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：2025/03/27

目录

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

$\mathbb{T} = \mathbb{N}_0$: 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 $\{\mu_t\}_{t\in\mathbb{T}}$？

0.1 事件概率

0.1.1 事件域

Definition 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 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 3
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 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 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 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 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

0.1.2 概率测度

Definition 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

$\mathbb{P}$ 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. $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$
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 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 $B_1 = A_1, B_2 = A_2 \cap (A_1)^c, \cdots, B_n = A_n \cap A_1^c \cap A_2^c \cap \cdots \cap A_{n-1}^c$
$B_n \subset A_n$，由可列可加性和推论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

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

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

Property 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

0.2 独立性

Definition 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 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 4
$A\perp\!\!\!\perp B\Rightarrow A\perp\!\!\!\perp B^c, A^c\perp\!\!\!\perp B, A^c\perp\!\!\!\perp B^c$

Definition 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 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 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 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

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

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 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 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 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

$X \perp\!\!\!\perp Y \Leftrightarrow \forall x \in S_X, y \in S_Y$ 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 2

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

用定理1证明

$\Rightarrow$ 已知 $X\perp\!\!\!\perp Y$，由定义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 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 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5，$\sigma(\mathbb{I}_{A_n})=\sigma(A_n)$，所以 (2) $\Leftrightarrow$ (3)
2. 下证 (2) $\to$ (1)，一般到特殊，$A_n\subseteq\sigma(A_n)$
3. 下证 (1) $\to$ (2)，$\sigma(A_n)=\{A_n,A_n^c,\varnothing,\Omega\}$，$\varnothing\perp\!\!\!\perp A_n,\Omega\perp\!\!\!\perp A_n$，由性质4，$\varnothing\perp\!\!\!\perp A_n^c,\Omega\perp\!\!\!\perp A_n^c$

由定理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 □

0.3 条件概率与条件独立

Definition 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 4 (乘法公式)

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

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

Theorem 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 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13

Theorem 7

$\mathbb{P}_B : \mathcal{F} \to [0,1]$ 也是 $(\Omega, \mathcal{F})$ 上的概率测度 [定义5]

Property 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13

Definition 14

称 C 条件发生下，A 与 B 独立，若

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

记为 $A \perp\!\!\!\perp_C B$（条件独立）

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

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

0.4 期望与条件期望

0.4.1 离散随机变量的期望

Definition 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 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 6

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

Example 7

X 是离散随机变量，由定义10，$X = \sum_{x \in S} x\mathbb{I}_{A_x}$，其中 $A_x := \{X = x\}$。B 是任意的，求 $\mathbb{E}(\mathbb{I}_B X)$

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 $A_x = \{A_{x_k}\}_{k \geqslant 1}$ 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 $\mathbb{E}(|\mathbb{I}_B X|) < \infty$ 说明期望存在

对 $\forall \omega \in B$

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

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4

对于 $\omega\in\Omega$，由划分

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

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

最后一个等号参考期望的定义15

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

Theorem 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10和练习7所示

$$X = \sum_{x\in S} x\mathbb{I}_{\{X=x\}} = \sum_{k\geqslant 1} x_k\mathbb{I}_{\Lambda_k}$$

$$\sum_{x\in S} x\mathbb{P}(X=x) = \sum_{k\geqslant 1}\mathbb{P}(X=x_k)$$

只有在“求和绝对收敛”（见定义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. $X\perp\!\!\!\perp Y,\quad g,h:\mathbb{R}\to\mathbb{R}$，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 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 11

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

证明：

0.4.2 条件期望

1° 关于“给定集合”的条件期望

Definition 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 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 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 9

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

Property 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7

至此没有用到独立性，可以得到以下推论

Corollary 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 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 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8成立。

Example 10

$$
\begin{aligned}
\mathbb{E}(X|\{\varnothing,\Omega\}) &= \mathbb{E}(X|\sigma(\Omega)) \\
&= \mathbb{I}_\Omega \mathbb{E}(X|\Omega) \qquad [\text{定义}(18)] \\
&= \sum_{x\in S} x\mathbb{P}(X=x|\Omega) \qquad [\text{定义}(17), \Omega \perp\!\!\!\perp X] \\
&= \sum_{x\in S} x\mathbb{P}(X=x) \\
&= \mathbb{E}(X)
\end{aligned}
$$

独立可以理解为：什么信息也没提供

Example 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 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 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 $\mathbb{I}_X = \{\{X = x\}|x \in S\}$

$$\begin{aligned}
\mathbb{E}(X|\sigma(X)) &= \sum_{x\in S} \mathbb{I}_{\{X=x\}}\mathbb{E}(X|X=x) \\
&= \sum_{x\in S} \mathbb{I}_{\{X=x\}}\mathbb{E}(X\mathbb{I}_{\{X=x\}})/\mathbb{P}(X=x) \quad [\text{推论}(2)] \\
&= \sum_{x\in S} \mathbb{I}_{\{X=x\}} \cdot \frac{x\cdot\mathbb{P}(X=x) + 0\cdot\mathbb{P}(X\neq x)}{\mathbb{P}(X=x)} \\
&= \sum_{x\in S} x\mathbb{I}_{\{X=x\}} = X \quad \square
\end{aligned}$$

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

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

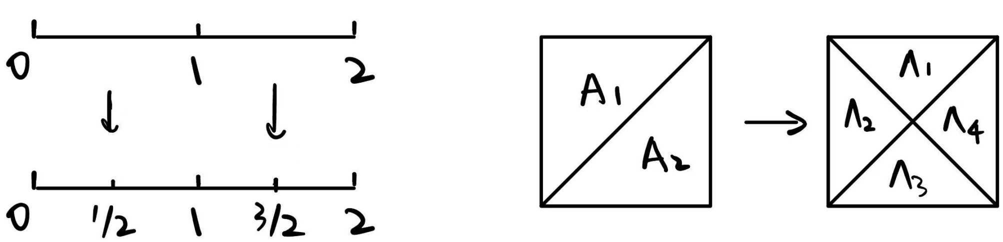


图 1: 加细划分

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

$$\begin{aligned}
\mathbb{E}(XY|\sigma(\Pi))(\omega) &= \mathbb{E}(\sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k} Y|\sigma(\Pi))(\omega) \qquad [X = \sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k}] \\
&= \mathbb{E}(\sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k} Y|\Lambda_j) \qquad [\sigma(\Pi)\text{定义}] \\
&= \mathbb{E}(\sum_{k\geqslant 1} x_k \mathbb{I}_{\Lambda_k} Y \mathbb{I}_{\Lambda_j})/\mathbb{P}(\Lambda_j) \qquad [\text{推论}(2)] \\
&= \mathbb{E}(Y x_j \mathbb{I}_{\Lambda_j})/\mathbb{P}(\Lambda_j) \qquad [\mathbb{I}_{\Lambda_k}\mathbb{I}_{\Lambda_j}\text{当}\Lambda_k \neq \Lambda_j\text{时} = 0] \\
&= x_j \mathbb{E}(Y \mathbb{I}_{\Lambda_j})/\mathbb{P}(\Lambda_j) \\
&= x_j \mathbb{E}(Y|\mathbb{I}_{\Lambda_j}) \\
&= X(\omega)\mathbb{E}(Y|\mathbb{I}_{\Lambda_j})
\end{aligned}$$

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

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 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 ([4].P280. 推论)
从 273-280

Corollary 3 (来自定理12(1)) 1.（重期望公式）$\mathbb{E}|\mathbb{E}(X|\sigma(\Pi))| = \mathbb{E}|X|, \mathbb{E}(\mathbb{E}(X|\sigma(\Pi))) = \mathbb{E}(X)$
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) $\mathbb{E}(Y\mathbb{I}_A) = \mathbb{E}(\mathbb{I}_A \mathbb{E}(X|\sigma(\Pi)))$

$$\begin{aligned}\mathbb{E}(Y\mathbb{I}_A) &= \mathbb{E}(\mathbb{I}_A \mathbb{E}(X|\sigma(\Pi))) \\ &= \mathbb{E}(\mathbb{E}(X\mathbb{I}_A|\sigma(\Pi))) \qquad [A \in \sigma(\Pi), \text{性质}(12)] \\ &= \mathbb{E}(X\mathbb{I}_A) \qquad [\text{重期望-推论}(3)]\end{aligned}$$

3° 关于离散随机变量的条件期望

Definition 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 13

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

Example 14

$\mathbb{I}_A \perp\!\!\!\perp \mathbb{I}_B \Rightarrow \mathbb{E}(\mathbb{I}_A|\mathbb{I}_B) = [\text{Exa}(11)]\mathbb{E}(\mathbb{I}_A)$

Example 15

$\mathbb{E}(X|X) = \mathbb{E}(X|\sigma(X)) = X[\text{Exa } 12]$

Property 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 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 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20是有意义的，well-defined，make sense，良定义）

Problem 4 (作业 2-2)

证明定理13在 $n=2$ 时成立

Definition 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 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：

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

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 14

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

证明：在定义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 4

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

0.5 随机过程

0.5.1 什么是随机过程

Definition 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

0.5.2 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 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 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 26 (过程的有限维分布族)
定义

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

0.5.3 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 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 14 (Kolmogorov 定理)
设分布函数族

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

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

0.5.4 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 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

1 马氏链

1.1 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 28 ((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：$\forall n \geqslant 0, x_0, x_1, \cdots, x_n, y \in S$，

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

其中 X_0 的分布称为 X 的初始分布

Definition 29
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 (马氏性的等价表示)
[Grimmett [2]] 下面三个命题等价

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

$$\mathbb{P}(X_{n+1} = y | X_{n_1} = x_{n_1}, \cdots, X_{n_k} = x_{n_k}) = \mathbb{P}(X_{n+1} = y | X_{n_k} = x_{n_k}) \tag{M_2}$$

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

$$\mathbb{P}(X_{n+m} = y | X_0 = x_0, \cdots, X_n = x_n) = \mathbb{P}(X_{n+m} = y | X_n = x_n) \tag{M_3}$$

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

$$
\{X_{n+k} = y\} \perp\!\!\!\perp_{\{X_n = x_n\}} \{X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\} \xrightarrow{\text{性质}(14)} \{X_{n+k} = y\} \perp\!\!\!\perp_{\{X_n = x_n\}} \{X_{n-1} = x_{n-1}\}
$$

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

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

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

$$
\begin{aligned}
\tilde{\mathbb{P}}_n(X_{n+k+1} = y) = &\sum_{x_{n+1} \in S} \tilde{\mathbb{P}}_n(X_{n+k+1} = y | X_{n+1} = x_{n+1}) \cdot \tilde{\mathbb{P}}_n(X_{n+1} = x_{n+1}) \quad [\text{定理}(5)] \\
= &\sum_{x_{n+1} \in S} \mathbb{P}(X_{n+k+1} = y | X_{n+1} = x_{n+1}, X_n = x_n) \cdot \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \quad [(*), \text{归纳法假设}] \\
= &\sum_{x_{n+1} \in S} \mathbb{P}(X_{n+k+1} = y, X_{n+1} = x_{n+1}, X_n = x_n) / \mathbb{P}(X_n = x_n) \quad [\text{乘法公式-定理}(4)] \\
= &\,\mathbb{P}(X_{n+k+1} = y, X_n = x_n) / \mathbb{P}(X_n = x_n) \\
= &\,\mathbb{P}(X_{n+k+1} = y | X_n = x_n)
\end{aligned}
$$

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14和全概公式5凑出乘法公式4，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 5

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 16 (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

$$\begin{aligned}&\mathbb{P}_{\{X_n=x_n\}}((X_0,\cdots,X_{n-1})\in A,(X_{n+1},\cdots,X_{n+m})\in B)\\=&\mathbb{P}_{\{X_n=x_n\}}((X_0,\cdots,X_{n-1})\in A)\cdot\mathbb{P}_{\{X_n=x_n\}}((X_{n+1},\cdots,X_{n+m})\in B)\end{aligned}$$

   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，证明见定理1
只证 $m = 2$，令

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

$\Rightarrow$

$$\begin{aligned}
\tilde{\mathbb{P}}_n((X_{n+1}, X_{n+2}) = (x_{n+1}, x_{n+2})) &= \tilde{\mathbb{P}}_n(X_{n+1} = x_{n+1}) \cdot \tilde{\mathbb{P}}_n(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}) \\
&= \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \cdot \mathbb{P}(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}) \qquad [M_1] \\
&= \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n) \cdot \mathbb{P}(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}, X_n = x_n) \qquad [\text{推论}(5)] \\
&= \mathbb{P}_{\{X_n=x_n\}}(X_{n+1} = x_{n+1}) \cdot \mathbb{P}_{\{X_n=x_n\}}(X_{n+2} = x_{n+2} | X_{n+1} = x_{n+1}) \\
&= \mathbb{P}_{\{X_n=x_n\}}((X_{n+1}, X_{n+2}) = (x_{n+1}, x_{n+2})) \qquad [\text{乘法公式-定理}(4)]
\end{aligned}$$

Corollary 6
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

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

1.2 时齐马氏链与转移概率

Definition 30 (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 31
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

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

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 15 (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 32 (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28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32？因为该定义决定了马氏链的有限维分布

Example 16 (Gambler's Ruin)

[Durrett [1],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 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 17 (Two-Stage Markov Chains)

[Durrett [1], 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. $Y_n = (X_n, X_{n-1}), n \geqslant 1$ 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 1

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 值离散时间的随机过程，则 $X \sim \text{Markov}(\mu, P)$ 当且仅当 X 有有限维分布，

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = \mu_{i_0} P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{n-1},i_n} \quad (\forall n \geqslant 0, i_j \in S)$$

证明：$\Rightarrow$

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

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 [3], prop 2.1.1 (PDF 链接)

$\Leftarrow$

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

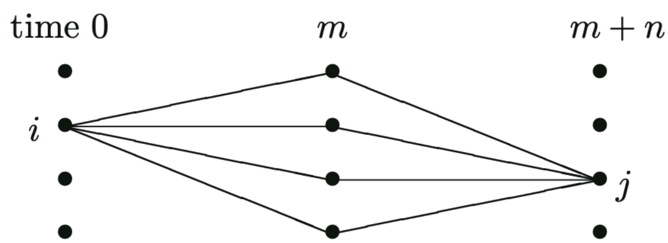
2.
$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \cdots, X_n = x_n) = \frac{\mathbb{P}(X_0 = i_0, \cdots, X_{n+1} = i_{n+1})}{\mathbb{P}(X_0 = i_0, \cdots, X_n = i_n)} = P_{i_n, i_{n+1}}$$

由时齐马氏链定义，初始分布和转移矩阵都符合定义32

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

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

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



1.3 多步转移概率与矩阵乘法

Definition 33

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？

1.3.1 Chapman-Kolmogorov 方程

Theorem 16 (C-K 方程)

设 $\{X_n, x \geqslant 0\}$ 为马氏链

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

其中 $i, j \in S, m, n, r \geqslant 0$，即

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

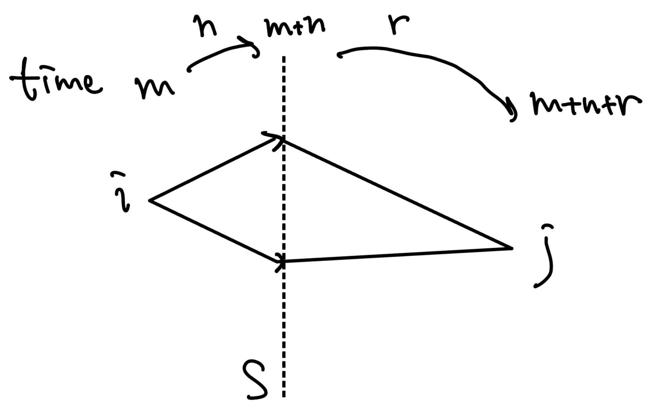


图 5: Multi-steps

证明：

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

Corollary 7

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 $P^{m+n} = P^m P^n$

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

$\forall n \geqslant 0, P(n) = P^n$ 仍是一个随机矩阵（定理15）

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）

1.3.2 马氏链的任意有限维分布

Proposition 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 3

$\mu^{(m+n)}=\mu^{(n)}P^m(\forall m,n\geqslant 0)$，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 17 (任意有限维分布 II)

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

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

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，$\mu,P\xrightarrow{\text{定理(14)}}$ 有限维分布族 $\to X\sim\text{Markov}(\mu,P)$，$\mu,P$ 可以刻画具备对称性、相容性的有限维分布

具体地，参考 Resnick [3], P62, Section 2.1 (PDF 链接)

1.4 （从固定点出发的）马氏链

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

1.4.1 链的状态：常返和暂留

Definition 34

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 $\forall k, \exists n_k, s.t.\{X_{n_k} = i\}$ 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 18

$1, 2, 3, 4, \cdots$ 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 19

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

1.4.2 从数学角度：并改写成不交并

$$\begin{aligned}
i \text{ 常返} &\Leftrightarrow \mathbb{P}_i(\cup_{n\geqslant 1}\{X_n=i\})=1\\
&\Leftrightarrow \mathbb{P}_i(\text{有限步到达}i)=1\\
&\Leftrightarrow \mathbb{P}(\text{从}i\text{出发条件下，有限时间内回到}i)=1
\end{aligned}$$

$B_1(i)=\{X_1=i\}, B_2(i)=\{X_2=i\}\backslash\{X_1=i\}=\{X_2=i, X_1\neq i\}, \cdots, B_n(i)=\{X_n=i, X_{n-1}\neq i\cdots, X_1\neq i\}$

$$\Rightarrow \sum_{n\geqslant 1} B_n(i)=\bigcup_{n\geqslant 1}\{X_n=i\}[\text{练习}(2)]$$

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

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

记号

$$f_{ij}:=\mathbb{P}_i(\text{首次访问}j\text{的时刻为有限时间})$$

$$f_{ij}(n):=\mathbb{P}_i(B_n(j))=\mathbb{P}_i(\text{首次访问}j\text{的时刻为}n)$$

Proposition 4
常返和暂留的等价命题

1. i 常返 $\Leftrightarrow 1=f_{ii}=\sum_{n\geqslant 1}f_{ii}(n)$
2. i 暂留 $\Leftrightarrow 1>f_{ii}=\sum_{n\geqslant 1}f_{ii}(n)$

1.4.3 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 4 (Grimmett [2], Thm 6.3.3)
设 $|s|<1$，则

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

其中

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

证明：构造不交并，$B_m(i) = \{X_m = i, X_{m-1}\neq i, \cdots, X_1 \neq i\}, m\geqslant 1$
$\Rightarrow \sum_{m\geqslant 1} B_m(i) = \cup_{n\geqslant 1}\{X_n = i\}, B_m(i) \subseteq \{X_n \neq i\}, m \geqslant n+1$

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

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

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

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（引理2M_3）

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

当 $n \geqslant 1$ 时，

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

把 $\mathbb{I}_{\{1\leqslant m\leqslant n\}}(s^{n-m}p_{jj}(n-m))(s^m f_{ij}(m))$ 看作 $a_{n,m}$，由推论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/∞

$$
\begin{aligned}
&\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mathbb{I}_{\{1\leqslant m\leqslant n\}} s^{n-m} p_{jj}(n-m) s^m f_{ij}(m) \\
=&\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathbb{I}_{\{1\leqslant m\leqslant n\}} s^{n-m} p_{jj}(n-m) s^m f_{ij}(m) \\
=&\sum_{m=1}^{\infty} (\sum_{n=m}^{\infty} s^{n-m} p_{jj}(n-m)) s^m f_{ij}(m) \\
=&(\sum_{n=0}^{\infty} s^n p_{jj}(n))(\sum_{m=1}^{\infty} s^m f_{ij}(m)) < \infty \quad [\text{变量代换} n \leftarrow n-m]
\end{aligned}
$$

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 5

(1) j 常返 $\Leftrightarrow$

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

(2) j 暂留 $\Leftrightarrow$

$$
\begin{aligned}
1 > f_{jj} &\Leftrightarrow \sum_{n\geqslant 0} p_{jj}(n) < \infty \\
&\Rightarrow \sum_{n\geqslant 0} p_{ij}(n) < \infty, \forall i \in S \\
&\Rightarrow \lim_{n\to\infty} p_{ij}(n) = 0, \forall i \in S
\end{aligned}
$$

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

$\Rightarrow P_{jj}(s) = 1 + P_{jj}(s) F_{jj}(s)$

$\Rightarrow$

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

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

Problem 6 (作业 5-1)
证明：j 暂留 $\Rightarrow \sum_{n\geqslant 0} p_{ij}(n) < \infty, \forall i \in S$

1.4.4 从“首次回访时间”角度判别

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

Definition 35
首次回访 j 的时刻

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

约定 $\min \varnothing = +\infty$

注：$\{T_j = \infty\} \Leftrightarrow \{\omega | \{n \geqslant 1 | X_n(\omega) = j\} = \varnothing\}$
$\Leftrightarrow \{\omega | X_n(\omega) \neq j, \forall n \geqslant 1\} = \cap_{n\geqslant 1}\{X_n \neq j\}$

Property 17
$f_{jj} = \mathbb{P}_j(T_j < \infty), f_{jj}(n) = \mathbb{P}_j(T_j = n)$

定义 $\mathbb{P}_j(T_j < \infty) = \rho_{jj}$

Proposition 6
联系命题5

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

Definition 36
j 的平均回访时间

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

Theorem 18

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

证明：

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

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

(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

$$\begin{aligned}
\mathbb{E}_j T_j &= \mathbb{E}_j T_j \mathbb{I}_{\{T_j<\infty\}} \\
&= \mathbb{E}_j T_j \mathbb{I}_{\sum_{n\geqslant 1}\{T_j=n\}} \\
&= \mathbb{E}_j \sum_{n\geqslant 1} T_j \mathbb{I}_{\{T_j=n\}} \quad [\text{定义}(10)] \\
&= \sum_{n\geqslant 1} n \mathbb{P}_j(T_j = n) \\
&= \sum_{n\geqslant 1} n f_{jj}(n)
\end{aligned}$$

Definition 37

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

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

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

Definition 38

链在时刻 0 之后，访问 j 的次数

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

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

至此，做个阶段性小结，回顾 i 常返的几种等价表示

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

无数次地回访 $\leftrightarrow$ 访问次数 $=\infty\leftrightarrow\mathbb{P}_i(N(i):=\sum_{n\geqslant 1}\mathbb{I}_{\{X_n=i\}}=\infty)=1$
两种表述的等价条件互相等价吗？即

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

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，

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

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

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

Definition 39
第 k 次访问概率

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

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

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

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 5
(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

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

证明：$\Leftarrow$ 显然，因为 $\{N(i) = \infty\}$ 相对 $N(i) \geqslant 1$ 是小集合
$\Rightarrow$

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

暂留的证明同理：

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

1.4.5 从“平均回访次数”角度判别

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

Lemma 6 (Durrett [1], lem 1.11)

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

证明：

$$\begin{aligned} \mathbb{E}_y N(y) &= \mathbb{E}_y \sum_{k\geqslant 1} \mathbb{I}_{\{N(y)\geqslant k\}} \overset{\text{Exa}(7)}{=} \sum_{k\geqslant 1} \mathbb{P}(N(y) \geqslant k) = \sum_{k\geqslant 1} \mathbb{P}_y(T_y^{(k)} < \infty) \\ &= \sum_{k\geqslant 1} \rho_{yy}^{(k)} \overset{\text{SMP}}{=} \sum_{k\geqslant 1} \rho_{yy}^{k} \end{aligned}$$

$\rho_{yy} = 1 \Rightarrow \mathbb{E}_y N(y) = \infty$
$\rho_{yy} < 1 \Rightarrow \mathbb{E}_y N(y) = \rho_{yy}/(1-\rho_{yy})$
下面证：i 常返 $\Leftrightarrow \mathbb{E}_i N(i) = \infty$，也就是证 $\mathbb{P}_i(N(i) = \infty) = 1 \Leftrightarrow \mathbb{E}_i N(i) = \infty$
证明：$\Rightarrow$ 显然

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

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

将上面几个角度总结成下面定理

Theorem 19 (链的状态：等价表述)

i常返 $\overset{\text{Def}}{\Leftrightarrow} 1=\mathbb{P}_i(\bigcup_{n\geqslant 1}\{X_n=i\})$ [回访发生的概率]

$$\Leftrightarrow 1=f_{ii}:=\sum_{n\geqslant 1}f_{ii}(n)=\sum_{n\geqslant 1}\mathbb{P}_i(X_1\neq i,\cdots,X_{n-1}\neq i,X_n=i)\Leftrightarrow\sum_{n\geqslant 1}p_{ii}(n)=\infty \quad \text{[首次回访发生]}$$

$$\Leftrightarrow 1=\rho_{ii}=\mathbb{P}_i(T_i^{(1)}:=T_i<\infty)=\mathbb{P}_i\{N(i)\geqslant 1\}=1$$

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

$$\Leftrightarrow \mathbb{E}_iN(i)=\infty$$

1.4.6 停时与强马氏性

Definition 40 (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 20

$T_y^{(1)}:=T_y:=\min\{n\geqslant 1|X_n=y\}$

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

Definition 41 (停时 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：$B\in\mathcal{F}_i \Leftrightarrow B$ 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）

$\Leftrightarrow B\cap\{T=n\}\in\sigma(X_0,\cdots,X_n),\forall n$

Proposition 7 (强马氏性)

$X:=\{X_n,n\geqslant 0\}\sim\text{Markov}(\mu,P)$，$\tau$ 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：只证 (1)，(2) 留作作业

$$\begin{aligned}
&\mathbb{P}(X_{\tau+0}=i_0,\cdots,X_{\tau+n}=i_n,\tau=m,X_\tau=x)\\
=&\mathbb{P}(X_m=i_0,X_{m+1}=i_1,\cdots,X_{m+n}=i_n,\tau=m)\\
=&\mathbb{P}(X_{m+1}=i_1,\cdots,X_{m+n}=i_n,\tau=m|X_m=i)\mathbb{P}(X_m=i)\\
=&\mathbb{P}(X_{m+1}=i_1,\cdots,X_{m+n}=i_n|X_m=i)\mathbb{P}(\tau=m,X_m=i)\\
=&\frac{\delta_{xi_0}p_{i_0,i_1}p_{i_1,i_2}\cdots p_{i_{n-1},i_n}}{\delta_{xi_0}}\mathbb{P}(\tau=m,X_\tau=i)\\
=&p_{i_0,i_1}p_{i_1,i_2}\cdots p_{i_{n-1},i_n}\mathbb{P}(\tau=m,X_\tau=i)
\end{aligned}$$

其中第二个等号位置，$X_{m+1}=i_1,\cdots,X_{m+n}=i_n\in\sigma(X_{m+1},\cdots,X_{m+n})$ 为未来，$X_m=i$ 为现在，$\tau=m\in\sigma(X_0,\cdots,X_m)$ 中多出来个“现在”(X_m)。但不影响，因为可证

$$X_m\perp\!\!\!\perp_{\{X_m=i\}}X_{m+1}$$

对 m 求和，两边同除 $\mathbb{P}(\tau<\infty,X_\tau=i)$

$$\Rightarrow\mathbb{P}(X_{\tau+0}=i_0,\cdots,X_{\tau+n}=i_n|\tau<\infty,X_\tau=i)=\delta_{xi_0}p_{i_0,i_1}\cdots p_{i_{n-1},i_n}$$

$\Rightarrow(X_{\tau+n})_{n\geqslant 0}\sim\text{Markov}(\delta_x,P)$ under $\mathbb{P}(\cdot|\tau=\infty,X_\tau=x)$

$T_y^{(0)} := 0, T_y^{(1)} := T_y$，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 7

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

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

Corollary 8

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

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

1.5 类结构

1.5.1 状态 i 间的关系：可达与互通

Definition 42 (可达)

$i, j \in S$，若 $\exists n \geqslant 0, s.t. p_{ij}(n) > 0$，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 43 (互通)

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

Theorem 20

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 [1]]
3. $\exists$ 某些状态，$i_0 = i, i_1, \cdots, i_{n-1}, i_n = j, s.t. p_{i_0,i_1} \cdots p_{i_{n-1},i_n} > 0$
4. $\mathbb{P}_i(\exists n \geqslant 0, X_n = j) > 0$

Problem 7 (作业 6-1)
证明：定理20命题的等价性，即 $1 \Leftrightarrow 2, 1 \Leftrightarrow 3, 1 \Leftrightarrow 4$

Problem 8 (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 18 (Durrett [1], lem 1.4)
若 $i \to j, j \to k$，则 $i \to k$

证明：$i \to j \Leftrightarrow \exists n \geqslant 0, s.t. p_{ij}(n) > 0$
$j \to k \Leftrightarrow \exists n \geqslant 0, s.t. p_{jk}(n) > 0$

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

Property 19
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 $i \leftrightarrow k$

1.5.2 常返与暂留是类性质

Lemma 8 (Durrett [1], Thm 1.5&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：$i \neq j, \rho_{ji} < 1 \Rightarrow 0 < 1 - \rho_{ji} = 1 - \mathbb{P}_j(T_i < \infty) = \mathbb{P}_j(T_i = \infty)$
为了证 i 暂留，即证 $\rho_{ii} < 1$，即 $\mathbb{P}_i(T_i = \infty) > 0$
(Step 1) $\rho_{ij} > 0, i \to j \Rightarrow \exists k \geqslant 1, s.t. p_{ij}(k) > 0$

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

由 C-K 方程（定理16）知，$\exists$ 与 i,j 不同的状态 $i_1,\cdots,i_{k-1},s.t.$

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

(Step 2)

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

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

$\tau_1=0,\tau_2=K$ 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 $\mathbb{P}_j(\cdot):=\mathbb{P}(\cdot|X_K=j)$ 下，$(X_n)_{n\geqslant 0}\sim\text{Markov}(\delta_j,P)$

$$\begin{aligned}
\lim_{m\to\infty}\mathbb{P}(\bigcap_{n=K+1}^{m}\{X_n\neq i\}|X_K=j)&=\mathbb{P}(\bigcap_{n=K+1}^{K+m}\{X_n\neq i\}|X_K=j)\\
&=\tilde{\mathbb{P}}(X_{K+n}\neq i,1\leqslant n\leqslant m)\\
&=\mathbb{P}_j(X_n\neq i,1\leqslant n\leqslant m)\xrightarrow{m\to\infty}\mathbb{P}_j(\bigcap_{n\geqslant 1}\{X_n\neq i\})=\mathbb{P}_j(T_j=\infty)=1-p_{ji}>0
\end{aligned}$$

因此 $\mathbb{P}_i(T_i=\infty)>0$

Corollary 9

$i\to j$, i 常返 $\Rightarrow j$ 常返

证明：$i\neq j,i\to j$，i 常返，由推论8，知 $\rho_{ji}=1>0$，所以 $j\to i,i\leftrightarrow j$

$\exists m,n\geqslant 0,s.t.p_{ij}(m)>0,p_{ji}(n)>0$

$\forall r\geqslant 0,p_{jj}(n+r+m)\overset{\text{C-K}}{\geqslant}p_{ji}(n)p_{ii}(r)p_{ij}(m)$

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)

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

Corollary 10

$i \leftrightarrow j$，则 i 常返 $\Leftrightarrow j$ 常返

Definition 44 (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 45 (链的不可约)

若 S 不可约，则称链不可约

Theorem 21

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

1.5.3 状态空间分解

Definition 46 (闭集)

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

Problem 9 (作业 6-3)

$C \subseteq S$ 闭集等价于

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

Example 21

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

Theorem 22

每一个有限的不可约闭集都是常返的

参考文献

[1] Rick Durrett. Essentials of Stochastic Processes. 01 1999.

[2] Geoffrey Grimmett and David Stirzaker. Probability and random processes. Oxford University Press, Oxford; New York, 2001.

[3] Sidney I. Resnick. Adventures in stochastic processes. Birkhauser Verlag, CHE, 1992.

[4] 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.