb{E}|X_1| < +\infty$, 令 $S_n = \sum_{k=1}^{n} X_k$, 则 $\frac{S_n}{n} \xrightarrow[n\to\infty]{\text{a.s.}} \mathbb{E}X_1$, 即存在零测集 $\tilde{\Omega}^c$, s.t. $\forall \omega \in \tilde{\Omega}$, 有 $\lim\limits_{n\to\infty} \frac{S_n}{n}(\omega) \xrightarrow[n\to\infty]{\text{a.s.}} \mathbb{E}X_1$, 即 $\mathbb{P}(\lim\limits_{n\to\infty} \frac{S_n}{n} = \mathbb{E}X_1) = 1$.

注: 可列 r.v. 的极限也是 r.v., 而不可列 r.v. 的极限不一定是 r.v.

证明 *Lem 4.1*. 应用 SLLN 知, $T_n/n \xrightarrow[n\to\infty]{\text{a.s.}} \mathbb{E}\tau_1 \in (0, +\infty]$.
存在零测集 $\tilde{\Omega}^c$, s.t. $\forall \omega \in \tilde{\Omega}$, 有

$$\lim_{n\to\infty} \frac{T_n}{n}(\omega) = \mathbb{E}\tau_1$$

$\because \mathbb{E}\tau_1 > 0, T_n(\omega) \approx n\mathbb{E}\tau_1$. $\therefore 0 \leqslant T_n \uparrow \Rightarrow \forall \omega \in \tilde{\Omega}$, 有 $\lim_{n\to+\infty} T_n(\omega) = +\infty$.
$N_t(\omega) = \max\{n \geqslant 0 | T_n(\omega) \leqslant t\}$, 由于 $T_n(\omega) \to +\infty, \therefore \forall \omega \in \tilde{\Omega}, \forall t \geqslant 0$, 至多只有有限个 n, 使 $T_n(\omega) \leqslant t$, 即至多只有有限个 $T_n(\omega)$ 落在 $[0,t]$ 上.
$\Rightarrow \forall \omega \in \tilde{\Omega}, \forall t \geqslant 0, N_t(\omega) < +\infty, \Rightarrow \forall t \geqslant 0, N(t) < +\infty$ (a.s.) $\square$

4.2 极限定理

4.2.1 更新过程的大数定律

先讲一下过程本身的极限.

Lemma 4.2

$$N(t) \xrightarrow[t\to\infty]{\text{a.s.}} +\infty$$

即存在零测集 $\tilde{\Omega}^c$, s.t. $\forall \omega \in \tilde{\Omega}$, 有 $\lim_{t\to\infty} N(t)(\omega) = +\infty$. 即 $\mathbb{P}(\lim_{t\to\infty} N(t) = +\infty) = 1$.

注: $N(t)$ 非可列 r.v., 但 $\lim_{t\to\infty} N(t)$ 仍为 r.v., 因为 $N(t)$ 关于 t 右连左极 (*cadlag*), 即右连续存在, 左极限存在.

证明. $\lim_{t\to\infty} N(t) < +\infty$, $N(t)$ 关于 $t\uparrow$ 单调上升.

$\Rightarrow \exists M > 0, \text{s.t. } \forall t \geqslant 0, N(t) \leqslant M \Rightarrow T_M \leqslant t < T_{M+1}$

$\Rightarrow \forall t \geqslant 0, T_{M+1} > t$

$\Rightarrow T_{M+1} = +\infty$

$$\begin{aligned}
\mathbb{P}(\lim_{t\to\infty} N(t) < +\infty) &\leqslant \mathbb{P}(\exists n, T_n = +\infty) = \mathbb{P}(\exists n, \tau_n = +\infty) \\
&\leqslant \sum_{n\geqslant 1} \mathbb{P}(\tau_n = +\infty) = \sum_{n\geqslant 1}(1 - \mathbb{P}(\tau_n < +\infty)) \\
&= \sum_{n\geqslant 1}(1 - \lim_{t_n\to\infty} F(t_n)) = 0
\end{aligned}$$

□

Theorem 4.2 (更新过程的 LLN)

$$\frac{N(t)}{t} \xrightarrow[t\to\infty]{\text{a.s.}} \frac{1}{\mathbb{E}\tau_1}$$

即 $\mathbb{P}(\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mathbb{E}\tau_1}) = 1$, (因 $\frac{N(t)}{t}$ cadlag).

证明. 当 $N(t) < +\infty$ 时, $T_{N(t)} \leqslant t < T_{N(t)+1}$

当 $0 < N(t) < +\infty$ 时,

$$\begin{aligned}
\boxed{\frac{T_{N(t)}}{N(t)}} \leqslant \frac{t}{N(t)} &< \frac{T_{N(t)+1}}{N(t)} \\
&= \boxed{\frac{T_{N(t)+1}}{N(t+1)}} \cdot \frac{N(t+1)}{N(t)}
\end{aligned} \tag{*1}$$

方框内极限相同.

1. 由 Lem 4.2 知, $\exists \tilde{\Omega}_1^c, \mathbb{P}(\tilde{\Omega}_1^c) = 0, \text{s.t. } \forall \omega \in \tilde{\Omega}_1, \lim_{t\to\infty} N(t)(\omega) = +\infty$.

$$\omega \in \tilde{\Omega}_1, \lim_{t\to\infty} \frac{N(t)+1}{N(t)}(\omega) = 1$$

2. 由 SLLN 知, $\frac{1}{n}T_n \xrightarrow[n\to\infty]{\text{a.s.}} \mathbb{E}\tau_1$
   $\Rightarrow \exists \tilde{\Omega}_2^c, \mathbb{P}(\tilde{\Omega}_2^c) = 0, \text{s.t. } \forall \omega \in \tilde{\Omega}_2, \lim_{n\to\infty} \frac{T_n}{n}(\omega) = \mathbb{E}\tau_1$
   $\Rightarrow \forall \omega \in \tilde{\Omega}_1 \cap \tilde{\Omega}_2 = (\tilde{\Omega}_1^c \cup \tilde{\Omega}_2^c)^c$, 有
   $$\lim_{t\to\infty} \frac{T_{N(t)}}{N(t)}(\omega) = \mathbb{E}\tau_1$$
   (这里用 $\epsilon - \delta$ 语言自己写一下)

由 1, 2, (*1) 知, $\frac{t}{N(t)} \xrightarrow[t\to\infty]{\text{a.s.}} \mathbb{E}\tau_1$. □

4.2.2 更新报酬过程及 LLN

Definition 4.2

设 $\{N(t), t \geqslant 0\}$ 为一更新过程. $\{\tau_k, k \geqslant 1\}$ 为其时间间隔序列, 设 $\{r_k, k \geqslant 1\}$ 为一 iid 随机变量列, 则称由 $R(t) := \sum_{k=1}^{N(t)} r_k$ 定义的过程 $\{R(t), t \geqslant 0\}$ 为更新报酬过程.

注:

1. r_k: 第 k 次更新时刻 (T_k) 的报酬/花费
2. $R(t)$: $[0, t]$ 中总报酬 ($[0, T_{N(t)}]$, 忽略 $(T_{N(t)}, t]$).

Theorem 4.3 (更新报酬过程的 LLN)

设 $\mathbb{E}|r_1| < \infty, \mathbb{E}|\tau_1| \in (0, +\infty)$, 则
$$\frac{R(t)}{t} \xrightarrow[t\to\infty]{\text{a.s.}} \frac{\mathbb{E}r_1}{\mathbb{E}\tau_1}$$

证明.
$$\frac{R(t)}{t} = \frac{\sum_{k=1}^{N(t)} r_k}{N(t)} \cdot \frac{N(t)}{t} \tag{*}$$

1. 由 SLLN 知, $\boxed{\frac{1}{n}\sum_{k=1}^{n} r_k} \xrightarrow[n\to\infty]{\text{a.s.}} \mathbb{E}r_1$. 又 Lem 4.2 知, $\boxed{N(t)} \xrightarrow[t\to\infty]{\text{a.s.}} +\infty$. 所以 $\sum_{k=1}^{N(t)} r_k / N(t) \xrightarrow[t\to\infty]{\text{a.s.}} \mathbb{E}r_1$
2. 由 Thm 4.2 知,
$$\frac{N(t)}{t} \xrightarrow[t\to\infty]{\text{a.s.}} \frac{1}{\mathbb{E}\tau_1}$$

由 1, 2, (*) 知
$$\frac{R(t)}{t} \xrightarrow[t\to\infty]{\text{a.s.}} \frac{\mathbb{E}r_1}{\mathbb{E}\tau_1}$$
□

Example 4.1 (长远看汽车的费用)

1. 一辆车的寿命 (首次发生故障的事件) X_k: 密度函数 h, $\{X_k, k \geqslant 1\}$ iid
2. 更新间隔时间 (买新车): $\tau_k = X_k \wedge T$
3. 第 k 次更新产生的花费: $r_k = A(\text{买新车}) + B\mathbb{I}_{\{\tau_k \leqslant T\}}$

问: 长远来看, 单位时间的花费是多少? (用 LLN)

$$\frac{\sum_{k=1}^{N(t)} r_k}{t} \leqslant \frac{[0,t]\text{间的花费}}{t} \leqslant \frac{\sum_{k=1}^{N(t)+1} r_k}{t}$$

从左到右分别对应时间 $[0, T_{N(t)}], [0,t], [0, T_{N(t)+1}]$

解. 令 $R(t) := \sum_{k=1}^{N(t)} r_k$, 其中 $N(t) = \max\{n | \sum_{k=1}^{n} \tau_k \leqslant t\}$ (买新车的更新过程). 欲求 $\lim_{t\to\infty} \frac{R(t)}{t}$.

$$\begin{aligned}\mathbb{E}\tau_1 = \mathbb{E}(X_1 \wedge T) &= \int_0^{+\infty} (x \wedge T) h(x) dx \\ &= \int_0^T x h(x) dx + \int_T^{+\infty} T h(x) dx\end{aligned}$$

$$\mathbb{E}r_1 = A + B\mathbb{E}\mathbb{I}_{\{\tau_1 < T\}} = A + B\mathbb{E}\mathbb{I}_{\{X_1 < T\}} = A + B\int_0^T h(x) dx$$

由 Thm 4.3 知,

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{A + B\int_0^T h(x)dx}{\int_0^{+\infty} (x \wedge T) h(x) dx}, \text{a.s.}$$

□

Theorem 4.4 (Ross, Thm 3.6.1)

设 $\{(r_k, \tau_k), k \geqslant 1\}$ 是 iid 的,

$$\lim_{t\to\infty} \frac{\mathbb{E}R(t)}{t} = \frac{\mathbb{E}r_1}{\mathbb{E}\tau_1}$$

LHS 是数的极限, 不是随机变量的极限. 其中 $\frac{\mathbb{E}R(t)}{t}$ 被称为更新报酬函数.

注: $\{(\tau_k, r_k), k \geqslant 1\}$ 相互独立 $\Rightarrow$

1. $\{\tau_k, k \geqslant 1\}$ 相互独立, $\{r_k, k \geqslant 1\}$ 相互独立
2. $\{\tau_k, 1 \leqslant k \leqslant n\} \perp\!\!\!\perp \{r_k, k \geqslant n+1\}$
3. $\{\tau_k\} \perp\!\!\!\perp \{r_j, j \neq k\}$ (r_k 取全集即可)

4.2.3 交替更新过程及 LLN

Definition 4.3

设 $\{(s_k, u_k), k \geqslant 1\}$ 为 iid 的随机变量向量列, $s_k \geqslant 0, u_k \geqslant 0, \forall k \geqslant 1$. 令 $\tau_k := s_k + u_k, k \geqslant 1$. 定义

$$N(t) = \begin{cases} \max\{n | \sum_{k=1}^{n} \tau_k \leqslant t\} & t > 0 \\ 0 & t = 0 \end{cases}$$

称 $\{N(t), t \geqslant 0\}$ 为交替更新过程.

Theorem 4.5 (交替更新过程的 LLN)

设存在分布函数 H 使得 $\tau_k \sim H, \mathbb{E}S_1 \in (0,+\infty), \mathbb{E}u_1 \in (0,+\infty)$, 则

1. $\dfrac{\sum_{k=1}^{N(t)} s_k}{t} \xrightarrow[t\to\infty]{\text{a.s.}} \dfrac{\mathbb{E}s_1}{\mathbb{E}s_1+\mathbb{E}u_1}$
2. $\dfrac{[0,t]\text{中系统处于状态 1 的时长}}{t} \xrightarrow[t\to\infty]{\text{a.s.}} \dfrac{\mathbb{E}s_1}{\mathbb{E}s_1+\mathbb{E}u_1}$, 即系统处于状态 1 的事件比例的极限为 $\dfrac{\mathbb{E}s_1}{\mathbb{E}s_1+\mathbb{E}u_1}$

证明. 1. Thm 4.3 中取 $r_k = s_k$.

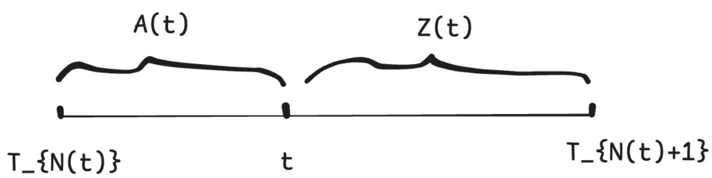
2. 令 $U(t) := [0,t]$ 中系统处于状态 1 的时长

$$\frac{\sum_{k=1}^{N(t)} s_k}{t} \leqslant \frac{u(t)}{t} \leqslant \frac{\sum_{k=1}^{N(t)+1} s_k}{t} = \frac{\sum_{k=1}^{N(t)+1} s_k}{N(t)+1} \cdot \frac{N(t)+1}{N(t)}$$

由 Thm 4.3 得到想要的结果.

□

4.2.4 使用年龄和剩余寿命



$A(t)$ 为使用年龄, $A(t) = t - T_{N(t)}$. $Z(t)$ 为剩余寿命, $Z(t) = T_{N(t)+1} - t$.

Example 4.2

某零件按更新过程 $\{N(t), t \geqslant 0\}$ 替换, 其更新间隔时间序列 $\{\tau_k, k \geqslant 1\}$. 设 $\mathbb{E}\tau_1^2 < \infty$, 且 $0 < \mathbb{E}\tau_1 < \infty$, 则零件的长程平均使用年龄

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t A(s)ds = \frac{\mathbb{E}\tau_1^2}{2\mathbb{E}\tau_1}$$

若是长程平均寿命即把 $A(s)$ 换成 $Z(s)$.

注: $\int_0^t A(s)ds$ 有无意义? 有的, 因为 $N(t)$ 是几乎处处右连左极, 在 $[0,t]$ 内有有限多跳. 并非闭区间内的连续函数才可积 (参考梅加强 [6]).

定义 6.1.1 (Riemann 积分)**.** 设 f 如上, 如果存在实数 I, 使得任给 $\varepsilon > 0$, 均存在 $\delta > 0$, 对任何分割 π, 只要 $\|\pi\| < \delta$, 就有

$$\Big|\sum_{i=1}^{n} f(\xi_i)\Delta x_i - I\Big| < \varepsilon, \quad \forall\ \xi_i \in [x_{i-1}, x_i],\ i = 1, \cdots, n,$$

则称 f 在 $[a, b]$ 上 *Riemann* 可积或可积, I 为 f 在 $[a, b]$ 上的 (定) 积分, 记为

$$I = \int_a^b f(x)\, dx = \lim_{\|\pi\|\to 0} \sum_{i=1}^{n} f(\xi_i)\Delta x_i.$$

其中 f 称为被积函数, $[a, b]$ 称为积分区间, a, b 分别称为积分下限与积分上限.

图 10: 定义 6.1.1-Riemann 积分

证明. 由 $T_{N(t)} \leqslant t$,

$$\begin{aligned}
\int_0^t A(s)ds &\geqslant \int_0^{T_{N(t)}} A(s)ds \\
&= \sum_{k=1}^{N(t)} \int_{T_{k-1}}^{T_k} (s - T_{N(s)})ds \\
&= \sum_{k=1}^{N(t)} \int_{T_{k-1}}^{T_k} (s - T_{k-1})ds \\
&= \sum_{k=1}^{N(t)} \frac{1}{2}(s - T_{k-1})^2 \Big|_{T_{k-1}}^{T_k} \\
&= \sum_{k=1}^{N(t)} \frac{1}{2}(T_k - T_{k-1})^2 = \frac{1}{2}\sum_{k=1}^{N(t)} \tau_k^2
\end{aligned}$$

由 Thm 4.3 知,

$$\frac{1}{2t}\sum_{k=1}^{N(t)} \tau_k^2 \xrightarrow[t\to\infty]{\text{a.s.}} \frac{\mathbb{E}\tau_1^2}{2\mathbb{E}\tau_1}$$

同理,

$$\begin{aligned}
\int_0^t A(s)ds &\leqslant \int_0^{T_{N(t)+1}} A(s)ds \\
&= \sum_{k=1}^{N(t)+1} \frac{1}{2}(T_k - T_{k-1})^2 \\
&= \frac{1}{2}\sum_{k=1}^{N(t)+1} \tau_k^2
\end{aligned}$$

由 Thm 4.3 知,

$$\frac{1}{2t}\sum_{k=1}^{N(t)} \tau_k^2 \xrightarrow[t\to\infty]{\text{a.s.}} \frac{\mathbb{E}\tau_1^2}{2\mathbb{E}\tau_1}$$

□

5 连续时间马氏链

5.1 定义

S 至多为可数集 $(S \subseteq \mathbb{N})$

Definition 5.1 (马氏性)
称 S 值的过程 $\{X_t, t \geqslant 0\}$ 具有马氏性, 若对 $\forall n \geqslant 1, i, j, i_0, \cdots, i_n \in S, 0 \leqslant s_0 < s_1 < \cdots < s_n < s, t > 0$, 有

$$\mathbb{P}(X_{t+s} = j | X_s = i, X_{s_k} = i_k, 0 \leqslant k \leqslant n) = \mathbb{P}(X_{t+s} = j | X_s = i) \tag{5.1}$$

其中, 称 $\mathbb{P}(X_{t+s} = j | X_s = i) = p_{s,s+t}(i, j)$ 为转移概率.

Definition 5.2 (马氏链)
称 $\{X_t, t \geqslant 0\}$ 是一个连续时间马氏链 (CTMC), 若其具有马氏性, 且轨道右连续.

Definition 5.3 (时齐性)
若 CTMC, $\{X_t, t \geqslant 0\}$ 的转移概率具有时齐性,

$$p_{s,s+t}(i, j) = p_{0,t}(i, j) \quad \forall i, j \in S, s, t \geqslant 0$$

称过程是时齐的, 并简记 $p_t(i, j) := p_{s,s+t}(i, j)$.

Definition 5.4 (正则性/标准的)
称 $\{X_t, t \geqslant 0\} \sim$ CTMC 具有正则性, 若

$$\lim_{t \to 0^+} p_t(i, j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

DTMC 的最小步长为 1, 因此称一步转移概率矩阵为转移概率矩阵. 但 CTMC 不存在最小步长, 转移概率矩阵如何给出?

Definition 5.5
称矩阵族 $\{P_t = (p_t(i, j))_{i,j \in S}, t \geqslant 0\}$ 为过程的 "转移半群".

Theorem 5.1

$\{P_t, t \geqslant 0\}$ 为随机半群, 即

1. $\lim_{t\to 0^+} P_t = P_0 = \mathbb{I}_{S\times S}$ (单位阵)
2. 对每一个 P_t 都是一个随机矩阵
3. (C-K 方程/半群性质) 对 $\forall s,t \geqslant 0, P_{s+t} = P_s P_t$, 即

$$p_{s+t}(i,j) = \sum_{k\in S} p_s(i,k)p_t(k,j)$$

证明. 证明结论 (3).

$$\begin{aligned} p_{t+s}(i,j) &= \mathbb{P}(X_{t+s} = j | X_0 = i) \\ &= \sum_{k\in S} \mathbb{P}(X_{t+s} = j, X_s = k | X_0 = i) \\ &= \sum_{k\in S} \mathbb{P}(X_{t+s} = j | X_s = k)\mathbb{P}(X_s = k | X_0 = i) \\ &= \sum_{k\in S} p_s(i,k)p_t(k,j) \end{aligned}$$

□

Example 5.1

Poisson 过程是时齐 CTMC.

$\mathbb{P}(N_{t+s} = j | N_s = i) = \mathbb{P}(N_{t+s} - N_s = j - i | N_s = i) = \mathbb{P}(N_t = j - i)$

$$p_t(i,j) = \begin{cases} e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} & j \geqslant i \\ 0 & j < i \end{cases}$$

Example 5.2

$\{N(t), t \geqslant 0\} \sim \mathrm{PP}(\lambda) \perp\!\!\!\perp \{Y_n, n \geqslant 1\} \sim \mathrm{DTMC}(u((i,j))_{i,j\in S})$, $\{X_t := Y_{N(t)}, t \geqslant 0\} \sim \mathrm{CTMC}$

证明. Claim 1: $\mathbb{P}(X_{t+s} = j | X_s = i) = \sum_{n\geqslant 0} e^{-\lambda t}\frac{(\lambda t)^n}{n!} \cdot u^{(n)}(i,j)$

先固定 $N(t)$. $\mathbb{P}(X_{t+s}=j,X_s=i|N(s)=m,N(t+s)=\tilde{m})=\mathbb{P}(Y_{\tilde{m}}=j,Y_m=i),0\leqslant m\leqslant\tilde{m}$.

$$\begin{aligned}\mathbb{P}(X_{t+s}=j,X_s=i) &= \sum_{0\leqslant m\leqslant\tilde{m}}\mathbb{P}(Y_{\tilde{m}},Y_m=i)\cdot\mathbb{P}(N(s)=m,N(t+s)=\tilde{m})\\ &= \sum_{0\leqslant m\leqslant\tilde{m}}\mathbb{P}(Y_{\tilde{m}}|Y_m=i)\mathbb{P}(Y_m=i)\mathbb{P}(N(s)=m)\mathbb{P}(N(t+s)-N(s)=\tilde{m}-m)\\ &= \sum_{0\leqslant m\leqslant\tilde{m}}\left[u^{(\tilde{m}-m)}(i,j)e^{-\lambda t}\frac{(\lambda t)^{\tilde{m}-m}}{(\tilde{m}-m)!}\right]\mathbb{P}(Y_{N(s)}=i,N(s)=m)\\ &= \sum_{m\geqslant 0}\left(\sum_{n\geqslant 0}u^{(n)}(i,j)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\right)\mathbb{P}(Y_{N(s)}=i,N(s)=m)\\ &= \sum_{n\geqslant 0}u^{(n)}(i,j)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\mathbb{P}(X_s=i)\end{aligned}$$

Claim 2: $\forall n\geqslant 1,i,j,i_0,\cdots,i_n\in S,0\leqslant s_0<s_1<\cdots<s_n<s,t>0$, 有

$$\mathbb{P}(X_{t+s}=j|X_s=i,X_{s_k}=i_k,0\leqslant k\leqslant n)=\sum_{n\geqslant 0}e^{-\lambda t}\frac{(\lambda t)^n}{n!}u^{(n)}(i,j)$$

□

5.2 转移速率矩阵与转移概率的计算

5.2.1 转移概率的连续性

Proposition 5.1

设 $\{P_t, t \geqslant 0\}$ 是 S 上的随机半群, 则

1. $\forall t \geqslant 0, i \in S$, 有 $p_t(i,i) > 0$
2. 若 $\exists s > 0$, 使 $p_s(i,i) = 1$, 则 $\forall t \geqslant 0, p_t(i,i) = 1$
3. 若 $\exists s > 0$, 使 $p_s(i,i) > 0$, 则 $\forall t \geqslant s, p_t(i,i) > 0$

证明. 1. $\lim_{t\to 0^+} p_t(i,i) = 1 \Rightarrow \exists \delta > 0, \text{s.t. } \forall t \in [0,\delta], p_t(i,i) > 0$

对 $\forall t \geqslant 0, \exists s \in [0,\delta), n \in \mathbb{N}$, 有 $t = s + n\delta$. 故由 C-K 方程,

$$\begin{aligned} p_t(i,i) &\geqslant p_{n\delta}(i,i)p_s(i,i) \\ &\overset{\text{C-K}}{\geqslant} (p_\delta(i,i))^n p_s(i,i) > 0 \end{aligned}$$

2. $p_s(i,i) = 1 \Rightarrow \forall n \in \mathbb{N}$, 有 $p_{ns}(i,i) \overset{\text{C-K}}{\geqslant} (p_s(i,i))^n = 1\ (*1)$

(反证法) 假设 $\exists t > 0, \text{s.t. } p_t(i,i) < 1 \Rightarrow \sum_{k\neq i} p_t(i,i) > 0. \exists j \neq i, \text{s.t. } p_t(i,j) > 0(*2)$.

取 n 使 $ns \geqslant t$, 则

$$\begin{aligned} p_{ns}(i,i) &= 1 - \sum_{k\neq i} p_{ns}(i,k) \\ &\leqslant 1 - p_{ns}(i,j) \\ &\overset{\text{C-K}}{\leqslant} 1 - p_t(i,j)\underbrace{p_{ns,t}(j,j)}_{>0,\ \text{by (1)}} \\ &< 1 \quad \text{by } (*2),(1) \end{aligned}$$

与 $(*1)$ 矛盾.

3. $p_s(i,j) > 0, p_t(i,j) \geqslant p_s(i,j)p_{t-s}(j,j) \overset{(1)}{>} 0, t \geqslant s$.

□

Theorem 5.2

$p_s(i,j)$ 关于 $t \geqslant 0$ 一致连续, 且对 $t \geqslant s \geqslant 0, i, j \in S$.

$$|p_t(i,j) - p_s(i,j)| \leqslant 1 - p_{t-s}(i,i) \tag{5.2}$$

证明.

$$\begin{aligned} p_t(i,j) - p_s(i,j) &\xlongequal{\text{C-K}} \sum_{k\in S} p_{t-s}(i,k)p_s(k,j) - p_s(i,j) \\ &= \sum_{k\neq i} p_{t-s}(i,k)p_s(k,j) - [1 - p_{t-s}(i,i)]p_s(i,j) \\ &=: I_1 - I_2 \end{aligned}$$

$I_1 \leqslant \sum_{k\neq i} p_{t-s}(i,k) = 1 - p_{t-s}(i,i). I_2 \leqslant 1 - p_{t-s}(i,i)$.

$\Rightarrow |\,\mathrm{LHS}\,| \leqslant I_1 \vee I_2 \leqslant 1 - p_{t-s}(i,i)$ □

5.2.2 转移概率的可微性与 Kolmogorov 方程

由 Prop 和 Thm 可证可微性, 但即便如此证明也不是件简单的事, 因此只需要知道下述定理存在即可.

Theorem 5.3

P_t 在 $t=0$ 处右导数存在, 具体地,

1. $\forall i \in S$, 下列极限存在.

$$\lim_{t\to 0^+} \frac{p_t(i,i)-1}{t} = q(i,i) := -\sup_{t\geqslant 0} \frac{1-p_t(i,i)}{t} \in [-\infty, 0]$$

2. $\forall j \neq i$, 下列极限存在

$$\lim_{t\to 0^+} \frac{p_t(i,j)}{t} = q(i,j) \in [0, +\infty)$$

3. $\forall i \in S, \sum_{k\neq i} q(i,k) \leqslant -q(i,i)$

Definition 5.6 (密度矩阵/Q 矩阵)

称 S 上的矩阵 $Q = (q(i,j))_{i,j\in S}$ 为密度矩阵, 若

1. $\forall i \in S, q_i := -q(i,i) \in [0, +\infty]$
2. $\forall i,j \in S, q(i,j) \in [0, +\infty)$
3. $\forall i \in S, \sum_{k\neq i} q(i,k) \leqslant q_i$

Definition 5.7

由 Thm 5.3 知, $\{P_t, t\geqslant 0\}$ 在 0 处的右导数矩阵. $Q := P_0' = (p_0'(i,j))_{i,j\in S}$ 存在且其为密度矩阵. 称此 Q 为 $\{P_t, t\geqslant 0\}$ 或 X 的转移速率矩阵/无穷小生成元.

Definition 5.8

若 $\forall i, q_i = |q(i,i)| = \sum_{k\neq i} q(i,j) < \infty$ 称密度矩阵 Q 为保守的.

CTMC $\to$ 转移半群 $\to$ 转移速率矩阵. 问: 已知 Q, 能否得到 CTMC? 这个问题类比 DTMC 则为: 已知转移矩阵, 能否得到 MC?

Theorem 5.4 (Kolmogorov 向前/向后方程)

对具有保守的 Q 的随机半群 $\{P_t, t\geqslant 0\}$ 有

$$\begin{cases} P_t' = QP_t & \text{向后} \\ P_0 = I \end{cases} \qquad \begin{cases} P_t' = P_tQ & \text{向前} \\ P_0 = I \end{cases}$$

注: $AB \neq BA, QP_t = P_tQ$.

$$\text{CTMC} \overset{+\mu_0}{\Longleftrightarrow} \{P_t, t \geqslant 0\} \overset{P_0'}{\Rightarrow} Q \text{ (保守的)}$$

反之, 保守的 $Q \Rightarrow \{P_t, t \geqslant 0\}$
Kolmogorov 方程存在唯一性的解, 可以由 Q 构造 CTMC. 侯氏定理.

Theorem 5.5
在适当的正则性条件下, 有 Q 的向前方程与向后方程存在唯一解, 即 CTMC, $\{P_t, t \geqslant 0\}$ + 初始分布μ_0, Q 一一对应.

Example 5.3 (例 4.7)
$\{N(t), t \geqslant 0\}\,\mathrm{PP}(\lambda)$.

$$p_t(i,j) = \begin{cases} e^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!} & j \geqslant i \\ 0 & j < i \end{cases}$$

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & \cdots & 0 \\ 0 & -\lambda & \lambda & 0\cdots & 0 & \\ \vdots & \vdots & -\lambda & \lambda & \cdots & \\ \vdots & \vdots & \vdots & \cdots & \vdots & \end{pmatrix}$$

即 $Q(i,i) = -\lambda, Q(i,i+1) = \lambda$.

此前并未限定状态空间有限, 当状态空间有限时:

Corollary 5.1
设状态空间 S 有限, 则

$$P_t = e^{tQ} = \sum_{n \geqslant 0} \frac{(tQ)^n}{n!}$$

Example 5.4 (两状态的 MC)

$$\mathbf{Q} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \end{array} \begin{array}{c} \begin{array}{cc} \mathbf{1} & \mathbf{2} \end{array} \\ \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \end{array}$$

$\lambda, \mu > 0$, 求 P_t.

解. 由 Kolmogorov 方程, $P_t' = QP_t$, 即

$$\begin{pmatrix} p_t'(1,1) & p_t'(1,2) \\ p_t'(2,1) & p_t'(2,2) \end{pmatrix} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \begin{pmatrix} p_t(1,1) & p_t(1,2) \\ p_t(2,1) & p_t(2,2) \end{pmatrix}$$

$$\begin{cases} p_t'(1,1) = -\lambda p_t(1,1) + \lambda p_t(2,1) \\ p_t'(2,1) = \mu p_t(1,1) - \mu p_t(2,1) \end{cases}$$

$$\begin{cases} (p_t(1,1) - p_t(2,1))' = -(\lambda+\mu)(p_t(1,1) - p_t(2,1)) \\ p_0(1,1) - p_0(2,1) = 1 - 0 = 1 \end{cases}$$

$p_t(1,1) - p_t(2,1) = e^{-(\lambda+\mu)t}$. 代回

$$\begin{cases} p_t'(1,1) = -\lambda e^{-(\lambda+\mu)t} \\ p_0(1,1) = 1 \end{cases}, \quad \begin{cases} p_t'(2,1) = -\mu e^{-(\lambda+\mu)t} \\ p_0(2,1) = 0 \end{cases}$$

$$\begin{aligned} p_t(1,1) &= \int_0^t -\lambda e^{-(\lambda+\mu)s} ds + 1 \\ &= \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)s}\Big|_0^t + 1 = \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu} \end{aligned}$$

(书上答案有误, 注意)

$$\begin{aligned} p_t(2,1) &= \int_0^t \mu e^{-(\lambda+\mu)s} ds + 0 \\ &= \frac{-\mu}{\lambda+\mu} e^{-(\lambda+\mu)s}\Big|_0^t = \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} \end{aligned}$$

□

5.2.3 轨道的跳跃性质

探讨 Q 表示了马氏链的什么.

Theorem 5.6

设 S 值右连续马氏链 $X := \{X_t, t \geqslant 0\}$ 具有保守的转移速率矩阵 $Q = (q(i,j))_{i,j\in S}$. 定义首跳时间 $\eta := \inf\{t > 0 | X_t \neq X_0\}$. 其中 $\inf\varnothing = +\infty$. 那么, 对 $\forall t \geqslant 0$, 有 $\mathbb{P}_i(\eta > t) = \mathbb{P}(\eta > t | X_0 = i) = e^{-q_i \cdot t}$, 即在 $\mathbb{P}_i = \mathbb{P}(\cdot | X_0 = i)$ 下, $\eta \sim \mathrm{EXP}(q_i)$, 其中 $q_i = -q(i,i)$.

Corollary 5.2

若 $q(i,i) = 0$, 则 $\mathbb{P}_i(\eta > t) = 1, \forall t \geqslant 0$.

由 $\{\eta = \infty\} = \cap_{n\geqslant 1}\{\eta > n\}, \mathbb{P}_i(\eta = \infty) = 1$.

$$1 = \mathbb{P}(\eta = \infty) = \mathbb{P}(X_t = X_0, \forall t > 0 | X_0 = i) = \mathbb{P}(X_t = i, \forall t \geqslant 0 | X_0 = i)$$

即 i 为吸收态.

证明 *Thm 5.6*. (Step 1) 先证明一个数分结论.

$$\begin{aligned}\lim_{s\to 0^+}(p_{st}(i,i))^{1/s} &= \lim_{s\to 0^+}\exp\left(\frac{\ln p_{st}(i,i)}{st}\cdot t\right)\\ &\xlongequal{\ln x\sim x-1(x\to 1)} \exp\left(\lim_{s\to 0^+}\frac{p_{st}(i,i)-1}{st}\cdot t\right)\\ &= \exp(tq(i,i)) = e^{-q_i t}\end{aligned}$$

(Step 2) Claim: $\mathbb{P}(\eta > t|X_0 = i) = \mathbb{P}(X_s = i, \forall s \in [0,t]|X_0 = i)$

"$\Rightarrow$" $t < \eta = \inf\{t > 0|X_t \neq i\}, X_0 = i.\forall s \in [0,t], X_s = i$

"$\Leftarrow$" $X_s = i, \forall s \in [0,t] \Rightarrow X_t = i \xrightarrow{\text{右连续}} \exists \delta > 0, \forall u \in [t, t+\delta], \text{s.t. } X_u = i, \eta \geqslant t+\delta > t$.

(Step 3)

$$\begin{aligned}\mathbb{P}_i(\eta > t) &= \mathbb{P}_i(X_s = i, \forall s \in [0,t])\\ &\xlongequal{\text{右连续}} \mathbb{P}_i\Big(\bigcap_{n\geqslant 1}\{X_{kt/2^n} = i, k = 0,1,\cdots,2^n\}\Big) \quad (*)\\ &= \lim_{n\to\infty}\mathbb{P}_i(X_{kt/2^n} = i, k = 0,1,\cdots,2^n)\\ &= \lim_{n\to\infty}(p_{t/2^n}(i,i))^{2^n} = e^{-q_i t}\end{aligned}$$

$(*)$ 是一个数分结论, 回去证明. $\forall s$ 可找到一列 n 逼近. □

令 $T_0 = 0$, 归纳定义 $T_n := \inf\{t > 0|X_t \neq X_{T_{n-1}}\}, n \geqslant 1$. T_1 即上面的 η.

T_n: 第 n 次跳跃时间 $(T_1 = \eta)$

Lemma 5.1

设右连续马氏链 X 具有保守的 Q, 则

1. 在 $[0, \lim_{n\to\infty} T_n)$ 上, X 的轨道为阶梯函数, 即 $X_t = X_{T_n}, t \in [T_n, T_{n+1}]$.
2. 若 $q_i > 0$, 则 $\{X_{T_n}, n \geqslant 1\} \sim \text{DTMC}(\hat{P})$, 其中

$$(\hat{P})_{ij} = \hat{p}_{ij} = \frac{q(i,j)}{q(i)}(1-\delta_{ij})$$

称 $\{X_{T_n}, n \geqslant 1\}$ 为 X 的嵌入链.

(走神)

5.2.4 过程的构造

设 $Q = (q(i,j))_{i,j\in S}$ 为保守的密度矩阵, 其中

$$q_i = -q(i,i) = \sum_{k\neq i} q(i,k)$$

假定 $q(i,i) > 0, \forall i \in S$, 假设没有吸收态. (实际上无需此假设也成立, 但此处为了和教材一致)

定义 $\hat{P} := (\hat{p}_{i,j})_{i,j\in S}$, 其 $\hat{p}_{ij} = \frac{q(i,j)}{q(i)}(1-\delta_{ij})$, 则 $\hat{P}$ 为随机矩阵, 并称其路径矩阵.

两种看过程的角度:

1. $X(t,\cdot)$ r.v. $\forall t \geqslant 0$

2. $X(\cdot,\omega)\in C([0,+\infty))$

设

1. $\{Y_n, n\geqslant 0\}\sim \mathrm{DTMC}(\mu_0,\hat{P})$
2. $\tau_0,\tau_1,\cdots \overset{\text{iid}}{\sim} \mathrm{EXP}(1)$, 注: $\tau_n/q_i\sim \mathrm{EXP}(q_i)$
3. $\{\tau_k, k\geqslant 0\}\perp\!\!\!\perp\{Y_n, n\geqslant 0\}$

下面构造 CTMC.

1. $t=n$ 时, 处于状态 Y_0, $\eta_1=\tau_0/q(Y_0)\sim \mathrm{EXP}(q(i))$ 在 $\mathbb{P}(\cdot|Y_0=i)$, 记 $\eta_1\sim \mathrm{EXP}(Y_0)$.
2. $t=T_1=\eta_1$ 时, 跳到状态 Y_1, 在 Y_1 待了 $\eta_2=\tau_1/q(Y_1)\sim \mathrm{EXP}(Y_1)$
3. $t=T_2=\eta_1+\eta_2$ 时, 跳到状态 Y_2, 在 Y_2 待了 $\eta_3=\tau_2/q(Y_2)\sim \mathrm{EXP}(Y_2)$

由此类推, $\eta_n=\tau_{n-1}/q(Y_{n-1}),\forall n\geqslant 1.T_n=\sum_{k=1}^n\eta_k, T_0=0$.
令 $X_t=Y_n$ (当 $t\in[T_n,T_{n+1}]$), 若

$$\mathbb{P}(\lim_{n\to\infty}T_n=+\infty|X_0=Y_0=i)=1 \tag{5.3}$$

称 X 为以 Q 为速率矩阵的跳过程, 称 Y 为 X 的嵌入链, $\{\eta_k,k\geqslant 1\}$ 为 X 的等待时间序列, T_n 为第 n 次跳的时刻.
注: $\forall t\geqslant 0$, 至多有限个 n, s.t. $T_n\leqslant t$.

Lemma 5.2
以 Q 为速率矩阵的跳过程 X 是一个 CTMC, 且 $X\sim \mathrm{CTMC}(Q)$ 以及 $P_t'=QP_t, P_t'=P_tQ$. 其中 $\{P_t,t\geqslant 0\}$ 是 X 的转移半群.

注: $X\sim\mathrm{CTMC}(\mu_0,(P_t)_{t\geqslant 0})$ (有限维分布族)
$\Rightarrow X\sim \mathrm{CTMC}(\mu_0,Q)$
$\Leftarrow$ Kolmogorov 方程的适定性
$\tilde{X}\sim$ 跳过程$(\mu_0,Q)\sim \mathrm{CTMC}(\mu_0,(\tilde{P}_t)_{t\geqslant 0})$
且 Kolmogorov 方程 $\tilde{P}_t'=\tilde{P}_tQ=Q\tilde{P}_t,\Rightarrow P_t=\tilde{P}_t$, 即 $X\overset{(d)}{=}\tilde{X}$.
(5.3) 不好验证, 有什么好验证的充分条件使其成立吗?

Lemma 5.3
若 $\{q_i,i\in S\}$ 有界, 则条件 (5.3) 成立. 特别地, S 有限, 则条件 (5.3) 成立.

Example 5.5 (纯生过程)
$S=\{1,2,3,\cdots\},\lambda_1,\lambda_2,\cdots$ 为一列正数. 令 $q_i=q(i,i+1)=\lambda_i,\forall i\geqslant 1$, 称出生速率.

$$Q=\begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \cdots & \cdots & \cdots \\ 0 & -\lambda_2 & \lambda_2 & 0 & \cdots & \cdots \\ 0 & 0 & -\lambda_3 & \lambda_3 & 0 & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \end{pmatrix}$$

转移速率图:

$$1 \xrightarrow{\lambda_1} 2 \xrightarrow{\lambda_2} 3 \xrightarrow{\lambda_3} 4 \cdots$$

Claim: (5.3) 成立 $\iff \sum_{n=1}^{\infty} 1/\lambda_n = \infty$.

1. Poisson 过程 PP(λ) 为纯生过程
2. (Durrett Exa 4.5) $q_i = q(i, i+1) = \lambda i^p, \forall i \geqslant 1$

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \frac{1}{\lambda} \sum_{n=1}^{\infty} \frac{1}{n^p} = \begin{cases} < \infty & p > 1 \\ \infty & p \leqslant 1 \end{cases}$$

   $p \in [0, 1]$ 时, 可定义由 Q 为速率矩阵的跳过程. $p = 0$ 时为 Poisson 过程. 特别地, $p = 1$ 时称为 Yule 过程.

3. $\sum_{n=1}^{\infty} 1/\lambda_n = \infty$ 时, 嵌入链 $\{Y_n, n \geqslant 0\} \sim \text{DTMC}(\hat{P})$, 其中

$$(\hat{P})_{i,j} = \hat{p}_{i,j} = \frac{q(i,j)}{q(i)}(1 - \delta_{ij}) \Rightarrow \hat{p}_{i,i+1} = 1$$

Example 5.6 (生灭过程)
$S = \{0, 1, 2, 3, \cdots\}$ 令 $q(i, i+1) = \lambda_i, \forall i \geqslant 0$ (出生速率). $q(i, i-1) = \alpha_i, \forall i \geqslant 1$ (死亡速率), 其他为 0.

$$q_i = -q(i,i) = \sum_{k \neq i} q(i,k) = \begin{cases} \lambda_i + \alpha_i & i \geqslant 1 \\ \lambda_i & i = 0 \end{cases}$$

转移速率图:

参考文献

[1] 强马氏性 v2. 03 2025. [Link].

[2] Rick Durrett. Essentials of Stochastic Processes. 01 1999. doi:10.1007/978-1-4614-3615-7.

[3] Geoffrey Grimmett and David Stirzaker. Probability and random processes. Oxford University Press, Oxford; New York. doi:10.1017/mag.2022.154.

[4] Sidney I. Resnick. Adventures in stochastic processes. Birkhauser Verlag, CHE, 1992. [Link].

[5] Sheldon M Ross. Stochastic processes. John Wiley & Sons, 1995.

[6] 梅加强. 数学分析讲义. 2010.

[7] 菲赫金哥尔茨. 微积分学教程, volume 2. 高等教育出版社, 北京, 8 edition, 2006.