随机过程笔记

Stochastic Process

Second Edition

董晟渤

前言

随机过程是统计学专业的一门重要必修课. 在概率论对随机变量的研究的基础上, 随机过程考虑一组随机变量 (通常是随时间变化的), 由此进入了更加复杂的随机世界. 随机过程也有非常广泛的应用, 从物理学到金融学, 许多现象都可以用随机过程描述. 整理本份笔记, 是因为笔者对概率论与随机过程较感兴趣, 或许未来会把概率论与随机过程当作自己的方向 (当然也有很大的概率会被打脸).

本门课程所用的教材是何书元老师的随机过程 (见 [**1**]), 这是一本比较经典的教材, 内容非常丰富. 所需要的先修课程是数学分析和初等概率论 (建议的参考书是 [**2**] 与 [**3**]). 比较遗憾的是, 本课程只有 3 学分、48 学时, 在八周的课上, 仅学习了该书的一部分内容, 让人觉得意犹未尽. 除了教材上的内容以外, 本门课程的老师也补充了不少知识点, 讲解的顺序也更加合理. 本份笔记正是按照老师所讲的内容整理而成.

当然, 由于时间仓促, 以及笔者是随机过程的初学者, 水平有限, 在本份笔记中可能存在不少错误. 如读者发现错误, 可直接向笔者指出; 如果有任何建议或想法, 也欢迎来交流. 笔者的联系方式 (QQ) 是 198368289.

最后, 感谢程晓青老师在八周内对本门课程认真的授课和答疑; 也感谢笔者概率论课程的两位老师, 分别是段启宏老师和朱学虎老师, 他们的建议对笔者也有巨大的帮助; 除此之外, 还应该感谢信计 91 的张博闻同学对排版的帮助, 统计 91 的符露露同学、应数 91 的李惊晨同学提供的笔记和建议, 以及, 感谢包容笔记的错误与不足的各位读者.

董晟渤
2021 年 11 月
于西安交通大学

目录

第 1 章 随机过程入门

1.1 随机过程的定义

随机过程是依赖于参数的一组随机变量的全体, 在各个领域都有广泛的应用. 一般情况下, “参数”指的是时间. 在这里用 t 表示时间. 定义1.1说明了什么是随机过程.

定义 1.1 (随机过程)

设 $(\Omega, \mathscr{F}, \mathbb{P})$ 为概率空间, 参数 $T \subset (-\infty, +\infty)$, 若对任意的 $t \in T$, $X(\omega, t)$ 是一随机变量, 则称

$$\{X(\omega, t), t \in T\}$$

为 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的一个随机过程. ♣

需要注意, 随机过程是定义在相同的概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的. 为了方便, 以下说明清楚随机过程的定义中涉及到的概念.

(1) T 称为参数集或指标集. 例如可以取 $T = \{0, 1, 2, \cdots\}$, $T = \{0, \pm 1, \pm 2, \cdots\}$ 或 $T = [a, b]$. 若 T 是可列集, 则称该随机过程为随机序列.

(2) 需要区分随机过程和随机变量: $X : \Omega \to \mathbb{R}$ 是随机变量, 通常记作 $X(\omega)$, 而 $X : \Omega \times T \to \mathbb{R}$ 是随机过程, 通常记作 $X(\omega, t)$;

(3) 固定 $t \in T$, 则 $X(\omega, t)$ 是随机变量, 称为在时刻 t 的状态. $X(\omega, t) = x$ 表示在时刻 t 的状态为 x. $X(\omega, t)$ 的所有取值的全体称为状态空间.

(4) 固定 $\omega \in \Omega$, 则 $X(\omega, t) = g(t)$ 称为样本函数, 也称为该随机过程的一次实现或一个轨道.

(5) 随机过程是有限维随机变量的推广, 可以记作

$$X(\omega, t), \quad \{X(t)\}, \quad X(t), \quad \text{或} \quad \{X(t), t \in T\}.$$

以下考虑一些具体的随机过程的例子, 来帮助理解随机过程的概念.

例 1.1 抛硬币 考虑抛硬币的过程, $\Omega = \{H, T\}$. 定义

$$X(\omega, t) = \begin{cases} \cos \pi t, & \omega = H, \\ 2t, & \omega = T, \end{cases} \quad t \in (0, +\infty),$$

设 $\mathbb{P}(H) = \mathbb{P}(T) = \dfrac{1}{2}$, 则 $\{X(\omega, t)\}$ 的样本函数为

$$\begin{cases} X(H, t) = \cos \pi t, \\ X(T, t) = 2t, \end{cases}$$

状态空间为 $[-1,+\infty)$.

例 1.2 设 $N(t)$ 表示某服务站在 $[0,t]$ 内到达的顾客数. 则样本函数为单调不减的右连续阶梯函数.

后面将利用某种特殊的分布, 给出例1.2的具体刻画.

1.2 随机过程的概率特性

1.2.1 分布函数

接下来, 为了更深入地研究随机过程, 我们需要回答的两个问题是:

(1) 既然随机过程是一组随机变量, 那么随机过程是否像随机变量一样, 用分布去刻画?

(2) 以及, 在上面的基础上, 随机过程是否可以用有限维的分布刻画?

在回答第一个问题的时候, 我们引入分布函数簇的概念. 而第二个问题, 则涉及到 Kolmogorov 存在性定理 [1]. 首先给出随机过程的分布函数的定义.

定义 1.2 (分布函数)

设 $\{X(t)\}$ 为随机过程, 对任意的正整数 n, 及对任意的 $t_1, t_2, \cdots, t_n \in T$, 称

$$F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n) = \mathbb{P}(X(t_1) \le x_1, X(t_2) \le x_2, \cdots, X(t_n) \le x_n),$$

为 $\{X(t)\}$ 的n 维分布函数. 当 n 与 $t_1,t_2,\cdots,t_n$ 改变时, 上面的函数构成分布函数簇. ♣

特别地, 当 $n=1$ 时, 随机过程 $\{X(t)\}$ 的一维分布函数为

$$F_X(x,t) = \mathbb{P}(X(t) \le x),$$

这就是随机变量 $X(t)$ 的分布函数. 一维分布函数是较容易求出的, 例1.3是一个较为复杂的求一维分布函数的例子.

例 1.3 设 $X(t) = te^Y$, 其中 $t>0$, $Y \sim \mathrm{Exp}(\lambda)$, 求 $\{X(t)\}$ 的一维分布函数.

解答 当 $x>t$ 时, 有

$$\mathbb{P}(X(t) \le x) = \mathbb{P}\left(Y \le \ln\frac{x}{t}\right) = 1 - \exp\left(-\lambda \ln\frac{x}{t}\right) = 1 - \left(\frac{t}{x}\right)^{\lambda}.$$

---

[1]对任何一列概率空间 $\{(X_k,\mathscr{F}_k,\mathbb{P}_k), k=1,2,\cdots\}$, 在 $\left(\prod_{k=1}^{\infty} X_k, \prod_{k=1}^{\infty}\mathscr{F}_k\right)$ 上有唯一的概率测度 $\mathbb{P}$, 使得对每个 $n=1,2,\cdots$ 和每一组 $A_1\in\mathscr{F}_1, A_2\in\mathscr{F}_2,\cdots,A_n\in\mathscr{F}_n$, 有 $\mathbb{P}\left\{\pi^{-1}\left(\prod_{k=1}^{n}A_k\right)\right\} = \prod_{k=1}^{n}\mathbb{P}_k(A_k)$. 这超出了课程的范围. 关于该定理的介绍和证明, 详见 **[3]**.

因此 $\{X(t)\}$ 的一维分布函数

$$F_X(x,t)=\begin{cases}1-\left(\dfrac{t}{x}\right)^{\lambda}, & x>t,\\ 0, & x\le t.\end{cases}$$

1.2.2 独立性

概率论中, 有一个独有且重要的概念叫做独立性. 我们知道, 若两个随机变量独立, 则它们的联合分布函数可以写成各自的分布函数的乘积. 在这样的想法之下, 给出随机过程的独立性的定义.

定义 1.3 (独立性)

设 $\{X(t)\},\{Y(t)\}$ 是随机过程, 若对任意的正整数 m,n, 及对任意的 $t_1,\cdots,t_n$, $t_1',\cdots,t_m'\in T$, 都有

$$F_{t_1,\cdots,t_n,t_1',\cdots,t_m'}(x_1,\cdots,x_n,y_1,\cdots,y_m)=F_{t_1,\cdots,t_n}(x_1,\cdots,x_n)\cdot F_{t_1',\cdots,t_m'}(y_1,\cdots,y_m),$$

则称 $\{X(t)\}$ 和 $\{Y(t)\}$ 独立. ♣

另外, 对于某个随机过程, 考虑增量的独立性, 则有定义1.4.

定义 1.4 (独立增量过程)

设 $\{X(t)\}$ 是随机过程, 若对任意有限个 $t_1<t_2<\cdots<t_n$,

$$X(t_2)-X(t_1),\quad X(t_3)-X(t_2),\quad \cdots,\quad X(t_n)-X(t_{n-1})$$

是相互独立的, 则称 $\{X(t)\}$ 为独立增量过程. ♣

除了独立以外, 概率论中“同分布”的特性也非常重要.

定义 1.5 (平稳增量过程)

设 $\{X(t)\}$ 是随机过程, 若对任意的 $s>0$, $t_2>t_1>0$, $X(t_2)-X(t_1)$ 与 $X(t_2+s)-X(t_1+s)$ 同分布, 则称 $\{X(t)\}$ 为平稳增量过程. ♣

以下是独立增量的过程的一个例子.

例 1.4 独立增量过程 设 $\{X(n),n=0,1,\cdots\}$ 相互独立, 令 $S(i)=\sum\limits_{n=0}^{i}X(n)$, 则 $\{S(i),i=0,1,\cdots\}$ 为独立增量过程.

证明 对任意的正整数 $0 \le n_0 < n_1 < \cdots < n_m$, 都有

$$S(n_1) - S(n_0) = X(n_0 + 1) + \cdots + X(n_1),$$
$$S(n_2) - S(n_1) = X(n_1 + 1) + \cdots + X(n_2),$$
$$\cdots$$
$$S(n_m) - S(n_{m-1}) = X(n_{m-1} + 1) + \cdots + X(n_m),$$

其中右边每一项都是相互独立的.

1.2.3 数字特征

在概率论中, 我们研究了某个随机变量的数字特征, 包括期望、方差等等. 对于随机过程, 如果取定 $t \in T$, 则 $X(t)$ 也是一个随机变量. 在此基础上, 我们也可以定义随机过程的数字特征.

(1) 对于随机过程 $\{X(t)\}$, 定义均值函数

$$m(t) = \mathbb{E}X(t);$$

(2) 对于随机过程 $\{X(t)\}$, 定义方差函数

$$D(t) = \mathbb{E}(X(t) - m(t))^2;$$

同时定义标准差函数

$$\sigma(t) = \sqrt{D(t)};$$

(3) 对于随机过程 $\{X(t)\}$, 定义自协方差

$$C(t_1, t_2) = \mathrm{Cov}(X(t_1), X(t_2)) = \mathbb{E}(X(t_1) - m(t_1))(X(t_2) - m(t_2));$$

(4) 对于随机过程 $\{X(t)\}$ 与 $\{Y(t)\}$, 定义互协方差

$$C_{X,Y}(t_1, t_2) = \mathrm{Cov}(X(t_1), Y(t_2)) = \mathbb{E}(X(t_1) - m_X(t_1))(Y(t_2) - m_Y(t_2)).$$

在此基础上, 若对任意的 $t_1, t_2 \in T$, 都有

$$C_{X,Y}(t_1, t_2) = 0,$$

则称 $\{X(t)\}$ 与 $\{Y(t)\}$ 互不相关;

(5) 对于随机过程 $\{X(t)\}$, 定义自相关函数

$$R(t_1, t_2) = \mathbb{E}X(t_1)X(t_2).$$

在此基础上, 容易推导出

$$C(t_1, t_2) = R(t_1, t_2) - m(t_1)m(t_2).$$

(6) 对于随机过程 $\{X(t)\}$ 与 $\{Y(t)\}$, 定义互相关函数

$$R_{X,Y}(t_1, t_2) = \mathbb{E}X(t_1)Y(t_2).$$

在此基础上, 容易推导出

$$C_{X,Y}(t_1,t_2)=R_{X,Y}(t_1,t_2)-m_X(t_1)m_Y(t_2).$$

例1.5是一个计算随机过程的数字特征的例子.

例 1.5 设 $\{X(t),t\in T\}$ 是独立增量过程, $X(0)=0$, $m(t)=\mathbb{E}X(t)$, $D(t)=\mathrm{Var}X(t)$.

(1) 求 $\{X(t)\}$ 的自协方差函数 $C(t_1,t_2)$;

(2) 求 $\{X(t)\}$ 的自相关函数 $R(t_1,t_2)$.

解答 若 $0<t_1<t_2$, 根据定义得

$$X(t_1)-X(0)=X(t_1),\quad X(t_2)-X(t_1)$$

相互独立, 从而它们不相关, 计算得

$$\begin{aligned}0&=\mathrm{Cov}(X(t_1),X(t_2)-X(t_1))\\&=\mathrm{Cov}(X(t_1),X(t_2))-\mathrm{Var}X(t_1)\\&=C(t_1,t_2)-D(t_1),\end{aligned}$$

从而 $C(t_1,t_2)=D(t_1)$; 同理当 $0<t_2<t_1$ 时有 $C(t_1,t_2)=D(t_2)$, 因此

$$C(t_1,t_2)=D(\min\{t_1,t_2\}).$$

根据自协方差函数与自相关函数之间的关系, 计算得

$$R(t_1,t_2)=C(t_1,t_2)+m(t_1)m(t_2)=D(\min\{t_1,t_2\})+m(t_1)m(t_2).$$

本例中的随机过程较为特殊, 满足 $X(0)=0$, 这导致了它的自协方差 $C(t_1,t_2)=D(\min\{t_1,t_2\})$. 若设 $s<t$, 则 $C(s,t)=D(s)$. 在后面将会发现, 有许多的随机过程都具有 $X(0)=0$ 的性质, 从而例1.5的结论十分有用.

1.3 课后习题

本节的习题主要为概率论的复习.

问题 1.1 若 $X_1,X_2,\cdots,X_n$ 相互独立, $X_i\sim\mathcal{P}(\lambda_i)$, 则

(1) $\mathbb{E}X_i=\lambda_i$, $\mathrm{Var}X_i=\lambda_i$;

(2) $X_1+X_2+\cdots+X_n\sim\mathcal{P}(\lambda_1+\lambda_2+\cdots+\lambda_n)$.

问题 1.2 当 $P(A)>0$ 时, 推导乘法公式

$$\mathbb{P}(B_1B_2\cdots B_n|A)=\mathbb{P}(B_1|A)\mathbb{P}(B_2|B_1A)\cdots\mathbb{P}(B_n|B_1B_2\cdots B_{n-1}A).$$

问题 1.3 设 $X_1,X_2,\cdots$ 是来自总体 X 的随机变量, $\mu=\mathbb{E}X$, $\sigma^2=\mathrm{Var}X<\infty$. 对于和总体 X 独立的取非负整数值的随机变量 N, 当 $\sigma_N^2=\mathrm{Var}N<\infty$ 时, 计算

$$\mathbb{E}(X_1+X_2+\cdots+X_N),\quad \mathrm{Var}(X_1+X_2+\cdots+X_N).$$

问题 1.4 设 X 是非负随机变量, $X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本, $k \le n$.

(1) 计算 $\mathbb{E}(X_1 + X_2 + \cdots + X_k | X_1 + X_2 + \cdots + X_n = t)$;

(2) 如果 U 在 $[0, t]$ 上均匀分布, 且与 X_1, X_2 独立, 计算

$$\mathbb{P}(U < X_1 | X_1 + X_2 = t).$$

第 2 章 Poisson 过程

2.1 计数过程与 Poisson 过程

2.1.1 计数过程

在生活中, “计数”无处不在. 随着时间的推移, “计数”的过程将成为随机过程. 定义2.1给出了该随机过程的定义.

定义 2.1 (计数过程)

设 $t \geq 0$, $N(t)$ 表示 $[0, t]$ 内某类事件发生的个数, 则随机过程 $\{N(t)\}$ 称为计数过程. ♣

在定义2.1的基础上, 容易看出计数过程有如下的性质.

(1) 对任意的 $t \geq 0$, $N(t)$ 的取值为非负整数;

(2) 对任意的 $t > s \geq 0$, $N(t) \geq N(s)$;

(3) 对任意的 $t > s \geq 0$, $N(t) - N(s)$ 表示时间段 $(s, t]$ 内发生的事件数;

(4) $\{N(t)\}$ 的轨迹是单调不减的右连续阶梯函数.

或许读者会发现, 在上一节例1.2中所定义的随机过程 $\{N(t)\}$ 就是一个计数过程. 以下为了方便, 对 $t, s \geq 0$, 记 $N(s, s+t] = N(s+t) - N(s)$.

2.1.2 Poisson 过程的定义

在概率论中, 我们接触到了 Poisson 分布. 我们用 $\mathcal{P}(\lambda)$ 表示强度为 λ 的 Poisson 过程, 设随机变量 $X \in \mathcal{P}(\lambda)$, 则它的取值只可能为非负整数, 且

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda}, \quad k = 0, 1, \cdots.$$

在此基础上, 定义2.2给出了 Poisson 过程的定义. 后面我们将会发现, Poisson 过程是最简单的也是应用极多的一个计数过程.

定义 2.2 (Poisson 过程)

设计数过程 $\{N(t)\}$ 满足:

(1) $N(0) = 0$;

(2) $\{N(t)\}$ 是独立增量过程;

(3) 对任意的 $t, s \geq 0$, $N(s, s+t] \sim \mathcal{P}(\lambda t)$, 也即

$$\mathbb{P}(N(s, s+t] = k) = \frac{(\lambda t)^k}{k!} \cdot \mathrm{e}^{-\lambda t},$$

则称 $\{N(t)\}$ 为强度为 λ 的 *Poisson* 过程.

结合 Poisson 分布的性质, 我们可以看出 Poisson 过程具有以下的性质.

(1) Poisson 过程是平稳增量过程;

(2) Poisson 过程的均值函数 $m(t) = \mathbb{E}N(t) = \lambda t$, 并且
$$\lambda = \frac{\mathbb{E}N(t)}{t},$$
从而参数 λ 表示单位时间 t 内事件发生个数的平均值, 从而称为强度, 这是 Poisson 过程的一个很重要的特性;

(3) Poisson 过程的方差函数 $D(t) = \mathrm{Var}N(t) = \lambda t$;

(4) 根据例1.5所得结论计算得, Poisson 过程的自协方差函数
$$C(s,t) = D(\min\{t,s\}) = \lambda \min\{t,s\};$$
自相关函数
$$R(s,t) = C(s,t) + m(s)m(t) = \lambda \min\{t,s\} - \lambda^2 ts.$$

以上的定义2.2可以认为是“宏观”的定义. 它对计数过程 $\{N(t)\}$ 的要求是, 在一个“宏观”的区间 $(s, s+t]$ 上, 事件发生的过程服从 Poisson 分布 $\mathcal{P}(\lambda t)$. 以下给出的定义2.3, 是对 Poisson 过程的另外一个角度的刻画.

定义 2.3 (Poisson 过程)

设计数过程 $\{N(t)\}$ 满足:

(1) $N(0) = 0$;

(2) $\{N(t)\}$ 是独立增量过程, 并且有平稳增量性;

(3) 对任意的 $t \geq 0$ 及对充分小的 Δt, 有
$$\begin{cases} \mathbb{P}(N(t, t+\Delta t] = 1) = \lambda \cdot \Delta t + o(\Delta t), \\ \mathbb{P}(N(t, t+\Delta t] \geq 2) = o(\Delta t), \end{cases}$$

则称为 $\{N(t)\}$ 强度为 λ 的 *Poisson* 过程.

定义2.3可以认为是“微观”的定义. 第一式对计数过程 $\{N(t)\}$ 的要求是, 在极短的时间 Δt 内, 恰有一次事件发生的概率和 $\lambda \cdot \Delta t$ 是同阶无穷小. 事件在 $(t, t+\Delta t]$ 时间段内发生的概率与 Δt 成正比, 这与我们对参数 λ 的理解是一致的. 而第二式说明了同一时刻内不会发生两个或两个以上的事件.

接下来还需要验证, 上面两个定义是等价的, 从而不会导致矛盾.

定理 2.1

定义2.2和定义2.3是等价的.

证明 定义2.2 $\Longrightarrow$ 定义2.3. 首先计算得

$$\mathbb{P}(N(t,t+\Delta t]=1)=\lambda\cdot\Delta t\mathrm{e}^{-\lambda\cdot\Delta t}=\lambda\cdot\Delta t+o(\Delta t);$$

同理有

$$\begin{aligned}\mathbb{P}(N(t,t+\Delta t]\geq 2)&=1-\mathbb{P}(N(t,t+\Delta t]=0)-\mathbb{P}(N(t,t+\Delta t]=1)\\&=1-(1-\lambda\cdot\Delta t+o(\Delta t))-(\lambda\cdot\Delta t+o(\Delta t))\\&=o(\Delta t).\end{aligned}$$

定义2.3 $\Longrightarrow$ 定义2.2. [1] 在 $[0,t]$ 插入分点 $0=x_0<x_1<\cdots<x_n=t$, 其中 $x_i=\dfrac{it}{n}$, 并对 $1\leq i\leq n$, 记 $Y_i=N(x_{i-1},x_i]$, 则 $Y_1,Y_2,\cdots,Y_n$ 相互独立, 且

$$\begin{cases}\mathbb{P}(Y_i=0)=1-\lambda\cdot\dfrac{t}{n}+o\left(\dfrac{t}{n}\right),\\\mathbb{P}(Y_i=1)=\lambda\cdot\dfrac{t}{n}+o\left(\dfrac{t}{n}\right),\\\mathbb{P}(Y_i\geq 2)=o\left(\dfrac{t}{n}\right).\end{cases}$$

对任意的正整数 k, 记事件

$$\begin{aligned}A_n&=\left(\sum_{i=1}^n Y_i=k,\text{其中 }k\text{ 个 }Y_i=1,\text{其余为 }0\right),\\B_n&=\left(\sum_{i=1}^n Y_i=k,\text{其中存在 }Y_i\geq 2\right).\end{aligned}$$

对于事件 A_n, 有

$$\mathbb{P}(A_n)=\binom{n}{k}\cdot\mathbb{P}^k(Y_i=1)\cdot\mathbb{P}^{n-k}(Y_i=0),$$

在上式中令 $n\to\infty$, 依 Poisson 定理[2] 知 $\lim\limits_{n\to\infty}\mathbb{P}(A_n)=\dfrac{(\lambda t)^k}{k!}\cdot\mathrm{e}^{-\lambda t}$; 对于事件 B_n, 有

$$\mathbb{P}(B_n)\leq\mathbb{P}\left\{\bigcup_{i=1}^n(Y_i\geq 2)\right\}\leq\sum_{i=1}^n\mathbb{P}(Y_i\geq 2)=n\cdot o\left(\frac{t}{n}\right),$$

在上式中令 $n\to\infty$, 得 $\lim\limits_{n\to\infty}\mathbb{P}(B_n)=0$; 最后, 对任意的 n, 都有

$$\mathbb{P}(N(t)=k)=\mathbb{P}(A_n\cup B_n)=\mathbb{P}(A_n)+\mathbb{P}(B_n),$$

在上式中令 $n\to\infty$, 则

$$\mathbb{P}(N(t)=k)=\frac{(\lambda t)^k}{k!}\cdot\mathrm{e}^{-\lambda t},$$

再根据 $\{N(t)\}$ 是平稳增量过程知, 对任意的 $t,s\geq 0$, 都有 $N(s,s+t]\sim\mathcal{P}(\lambda t)$.

[1] M. Ross 的书 [4] 采用微分方程的方法来推导.

[2] 设 $\lim\limits_{n\to\infty}np_n=\lambda$, 则 $\lim\limits_{n\to\infty}\dbinom{n}{k}p_n^k(1-p_n)^{n-k}=\dfrac{\lambda^k\cdot\mathrm{e}^{-\lambda}}{k!}$, 该定理可以说明, 满足一定条件的二项分布的极限是 Poisson 分布. 详见 [2].

2.1.3 Poisson 过程的应用

Poisson 过程在实际生活中具有广泛的应用, 以下是几个例子.

例 2.1 上海证券交易所开盘后, 股票买卖的依次成交构成一个 Poisson 过程. 如果每 10 分钟平均有 12 万次买卖成交, 计算该 Poisson 过程的强度 λ 和 1 秒内成交 100 次的概率.

解答 用 $\{N(t)\}$ 表示所述的 Poisson 过程, 10 分钟内的平均成交次数是

$$\mathbb{E}N(t,t+10] = 10\lambda = 120000,$$

于是 $\lambda = 12000$ 次/分钟. 用 $\{N_1(t)\}$ 表示以秒为单位的 Poisson 过程时, 强度是 $\lambda_1 = \dfrac{\lambda}{60} = 200$. 于是 1 秒内成交 100 次的概率是

$$\mathbb{P}(N_1(1) = 100) = \frac{\lambda_1^{100}}{100!} \cdot e^{-\lambda_1} = \frac{200^{100}}{100!} \cdot e^{-200} \approx 1.88 \times 10^{-15}.$$

例 2.2 设车辆通过的数量是一个 Poisson 过程, 且在一分钟内有 $\mathbb{P}(N(1) = 0) = 0.2$.

(1) 求 $\mathbb{P}(N(2) > 1)$;
(2) 求 $\mathbb{E}N(5)$;
(3) 求 $\mathrm{Var}N(5)$;
(4) 求 $\mathbb{P}(N(5) \geq 1)$.

解答 首先由 $\mathbb{P}(N(1) = 0) = e^{-\lambda} = 0.2$, 解得 $\lambda = \ln 5$. 对于 (1), 计算得

$$\begin{aligned}\mathbb{P}(N(2) > 1) &= 1 - \mathbb{P}(N(2) = 0) - \mathbb{P}(N(2) = 1)\\ &= 1 - e^{-2\lambda} - 2\lambda \cdot e^{-2\lambda}\\ &= \frac{24 - 2\ln 5}{25};\end{aligned}$$

对于 (2) 和 (3), 容易得到 $\mathbb{E}N(5) = \mathrm{Var}N(5) = 5\lambda = 5\ln 5$; 对于 (4), 计算得

$$\mathbb{P}(N(5) \geq 1) = 1 - \mathbb{P}(N(5) = 0) = 1 - e^{-5\lambda} = \frac{3124}{3125}.$$

接下来的例子, 可以说明 Poisson 过程与二项分布之间的关系.

例 2.3 设某商场中, 男顾客平均每分钟有 1 人, 而女顾客平均每分钟有 2 人.

(1) 到达商场的总顾客人数服从什么分布?
(2) 已知 t 时刻商场中已有 50 人, 试求出商场中已有 30 个女性顾客的概率是多少? 平均有多少个女性顾客?

解答 (1) 设男顾客为强度为 λ_1 的 Poisson 过程 $\{N_1(t)\}$, 女顾客为强度为 λ_2 的 Poisson 过程 $\{N_2(t)\}$, 并且 $\mathbb{E}N_1(1) = \lambda_1 = 1, \mathbb{E}N_2(1) = \lambda_2 = 2$. 记 $N(t) = N_1(t) + N_2(t)$, 根据 Poisson 分布对参数 λ 的再生性知, $\{N(t)\}$ 为强度为 $\lambda = \lambda_1 + \lambda_2 = 3$ 的 Poisson 过程.

(2) 计算得

$$\begin{aligned}\mathbb{P}(N_2(t)=30|N(t)=50) &= \frac{\mathbb{P}(N_1(t)=20, N_2(t)=30)}{\mathbb{P}(N(t)=50)}\\ &= \frac{\mathbb{P}(N_1(t)=20)\mathbb{P}(N_2(t)=30)}{\mathbb{P}(N(t)=50)}\\ &= \frac{t^{20}(2t)^{30}}{(3t)^{50}}\cdot\frac{50!}{20!\cdot 30!}\cdot \mathrm{e}^{-(1+2-3)t}\\ &= \binom{50}{20}\left(\frac{2}{3}\right)^{30}\left(\frac{1}{3}\right)^{20}\\ &\approx 0.0705.\end{aligned}$$

在上面的基础上, 记 $n=50, p=\dfrac{2}{3}$, 容易验证 $N_2(t)\sim\mathcal{B}(n,p)$. 因此

$$\mathbb{E}N_2(t)=np=\frac{100}{3}\approx 33.33.$$

2.2 Poisson 呼叫流

在本节中, 我们研究的对象是 Poisson 过程中事件所发生的时刻.

定义 2.4 (呼叫时刻与呼叫流)

设 $\{N(t)\}$ 为强度为 λ 的 Poisson 过程, S_n 表示第 n 个事件发生的时刻, S_n 称作第 n 个呼叫时刻或到达时刻, $\{S_n\}$ 称为 $\{N(t)\}$ 的呼叫流. ♣

为了研究 Poisson 呼叫流, 我们自然想要研究其分布. 为此需要解决两个问题:

(1) 首先, 随机变量 S_n 的分布是什么, 是否可以求出来?

(2) 其次, 是否可以求出 $(S_1,S_2,\cdots,S_n)$ 的联合分布?

2.2.1 等待时间间隔与到达时刻的分布

记等待时间间隔 $X_n=S_n-S_{n-1}$, 则 $S_n=X_1+X_2+\cdots+X_n=\sum\limits_{n=1}^{\infty}X_n$. 我们首先尝试求出 $\{X_n\}$ 的分布.

定理 2.2

Poisson 过程 $\{N(t)\}$ 的等待时间间隔 $X_1,X_2,\cdots$ 是来自指数总体 $\mathrm{Exp}(\lambda)$ 的随机变量. ♡

证明 首先, 事件 $(X_1>t)$ 的含义是 $[0,t]$ 内无事件发生. 根据

$$\mathbb{P}(X_1>t)=\mathbb{P}(N(t)=0)=\mathrm{e}^{-\lambda t},$$

得到 $X_1 \sim \mathrm{Exp}(\lambda)$. 接下来, 根据 Poisson 过程的独立增量性, 有

$$\mathbb{P}(X_2 > t | X_1 = s) = \mathbb{P}(N(s, s+t] = 0 | X_1 = s) = \mathbb{P}(N(s, s+t] = 0) = \mathrm{e}^{-\lambda t},$$

从而 $X_2 \sim \mathrm{Exp}(\lambda)$. 以此类推, 即可证明该定理.

接下来, 我们再来尝试求出第 n 个呼叫时 $S_n = X_1 + X_2 + \cdots + X_n$ 所服从的分布. 一方面, 我们可以根据 $\{X_n\}$ 独立同分布, 且来自总体 $\mathrm{Exp}(\lambda)$, 直接得到 S_n 的分布是 $\Gamma(n, \lambda)$; 另外一方面, 我们也可以从 $\{N(t)\}$ 出发来得到 $\{S_n\}$ 的分布. 为此, 我们通过如下的方式来建立 $\{S_n\}$ 与 $\{N(t)\}$ 的联系:

- $\{N(t) \geq n\} \iff \{S_n \leq t\}$;
- $\{N(t) = n\} \iff \{S_n \leq t < S_{n+1}\}$.

上面两条是容易理解的. 如果 $N(t) \geq n$, 则第 n 个事件在 t 时刻前发生, 从而 $S_n \leq t$; 而如果 $N(t) = n$, 则在 $[0, t]$ 内恰有 n 个事件发生, 而第 $n+1$ 次发生位于 (t, ∞) 内, 从而 $S_n \leq t < S_{n+1}$.

定理 2.3 设 S_n 是 Poisson 过程的第 n 个到达时刻, 则 $S_n \sim \Gamma(n, \lambda)$.

证明 设 $F_n(t)$ 是 S_n 的分布函数. 根据 $\{S_n\}$ 与 $\{N(t)\}$ 之间的联系, 有

$$F_n(t) = \mathbb{P}(S_n \leq t) = \mathbb{P}(N(t) \geq n) = 1 - \sum_{k=0}^{n-1} \mathbb{P}(N(t) = k) = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t}.$$

考虑 S_n 的概率密度函数

$$\begin{aligned}
g_n(t) &= \frac{\mathrm{d}}{\mathrm{d}t} F_n(t) \\
&= \lambda \mathrm{e}^{-\lambda t} - \sum_{k=1}^{n-1} \frac{\lambda}{(k-1)!} (\lambda t)^{k-1} \mathrm{e}^{-\lambda t} + \sum_{k=1}^{n-1} \frac{\lambda}{k!} (\lambda t)^k \mathrm{e}^{-\lambda t} \\
&= \lambda \mathrm{e}^{-\lambda t} - \lambda \mathrm{e}^{-\lambda t} + \frac{\lambda^n}{(n-1)!} t^{n-1} \mathrm{e}^{-\lambda t} \\
&= \frac{\lambda^n}{(n-1)!} t^{n-1} \mathrm{e}^{-\lambda t},
\end{aligned}$$

以上的概率密度函数正是 $\Gamma(n, \lambda)$ 的概率密度函数, 因此 $S_n \sim \Gamma(n, \lambda)$.

2.2.2 到达时刻的联合分布

在求出了 S_n 的分布之后, 我们更进一步, 希望求出 $\boldsymbol{S} = (S_1, S_2, \cdots, S_n)$ 的分布, 也即到达时刻的联合分布. 为此, 需要先介绍引理2.1.

引理 2.1

设 $F(x_1,x_2,\cdots,x_n)$ 是 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的联合分布函数, 令

$$G_k(x_1,x_2,\cdots,x_n)=\mathbb{P}(X_1>x_1,X_2\le x_2,\cdots,X_{2k-1}>x_{2k-1},X_j\le x_j,2k\le j\le n),$$

则在 G_k 存在 n 阶连续混合偏导的区域内, F 存在 n 阶连续混合偏导数, 且

$$\frac{\partial^n F(x_1,x_2,\cdots,x_n)}{\partial x_n\cdots\partial x_1}=(-1)^k\frac{\partial^n G(x_1,\cdots,x_n)}{\partial x_n\cdots\partial x_1}.$$

我们不加证明地直接使用引理2.1. 为了方便理解, 考虑一个特殊的例子: 当 $n=2$ 时, 令

$$\begin{aligned}G(x,y)&=\mathbb{P}(X>x,y\le y)\\&=P(Y\le y)-\mathbb{P}(X\le x,Y\le y),\end{aligned}$$

则在 G 存在连续混合偏导的区域内, 有

$$\frac{\partial^2 G(x,y)}{\partial y\partial x}=-\frac{\partial^2 F(x,y)}{\partial y\partial x}.$$

借助引理2.1, 我们来求 $\boldsymbol{S}=(S_1,S_2,\cdots,S_n)$ 的分布.

定理 2.4

设 S_n 是 Poisson 过程的第 n 个到达时刻, 则 $\boldsymbol{S}=(S_1,S_2,\cdots,S_n)$ 的概率密度函数

$$g(s_1,\cdots,s_n)=\lambda^n\mathrm{e}^{-\lambda s_n},\quad 0<s_1<\cdots<s_n.$$

证明 首先, 设 $n=2k-1$, 并令 $\boldsymbol{s}=(s_1,s_2,\cdots,s_n)$, 其中 $0<s_1<\cdots<s_n$. 记事件

$$A_i=(N(s_{i-1},s_i]=0),\quad B_i=(N(s_{i-1},s_i]=2),$$

则当 $i\ne j$ 时, A_i 与 A_j, A_i 与 B_j, B_i 与 B_j 均独立, 且

$$\mathbb{P}(A_i)=\mathrm{e}^{-\lambda(s_i-s_{i-1})},\quad \mathbb{P}(B_i)=\frac{\lambda^2(s_i-s_{i-1})^2}{2}\mathrm{e}^{-\lambda(s_i-s_{i-1})},$$

从而

$$\begin{aligned}G_k(s_1,s_2,\cdots,s_n)&=\mathbb{P}(S_1>s_1,S_2\le s_2,\cdots,S_n>s_n)\\&=\mathbb{P}(A_1B_2A_3B_4\cdots A_n)\\&=\mathbb{P}(A_1)\mathbb{P}(B_2)\mathbb{P}(A_3)\mathbb{P}(B_4)\cdots\mathbb{P}(A_n)\\&=\frac{\lambda^{2(k-1)}\mathrm{e}^{-\lambda s_n}}{2^{k-1}}\cdot(s_2-s_1)^2(s_4-s_3)^2\cdots(s_{2k-2}-s_{2k-3})^2,\end{aligned}$$

对其求混合偏导得

$$\frac{\partial^n G_k(s_1,\cdots,s_n)}{\partial s_n\cdots\partial s_1}=\frac{\lambda^{2(k-1)}}{2^{k-1}}\cdot(-2)^{k-1}\cdot(-1)\cdot\mathrm{e}^{-\lambda s_n}\cdot\lambda.$$

根据引理2.1, 得到 $\boldsymbol{S}$ 的概率密度函数

$$g(s_1,\cdots,s_n)=(-1)^k\cdot\frac{\partial^n G_k(s_1,\cdots,s_n)}{\partial s_n\cdots\partial s_1}=\lambda^n\mathrm{e}^{-\lambda s_n}.$$

接下来, 设 $n=2k$, 同样可以得到, 对 $0<s_1<\cdots<s_n$, 有 $g(s_1,\cdots,s_n)=\lambda^n \mathrm{e}^{-\lambda s_n}$.

2.2.3 到达时刻的条件分布

在上面的基础上, 我们来求出给定 $N(t)$ 时 S_n 的分布. 例如, 若给定 $N(t)=1$, 考虑 S_1 的分布, 设 $s\le t$, 则有

$$\begin{aligned}\mathbb{P}(S_1\le s|N(t)=1)&=\frac{\mathbb{P}(S_1\le s,N(t)=1)}{\mathbb{P}(N(t)=1)}\\&=\frac{\mathbb{P}(N(s)=1,N(s,t]=0)}{\mathbb{P}(N(t)=1)}\\&=\frac{\lambda s\cdot\mathrm{e}^{-\lambda s}\cdot\mathrm{e}^{-\lambda(t-s)}}{\lambda t\cdot\mathrm{e}^{-\lambda t}}\\&=\frac{s}{t}.\end{aligned}$$

接下来, 进一步求出给定 $N(t)=n$ 时, $\boldsymbol{S}=(S_1,\cdots,S_n)$ 的联合分布. 为此, 我们需要用到求概率密度的“微元法”: 对于一维的情况, 有

$$f(y)=\lim_{h\to 0}\frac{\mathbb{P}(y<Y\le y+h)}{h}\implies\mathbb{P}(y<Y\le y+h)=f(y)h+o(h);$$

而对于 n 维情形, 则有

$$f(y_1,\cdots,y_n)=\lim_{h_1,\cdots,h_n\to 0}\frac{\mathbb{P}(y_1<Y_1\le y_1+h_1,\cdots,y_n<Y_n\le y_n+h_n)}{h_1\cdots h_n},$$

或者写成

$$\mathbb{P}(y_1<Y_1\le y_1+h_1,\cdots,y_n<Y_n\le y_n+h_n)=f(y_1,\cdots,y_n)h_1\cdots h_n+o(h_1\cdots h_n).$$

最后一个式子便是我们接下来要用的结论.

定理 2.5

设 S_n 是 Poisson 过程的第 n 个到达时刻, 则在条件 $N(t)=n$ 下, $\boldsymbol{S}=(S_1,S_2,\cdots,S_n)$ 的联合密度

$$h(s_1,s_2,\cdots,s_n)=\frac{n!}{t^n},\quad 0<s_1<\cdots<s_n<t.$$

证明 首先考虑条件概率

$$\begin{aligned}&\mathbb{P}(s_i<S_i\le s_i+h_i,1\le i\le n|N(t)=n)\\&=\frac{\mathbb{P}(N(s_i,s_i+h_i]=1,1\le i\le n,\text{且在 }[0,t]\text{ 内别处无事件发生})}{\mathbb{P}(N(t)=n)}\\&=\frac{1}{\frac{(\lambda t)^n}{n!}\cdot\mathrm{e}^{-\lambda t}}\cdot\prod_{i=1}^{n}\lambda h_i\cdot\mathrm{e}^{-\lambda h_i}\cdot\exp\left\{-\lambda\left(t-\sum_{i=1}^{n}h_i\right)\right\}\\&=\frac{n!}{t^n}\cdot h_1\cdots h_n,\end{aligned}$$

因此概率密度函数

$$h(s_1, s_2, \cdots, s_n) = \frac{n!}{t^n}.$$

2.2.4 到达时刻与均匀分布的联系

在上面的结果的基础上, 设 $U \sim \mathcal{U}[0, t]$, $U_1, U_2, \cdots, U_n$ 是来自总体 U 的随机变量, 考虑排序后的随机向量 $(U_{(1)}, U_{(2)}, \cdots, U_{(n)})$, 通过计算得到其概率密度函数也为 $\frac{n!}{t^n}$. 因此 $(U_{(1)}, U_{(2)}, \cdots, U_{(n)})$ 与 $(S_1, S_2, \cdots, S_n)|N(t) = n$ 同分布. 进一步, 可以得到如下的重要结论.

定理 2.6

设 $U \sim \mathcal{U}[0, t]$, $U_1, U_2, \cdots, U_n$ 是来自总体 U 的随机变量, $h(s)$ 是实函数, 则

(1) $\sum_{i=1}^{n} S_i|N(t) = n$ 和 $\sum_{i=1}^{n} U_i$ 同分布;

(2) $\sum_{i=1}^{n} h(S_i)|N(t) = n$ 和 $\sum_{i=1}^{n} h(U_i)$ 同分布;

(3) 当 $\mathbb{E}h(U)$ 存在时, $\mathbb{E}\left(\sum_{i=1}^{n} h(S_i)\middle|N(t) = n\right) = n\mathbb{E}h(U)$.

证明 因为 $(U_{(1)}, U_{(2)}, \cdots, U_{(n)})$ 与 $(S_1, S_2, \cdots, S_n)|N(t) = n$ 同分布, 所以

$$\sum_{i=1}^{n} S_i|N(t) = n \quad 与 \quad \sum_{i=1}^{n} U_i = \sum_{i=1}^{n} U_{(i)}$$

同分布, 同理可以得到

$$\sum_{i=1}^{n} h(S_i)|N(t) = n \quad 与 \quad \sum_{i=1}^{n} h(U_i) = \sum_{i=1}^{n} h(U_{(i)})$$

同分布, 再利用同分布的随机变量具有相同的数学期望, 即可得到

$$\mathbb{E}\left(\sum_{i=1}^{n} h(S_i)\middle|N(t) = n\right) = n\mathbb{E}h(U).$$

2.2.5 简单呼叫流

根据上一小节中的结论, 设 $\{N(t)\}$ 为 Poisson 过程, $\{X_n\}$ 为其等待时间间隔, 则 $X_i \sim \mathrm{Exp}(\lambda)$, 并且到达时刻 $S_n = X_1 + X_2 + \cdots + X_n \sim \Gamma(n, \lambda)$. 另外, 如果从指数分布出发, 可以给出以下的定义.

定义 2.5 (简单呼叫流)

设随机变量 $Y_i \sim \mathrm{Exp}(\lambda)$, 则称

$$\xi_n = Y_1 + Y_2 + \cdots + Y_n, \quad n = 1, 2, \cdots$$

是简单呼叫流或 *Poisson* 流.

Poisson 过程本质上是一个计数过程. 通过到达时刻, 我们将 Poisson 过程和一个随机序列联系在了一起. 反过来, 定义2.5所给出的随机序列也可以和一个计数过程 (记作 $\{M(t)\}$) 联系在一起. 我们自然会好奇, 这个计数过程 $\{M(t)\}$ 是否就是 Poisson 过程.

注意到 $M(t) = m$ 等价于 $[0, t]$ 内恰好有 m 次呼叫, 因此

$$M(t) = \sum_{i=1}^{\infty} I[\xi_j \le t], \quad t \ge 0,$$

其中 $I[x]$ 是示性函数. 接下来验证, $\{M(t)\}$ 就是一个参数为 λ 的 Poisson 过程. 此时 Poisson 过程 $\{N(t)\}$ 也可以写成

$$N(t) = \sum_{i=1}^{\infty} I[S_i \le t], \quad t \ge 0.$$

对任意的正整数 n, 以及对任意的 $0 < t_1 < t_2 < \cdots < t_n$, 都有

$$(N(t_1), N(t_2), \cdots, N(t_n)) \quad 和 \quad (M(t_1), M(t_2), \cdots, M(t_n))$$

同分布, 因此 $M(t)$ 是 Poisson 过程.

在本节中, 研究等待时间也是有意义的. 见下例.

例 2.4 等待时间的期望 设火车站顾客的数量为 λ 的 Poisson 过程 $\{N(t)\}$, 火车 t 时刻离开车站, 求 $[0, t]$ 内到达车站的顾客等待时间总和的期望.

解答 设 S_i 为第 i 个顾客的到达时间, 则 $t - S_i$ 为等待时间, 等待时间的总和

$$T = \sum_{i=1}^{N(t)} (t - S_i),$$

计算得在 $N(t) = n$ 时的条件期望

$$\mathbb{E}(T|N(t) = n) = \mathbb{E}\left(\sum_{i=1}^{n} (t - S_i) \middle| N(t) = n\right)$$

再设 $U_i \sim \mathcal{U}[0, t]$, 根据定理2.6得

$$\mathbb{E}\left(\sum_{i=1}^{n} (t - S_i) \middle| N(t) = n\right) = E\left(\sum_{i=1}^{n} (t - U_i)\right) = \sum_{i=1}^{n} \mathbb{E}U_i = \frac{1}{2}nt.$$

因此

$$\mathbb{E}T = \mathbb{E}\left(\mathbb{E}(T|N(t) = n)\right) = \frac{1}{2}t \cdot \mathbb{E}N(t) = \frac{1}{2}\lambda t^2.$$

例 2.5 汽车按照强度为 λ 的 Poisson 流通过广场, 第 i 辆汽车通过时造成的空气污染为 D_i. D_i 随着时间的推移而减弱, 经过时间 s 污染减弱为 $D_i \mathrm{e}^{-as}$, 其中正常数 a 是扩散常

数. 假设 $D_1, D_2, \cdots$ 是来自总体 D 的随机变量, 且与 Poisson 流独立. 计算 $[0,t]$ 内通过的汽车在 t 时造成的平均污染.

解答 用 $\{N(t)\}$ 表示所述的 Poisson 过程, 用 S_i 表示第 i 辆汽车的通过时间. $[0,t]$ 内通过了 $N(t)$ 辆汽车, 造成 t 时的污染是

$$D(t) = \sum_{i=1}^{N(t)} D_i \mathrm{e}^{-a(t-S_i)}.$$

注意 D_i 和 $N(t), S_i$ 独立, 从而

$$\begin{aligned}
\mathbb{E}(D(t)|N(t)=n) &= \sum_{i=1}^{n} \mathbb{E}\left(D_i \mathrm{e}^{-a(t-S_i)}\middle| N(t)=n\right) \\
&= \sum_{i=1}^{n} \mathbb{E}D_i \mathrm{e}^{-at} \cdot \mathbb{E}\left(\mathrm{e}^{aS_i}\middle| N(t)=n\right) \\
&= \mathbb{E}D\mathrm{e}^{-at} \cdot \mathbb{E}\left(\sum_{i=1}^{n} \mathrm{e}^{aS_n}\middle| N(t)=n\right),
\end{aligned}$$

再利用定理2.6得

$$\mathbb{E}\left(\mathrm{e}^{aS_i}\middle| N(t)=n\right) = \mathbb{E}\sum_{i=1}^{n} \mathrm{e}^{aU_i} = \frac{n}{at}\left(\mathrm{e}^{at}-1\right),$$

其中 $U \sim \mathcal{U}[0,t]$, 因此

$$\mathbb{E}(D(t)|N(t)=n) = n \cdot \frac{\mathbb{E}D}{at} \cdot \left(1-\mathrm{e}^{-at}\right) = N(t) \cdot \frac{\mathbb{E}D}{at} \cdot \left(1-\mathrm{e}^{-at}\right),$$

从而 $\mathbb{E}D(t) = \mathbb{E}\mathbb{E}(D(t|N(t))) = \dfrac{\lambda\mathbb{E}D}{a} \cdot \left(1-\mathrm{e}^{-at}\right)$.

2.3 年龄与剩余寿命

定义 2.6 (年龄与剩余寿命)

设 $\{N(t)\}$ 是 Poisson 过程, $\{S_n\}$ 是等待时刻, 称

$$A(t) = t - S_{N(t)}, \quad R(t) = S_{N(t)+1} - t$$

分别为年龄和剩余寿命. ♣

年龄和剩余寿命具有如下的性质.

(1) 剩余寿命的分布: $R(t) \sim \mathrm{Exp}(\lambda)$;

(2) 年龄的分布: $\mathbb{P}(A(t) \le u) = \begin{cases} 1-\mathrm{e}^{-\lambda u}, & u \in [0,t), \\ 1, & u \ge t; \end{cases}$

(3) $A(t)$ 和 $R(t)$ 独立.

我们来验证上面的性质. 首先, 计算得

$$\mathbb{P}(R(t)>u)=\mathbb{P}(S_{N(t)+1}>u+t)=\mathbb{P}(N(t,u+t]=0)=\mathrm{e}^{-\lambda u},$$

因此 $R(t)\sim \mathrm{Exp}(\lambda)$; 其次, 设 $u\in[0,t)$, 则

$$\mathbb{P}(A(t)>u)=\mathbb{P}(S_{N(t)}<t-u)=\mathbb{P}(N[t-u,t]=0)=\mathbb{P}(N(t-u,t]=0)=\mathrm{e}^{-\lambda u},$$

但是当 $u\geq t$ 时, 一定有 $A(t)\leq t\leq u$, 因此

$$\mathbb{P}(A(t)\leq u)=\begin{cases}1-\mathrm{e}^{-\lambda u}, & u\in[0,t),\\ 1, & u\geq t;\end{cases}$$

最后, 注意到对任意的 u,v, $N[t-u,t]$ 与 $N(t,t+v]$ 都独立, 因此 $A(t)$ 与 $R(t)$ 也独立.

在上面研究了 $A(t)$ 和 $R(t)$ 的分布. 在此基础上, 我们可以研究 $S_{N(t)}$ 和 $S_{N(t)+1}$ 的分布. 首先, 根据 $A(t)$ 的分布得

$$\mathbb{P}(S_{N(t)}\leq s)=\mathbb{P}(t-A(t)\leq s)=\mathbb{P}(A(t)\geq t-s)=\begin{cases}\mathrm{e}^{-\lambda(t-s)}, & 0\leq s\leq t,\\ 1, & s>t;\end{cases}$$

接下来, 根据 $R(t)$ 的分布得

$$\mathbb{P}(S_{N(t)+1}\leq s)=\mathbb{P}(R(t)+t\leq s)=\mathbb{P}(R(t)\leq s-t)=\begin{cases}1-\mathrm{e}^{-\lambda(s-t)}, & s>t,\\ 0, & s\leq t.\end{cases}$$

接下来, 记 $X(t)=A(t)+R(t)$, 我们来尝试求出 $X(t)$ 和 $A(t)$ 的期望. 首先对 $X(t)$ 的期望进行估计, 有

$$\mathbb{E}X(t)=\mathbb{E}A(t)+\mathbb{E}R(t)=\mathbb{E}A(t)+\frac{1}{\lambda}>\frac{1}{\lambda},$$

可以认为 t 时服役的部件平均使用寿命比同型号部件的平均使用寿命大; 接下来, 考虑 $A(t)$ 期望, 计算得

$$\mathbb{E}A(t)=\int_0^\infty u\mathrm{d}\mathbb{P}(A(t)\leq u)=\int_0^t \lambda u\mathrm{e}^{-\lambda u}\mathrm{d}u+t\mathrm{e}^{-\lambda t},$$

因此 $\lim\limits_{t\to\infty}\mathbb{E}A(t)=\dfrac{1}{\lambda}$, 结合上式便有 $\lim\limits_{t\to\infty}\mathbb{E}X(t)=\dfrac{2}{\lambda}$.

2.4 Poisson 过程的汇合、分流与复合

本节中, 我们研究以下几个问题:

(1) 两个 Poisson 过程的和是否还是 Poisson 过程?

(2) 某个 Poisson 过程是否可以按照某种方式, 拆分为两个或者多个 Poisson 过程?

(3) 以及, Poisson 过程是否可以按照某种方式复合?

这些问题的研究也具有非常重要的现实意义.

2.4.1 Poisson 过程的汇合

首先考虑两个 Poisson 过程的相加.

定理 2.7

设 $\{N_1(t)\}, \{N_2(t)\}$ 是相互独立的, 强度分别为 λ_1 和 λ_2 的 Poisson 过程, 则

$$N(t) = N_1(t) + N_2(t), \quad t \geq 0$$

是强度为 $\lambda_1 + \lambda_2$ 的 Poisson 过程.

证明 (1) 首先证明 $N(0) = 0$. 这是因为 $N(0) = N_1(0) + N_2(0) = 0$.

(2) 其次证明 $\{N(t)\}$ 是独立增量过程与平稳增量过程. 对于独立增量性, 对任意的 n, 以及对任意的 $0 < t_1 < t_2 < \cdots < t_n$, 根据 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 的独立增量性, 知

$$N_1(t_1, t_2], \quad N_1(t_2, t_3], \quad \cdots, \quad N_1(t_{n-1}, t_n]$$

和

$$N_2(t_1, t_2], \quad N_2(t_2, t_3], \quad \cdots, \quad N_2(t_{n-1}, t_n]$$

是相互独立的, 又 $N(t) = N_1(t) + N_2(t)$, 因此

$$N(t_1, t_2], \quad N(t_2, t_3], \quad \cdots, \quad N(t_{n-1}, t_n]$$

也是相互独立的; 对于平稳增量性, 根据 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 的平稳增量性, 对任意的 $s > 0, t_2 > t_1 > 0$, $N_1(t_1, t_2]$ 与 $N_1(t_1 + s, t_2 + s]$ 同分布, $N_2(t_1, t_2]$ 与 $N_2(t_1 + s, t_2 + s]$ 同分布, 又 $N(t) = N_1(t) + N_2(t)$, 因此 $N(t_1, t_2]$ 与 $N(t_1 + s, t_2 + s]$ 同分布.

(3) 最后记 $\lambda = \lambda_1 + \lambda_2$, 我们证明

$$\begin{cases} \mathbb{P}(N(t, t + \Delta t] = 0) = 1 - \lambda \cdot \Delta t + o(\Delta t), \\ \mathbb{P}(N(t, t + \Delta t] = 1) = \lambda \cdot \Delta t + o(\Delta t), \\ \mathbb{P}(N(t, t + \Delta t] \geq 2) = o(\Delta t). \end{cases}$$

先计算得

$$\begin{aligned} \mathbb{P}(N(t, t + \Delta t] = 0) &= \mathbb{P}(N(t + \Delta t) - N(t) = 0) \\ &= \mathbb{P}(N_1(t + \Delta t) - N_1(t) + N_2(t + \Delta t) - N_2(t) = 0) \\ &= \mathbb{P}(N_1(t + \Delta t) - N_1(t) = 0, N_2(t + \Delta t) - N_2(t) = 0) \\ &= (1 - \lambda_1 \cdot \Delta t + o(\Delta t))(1 - \lambda_2 \cdot \Delta t + o(\Delta t)) \\ &= 1 - (\lambda_1 + \lambda_2) \cdot \Delta t + o(\Delta t); \end{aligned}$$

用同样的方法, 可以计算得到

$$\begin{aligned}\mathbb{P}(N(t,t+\Delta t]=1) &= \mathbb{P}(N_1(t+\Delta t)-N_1(t)=1, N_2(t+\Delta t)-N_2(t)=0)\\ &\quad+\mathbb{P}(N_1(t+\Delta t)-N_1(t)=0, N_2(t+\Delta t)-N_2(t)=1)\\ &=\lambda_1\cdot\Delta t(1-\lambda_2\cdot\Delta t+o(\Delta t))+\lambda_2\cdot\Delta t(1-\lambda_1\cdot\Delta t+o(\Delta t))\\ &=(\lambda_1+\lambda_2)\Delta t+o(\Delta t);\end{aligned}$$

最后, 注意到 $\mathbb{P}(N(t,t+\Delta t]=0)+\mathbb{P}(N(t,t+\Delta t]=1)+\mathbb{P}(N(t,t+\Delta t]\geq 2)=1$, 根据上面的结果即可得到

$$\mathbb{P}(N(t,t+\Delta t]\geq 2)=o(\Delta t).$$

根据定理2.7, 容易推得以下结论, 这也可以称为 Poisson 过程的可加性.

定理 2.8

设 $\{N_i(t), i=1,2,\cdots,m\}$ 是相互独立的, 强度分别为 λ_i 是 Poisson 过程, 则

$$N(t)=N_1(t)+N_2(t)+\cdots+N_m(t)$$

是强度为 $\lambda_1+\lambda_2+\cdots+\lambda_m$ 的 Poisson 过程.

在这里, 或许读者会思考, 若 $\lambda_1>\lambda_2$, $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 分别为强度为 λ_1 和 λ_2 的 Poisson 过程, 那么 $N_1(t)-N_2(t)$ 是否是 Poisson 过程? 答案是否定的. 这是因为, 有可能 $N_2(t)>N_1(t)$, 从而 $N_1(t)-N_2(t)<0$. 这便说明了, $N_1(t)-N_2(t)$ 甚至不是一个计数过程.

2.4.2 Poisson 过程的分流

接下来, 考虑将一个 Poisson 过程分流为两个不同的计数过程. 在这里的要求是, "分流"的过程与 Poisson 过程本身是独立的. 为此, 引入一个与该过程相独立的服从两点分布的随机变量来刻画"分流"的过程.

定理 2.9

设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, 随机变量 $\{Y_i\}$ 独立同分布, 与 $\{N(t)\}$ 独立, 且服从参数为 p 的两点分布, 也即

$$\mathbb{P}(Y_i=1)=p,\quad \mathbb{P}(Y_i=0)=1-p,\quad i=1,2,\cdots,$$

定义

$$N_1(t)=\sum_{i=1}^{N(t)}Y_i,\quad N_2(t)=\sum_{i=1}^{N(t)}(1-Y_i),$$

则 $N(t)=N_1(t)+N_2(t)$, 且 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 分别为强度为 $\lambda_1=p\lambda$ 和 $\lambda_2=(1-p)\lambda$ 的 Poisson 过程.

定理2.9的证明较为复杂, 在此不再叙述. 另外, 以下的定理给出了 Poisson 定理的可分解性, 相当于是定理2.9的推广.

定理 2.10

设 $\{N(t)\}$ 为强度为 λ 的 Poisson 过程, 每次发生的事件有 p_i 的概率被分入 A_i 中, 其中 $1 \leq i \leq n$, 且该过程与发生的时间独立. 设 $N_i(t)$ 表示 $[0,t]$ 内 A_i 中事件的个数, 则 $\{N_i(t)\}$ 是强度为 $\lambda_i = p_i\lambda$ 的 Poisson 过程. 当 $p_1 + p_2 + \cdots + p_n = 1$ 时, 这 n 个 Poisson 过程独立.

对于 Poisson 过程的汇合与分流, 有如下的实际的例子.

例 2.6 汽车按 Poisson 流驶向立体交叉桥 A. 经过调查知道, 由东面每分钟平均驶入 6 辆汽车, 由南面每分钟平均驶入 6.5 辆汽车, 由西面每分钟平均驶入 9 辆汽车, 由北面每分钟平均驶入 8.5 辆汽车. 在桥 A 上, 每辆车向左或向右转向行驶的概率是 0.3, 直行的概率是 0.35, 调头行驶的概率是 0.05. 计算各个方向上, 离开立交桥的汽车流的车流强度.

解答 对于该过程, 作出如下的分流表.

方向	向东分流	向南分流	向西分流	向北分流
东面驶入 $\lambda_1 = 6.0$	0.05	0.30	0.35	0.30
南面驶入 $\lambda_2 = 6.5$	0.30	0.05	0.30	0.30
西面驶入 $\lambda_3 = 9.0$	0.35	0.30	0.05	0.30
北面驶入 $\lambda_4 = 8.5$	0.30	0.35	0.30	0.05
驶出强度	$\lambda_1' = 7.95$	$\lambda_2' = 7.80$	$\lambda_3' = 7.05$	$\lambda_4' = 7.20$

例 2.7 从 $t = 0$ 开始, 客户按强度为 λ 的 Poisson 流点击一个网站. 每个客户点击后的浏览时间是相互独立的, 有共同的分布函数 $G(t)$. 用 $N_1(t)$ 表示 t 时已经离线的客户数, 用 $N_2(t)$ 表示 t 时在线的客户数, 则 $N_1(t), N_2(t)$ 是两个相互独立的 Poisson 随机变量, 分别有数学期望

$$\mathbb{E}N_1(t) = \lambda \int_0^t G(s)\mathrm{d}s, \quad \mathbb{E}N_2(t) = \lambda \int_0^t \overline{G}(s)\mathrm{d}s,$$

其中 $\overline{G}(s) = 1 - G(s)$.

证明 对于一个客户来讲, 用 S 表示他进入网站的时间, 用事件 A 表示他 t 时已经离线, 用 Y 表示他的在线时间. 对于 $s \leq t$, 有

$$\mathbb{P}(A|S = s) = \mathbb{P}(Y \leq t - s) = G(t - s).$$

记 $\mathbb{P}_t(\cdot) = \mathbb{P}(\cdot|S \le t)$, 则 $\mathbb{P}_t(\cdot)$ 是概率. 当 $S \le t$ 时, S 在 $[0,t]$ 内均匀分布, 从而

$$\begin{aligned}
\mathbb{P}_t(A) &= \mathbb{P}(A|S \le t)\\
&= \int_0^t \mathbb{P}_t(A|S = s)\mathrm{d}\mathbb{P}_t(S \le s)\\
&= \int_0^t \mathbb{P}(A|S \le t, S = s)\mathrm{d}\mathbb{P}(S \le s|S \le t)\\
&= \frac{1}{t}\cdot\int_0^t \mathbb{P}(A|S = s)\mathrm{d}s\\
&= \frac{1}{t}\cdot\int_0^t G(t-s)\mathrm{d}s\\
&= \frac{1}{t}\cdot\int_0^t G(s)\mathrm{d}s,
\end{aligned}$$

记 $p = \frac{1}{t}\cdot\int_0^t G(s)\mathrm{d}s$, $q = 1 - p = \frac{1}{t}\cdot\int_0^t \overline{G}(s)\mathrm{d}s$, 则每个在 $[0,t]$ 内进入网站的人在 t 时刻离线的概率为 p, 在线的概率为 q, 与其他客户的行为独立. 用 $\{N(t)\}$ 表示强度为 λ 的 Poisson 过程, 利用二项分布得到

$$\mathbb{P}(N_1(t) = k, N_2(t) = j|N(t) = k+j) = \binom{k+j}{k}p^k q^j,$$

于是

$$\begin{aligned}
\mathbb{P}(N_1(t) = k, N_2(t) = j) &= \mathbb{P}(N(t) = k+j)\mathbb{P}(N_1(t) = k, N_2(t) = j|N(t) = k+j)\\
&= \frac{(\lambda t)^{k+j}}{(k+j)!}\cdot \mathrm{e}^{-\lambda t}\cdot\binom{k+j}{k}p^k q^j\\
&= \frac{(\lambda t p)^k}{k!}\cdot \mathrm{e}^{-\lambda t p}\cdot\frac{(\lambda t q)^k}{k!}\cdot \mathrm{e}^{-\lambda t q},
\end{aligned}$$

在上式中分别对 j 和 k 求和, 可得 $N_1(t)$ 和 $N_2(t)$ 的分布

$$N_1(t) \sim \mathcal{P}(\lambda t p), \quad N_2(t) \sim \mathcal{P}(\lambda t q),$$

从而 $\mathbb{E}N_1(t) = \lambda t p, \mathbb{E}N_2(t) = \lambda t q$.

2.4.3 Poisson 过程的复合

最后, 我们对 Poisson 过程的复合感兴趣.

定义 2.7 (复合 Poisson 过程)

设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, $\{Z_i\}$ 相互独立, 且 $\{Z_i\}$ 和 $\{N(t)\}$ 相互独立, 称

$$M(t) = \sum_{i=1}^{N(t)} Z_i$$

为复合 *Poisson* 过程. ♣

在此, 求出 $M(t)$ 的期望和方差.

定理 2.11

设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, $\{Z_i\}$ 是独立同分布的随机变量序列, 且 $\mathbb{E}Z_i = \mu$, $\mathrm{Var}Z_i = \sigma^2 < \infty$, 定义复合 Poisson 过程

$$M(t) = \sum_{i=1}^{N(t)} Z_i,$$

则 $\mathbb{E}M(t) = \mu\lambda t$, $\mathrm{Var}M(t) = \lambda t(\sigma^2 + \mu^2)$.

证明 首先, 计算得

$$\mathbb{E}(M(t)|N(t)=n) = \mathbb{E}\left(\sum_{i=1}^{n} Z_i\right) = n\mu = \mu \cdot N(t),$$

因此 $M(t)$ 的期望

$$\mathbb{E}M(t) = \mathbb{E}(\mathbb{E}(M(t)|N(t)=n)) = \mu\mathbb{E}N(t) = \mu\lambda t;$$

接下来, 计算得

$$\begin{aligned}
\mathbb{E}(M^2(t)|N(t)=n) &= \mathbb{E}\left(\sum_{i=1}^{n} Z_i\right)^2 \\
&= \sum_{i=1}^{n} \mathbb{E}Z_i^2 + \sum_{i\neq j} \mathbb{E}Z_i Z_j \\
&= \sum_{i=1}^{n} (\sigma^2 + \mu^2) + \sum_{i\neq j} \mu^2 \\
&= n^2\mu^2 + n\sigma^2 \\
&= \mu^2 \cdot N^2(t) + \sigma^2 \cdot N(t),
\end{aligned}$$

因此 $M(t)$ 的二阶矩

$$\mathbb{E}M^2(t) = \mathbb{E}(\mathbb{E}(M^2(t)|N(t)=n)) = \mu^2 \cdot (\lambda^2 t^2 + \lambda t) + \sigma^2 \cdot \lambda t,$$

最后计算得到

$$\mathrm{Var}M(t) = \mathbb{E}M^2(t) - (\mathbb{E}M(t))^2 = \lambda t(\mu^2 + \sigma^2).$$

例 2.8 在上海证券交易所, 宝钢股份的交易流是强度为 λ(笔/分钟) 的 Poisson 流. 设第 j 笔交易量是 Z_j 手, 如果 $\{Z_j\}$ 是来自总体 Z 的随机变量, $\mu = \mathbb{E}Z$, $\sigma^2 = \mathrm{Var}Z$, 计算宝钢股份一小时内的交易量的数学期望和标准差.

解答 用 $\{N(t)\}$ 表示强度为 λ 的 Poisson 过程, 则 60 分钟内的交易量

$$M(60)=\sum_{j=1}^{N(60)} Z_j.$$

根据定理2.11, 计算得一小时内的平均交易量 $\mathbb{E}M(60)=60\lambda\mu$, 标准差 $\sqrt{\mathrm{Var}M(60)}=\sqrt{60\lambda(\mu^2+\sigma^2)}$.

2.5 课后习题

问题 2.1 设某商场中, 男顾客平均每分钟有 1 人, 而女顾客平均每分钟有 2 人.

(1) 到达商场的总顾客人数服从什么分布?

(2) 已知 t 时刻商场中已有 50 人, 试求出商场中已有 30 个女性顾客的概率是多少? 平均有多少个女性顾客?

问题 2.2 设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, $0\le s<t$, 验证在条件 $N(t)=n$ 下, $N(s)$ 服从二项分布 $\mathcal{B}\left(n,\frac{s}{t}\right)$.

问题 2.3 对于 Poisson 过程 $\{N(t)\}$, 计算 $\mathbb{E}[N(t)N(t+s)]$ 和 $\mathbb{E}[N(t+s)|N(t)]$.

问题 2.4 对于强度为 λ 的 Poisson 过程, 用 $N(t-)$ 表示区间 $[0,t)$ 内发生的事件数, 则

(1) $N(t)-N(t-)=0$, a.s.;

(2) $N[s,t]=N(t)-N(s-)$ 是闭区间 $[s,t]$ 内发生的事件数;

(3) $N[s,t]=N(s,t]$, a.s..

问题 2.5 对于 $n=2k$, 验证

$$\begin{aligned}&\mathbb{P}(S_1>s_1,S_2\le s_2,\cdots,S_{n-1}>s_{n-1},S_n\le s_n)\\&=\lambda^{n-2}\frac{(s_2-s_1)^2(s_4-s_3)^2\cdots(s_{n-2}-s_{n-3})^2}{2^{k-1}}\mathrm{e}^{-\lambda s_n}\cdot\sum_{j=2}^{\infty}\frac{\lambda^j(s_n-s_{n-1})^j}{j!}.\end{aligned}$$

问题 2.6 求 $S_{N(t)}$ 和 $S_{N(t)+1}$ 的分布函数.

问题 2.7 若 $\lambda_1>\lambda_2$, $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 分别为强度为 λ_1 和 λ_2 的 Poisson 过程, 那么 $N(t)=N_1(t)-N_2(t)$ 是否是 Poisson 过程?

问题 2.8 设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, T 是和该 Poisson 过程独立的随机变量. 当 T 服从参数为 β 的指数分布时,

(1) 求 $N(T)$ 的概率分布;

(2) 计算 $\mathbb{E}N(T)$.

问题 2.9 在有很多鱼的湖中钓鱼时, 渔夫平均每小时钓到两条鱼. 如果渔夫每天的钓鱼时间 T 在 3 至 8 小时内均匀分布, 他平均每天钓多少条鱼? 方差是多少?

第 3 章 Brown 运动

3.1 自由扩散与 Brown 运动

3.1.1 自由扩散

设某个质点从初始时刻开始, 从坐标原点出发做自由扩散运动.

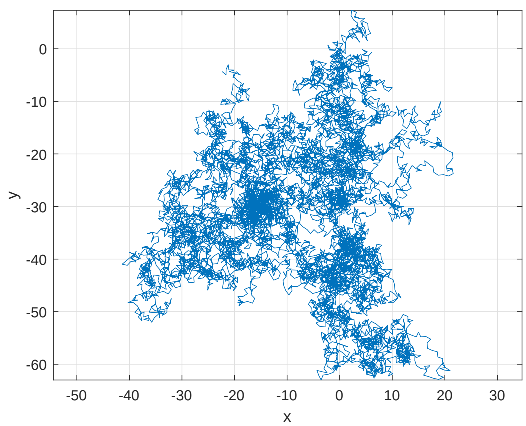


图 3.1: 利用 MATLAB 模拟自由扩散运动

建立直角坐标系, 设 t 时刻质点的位置为 $(X(t), Y(t))$, 考虑横坐标的位移 $X(t)$, 容易理解 $X(t)$ 具有以下的性质.

(1) 初始时刻位移为 0, 也即 $X(0) = 0$;

(2) 独立增量性, 也即在互不相交的时间段 $(t_{i-1}, t_i], i = 1, 2, \cdots, n$, 质点的位移

$$X(t_i) - X(t_{i-1}), \quad i = 1, 2, \cdots, n$$

是相互独立的;

(3) 空间对称性, 也即 $\mathbb{E}X(t) = 0$;

(4) 平稳增量性, 也即对任意的长度相等的时间段 $(s, t]$ 和 $(s+h, t+h]$, 质点的位移

$$X(t) - X(s), \quad X(t+h) - X(s+h)$$

有相同的分布;

(5) 有限性, 也即 $[0, t]$ 内位移的方差 $\sigma^2(t) = \mathrm{Var}X(t)$ 是 t 的连续函数.

将该运动与 Poisson 过程进行对比, 可以发现, Poisson 过程在时间上连续, 但是在空间上离散; 而该运动在时间上和在空间上都是连续的.

我们已经知道 $\mathbb{E}X(t)=0$. 另外, 我们来尝试推导方差 $\mathrm{Var}X(t)$ 的表达式. 计算得

$$\begin{aligned}\mathrm{Var}X(t+s)&=\mathrm{Var}[X(t+s)-X(t)+X(t)]\\&=\mathrm{Var}[X(t+s)-X(t)]+\mathrm{Var}X(t)\\&=\mathrm{Var}X(s)+\mathrm{Var}X(t),\end{aligned}$$

上面的结果说明了 $\mathrm{Var}X(t)$ 作为 t 的函数满足 Cauchy 方程[1], 因此 $\sigma^2(t)=\mathrm{Var}X(t)=Dt$. 在这里, D 被称为扩散常数.

然而, 计算得到期望和方差, 还不足以刻画出分布. 为此, 我们再进一步考虑求出自由扩散运动的分布函数. 在 $[0,t]$ 插入分点 $0=t_0<t_1<\cdots<t_n=t$, 其中 $t_i=\frac{it}{n}$, 并记 $Y_i=X(t_i)-X(t_{i-1})$, 则 $\{Y_i\}$ 独立同分布, 且有

$$X(t)=\sum_{i=1}^{n}Y_i.$$

进而计算出 $\mathbb{E}Y_i=0$, $\mathrm{Var}Y_i=\frac{Dt}{n}$. 考虑其特征函数

$$\phi_{Y_i}(u)=\mathbb{E}e^{iuZ_i}=1-\frac{Dt}{2n}\cdot u^2+o\left(\frac{u^2}{2n}\right),$$

因此

$$\phi_{X(t)}(u)=\prod_{i=1}^{n}\phi_{Y_i}(u)=\left(1-\frac{Dt}{2n}\cdot u^2+o\left(\frac{u^2}{2n}\right)\right)^n,$$

令 $n\to\infty$ 得 $\phi_{X(t)}(u)\to\exp\left\{-\frac{Dtu^2}{2}\right\}$, 而右边正是 $\mathcal{N}(0,Dt)$ 的特征函数, 因此 $X(t)\sim\mathcal{N}(0,Dt)$.

3.1.2 Brown 运动的定义

在上面的讨论的基础上, 我们给出 Brown 运动的定义.

定义 3.1 (Brown 运动)

若随机过程 $\{X(t),t\geq 0\}$ 满足条件

(1) $\mathbb{P}(\omega: t\to X_\omega(t)\text{连续})=1$, 也即轨迹连续的概率为 1, 且 $X(0)=0$;

(2) $\{X(t)\}$ 是独立增量过程;

(3) 对任意的 $t>s\geq 0$, $X(t)-X(s)\sim\mathcal{N}(0,D(t-s))$,

则称其为 *Brown* 运动. 特别地, 当 $D=1$ 时, 称 $\{X(t)\}$ 为标准 *Brown* 运动.

定义3.1描述了直线上的 Brown 运动. 将其推广, 可以得到平面上的 Brown 运动.

[1]形如 $f(x+y)=f(x)+f(y)$ 的函数方程称为 Cauchy 方程, 可以说它的解形如 $f(x)=ax$, 其中 a 为常数.

定义 3.2 (二维 Brown 运动)

若随机过程 $\{\boldsymbol{B}(t)=(X(t),Y(t)),t\geq 0\}$ 满足条件

(1) 轨迹连续的概率为 1;

(2) 具有独立增量性与平稳增量性;

(3) 对任意的 $t>0$, $X(t)$ 和 $Y(t)$ 相互独立, 且 $X(t),Y(t)\sim\mathcal{N}(0,\sigma^2 t)$,

则称其为二维 *Brown* 运动. ♣

3.2 Brown 运动的性质

3.2.1 Brown 运动与随机游走的联系

在这里指出, Brown 运动可以作为随机游走的极限. 对于质点随机游走的情形, 设其向左走和向右走的概率都为 $\frac{1}{2}$, 再设质点的步长为 Δx, 每次走动的时间间隔为 Δt. 记

$$X_i=\begin{cases}1, & \text{质点向右走,}\\ 0, & \text{质点向左走.}\end{cases}$$

则 $\mathbb{P}(X_i=1)=\mathbb{P}(X_i=-1)=\frac{1}{2}$, $\mathbb{E}X_i=0$, $\mathrm{Var}X_i=1$, 且

$$X(t)=\Delta x\cdot\sum_{i=1}^{\left[\frac{t}{\Delta t}\right]}X_i,$$

首先计算 $X(t)$ 的期望与方差. 计算得

$$\mathbb{E}X(t)=\Delta x\cdot\left[\frac{t}{\Delta t}\right]\cdot\mathbb{E}X_i=0,$$

$$\mathrm{Var}X(t)=(\Delta x)^2\cdot\left[\frac{t}{\Delta t}\right]\cdot\mathrm{Var}X_i=(\Delta x)^2\cdot\left[\frac{t}{\Delta t}\right].$$

在这里, 令 $\Delta x=\sqrt{D\cdot\Delta t}$, 则 $\mathrm{Var}X(t)=Dt$. 接下来研究 $X(t)$ 的分布, 依据 Levy 中心极限定理[2], 有

$$\frac{X(t)}{\sqrt{Dt}}=\frac{\cdot\sum_{i=1}^{\left[\frac{t}{\Delta t}\right]}X_i-0}{\sqrt{\left[\frac{t}{\Delta t}\right]}}\xrightarrow{d}\mathcal{N}(0,1),$$

因此 $X(t)\xrightarrow{d}\mathcal{N}(0,Dt)$, 此即说明 Brown 运动可以作为随机游走的极限.

[2]这是概率论中的重要定理. 设随机变量序列 $\{X,X_n,n=1,2,\cdots\}$ 独立同分布, $\mathbb{E}X=a$, $0<\mathrm{Var}X=\sigma^2<\infty$, 记 $S_n=\sum_{k=1}^{n}X_k$, 则 $\frac{S_n-na}{\sqrt{n}\sigma}\xrightarrow{d}\mathcal{N}(0,1)$. 详见 [2].

3.2.2 Brown 运动与 Gauss 过程的联系

首先, 根据例 1.5 可得

$$\mathrm{Cov}(B(s),B(t))=D\min\{s,t\},\quad \mathbb{E}(B(s)B(t))=D\min\{s,t\}.$$

通常设 $0\le s\le t$, 则上式即为 Ds. 同时, 在 Brown 运动的定义中, 我们会发现其与正态分布密切相关. 在这里, 我们给出 Gauss 过程的定义, 然后研究 Brown 运动与 Gauss 运动的关系.

定义 3.3 (Gauss 过程)

如果对任意的 $n\ge 1$, 以及对任意的 $t_1,t_2,\cdots,t_n$,

$$(X(t_1),X(t_2),\cdots,X(t_n))$$

服从 n 维正态分布, 则称随机过程 $\{X(t)\}$ 为 *Gauss* 过程. ♣

我们通过定理3.1来说明 Gauss 过程与 Brown 运动之间的联系. 该定理也给出了判断某种运动是 Brown 运动的方式.

定理 3.1

若 Gauss 过程 $\{X(t),t\ge 0\}$ 的轨迹在 $[0,\infty)$ 中连续的概率为 1, 且 $X(0)=0$, 则 $\{X(t)\}$ 是标准 Brown 运动当且仅当

$$\mathbb{E}X(t)=0,\quad \mathbb{E}X(t)X(s)=s,\quad 0\le s\le t.$$

♡

证明 一方面, 设 $\{X(t)\}$ 是标准 Brown 运动, 则 $X(t)\sim\mathcal{N}(0,t)$, 设 $t\ge s\ge 0$, 则

$$\begin{aligned}&\mathbb{E}X(t)=0,\\&\mathbb{E}X(t)X(s)=\mathbb{E}(X(t)-X(s)+X(s))X(s)=\mathbb{E}X^2(s)=s.\end{aligned}$$

另外一方面, 设 $\{X(t)\}$ 是满足条件的 Gauss 过程, 首先根据条件, 它的轨迹连续的概率为 1; 接下来, 设 $0=t_0<t_1<\cdots<t_n$, 要验证 $\{X(t)\}$ 是独立增量过程, 只需证明 $X(t_i)-X(t_{i-1})$ 与 $X(t_j)-X(t_{j-1})$ 独立. 考虑到 $\{X(t)\}$ 是 Gauss 过程, 设 $i<j$, 则

$$\begin{aligned}&\mathrm{Cov}(X(t_i)-X(t_{i-1}),X(t_j)-X(t_{j-1}))\\&=\mathbb{E}(X(t_i)-X(t_{i-1}))(X(t_j)-X(t_{j-1}))\\&=\mathbb{E}X(t_i)X(t_j)-\mathbb{E}X(t_i)X(t_{j-1})-\mathbb{E}X(t_{i-1})X(t_j)+\mathbb{E}X(t_{i-1})X(t_{j-1})\\&=0,\end{aligned}$$

因此 $X(t_i)-X(t_{i-1})$ 与 $X(t_j)-X(t_{j-1})$ 独立; 最后, 设 $t>s\ge 0$, 则

$$\begin{aligned}&\mathbb{E}(X(t)-X(s))=0,\\&\mathrm{Var}(X(t)-X(s))^2=\mathbb{E}X^2(t)+\mathbb{E}X^2(s)-2\mathbb{E}X(t)X(s)=t-s.\end{aligned}$$

因此 $\{X(t)\}$ 是标准 Brown 运动.

利用该定理, 可以判断某种运动是否是 Brown 运动. 以下是几个例子.

例 3.1 设 $B(t)$ 是标准 Brown 运动, a 是正常数, 则以下的随机过程都是标准 Brown 运动:

(1) $W(t) = -B(t), t \geq 0$;

(2) $W(t) = B(t+a) - B(a), t \geq 0$;

(3) $W(t) = B(at)/\sqrt{a}, t \geq 0$;

(4) $W(0) = 0, W(t) = tB(1/t), t > 0$;

(5) 对于正数 T, $W(t) = B(T-t) - B(T)$ 是时间段 $[0, T]$ 内的 Brown 运动.

证明 (1) 首先 $W(t)$ 是 Gauss 过程, 且

$$\mathbb{E}W(t) = -\mathbb{E}B(t) = 0;$$

再设 $t \geq s \geq 0$, 则

$$\mathbb{E}W(t)W(s) = \mathbb{E}B(t)B(s) = s.$$

因此 $W(t)$ 是标准 Brown 运动.

(2) 首先 $W(t)$ 是 Gauss 过程, 且

$$\mathbb{E}W(t) = \mathbb{E}B(t+a) - \mathbb{E}B(a) = 0;$$

再设 $t \geq s \geq 0$, 则

$$\begin{aligned}
\mathbb{E}W(t)W(s) &= \mathbb{E}(B(t+a) - B(a))(B(s+a) - B(a)) \\
&= \mathbb{E}B(t+a)B(s+a) - \mathbb{E}B(a)B(s+a) - \mathbb{E}B(a)B(t+a) + \mathbb{E}B^2(a) \\
&= s + a - a - a + a \\
&= s.
\end{aligned}$$

因此 $W(t)$ 是标准 Brown 运动.

(3) 首先 $W(t)$ 是 Gauss 过程, 且

$$\mathbb{E}W(t) = \mathbb{E}\frac{B(at)}{\sqrt{a}} = 0;$$

再设 $t \geq s \geq 0$, 则

$$\mathbb{E}W(t)W(s) = \frac{1}{a}\mathbb{E}B(at)B(st) = \frac{as}{a} = s.$$

因此 $W(t)$ 是标准 Brown 运动.

(4) 首先 $W(t)$ 是 Gauss 过程, 且

$$\mathbb{E}W(t) = \mathbb{E}tB\left(\frac{1}{t}\right) = 0;$$

再设 $t \geq s \geq 0$, 则

$$\mathbb{E}W(t)W(s) = ts\mathbb{E}B\left(\frac{1}{t}\right)B\left(\frac{1}{s}\right) = \frac{ts}{t} = s.$$

因此 $W(t)$ 是标准 Brown 运动.

(5) 首先 $W(t)$ 是 Gauss 过程, 且

$$\mathbb{E}W(t) = \mathbb{E}B(T-t) - B(T) = 0;$$

再设 $t \geq s \geq 0$, 则

$$\begin{aligned}
\mathbb{E}W(t)W(s) &= \mathbb{E}(B(T-t) - B(T))(B(T-s) - B(T)) \\
&= \mathbb{E}B(T-t)B(T-s) - \mathbb{E}B(T)B(T-s) - \mathbb{E}B(T)B(T-t) + \mathbb{E}B^2(T) \\
&= T - t - (T-s) - (T-t) + T \\
&= s.
\end{aligned}$$

因此 $W(t)$ 是 $[0,T]$ 上的标准 Brown 运动.

例 3.2 设 $W(t)$ 和 $B(t)$ 是独立的标准 Brown 运动.

(1) 若 $X(t) = aB(t) + bW(t)$ 是标准 Brown 运动, 求 a, b 所满足的条件?

(2) 若 $Y(t) = B(2t) - B(t)$, 其是否是标准 Brown 运动?

解答 (1) 计算得 $\mathbb{E}X(t) = a\mathbb{E}B(t) + b\mathbb{E}W(t) = 0$, 再设 $t \geq s \geq 0$, 则

$$\begin{aligned}
\mathbb{E}X(t)X(s) &= \mathbb{E}(aB(t) + bW(t))(aB(s) + bW(s)) \\
&= a^2\mathbb{E}B(t)B(s) + ab(\mathbb{E}B(t)W(s) + \mathbb{E}B(s)W(t)) + b^2\mathbb{E}W(t)W(s) \\
&= (a^2 + b^2)s,
\end{aligned}$$

因此 $a^2 + b^2 = 1$;

(2) 计算得 $\mathbb{E}Y(t) = \mathbb{E}B(2t) - \mathbb{E}B(t) = 0$, 再设 $t \geq s \geq 0$, 则

$$\begin{aligned}
\mathbb{E}Y(t)Y(s) &= \mathbb{E}(B(2t) - B(t))(B(2s) - B(s)) \\
&= \mathbb{E}B(2t)B(2s) - \mathbb{E}B(t)B(2s) - \mathbb{E}B(2t)B(s) + \mathbb{E}B(t)B(s) \\
&= \begin{cases} 2s - t - s + s = 2s - t, & t \leq 2s, \\ 2s - 2s - s + s = 0, & t > 2s. \end{cases}
\end{aligned}$$

因此 $Y(t) = B(2t) - B(t)$ 不是标准 Brown 运动.

3.3 首中时、最大值与 Arcsin 律

以下设 $\{B(t), t \geq 0\}$ 是标准 Brown 运动. 对于标准 Brown 运动而言, $D = 1$, 这使得问题的研究更加容易. 而标准 Brown 运动的性质也容易推广到一般的 Brown 运动上.

3.3.1 首中时及其分布

考虑标准 Brown 运动首次到达 a 的时刻.

定义 3.4 (首中时)

对于常数 a, 称

$$T_a = \inf\{t : t \geq 0, B(t) = a\},$$

为 a 的首中时.

首先设 $a > 0$, 注意到

$$\begin{cases} \mathbb{P}(B(t) \geq a | T_a > t) = 0, \\ \mathbb{P}(B(t) \geq a | T_a \leq t) = \mathbb{P}(B(t) < a | T_a \leq t) = \dfrac{1}{2}, \end{cases}$$

利用 $(T_a \leq t)$ 与 $(T_a > t)$ 进行分类, 得

$$\begin{aligned} \mathbb{P}(B(t) \geq a) &= \mathbb{P}(B(t) \geq a | T_a \leq t)\mathbb{P}(T_a \leq t) \\ &\quad + \mathbb{P}(B(t) \geq a | T_a > t)\mathbb{P}(T_a > t) \\ &= \frac{1}{2}\mathbb{P}(T_a \leq t), \end{aligned}$$

由此得到 $\mathbb{P}(T_a \leq t) = 2\mathbb{P}(B(t) \geq a)$, 又 $B(t) \sim \mathcal{N}(0, t)$, 设 $\Phi(t)$ 为标准正态分布 $\mathcal{N}(0, 1)$ 的分布函数, 则

$$\mathbb{P}(T_a \leq t) = 2\left[1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right], \quad a > 0.$$

又当 $a < 0$ 时, 情况是完全类似的, 据此可以得到

$$\mathbb{P}(T_a \leq t) = 2\left[1 - \Phi\left(\frac{|a|}{\sqrt{t}}\right)\right].$$

定理 3.2

对于标准布朗运动 $\{B(t)\}$ 和 $a \neq 0$.

(1) 质点最终到达 a 的概率为 1;

(2) 质点到达 a 的平均时间为 ∞.

证明 对于给定的 a, 质点最终到达 a 的概率

$$p = \lim_{t\to\infty} \mathbb{P}(T_a \leq t) = \lim_{t\to\infty} 2\left[1 - \Phi\left(\frac{|a|}{\sqrt{t}}\right)\right] = 1.$$

再考虑到达 a 所需的平均时间

$$\mathbb{E}T_a = \int_0^\infty t\mathrm{d}\mathbb{P}(T_a \leq t) = \int_0^\infty \frac{|a|}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{a^2}{2t}}\mathrm{d}t = \infty.$$

3.3.2 最大值及其分布

再考虑标准 Brown 运动在一定时间内能到达的最远距离.

定义 3.5 (最大值)

称

$$M_t = \sup_{0\le s\le t} B(s)$$

为 $\{B(t)\}$ 在 $[0,t]$ 上的最大值. ♣

根据 $\{M_t \ge a\} \iff \{T_a \le t\}$, 得

$$\mathbb{P}(M_t \le a) = 1 - \mathbb{P}(M_t \ge a) = 1 - \mathbb{P}(T_a \le t) = 1 - 2\left[1 - \Phi\left(\frac{|a|}{\sqrt{t}}\right)\right] = 2\Phi\left(\frac{|a|}{\sqrt{t}}\right) - 1.$$

此即为最大值的分布.

根据对称性, 我们知道 $\mathbb{E}M_t = 0$, 这是没有意义的. 接下来, 我们考虑计算当 $a \ge 0$ 时 M_t 的“单边期望”. 计算得

$$\mathbb{E}M_t^+ = \int_0^\infty a\mathrm{d}\mathbb{P}(M_t \le a) = \int_0^\infty \frac{2a}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{a^2}{2t}}\mathrm{d}a = \sqrt{\frac{2t}{\pi}}.$$

3.3.3 Arcsin 律

记 $N(a,a+b]$ 表示质点在 $(a,a+b]$ 内访问 0 的次数, 接下来考虑其分布. 根据平稳增量性, 计算得

$$\begin{aligned}
\mathbb{P}(N(a,a+b] \ge 1) &= \int_{-\infty}^{\infty} \mathbb{P}(N(a,a+b] \ge 1 | B(a) = x)\mathrm{d}\Phi(x) \\
&= \int_{-\infty}^{\infty} \mathbb{P}(N[0,b] \ge 1 | B(0) = x)\mathrm{d}\Phi(x) \\
&= \int_{-\infty}^{\infty} \mathbb{P}(T_{-x} \le b)\mathrm{d}\Phi(x) \\
&= \int_{-\infty}^{\infty} \mathbb{P}(T_x \le b)\mathrm{d}\Phi(x),
\end{aligned}$$

再根据 $\mathbb{P}(T_a \le t) = 2\left[1 - \Phi\left(\frac{|a|}{\sqrt{t}}\right)\right]$, [3]代入计算得

$$\begin{aligned}
\mathbb{P}(N(a,a+b] \ge 1) &= 2\int_{-\infty}^{\infty}\left[1 - \Phi\left(\frac{|x|}{\sqrt{b}}\right)\right]\mathrm{d}\Phi(x) \\
&= 4\int_0^{\infty}\left[1 - \Phi\left(\frac{x}{\sqrt{b}}\right)\right]\mathrm{d}\Phi(x) \\
&= 4\int_0^{\infty}\int_{\frac{x}{\sqrt{b}}}^{\infty} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v \frac{1}{\sqrt{2\pi a}}\mathrm{e}^{-\frac{x^2}{2a}}\mathrm{d}x.
\end{aligned}$$

[3]上课讲的过程相当于重新推导了该公式, 在这里为了简化, 直接应用该公式.

令 $y=\frac{x}{\sqrt{b}}$, 则 $\mathrm{d}y=\frac{1}{\sqrt{b}}\mathrm{d}x$, 代入得

$$\mathbb{P}(N(a,a+b]\geq 1)=\frac{2}{\pi}\sqrt{\frac{b}{a}}\int_0^\infty\int_y^\infty \mathrm{e}^{-\frac{av^2+by^2}{2a}}\mathrm{d}v\mathrm{d}y,$$

在此考虑三角代换, 令

$$\begin{cases} v=r\sin\theta, \\ y=\sqrt{\frac{a}{b}}r\cos\theta, \end{cases} \qquad 则 \quad \mathrm{d}v\mathrm{d}y=\begin{vmatrix} \sin\theta & r\cos\theta \\ \sqrt{\frac{a}{b}}\cos\theta & -\sqrt{\frac{a}{b}}r\sin\theta \end{vmatrix}\mathrm{d}r\mathrm{d}\theta=\sqrt{\frac{a}{b}}r\mathrm{d}r\mathrm{d}\theta,$$

且根据 $v\geq y\geq 0$ 得 $r\geq 0$, 以及 $\arcsin\sqrt{\frac{a}{a+b}}\leq\theta\leq\frac{\pi}{2}$, 因此

$$\begin{aligned}\mathbb{P}(N(a,a+b]\geq 1)&=\frac{2}{\pi}\int_{\arcsin\sqrt{\frac{a}{a+b}}}^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^\infty r\mathrm{e}^{-\frac{r^2}{2}}\mathrm{d}r \\ &=\frac{2}{\pi}\left(\frac{\pi}{2}-\arcsin\sqrt{\frac{a}{a+b}}\right) \\ &=1-\frac{2}{\pi}\arcsin\sqrt{\frac{a}{a+b}}.\end{aligned}$$

在上面的计算过程的基础上, 设 $t>0$, $x\in(0,1)$, 则有

$$\begin{aligned}\mathbb{P}(N(tx,t]=0)&=1-\mathbb{P}(N(tx,t]\geq 1) \\ &=1-\left(1-\frac{2}{\pi}\arcsin\sqrt{x}\right) \\ &=\frac{2}{\pi}\arcsin\sqrt{x},\end{aligned}$$

这便是所谓的 *Arcsin* 律.

对本节中所得的 Brown 运动的性质做一个总结, 即可得到以下定理.

定理 3.3

设 $\{B(t),t\geq 0\}$ 为标准 Brown 运动, 对于常数 a, 设 T_a 为 a 的首中时, M_t 为质点在 $[0,t]$ 内到达的最大值, $N(a,b]$ 为质点在 $(a,b]$ 内访问 0 的次数.

(1) $\mathbb{P}(T_a\leq t)=\mathbb{P}(T_{|a|}\leq t)=2\mathbb{P}(B(t)\geq|a|)=2\mathbb{P}\left(\frac{B(t)}{\sqrt{t}}\geq\frac{|a|}{\sqrt{t}}\right)=2\left[1-\Phi\left(\frac{|a|}{\sqrt{t}}\right)\right]$;

(2) 对 $a\geq 0$, 有 $\mathbb{P}(M_t\geq a)=\mathbb{P}(T_a\leq t)$;

(3) 对 $a\neq 0$, 有 $\mathbb{E}T_a=\infty$;

(4) 对 $b>a>0$, 有 $\mathbb{P}(N(a,b]=0)=\frac{2}{\pi}\arcsin\sqrt{\frac{a}{b}}$.

3.4 Brown 桥与经验过程

3.4.1 Brown 桥

首先在标准 Brown 运动的基础上, 给出 Brown 桥的定义.

定义 3.6 (Brown 桥)

设 $\{B(t), 0 \le t \le 1\}$ 为标准 Brown 运动, 称

$$X(t) = B(t) - tB(1), \quad 0 \le t \le 1$$

为 *Brown* 桥. ♣

根据定义, 容易得到如下的性质.

(1) $X(0) = B(0) = 0$, $X(1) = B(1) - B(1) = 0$, 因此可以称之为“桥”;

(2) $X(t) = B(t) - tB(1)$ 是 Gauss 过程;

(3) $\mathbb{E}X(t) = \mathbb{E}(B(t) - tB(1)) = 0$;

(4) 对 $0 \le s < t \le 1$, 有

$$\begin{aligned}\mathbb{E}X(t)X(s) &= \mathbb{E}(B(t) - tB(1))(B(s) - sB(1)) \\ &= \mathbb{E}B(t)B(s) - t\mathbb{E}B(1)B(s) - s\mathbb{E}B(t)B(1) + st\mathbb{E}B^2(1) \\ &= s - st - st + st \\ &= s(1-t),\end{aligned}$$

因此 Brown 桥不是标准 Brown 运动.

另外, 从 Gauss 过程出发, Brown 桥还有如下的定义.

定义 3.7 (Brown 桥)

设 $\{X(t), 0 \le t \le 1\}$ 是均值为 0 的 Gauss 过程, 且

$$\mathrm{Cov}(X(t), X(s)) = s(1-t), \quad 0 \le s \le t \le 1,$$

则称 $\{X(t), 0 \le t \le 1\}$ 为 *Brown* 桥. ♣

根据该定义, 可以计算得到 $\mathbb{E}X(0) = \mathrm{Var}X(0) = 0$, 并且考虑到 Gauss 分布可以由均值和方差唯一决定, 因此 $X(0) = 0$; 同理, 根据 $\mathbb{E}X(1) = \mathrm{Var}X(1) = 0$, 也可以得到 $X(1) = 0$. 最后, 从条件随机过程出发, 还可以得到 Brown 桥的第三个定义.

定义 3.8 (Brown 桥)

设 $\{B(t), 0 \le t \le 1\}$ 为标准 Brown 运动, 称

$$\{X(t), 0 \le t \le 1\} = \{B(t), 0 \le t \le 1 | B(1) = 0\}$$

为 *Brown* 桥. ♣

在这个定义下, 对 $0 \le s \le t \le 1$, 尝试求出 $B(s)$ 在 $B(t)=B$ 的条件下的分布, 记所需求的分布的概率密度为 $f_{s|t}(x|B)$, 则有

$$f_{s|t}(x|B) = \frac{f_s(x)f_{t-s}(B-x)}{f_t(B)} = \frac{1}{\sqrt{2\pi\cdot\dfrac{s(t-s)}{t}}}\exp\left\{-\frac{(tx-Bs)^2}{2ts(t-s)}\right\},$$

因此 $B(s)|B(t)=B \sim \mathcal{N}\left(\dfrac{Bs}{t}, \dfrac{s(t-s)}{t}\right)$, 其中

$$\mathbb{E}(B(s)|B(t)=B) = \frac{Bs}{t}, \quad \mathrm{Var}(B(s)|B(t)=B) = \frac{s(t-s)}{t}.$$

取 $t=1, B=0$ 得 $X(t)\sim\mathcal{N}(0, s(1-s))$, 从而 $\{X(t)\}$ 为 Gauss 过程, 且

$$\mathbb{E}X(s) = 0, \quad \mathrm{Var}X(s) = s(1-s).$$

另外, 对 $0 \le s \le t \le 1$, 考虑在 $B(t)=x$ 的条件下进行计算, 得到

$$\begin{aligned}
\mathbb{E}X(t)X(s) &= \mathbb{E}(B(t)B(s)|B(1)=0) \\
&= \mathbb{E}\left[\mathbb{E}(B(t)B(s)|B(1)=0, B(t))\right] \\
&= \int_0^1 \mathbb{E}(B(t)B(s)|B(1)=0, B(t)=x)\mathrm{d}F_{B(1)=0}(x),
\end{aligned}$$

其中计算得

$$\mathbb{E}(B(t)B(s)|B(1)=0, B(t)=x) = x\cdot\mathbb{E}(B(s)|B(1)=0, B(t)=x).$$

记 $B(s)|B(1)=0, B(t)=x$ 的概率密度函数为 $f_{s|1,t}(y|0,x)$[4], 则

$$\begin{aligned}
f_{s|1,t}(y|0,x) &= \frac{f_s(y)f_{t-s}(x-y)f_{1-t}(0-x)}{f_t(x)f_{1-t}(0-x)} \\
&= \frac{f_s(y)f_{t-s}(x-y)}{f_t(x)} \\
&= f_{s|t}(y|x),
\end{aligned}$$

这个式子右端即为 $B(s)|B(t)=x$ 的概率密度函数, 因此 $B(s)|B(1)=0, B(t)=x$ 与 $B(s)|B(t)=x$ 这两个随机变量同分布, 返回计算得到

$$\begin{aligned}
\mathbb{E}(B(t)B(s)|B(1)=0, B(t)=x) &= x\cdot\mathbb{E}(B(s)|B(t)=x) \\
&= x\cdot\frac{sx}{t} \\
&= \frac{s}{t}\cdot x^2.
\end{aligned}$$

[4]核心的思想仍然是“求分布”.

将上面的计算结果代入原式中得到

$$\begin{aligned}\mathbb{E}X(t)X(s) &= \frac{s}{t}\int_0^1 x^2 \mathrm{d}F_{B(1)=0}(x)\\ &= \frac{s}{t}\cdot\mathbb{E}(B^2(t)|B(1)=0)\\ &= \frac{s}{t}\cdot\mathrm{Var}X(t)\\ &= \frac{s}{t}\cdot t(1-t) = s(1-t).\end{aligned}$$

以上的结果, 说明了定义3.8和定义3.7是等价的.

3.4.2 经验过程

定义 3.9 (经验过程)

设 $U\sim\mathcal{U}[0,1]$, $F(t)=\mathbb{P}(U\le t)=t$, $U_1,\cdots,U_n$ 是来自 U 的随机变量, 称

$$D_n(t)=\sqrt{n}\cdot\left(\frac{1}{n}\sum_{j=1}^n I[U_j\le t]-F(t)\right),\quad t\in[0,1]$$

是经验过程.

♣

接下来, 我们来探究经验过程与 Brown 桥 $\{X(t),0\le t\le 1\}$ 之间的关系.

(1) $\mathbb{E}D_n(t)=\mathbb{E}X(t)=0$. 计算得

$$\begin{aligned}\mathbb{E}D_n(t) &= \sqrt{n}\cdot\mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n I[U_j\le t]-F(t)\right)\\ &= \sqrt{n}\left(\frac{1}{n}\cdot t-t\right)\\ &= 0;\end{aligned}$$

(2) 对 $0\le s\le t\le 1$, $\mathbb{E}D_n(s)D_n(t)=\mathbb{E}X(s)X(t)=s(1-t)$. 计算得

$$\begin{aligned}\mathbb{E}D_n(s)D_n(t) &= n\cdot\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n I[U_i\le s]-F(s)\right)\left(\frac{1}{n}\sum_{j=1}^n I[U_j\le t]-F(t)\right)\\ &= n\cdot\frac{1}{n^2}\cdot\mathbb{E}\left(\sum_{i=1}^n\sum_{j=1}^n I[U_i\le s]\cdot I[U_j\le t]\right)-n\cdot\frac{s}{n}\cdot\mathbb{E}\left(\sum_{i=1}^n I[U_i\le t]\right)\\ &\quad -n\cdot\frac{t}{n}\cdot\mathbb{E}\left(\sum_{i=1}^n I[U_i\le s]\right)+n\cdot\mathbb{E}F(s)F(t)\\ &= s+(n-1)st-nst-nst+nst\\ &= s(1-t);\end{aligned}$$

(3) 考虑随机变量 $I[U_j\le t]$, 则 $\mathbb{E}I[U_j\le t]=t$, $\mathrm{Var}I[U_j\le t]=t-t^2$, 依据 Levy 中心极

限定理得

$$\frac{\sum_{j=1}^{n} I[U_j \le t] - nt}{\sqrt{n\cdot t(1-t)}} \xrightarrow{d} \mathcal{N}(0,1),$$

因此

$$D_n(t) = \sqrt{n}\cdot\left(\frac{1}{n}\sum_{j=1}^{n} I[U_j \le t] - F(t)\right) \xrightarrow{d} \mathcal{N}(0, t(1-t)).$$

从而 $D_n(t)$ 渐进正态过程.

根据以上三点可以得到以下结论, 也被称为不变原理或泛函中心极限定理.

(1) $\{D_n(t), 0\le t\le 1\}$ 与 $\{X(t), 0\le t\le 1\}$ 具有相同的期望与方差;

(2) $\{D_n(t), 0\le t\le 1\}$ 依分布收敛到 $\{X(t), 0\le t\le 1\}$.

3.5 Brown 运动的变式

设 $\{X(t), t\ge 0\}$ 是标准 Brown 运动, 以下是几个 Brown 运动的变式.

(1) 原点反射的 Brown 运动: $Z(t) = |X(t)|, t\ge 0$;

(2) 几何 Brown 运动: $Y(t) = \mathrm{e}^{X(t)}, t\ge 0$;

(3) O-V 过程: $V(t) = \mathrm{e}^{-t}X(\mathrm{e}^{2t})$.

此时 $X(t)\sim\mathcal{N}(0,t)$, 概率密度函数 $p_X(t) = \dfrac{1}{\sqrt{2\pi t}}\cdot\mathrm{e}^{-\frac{x^2}{2t}}$. 以下尝试求出这几个运动的期望与方差.

(1) 对于 $Z(t) = |X(t)|, t\ge 0$, 计算得期望

$$\mathbb{E}Z(t) = \int_{-\infty}^{\infty}\frac{|x|}{\sqrt{2\pi t}}\cdot\mathrm{e}^{-\frac{x^2}{2t}}\mathrm{d}x = 2\int_{0}^{\infty}\frac{x}{\sqrt{2\pi t}}\cdot\mathrm{e}^{-\frac{x^2}{2t}}\mathrm{d}x = \sqrt{\frac{2t}{\pi}};$$

又注意到 $\mathbb{E}Z^2(T) = \mathbb{E}|X(t)|^2 = \mathbb{E}X^2(t) = \mathrm{Var}X(t) = t$, 因此方差

$$\mathrm{Var}Z(t) = \mathbb{E}Z^2(t) - (\mathbb{E}Z(t))^2 = \left(1-\frac{2}{\pi}\right)\cdot t.$$

(2) 对于 $Y(t) = \mathrm{e}^{X(t)}, t\ge 0$, 计算得期望

$$\mathbb{E}Y(t) = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi t}}\cdot\mathrm{e}^{x-\frac{x^2}{2t}}\mathrm{d}x = \mathrm{e}^{\frac{t}{2}};$$

又计算得

$$\mathbb{E}Y^2(t) = \mathbb{E}\mathrm{e}^{2X(t)} = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi t}}\cdot\mathrm{e}^{2x-\frac{x^2}{2t}}\mathrm{d}x = \mathrm{e}^{2t},$$

因此方差

$$\mathrm{Var}Y(t) = \mathbb{E}Y^2(t) - (\mathbb{E}Y(t))^2 = \mathrm{e}^{2t} - \mathrm{e}^{t} = \mathrm{e}^{t}(\mathrm{e}^{t}-1).$$

(3) 对于 $V(t) = e^{-t}X(e^{2t})$, 其中 $X(e^{2t}) \sim \mathcal{N}(0, e^{2t})$, 计算得期望

$$\mathbb{E}V(t) = e^{-t}\mathbb{E}X(e^{2t}) = 0;$$

方差

$$\mathrm{Var}V(t) = e^{-2t}\mathrm{Var}X(e^{2t}) = e^{-2t} \cdot e^{2t} = 1.$$

3.6 课后习题

问题 3.1 设 $B(t)$ 是标准 Brown 运动, a 是正常数, 则以下的随机过程都是标准 Brown 运动:

(1) $W(T) = -B(t), t \geq 0$;

(2) $W(T) = B(t+a) - B(a), t \geq 0$;

(3) $W(T) = B(at)/\sqrt{a}, t \geq 0$;

(4) $W(0) = 0, W(T) = tB(1/t), t > 0$;

(5) 对于正数 T, $W(t) = B(T-t) - B(T)$ 是时间段 $[0, T]$ 内的 Brown 运动.

问题 3.2 设 $W(t)$ 和 $B(t)$ 是独立的标准 Brown 运动.

(1) 若 $X(t) = aB(t) + bW(t)$ 是标准 Brown 运动, 求 a, b 所满足的条件?

(2) 若 $Y(t) = B(2t) - B(t)$, 其是否是标准 Brown 运动?

问题 3.3 用 $(X(t), Y(t))$ 表示二维标准 Brown 运动, 证明对任何常数 θ,

$$W(t) = X(t)\cos\theta + Y(t)\sin\theta, \quad t \geq 0$$

是标准 Brown 运动.

问题 3.4 对于 Brown 桥 $X(t), 0 \leq t \leq 1$, 验证

$$W(t) = (t+1)X\left(\frac{t}{1+t}\right), \quad t \geq 0$$

是标准 Brown 运动.

问题 3.5 设 $\{X(t)\}$ 是标准 Brown 运动, 求下列几个随机过程的期望与方差.

(1) $Z(t) = |X(t)|, t \geq 0$;

(2) $Y(t) = e^{X(t)}, t \geq 0$;

(3) $V(t) = e^{-t}X(e^{2t})$.

第 4 章 离散时间 Markov 链

在本节中, Markov 链指的是离散时间 Markov 链.

4.1 Markov 链与 Markov 性

4.1.1 Markov 链的定义

Markov 过程是最重要的随机过程, 而 Markov 链是 Markov 过程的特例.

定义 4.1 (Markov 链)

设状态空间 $I=\{1,2,\cdots\}$, 随机变量序列 $\{X_n\}$ 在 I 中取值, 若对任意的正整数 n, 及对 I 中任意的 $i,j,i_0,i_1,\cdots,i_{n-1}$, 都有

$$\begin{aligned}&\mathbb{P}(X_{n+1}=j|X_n=i,X_{n-1}=i_{i-1},\cdots,X_0=i_0)\\&=\mathbb{P}(X_{n+1}=j|X_n=i)\\&=\mathbb{P}(X_1=j|X_0=i),\end{aligned}$$

则称 $\{X_n\}$ 为时齐的 *Markov* 链, 简称为 *Markov* 链. 称

$$p_{ij}=\mathbb{P}(X_1=j|X_0=i),\quad i,j\in I$$

为 $\{X_n\}$ 从第 i 个状态到第 j 个状态的转移概率, 并称矩阵 $\boldsymbol{P}=(p_{ij})_{i,j\in I}$ 为 $\{X_n\}$ 的一步转移矩阵, 简称为转移矩阵. ♣

根据 Markov 链的定义, 在已知现在的情况下, 将来和过去无关. 这一点将在下一小节中详细说明.

4.1.2 Markov 链的性质

以下是 Markov 链和转移矩阵的基本性质.

(1) 转移概率 $0\le p_{ij}\le 1$, 且

$$\sum_{j\in I}p_{ij}=\sum_{j\in I}\mathbb{P}(X_1=j|X_0=i)=\mathbb{P}\left(\bigcup_{j\in I}(X_1=j)\middle|X_0=i\right)=1,$$

也即转移矩阵 $\boldsymbol{P}$ 的行和为 1, 这样的矩阵也被称为随机矩阵;

(2) 已知现在 $B=(X_n=i)$ 的情况下, 将来 $A=(X_{n+1}=j)$ 和过去 $C=(X_{n-1}=i_{n-1},\cdots,X_0=i_0)$ 独立, 这一性质也被称为 *Markov* 性. 这个性质的验证需要用到以下定理.

定理 4.1

对于事件 A, B, C, 当 $\mathbb{P}(AB) > 0$ 时, 有

$$\mathbb{P}(C|BA) = \mathbb{P}(C|B) \iff \mathbb{P}(AC|B) = \mathbb{P}(A|B) \cdot \mathbb{P}(C|B).$$

证明 定义 $\mathbb{P}_B(\cdot) = \mathbb{P}(\cdot|B)$, 则

$$\begin{aligned}
\mathbb{P}(C|AB) = \mathbb{P}(C|B) &\iff \mathbb{P}_B(C|A) = \mathbb{P}_B(C) \\
&\iff \frac{\mathbb{P}_B(AC)}{\mathbb{P}_B(A)} = \mathbb{P}_B(C) \\
&\iff \mathbb{P}_B(AC) = \mathbb{P}_B(A)\mathbb{P}_B(C) \\
&\iff \mathbb{P}(AC|B) = \mathbb{P}(A|B) \cdot \mathbb{P}(C|B).
\end{aligned}$$

下面的定理将 Markov 链的 Markov 性进行了扩充.

定理 4.2

设 I 是状态空间, $A, A_j \subset I, j = 0, 1, 2, \cdots$.

(1) 已知 $X_n = i$ 的情况下, 将来 $(X_m : m \geq n+1)$ 与过去 $(X_j : j \leq n-1)$ 独立;

(2) $\mathbb{P}(X_{n+k} = j|X_n = i) = \mathbb{P}(X_k = j, X_0 = i)$;

(3) $\mathbb{P}(X_{n+k} = j|X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) = \mathbb{P}(X_k = j|X_0 = i)$;

(4) $\mathbb{P}(X_{n+k} \in A|X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) = \mathbb{P}(X_k \in A|X_0 = i)$;

证明 (1) 与 (2) 的证明较麻烦, 在此不作要求. 在已知 (1) 与 (2) 的情况下, 以下证明 (3) 和 (4).

(3) 根据 (1) 知, $(X_{n+k} = j)$ 与 $(X_{n-1} = i_{n-1}), \cdots, (X_0 = i_0)$ 独立, 因此

$$\begin{aligned}
&\mathbb{P}(X_{n+k} = j|X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) \\
&= \sum_{i_0 \in A_0} \cdots \sum_{i_{n-1} \in A_{n-1}} \mathbb{P}(X_{n+k} = j|X_n = i, X_{n-1} = i_{n-1}, \cdots, X_0 = i_0) \\
&= \sum_{i_0 \in A_0} \cdots \sum_{i_{n-1} \in A_{n-1}} \mathbb{P}(X_{n+k} = j|X_n = i) \\
&= \mathbb{P}(X_{n+k} = j|X_n = i) \\
&= \mathbb{P}(X_k = j|X_0 = i).
\end{aligned}$$

(4) 根据 (3) 知

$$\begin{aligned}
&\mathbb{P}(X_{n+k} \in A|X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) \\
&= \sum_{j \in A} \mathbb{P}(X_{n+k} = j|X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) \\
&= \sum_{j \in A} \mathbb{P}(X_k = j|X_0 = i) \\
&= \mathbb{P}(X_k \in A|X_0 = i).
\end{aligned}$$

4.1.3 Markov 链的例子

我们接下来通过几个例子来理解 Markov 过程.

例 4.1 简单随机游走 假设质点向右走的概率为 p, 向左走的概率为 q, 其中 $p+q=1$, 则转移概率

$$p_{ij}=\begin{cases}p, & j=i+1,\\ q, & j=i-1,\\ 0, & \text{其他},\end{cases}$$

此时转移矩阵 $\boldsymbol{P}$ 是一个无限维的矩阵.

例 4.2 两端有吸收壁的随机游走 在例4.1的基础上, 假设质点在 0 处不能向左走, n 处不能向右走, 则 $I=\{0,1,\cdots,n\}$, 转移概率

$$p_{ij}=\begin{cases}p, & j=i+1, i=\{1,2,\cdots,n-1\},\\ q, & j=i-1, i=\{1,2,\cdots,n-1\},\\ 1, & j=i\in\{0,n\},\\ 0, & \text{其他}.\end{cases}$$

从而转移矩阵

$$\boldsymbol{P}=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ q & 0 & p & \cdots & 0\\ 0 & q & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{bmatrix}_{(n+1)\times(n+1)}.$$

例 4.3 两端有反射壁的随机游走 在例4.1的基础上, 假设质点在 0 处一定会向右走, n 处一定会向左走, 则 $I=\{0,1,\cdots,n\}$, 转移概率

$$p_{ij}=\begin{cases}p, & j=i+1, i=\{1,2,\cdots,n-1\},\\ q, & j=i-1, i=\{1,2,\cdots,n-1\},\\ 1, & (i,j)\in\{(0,1),(n,n-1)\},\\ 0, & \text{其他}.\end{cases}$$

从而转移矩阵

$$P = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ q & 0 & p & \cdots & 0 \\ 0 & q & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{(n+1)\times(n+1)}.$$

4.2 Markov 链的多步转移

4.2.1 Kolmogorov-Chapman 方程

对于 Markov 链 $\{X_n\}$, 可以定义

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

$$p_{ij}^{(k)} = \mathbb{P}(X_{n+k} = j | X_n = i) = \mathbb{P}(X_k = j | X_0 = i),$$

并且称 $p_{ij}^{(k)}$ 为 $\{X_n\}$ 的k 步转移概率. 称矩阵 $\boldsymbol{P}^{(k)} = (p_{ij}^{(k)})$ 为 $\{X_n\}$ 的k 步转移矩阵. 以下的方程给出了 k 步转移矩阵的计算方法.

定理 4.3 (Kolmogorov-Chapman 方程)

对任何 $m, n \geq 0$, 有

$$\begin{cases} p_{ij}^{(n+m)} = \sum_{k \in I} p_{ik}^{(n)} p_{kj}^{(m)}, \\ \boldsymbol{P}^{(n+m)} = \boldsymbol{P}^{n+m}. \end{cases}$$

证明 定义 $\mathbb{P}_{X_0}(\cdot) = \mathbb{P}(\cdot | X_0 = i)$, 依据全概率公式得

$$\begin{aligned} p_{ij}^{(n+m)} &= \mathbb{P}_{X_0}(X_{n+m} = j) \\ &= \sum_{k \in I} \mathbb{P}_{X_0}(X_{n+m} = j | X_n = k) P_{X_0}(X_n = k) \\ &= \sum_{k \in I} \mathbb{P}(X_{n+m} = j | X_n = k, X_0 = i) P(X_n = k | X_0 = i) \\ &= \sum_{k \in I} \mathbb{P}(X_{n+m} = j | X_n = k) P(X_n = k | X_0 = i) \\ &= \sum_{k \in I} p_{ik}^{(n)} p_{kj}^{(m)}. \end{aligned}$$

在上式的基础上, 考虑矩阵 $\boldsymbol{P}^{(n+m)}$, $\boldsymbol{P}^{(n)}$ 和 $\boldsymbol{P}^{(m)}$, 根据元素之间的关系得

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

取 $m = 1$, 迭代后即可得到 $\boldsymbol{P}^{(n+m)} = \boldsymbol{P}^{n+m}$.

定理4.3告诉我们, k 步转移等价于 k 次一步转移. 下面的定理是定理4.3的推论.

定理 4.4

对于正整数 $n, m, k, n_1, n_2, \cdots, n_k$ 和状态 i, j, l, 总有

(1) $p_{ij}^{(n+m)} \geq p_{il}^{(n)} p_{lj}^{(m)}$;

(2) $p_{ii}^{(n+k+m)} \geq p_{ij}^{(n)} p_{jl}^{(k)} p_{li}^{(m)}$;

(3) $p_{ii}^{(n_1+n_2+\cdots+n_k)} \geq p_{ii}^{(n_1)} p_{ii}^{(n_2)} \cdots p_{ii}^{(n_k)}$;

(4) $p_{ii}^{(nk)} \geq (p_{ii}^{(n)})^k$.

♡

证明 (1) 根据 Kolmogorov-Chapman 方程得

$$p_{ij}^{(n+m)} = \sum_{k\in I} p_{ik}^{(n)} p_{kj}^{(m)} \geq p_{il}^{(n)} p_{lj}^{(m)}.$$

(2) 根据 (1) 得

$$p_{ii}^{(n+k+m)} \geq p_{ij}^{(n)} p_{ji}^{(k+m)} \geq p_{ij}^{(n)} p_{jl}^{(k)} p_{li}^{(m)}.$$

(3) 根据 (1) 得

$$p_{ii}^{(n_1+n_2+\cdots+n_k)} \geq p_{ii}^{(n_1)} \cdot p_{ii}^{(n_2+\cdots+n_k)} \geq \cdots \geq p_{ii}^{(n_1)} p_{ii}^{(n_2)} \cdots p_{ii}^{(n_k)}.$$

(4) 在 (3) 中, 取 $n_1 = n_2 = \cdots = n_k = n$ 即可.

4.2.2 初始分布与 X_n 的分布

设状态空间 $I = \{1, 2, \cdots\}$, 记初始状态

$$\boldsymbol{\pi}^{(0)} = [\pi_1^{(0)}, \pi_2^{(0)}, \cdots], \quad 其中 \quad \pi_i^{(0)} = \mathbb{P}(X_0 = i),$$

并记 n 步转移之后的状态

$$\boldsymbol{\pi}^{(n)} = [\pi_1^{(n)}, \pi_2^{(n)}, \cdots], \quad 其中 \quad \pi_i^{(n)} = \mathbb{P}(X_n = i).$$

接下来, 我们希望得到 $\boldsymbol{\pi}^{(0)}$, $\boldsymbol{P}$ 和 $\boldsymbol{\pi}^{(n)}$ 之间的关系.

定理 4.5

设 Markov 链 $\{X_n\}$ 的初始分布为 $\boldsymbol{\pi}^{(0)}$, 概率转移矩阵为 $\boldsymbol{P} = (p_{ij})_{i,j\in I}$.

(1) 对任意的 $n_0 < n_1 < \cdots < n_m$, 有

$$\mathbb{P}(X_{n_0} = i_0, X_{n_1} = i_1, \cdots, X_{n_m} = i_m) = \pi_{i_0}^{(n_0)} p_{i_0 i_1}^{(n_1-n_0)} \cdots p_{i_{m-1}i_m}^{(n_m-n_{m-1})};$$

(2) 对任意的 $n \geq 1$ 及 $0 \leq k \leq n$, 有

$$\pi_j^{(n)} = \sum_{i \in I} \pi_i^{(0)} p_{ij}^{(n)}, \quad 或 \quad \boldsymbol{\pi}^{(n)} = \boldsymbol{\pi}^{(k)} \boldsymbol{P}^{n-k}.$$

证明 (1) 记事件 $A_k = (X_{n_k} = i_k)$, 依据概率的乘法公式得

$$\begin{aligned}
\mathbb{P}(A_0 A_1 \cdots A_m) &= \mathbb{P}(A_m | A_{m-1} \cdots A_0)\mathbb{P}(A_{m-1} | A_{m-2} \cdots A_0) \cdots \mathbb{P}(A_1 | A_0)\mathbb{P}(A_0) \\
&= \mathbb{P}(A_m | A_{m-1})\mathbb{P}(A_{m-1} | A_{m-2}) \cdots \mathbb{P}(A_1 | A_0)\mathbb{P}(A_0) \\
&= \pi_{i_0}^{(n_0)} p_{i_0 i_1}^{(n_1 - n_0)} \cdots p_{i_{m-1} i_m}^{(n_m - n_{m-1})},
\end{aligned}$$

其中 $\mathbb{P}(A_0) = \pi_{i_0}^{n_0}$, $\mathbb{P}(A_k | A_{k-1}) = p_{i_{k-1} i_k}^{(n_k - n_{k-1})}$.

(2) 依据全概率公式得

$$\begin{aligned}
\pi_j^{(n)} &= \mathbb{P}(X_n = j) \\
&= \sum_{i \in I} \mathbb{P}(X_n = j | X_0 = i)\mathbb{P}(X_0 = i) \\
&= \sum_{i \in I} \pi_i^{(0)} p_{ij}^{(n)}.
\end{aligned}$$

将其推广即可得到 $\boldsymbol{\pi}^{(n)} = \boldsymbol{\pi}^{(k)} \boldsymbol{P}^{n-k}$.

例 4.4 $\{X_n, n \in \mathbb{N}\}$ 是 Markov 链, 状态空间 $I = \{a, b, c\}$, 状态转移矩阵

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{3}{5} & \frac{2}{5} & 0 \end{bmatrix}.$$

(1) 计算 $\mathbb{P}(X_1 = b, X_2 = c, X_3 = a, X_4 = c, X_5 = a, X_6 = c, X_7 = b | X_0 = c)$;

(2) 计算 $\mathbb{P}(X_{n+2} = c | X_n = b)$.

解答 (1) 记事件 $A = (X_1 = b, X_2 = c, X_3 = a, X_4 = c, X_5 = a, X_6 = c, X_7 = b)$, 计算得

$$\begin{aligned}
\mathbb{P}(A | X_0 = c) &= \frac{\mathbb{P}(A, X_0 = c)}{\mathbb{P}(X_0 = c)} \\
&= \frac{\mathbb{P}(X_0 = c)}{\mathbb{P}(X_0 = c)} \cdot \frac{2}{5} \cdot \frac{1}{3} \cdot \frac{3}{5} \cdot \frac{1}{4} \cdot \frac{3}{5} \cdot \frac{1}{4} \cdot \frac{2}{5} \\
&= \frac{3}{2500}.
\end{aligned}$$

(2) 计算得

$$\begin{aligned}\mathbb{P}(X_{n+2}=c|X_n=b) &= \frac{\mathbb{P}(X_{n+2}=c,X_n=b)}{\mathbb{P}(X_n=b)}\\ &= \frac{\mathbb{P}(X_n=b)\cdot p_{bc}^{(2)}}{\mathbb{P}(X_n=b)}\\ &= \frac{1}{4}\cdot\frac{2}{3}+\frac{1}{3}\cdot 0+0\cdot\frac{1}{3}\\ &= \frac{1}{6}.\end{aligned}$$

或直接应用全概率公式得

$$\begin{aligned}\mathbb{P}(X_{n+2}=c|X_n=b) &= \sum_{i\in\{a,b,c\}}\mathbb{P}(X_{n+2}=c,X_{n+1}=i|X_n=b)\\ &= \frac{1}{4}\cdot\frac{2}{3}+\frac{1}{3}\cdot 0+0\cdot\frac{1}{3}\\ &= \frac{1}{6}.\end{aligned}$$

例 4.5 设甲胜的概率为 p, 负的概率为 q, 平的概率为 r, $p+q+r=1$. 胜得 $+1$ 分, 负得 -1 分, 平得 0 分. 设 X_n 为第 n 局时甲的分数.

(1) 求状态空间及转移矩阵 $\boldsymbol{P}$;

(2) 在甲得 1 分的情况下, 不超过两局可以结束比赛的概率.

解答 (1) 根据题意得 $I=\{-2,-1,0,1,2\}$, 且

$$\boldsymbol{P}=\begin{bmatrix}1&0&0&0&0\\ q&r&p&0&0\\ 0&q&r&p&0\\ 0&0&q&r&p\\ 0&0&0&0&1\end{bmatrix}.$$

(2) 在甲得 1 分的情况下, 不超过两局可以结束比赛的情况只可能是甲得 2 分. 因此只需计算 $\mathbb{P}(X_2=2|X_0=1)$. 计算得

$$\mathbb{P}(X_2=2|X_0=1)=0\cdot 0+0\cdot 0+q\cdot 0+r\cdot p+p\cdot 1=p(1+r).$$

4.3 状态的分类与命名

在这一节中我们希望研究当 $n\to\infty$ 时的 X_n 的分布. 在此之前, 需要对状态空间进行分类, 方便研究问题.

4.3.1 状态的连通性

定义 4.2 (状态的连通性)

设 I 是 $\{X_n\}$ 的状态空间.

(1) 如果 $p_{ii}=1$, 则称 i 为吸引状态;

(2) 如果存在 $n\geq 1$, 使得 $p_{ij}^{(n)}>0$, 则称i 通 j 或i 可达 j, 记作 $i\to j$;

(3) 如果 $i\to j, j\to i$, 则称 i 和 j 互通, 记作 $i\leftrightarrow j$. ♣

例如, 对于随机游走 (例4.1), 对任意的 i,j, 都有 $i\leftrightarrow j$; 而对于有吸收壁的随机游走 (例4.2), 0 和 n 为吸引状态, 对于 $i,j\in\{1,2,\cdots,n-1\}$, 都有 $i\leftrightarrow j$. 同时, $i\to 0,n$.

定义 4.3 (首达时与首达概率)

对任意的 $i,j\in I$, 称

$$T_{ij}=\min\{n: X_0=i, X_n=j, n\geq 1\}$$

为从状态 i 出发首次到达状态 j 的转移步数或时间, 简称为首达时,

$$f_{ij}^{(n)}=\mathbb{P}(T_{ij}=n|X_0=i)$$

为质点从状态 i 出发经过 n 步首次到达 j 的概率, 简称为首达概率. 特别地, $f_{ii}^{(n)}$ 为从状态 i 出发经过 n 步首次回到状态 i 的概率. ♣

对于首达概率, 当 $n=1$ 时, 有

$$f_{ij}^{(1)}=\mathbb{P}(T_{ij}=1|X_0=i)=\mathbb{P}(X_1=j|X_0=i)=p_{ij}.$$

定义 4.4 (迟早到达的概率)

对任意的 $i,j\in I$, 称

$$f_{ij}^{*}=\sum_{n=1}^{\infty}f_{ij}^{(n)}=\sum_{n=1}^{\infty}\mathbb{P}(T_{ij}=n|X_0=i)$$

为质点从状态 i 出发, 迟早会到达状态 j 的概率. ♣

以上定义的几个概率具有如下的性质.

定理 4.6

设 $f_{ij}^{(n)}$ 为质点从状态 i 出发经过 n 步首次到达 j 的概率, f_{ij}^{*} 为质点从状态 i 出发, 迟早会到达状态 j 的概率.

(1) 对任意的 $i,j\in I, n\geq 1, 0\leq f_{ij}^{(n)}\leq f_{ij}^{*}\leq 1$;

(2) 对任意的 $i,j\in I, n\geq 1$, $p_{ij}^{(n)}=\sum_{l=1}^{n}f_{ij}^{(l)}p_{jj}^{(n-l)}$;

(3) 对任意的 $i, j \in I$, $f_{ij}^* > 0 \iff i \to j$;
(4) 对任意的 $i, j \in I$, $f_{ij}^* > 0, f_{ji}^* > 0 \iff i \leftrightarrow j$.

证明 (1) 显然 $0 \le f_{ij}^{(n)} \le f_{ij}^*$, 再计算得

$$\begin{aligned} f_{ij}^* &= \sum_{n=1}^{\infty} f_{ij}^{(n)} \\ &= \sum_{n=1}^{\infty} \mathbb{P}(T_{ij} = n | X_n = i) \\ &= \mathbb{P}\left(\bigcup_{n=1}^{\infty} (T_{ij} = n) | X_n = i\right) \\ &\le 1, \end{aligned}$$

从而 $0 \le f_{ij}^{(n)} \le f_{ij}^* \le 1$.

(2) 由全概率公式得

$$\begin{aligned} p_{ij}^{(n)} &= \mathbb{P}(X_n = j | X_0 = i) \\ &= \sum_{l=1}^{\infty} \mathbb{P}(X_n = j | X_0 = i, T_{ij} = l) \mathbb{P}(T_{ij} = l | X_0 = i) \\ &= \sum_{l=1}^{\infty} \mathbb{P}(X_n = j | X_0 = i, X_l = j, X_k \neq j, 1 \le k \le l-1) f_{ij}^{(l)} \\ &= \sum_{l=1}^{n} \mathbb{P}(X_n = j | X_l = j) f_{ij}^{(l)} + 0 \\ &= \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}, \end{aligned}$$

其中当 $i \le n$ 时, 由 Markov 性知

$$\mathbb{P}(X_n = j | X_0 = i, X_l = j, X_k \neq j, 1 \le k \le l-1) = \mathbb{P}(X_n = j | X_l = j),$$

而当 $i > n$ 时, 有

$$\mathbb{P}(X_n = j | X_0 = i, T_{ij} = l) = 0.$$

从而 $p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$.

(3) 一方面, 设 $f_{ij}^* > 0$, 则存在 n_0, 使得 $f_{ij}^{(n_0)} > 0$, 计算得

$$p_{ij}^{(n_0)} = \sum_{l=1}^{n_0} f_{ij}^{(l)} p_{jj}^{(n_0-l)} \ge f_{ij}^{(n_0)} p_{jj}^{(n_0-n_0)} = f_{ij}^{(n_0)} > 0,$$

从而 $i \to j$. 另外一方面, 设 $i \to j$, 则存在 n, 使得 $p_{ij}^{(n)} > 0$. 取

$$n_0 = \min\{n : p_{ij}^{(n_0)} > 0\},$$

计算得

$$f_{ij}^{(n_0)} = \mathbb{P}(T_{ij} = n_0 | X_0 = i) = \mathbb{P}(X_{n_0} = j, X_k \neq j, 1 \leq k \leq n_0 - 1 | X_0 = i) > 0,$$

因此 $f_{ij}^* \geq f_{ij}^{(n_0)} > 0$.

(4) 由 (3) 即可得到该结论.

4.3.2 常返与非常返状态

定义 4.5 (常返与非常返)

设 $i \in I$, 若 $f_{ii}^* = 1$, 称状态 i 常返; 若 $f_{ii}^* < 1$, 称状态 i 非常返, 此时也称 i 为滑过态、瞬时态或暂态. ♣

记

$$u_{ij} = \mathbb{E}(T_{ij} | X_0 = i) = \sum_{n=1}^{\infty} n \cdot f_{ij}^{(n)},$$

其表示质点自 i 出发, 首次达到 j 的平均转移步数. 再记 $u_i = u_{ii}$, 其称为平均回转时间. 特别地, 对于非常返状态 j, 有 $u_j = \infty$.

定义 4.6 (正常返与零常返)

对于常返状态 $i \in I$, 若 $u_i < \infty$, 称状态 i 为正常返; 若 $u_i = \infty$, 称状态 i 为零常返. ♣

接下来, 给出判断状态的方法.

定理 4.7

设 I 是状态空间.

(1) $i \in I, i$ 是常返的 $\iff \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$;

(2) $i \in I, i$ 是非常返的 $\iff \sum_{n=0}^{\infty} p_{ii}^{(n)} = \dfrac{1}{1 - f_{ii}^*} < \infty$;

(3) $j \in I, j$ 是非常返的, 则对任意的 $i \in I$, $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$, 同时 $\lim_{n\to\infty} p_{ij}^{(n)} = 0$;

(4) $i \in I, i$ 是常返的, $i \to j$, 则 $i \leftrightarrow j$, 且 j 是常返的; i 是零常返的或是正常返的, $i \leftrightarrow j$, 则 i 是零常返的或是正常返的;

(5) $i \in I, i$ 是零常返的 $\iff \lim_{n\to\infty} p_{ii}^{(n)} = 0$;

(6) $j \in I, j$ 是零常返的, 则对任意的 $i \in I$, $\lim_{n\to\infty} p_{ij}^{(n)} = 0$. ♡

根据该定理, 得到的判断法则如下.

$$
\text{状态 } i \begin{cases} \text{常返} & \displaystyle\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty, \begin{cases} \text{零常返} & \displaystyle\lim_{n\to\infty} p_{ii}^{(n)} = 0, \\ \text{正常返} & \displaystyle\lim_{n\to\infty} p_{ii}^{(n)} \neq 0, \end{cases} \\ \text{非常返} & \displaystyle\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty. \end{cases}
$$

证明 (1) 根据 $p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$, 设 $0 < \rho < 1$, 计算得

$$
\begin{aligned}
\sum_{n=0}^{\infty} p_{ij}^{(n)} \rho^n &= p_{ij}^{(0)} + \sum_{n=1}^{\infty} \sum_{l=1}^{n} f_{ij}^{(l)} \rho^l p_{jj}^{(n-l)} \rho^{(n-l)} \\
&= \delta_{ij} + \sum_{l=1}^{\infty} f_{ij}^{(l)} \rho^l \sum_{n=l}^{\infty} p_{jj}^{(n-l)} \rho^{(n-l)} \\
&= \delta_{ij} + \sum_{l=1}^{\infty} f_{ij}^{(l)} \rho^l \cdot \sum_{n=0}^{\infty} p_{jj}^{(n)} \rho^{(n)}.
\end{aligned}
$$

在此, 记

$$
\begin{cases} F_{ij}(\rho) = \displaystyle\sum_{l=1}^{\infty} f_{ij}^{(l)} \rho^l, \\ G_{ij}(\rho) = \displaystyle\sum_{n=0}^{\infty} p_{ij}^{(n)} \rho^n, \end{cases}
$$

则上式等价于

$$
G_{ij}(\rho) = \delta_{ij} + F_{ij}(\rho) G_{jj}(\rho).
$$

令 $i = j$, 则 $G_{ii}(\rho) = 1 + F_{ii}(\rho) G_{ii}(\rho)$, 解得

$$
G_{ii}(\rho) = \frac{1}{1 - F_{ii}(\rho)},
$$

再令 $\rho \to 1$, 则 $F_{ii}(\rho) \to \sum_{l=1}^{\infty} f_{ii}^{(n)} = f_{ii}^*$, $G_{ii}(\rho) \to \sum_{n=0}^{\infty} p_{ii}^{(n)}$, 因此

$$
\sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1 - f_{ii}^*}.
$$

从而, i 是常返的 $\iff f_{ii}^* = 1 \iff \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

(2) 根据以上结论, i 是非常返的 $\iff f_i i^* < 1 \iff \sum_{n=0}^{\infty} p_{ii}^{(n)} = \dfrac{1}{1 - f_{ii}^*} < \infty$.

(3) 设 j 是非常返的, 根据以上结论得

$$
G_{ij}(\rho) = \delta_{ij} + F_{ij}(\rho) G_{jj}(\rho).
$$

设 $i \neq j$, 令 $\rho \to 1$, 则 $F_{ij}(\rho) \to \sum_{l=1}^{\infty} f_{ij}^{(n)} = f_{ij}^{*}$, $G_{ij}(\rho) = \sum_{n=0}^{\infty} p_{ij}^{(n)}$, 因此

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} = f_{ij}^{*} \cdot \sum_{n=0}^{\infty} p_{jj}^{(n)} < \infty.$$

再根据以上级数收敛得 $\lim_{n\to\infty} p_{ij}^{(n)} = 0$.

(4) 设 i 是常返的, $i \to j$, 则存在 n, 使得 $p_{ij}^{(n)} > 0$. 再设 $j \to i$, 则存在 m, 使得 $p_{ij}^{(m)} > 0$. 根据定理4.4得

$$p_{jj}^{(m+k+n)} \geq p_{ji}^{(m)} p_{ii}^{(k)} p_{ij}^{(n)},$$

因此

$$\sum_{k=0}^{\infty} p_{jj}^{(m+k+n)} \geq p_{ij}^{(m)} p_{ij}^{(n)} \cdot \sum_{k=0}^{\infty} p_{ii}^{(k)} = \infty,$$

此即说明 j 是常返的. 关于零常返和正常返的部分, 见 (5).

(5) i 是零常返的 $\iff \lim_{n\to\infty} p_{ii}^{(n)} = 0$, 证明见 [**5**]. 接下来, 设 i 是零常返的, $i \to j$, 则存在 n, 使得 $p_{ij}^{(n)} > 0$. 再设 $j \to i$, 则存在 m, 使得 $p_{ij}^{(m)} > 0$. 类似 (4), 有

$$0 = \lim_{k\to\infty} p_{ii}^{(n+k+m)} \geq \lim_{k\to\infty} p_{ij}^{(n)} p_{jj}^{(k)} p_{ji}^{(m)} = p_{ij}^{(n)} p_{ji}^{(m)} \cdot \lim_{k\to\infty} p_{jj}^{(k)},$$

因此 $\lim_{k\to\infty} p_{jj}^{(k)} = 0$, 此即说明 j 是零常返的.

(6) 设 j 是零常返的, 则 $\lim_{k\to\infty} p_{jj}^{(k)} = 0$. 从而

$$\begin{aligned} p_{ij}^{(n)} &= \sum_{l=1}^{n} f_{ij}^{l} p_{jj}^{(n-l)} \\ &= \sum_{l=1}^{m} f_{ij}^{l} p_{jj}^{(n-l)} + \sum_{l=m+1}^{n} f_{ij}^{l} p_{jj}^{(n-l)}. \end{aligned}$$

首先考虑前一项, 令 $n \to \infty$, 则 $\sum_{l=1}^{m} f_{ij}^{l} p_{jj}^{(n-l)} \to 0$, 从而对任意的 $\varepsilon > 0$, 存在 $N > 0$, 使得对任意的 $n > N$, 都有

$$\sum_{l=1}^{m} f_{ij}^{l} p_{jj}^{(n-l)} < \frac{\varepsilon}{2}.$$

再考虑后一项, 有 $\sum_{l=m+1}^{n} f_{ij}^{l} p_{jj}^{(n-l)} < \sum_{l=m+1}^{\infty} f_{ij}^{(l)}$. 根据 $f_{ij}^{*} = \sum_{n=1}^{\infty} f_{ij}^{(n)} \leq 1$, 知对上述 $\varepsilon > 0$, 存在 $M > 0$, 使得对任意的 $m > M$, 都有

$$\sum_{l=m+1}^{n} f_{ij}^{l} p_{jj}^{(n-l)} < \sum_{l=m+1}^{\infty} f_{ij}^{(l)} < \frac{\varepsilon}{2}.$$

取 $n > N, m > M$, 则有

$$p_{ij}^{(n)} = \sum_{l=1}^{m} f_{ij}^{l} p_{jj}^{(n-l)} + \sum_{l=m+1}^{n} f_{ij}^{l} p_{jj}^{(n-l)} < \varepsilon,$$

此即说明 $\lim\limits_{n\to\infty} p_{ij}^{(n)} = 0$.

例 4.6 简单随机游走 考虑简单随机游走, 质点向左走的概率为 q, 向右走的概率为 p, 状态空间 $I = \{0, \pm1, \pm2, \cdots\}$, 则对 $i, j \in I, i < j$, 根据概率论中的结论, 有

$$p_{ij}^{(n)} = \mathbb{P}(X_n = j|X_0 = i) = \begin{cases} \dbinom{n}{\frac{n-j+i}{n}} p^{\frac{n+j-i}{2}} q^{\frac{n-j+i}{n}}, & 2 \mid (n+j-i), \\ 0, & 2 \nmid (n+j-i). \end{cases}$$

当 $2 \mid (n+j-i), i = j$ 时, 若记 $n = 2k$, 则

$$p_{ii}^{(2k)} = \binom{2k}{k} p^k q^k.$$

考虑级数

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \sum_{k=0}^{\infty} p_{ii}^{(2k)} = \sum_{k=0}^{\infty} \frac{(2k)!}{k! \cdot k!} \cdot p^k q^k,$$

又当 $k \to \infty$ 时, 根据 Stirling 公式[1]有

$$\frac{(2k)!}{k! \cdot k!} \cdot p^k q^k \sim \frac{(4pq)^k}{\sqrt{\pi k}}.$$

(1) 当 $p = q = \dfrac{1}{2}$ 时, 有

$$\sum_{k=0}^{\infty} p_{ii}^{(2k)} = \sum_{k=0}^{\infty} \frac{1}{\sqrt{\pi k}} = \infty,$$

$$\lim_{k\to\infty} p_{ii}^{(2k)} = \lim_{k\to\infty} \frac{1}{\sqrt{\pi k}} = 0,$$

此时该过程是零常返的;

(2) 当 $p \neq q$ 时, 有 $pq < \dfrac{1}{4}$ 及 $\dfrac{1}{\sqrt{\pi k}} \leq \dfrac{1}{\sqrt{\pi}}$, 记 $\rho = 4pq < 1$, 则

$$\sum_{k=0}^{\infty} p_{ii}^{(2k)} = \sum_{k=0}^{\infty} \frac{\rho^k}{\sqrt{\pi}} < \infty,$$

此时该过程是非常返的.

例 4.7 平面随机游走 考虑平面随机游走, 质点往上下左右走的概率均为 $\dfrac{1}{4}$. 设一共走了

[1] 当 $n \to \infty$ 时, 阶乘 $n!$ 有渐进公式 $n! \sim \sqrt{2\pi n} \cdot \left(\frac{n}{\mathrm{e}}\right)^n$.

$2n$ 步, 其中 k 步向上, k 步向下, $n-k$ 步向左, $n-k$ 步向右, 则有

$$\begin{aligned}
p_{ii}^{(2n)} &= \sum_{k=0}^{n}\binom{2n}{k}\binom{2n-k}{k}\binom{n-k}{2n-2k}\cdot\left(\frac{1}{4}\right)^{2n}\\
&= \sum_{k=0}^{n}\frac{(2n)!}{(k!)^2((n-k)!)^2}\cdot\left(\frac{1}{4}\right)^{2n}\\
&= \sum_{k=0}^{n}\frac{2n!}{n!\cdot n!}\cdot\left(\frac{n!}{k!\cdot(n-k)!}\right)^2\cdot\left(\frac{1}{4}\right)^{2n}\\
&= \binom{2n}{n}\cdot\left(\frac{1}{4}\right)^{2n}\cdot\sum_{k=0}^{n}\binom{k}{n}^2,
\end{aligned}$$

其中

$$\sum_{k=0}^{n}\binom{k}{n}^2=\sum_{k=0}^{n}\binom{k}{n}\binom{n-k}{n}=\binom{n}{2n},$$

因此

$$p_{ii}^{(2n)}=\binom{2n}{n}^2\cdot\left(\frac{1}{4}\right)^{2n},$$

又当 $n\to\infty$ 时, 根据 Stirling 公式有

$$p_{ii}^{(2n)}\sim\frac{1}{\pi n},$$

因此 $\sum\limits_{n=0}^{\infty}p_{ii}^{(2n)}=\infty$, 且 $\lim\limits_{n\to\infty}p_{ii}^{(2n)}=0$, 这说明了该过程是零常返的.

例 4.8 转移矩阵为上三角矩阵 设 $I=\{1,2,3\}$, 转移矩阵

$$\boldsymbol{P}=\begin{bmatrix}\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\\0&\frac{3}{4}&\frac{1}{4}\\0&0&1\end{bmatrix}$$

(1) 求 T_{13} 的分布;

(2) 求 $f_{ii}^*, i=1,2,3$.

解答 (1) 设事件 $A_0=(X_0=1)$, $A_m=(X_m\neq 3)$, $1\le m\le n-1$, $A_n=(X_n=3)$, 则

$$\begin{aligned}
\mathbb{P}(T_{13}=n)&=\mathbb{P}(X_n=3,X_m\neq 3,1\le m\le n-1|X_0=1)\\
&=\mathbb{P}(A_1A_2\cdots A_n|A_0)\\
&=\mathbb{P}(A_1|A_0)\mathbb{P}(A_2|A_1)\cdots\mathbb{P}(A_n|A_{n-1}),
\end{aligned}$$

其中

$$\mathbb{P}(A_1|A_0)=\mathbb{P}(X_1\neq 3|X_0=1)=p_{11}+p_{12}=\frac{3}{4},$$

对 $2 \le m \le n-1$, 考虑计算 $\mathbb{P}(A_m|A_{m-1})$. 此时 X_{m-1} 的状态未定, 从而

$$\begin{aligned}
\mathbb{P}(A_m|A_{m-1}) &= \frac{\mathbb{P}(X_{m-1} \in \{1,2\}, X_m \in \{1,2\})}{\mathbb{P}(X_{m-1} \in \{1,2\})} \\
&= \frac{(p_{11}+p_{12}) \cdot \mathbb{P}(X_{m-1}=1) + (p_{21}+p_{22}) \cdot \mathbb{P}(X_{m-1}=2)}{\mathbb{P}(X_{m-1}=1) + \mathbb{P}(X_{m-1}=2)} \\
&= \frac{3}{4},
\end{aligned}$$

同样地, 有

$$\begin{aligned}
\mathbb{P}(A_n|A_{n-1}) &= \frac{\mathbb{P}(X_n = 3, X_{n-1} \in \{1,2\})}{\mathbb{P}(X_{n-1} \in \{1,2\})} \\
&= \frac{p_{13} \cdot \mathbb{P}(X_{n-1}=1) + p_{23} \cdot \mathbb{P}(X_{n-2}=3)}{\mathbb{P}(X_{n-1}=1) + \mathbb{P}(X_{n-1}=2)} \\
&= \frac{1}{4},
\end{aligned}$$

因此

$$\mathbb{P}(T_{13} = n) = \frac{3}{4} \cdot \left(\frac{3}{4}\right)^{n-2} \cdot \frac{1}{4} = \frac{3^{n-1}}{4^n}.$$

(2) 计算得

$$\begin{cases}
f_{11}^* = f_{11}^{(1)} = \dfrac{1}{2}, \\
f_{22}^* = f_{22}^{(1)} = \dfrac{3}{4}, \\
f_{33}^* = f_{33}^{(1)} = 1.
\end{cases}$$

另外, 也可以通过公式

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1 - f_{ii}^*}$$

来计算 f_{11}^*, 只需要计算出 $\boldsymbol{P}^n$ 即可.

例 4.9 设 $I = \{1,2,3,4\}$, 转移矩阵

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix},$$

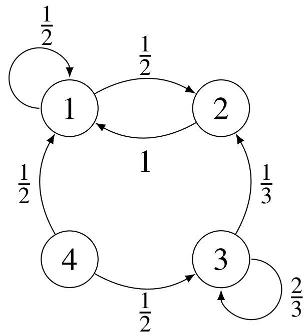
对每个状态进行分类.

解答 画出图像[2]后逐点分析即可.

(1) $f_{11}^{(1)} = \frac{1}{2}, f_{11}^{(2)} = \frac{1}{2}, f_{11}^{(k)} = 0, k \ge 3$, 此时

$$f_{11}^* = \frac{1}{2} + \frac{1}{2} = 1, \quad 且 \quad \mu_1 = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2} < \infty,$$

[2]特别感谢张博闻同学绘制的图像!



该状态正常返;

(2) $f_{22}^{(1)}=0,\ f_{22}^{(2)}=\frac{1}{2},\ f_{22}^{(k)}=\left(\frac{1}{2}\right)^{k-1},\ k\geq 3$, 此时

$$f_{22}^{*}=\sum_{k=1}^{\infty}\frac{1}{2^k}=1,\quad \text{且}\quad \mu_2=\sum_{k=1}^{\infty}\frac{k+1}{2^k}=3<\infty,$$

该状态正常返;

(3) $f_{33}^{(1)}=\frac{2}{3},\ f_{33}^{(k)}=0,\ k\geq 2$, 因此 $f_{33}^{*}=\frac{2}{3}<1$, 该状态非常返;

(4) $f_{44}^{(k)}=0,\ k\geq 1$, 因此 $f_{44}^{*}=0$, 该状态非常返.

4.3.3 周期和遍历状态

定义 4.7 (周期)

设 $i\in I$, 按如下方式定义 $\{X_n\}$ 的周期.

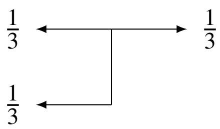
(1) 若 $\sum_{n=1}^{\infty}p_{ii}^{(n)}=0$, 则质点自 i 出发不会回到 i, 记此时周期为 ∞;

(2) 若存在正整数 d, 原点自 i 出发, 只能在 d 的整数倍回到 i, 也即 $p_{ii}^{(n)}>0 \implies n=kd$, 且 d 是满足此性质的最大整数, 记此时周期为 d;

(3) 若 $d=1$, 则称 i 为非周期的. ♣

例如, 对于简单随机游走, 每一个状态的周期 $d=2$. 以下是一个非周期的例子, 可以帮助理解以上定义.

例 4.10 设质点向左走和向右走的概率都是 $\frac{1}{3}$, 向下走两步再向左走的概率是 $\frac{1}{3}$.



则 $p_{ii}^{(2)}>\frac{1}{3}\cdot\frac{1}{3}>0$, 又 $p_{ii}^{(3)}>\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}>0$, 根据 $(2,3)=1$ 知 $d=1$, 从而每一个状态都是非周期的.

注意到 $(k,k+1)=1$, 若 $p_{ii}^{(k)}>0, p_{ii}^{(k+1)}>0$, 则由定义4.7知 i 是非周期的. 定理4.8是周期的基本性质. 在这里, 本定理的证明不做要求.

定理 4.8 (周期的性质)

设状态 i 的周期 $d_i<\infty$.

(1) d_i 是数集 $B_i=\{n|p_{ii}^{(n)}>0, n\geq 1\}$ 的最大公约数;

(2) 若 $i\leftrightarrow j$, 则 $d_i=d_j$; [a]

(3) 存在正整数 N_i, 对任意的 $n\geq N_i$, 都有 $p_{ii}^{(nd_i)}>0$.

[a] 注意到若 i,j 互通, 则 i 和 j 的状态 (常返、零常返与正常返) 也是相同的.

♡

定义 4.8 (遍历状态)

若 i 为正常返和非周期的, 则称 i 是遍历状态.

♣

此时, 非周期意味着存在正整数 N_i, 对任意的 $n>N_i$, 都有 $p_{ii}^{(n)}>0$. 另外, 遍历状态也可以通过互通来传递. 在此, 我们可以对状态空间进行分类.

定理 4.9

设常返状态 i 有周期 d_i, 平均回转时间

$$\mu_i=\mathbb{E}(T_i|X_0=i),$$

则 $\lim\limits_{n\to\infty} p_{ii}^{(nd_i)}=\dfrac{d_i}{\mu_i}$.

♡

定理 4.10 (周期分解定理)

设 $\{X_n\}$ 是一个周期为 d 的不可约 Markov 链, 状态空间 $I=\bigcup\limits_{r=1}^{d} B_r$, 其中 $B_i\cap B_j=\varnothing (1\leq i,j\leq d, i\neq j)$, 且 $\sum\limits_{j\in B_{r+1}} p_{ij}=1$, 其中 $i\in B_r$(约定 $B_{d+1}=B_1$). 将周期为 d 的不可约 Markov 链基于 $B_1,B_2,\cdots,B_d$ 重排后, 转移矩阵

$$P=\begin{bmatrix} 0 & P^{(1)} & 0 & \cdots & 0 \\ 0 & 0 & P^{(2)} & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P^{(d)} & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

♡

例 4.11 考虑简单随机游走, 设 $I=B_1\cup B_2$, 其中 B_1 为奇数点, B_2 为偶数点, 则对 $i\in B_1$,

有 $\sum_{j\in B_2} p_{ij} = 1$; 且对 $i \in B_2$, 有 $\sum_{j\in B_1} p_{ij} = 1$. 将转移矩阵进行重排, 可得

$$P = \begin{bmatrix} 0 & P^{(1)} \\ P^{(2)} & 0 \end{bmatrix}.$$

例 4.12 设 $I = \{0, 1, 2, 3\}$, 转移矩阵

$$P = \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}, \quad \text{从而} \quad P^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

记 $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, 则 $P = \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$, $P^2 = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$, 计算得

$$P^{(2k)} = (P^2)^k = \begin{bmatrix} A^k & 0 \\ 0 & A^k \end{bmatrix},$$

以及

$$P^{(2k+1)} = (P^2)^k P = \begin{bmatrix} A^k & 0 \\ 0 & A^k \end{bmatrix} \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix} = \begin{bmatrix} 0 & A^{k+1} \\ A^{k+1} & 0 \end{bmatrix},$$

因此 $p_{ii}^{(2k)} > 0, p_{ii}^{(2k+1)} = 0$, 此即说明 $d = 2$.

4.3.4 状态的等价

定义 4.9 (等价)

设 I 是 Markov 链 $\{X_n\}$ 的状态空间, 称集合

$$C = \{j | j \leftrightarrow i, j \in I\}$$

为一个等价类. 若 I 是一个等价类, 则称 Markov 链 $\{X_n\}$ 或状态空间 I 不可约. 设 $B \subset I$. 若质点不能从 B 中状态到达 $\overline{B} = I - B$ 中的状态, 则称 B 为闭集. ♣

(1) 结合定理4.7, 可以得到如下的等价类:
  (a). 设 $i \in C$, i 正常返, 则对任意的 $j \in C$, j 正常返. 此时 C 为正常返等价类;
  (b). 设 $i \in C$, i 正常返, 则对任意的 $j \in C$, j 零常返. 此时 C 为零常返等价类;
  (c). 设 $i \in C$, i 正常返, 则对任意的 $j \in C$, j 非常返. 此时 C 为非常返等价类;

(2) 设 B 是闭集, 则对任意的 $i \in B$, $\sum_{k\in B} p_{ik} = 1$;

(3) 考虑简单随机游走, 此时状态空间 $I = \{0, \pm1, \pm2, \cdots\}$, 则对任意的 $i, j \in I$, $i \leftrightarrow j$. 由例4.6知, 若 $p = q = \frac{1}{2}$, 则 I 为常返等价类; 若 $p \neq q$, 则 I 为非常返等价类;

(4) 考虑两端有吸收壁的随机游走, 此时状态空间 $I = \{0, 1, 2, \cdots, n\}$, 则 $\{0\}$ 和 $\{n\}$ 是常返等价类, $\{1, 2, \cdots, n-1\}$ 是非常返等价类.

定理 4.11

设 I 是状态空间, C 为一个等价类.

(1) 不同的等价类互不相交;

(2) C 中的状态有相同的类型, 或都是正常返的, 或都是零常返的, 或都是非常返的, 且 C 中所有状态的周期相同;

(3) 常返等价类是闭集, 也即质点不能走出常返等价类;

(4) 零常返等价类中有无穷多个状态;

(5) 非零常返等价类如果是闭集, 则有无穷多个状态.

证明 (1) 假设 C_1, C_2 是等价类, 且 $C_1 \cap C_2 \neq \varnothing$, 则存在 $i \in C_1 \cap C_2$. 对任意的 $j \in C_1$, 有 $i \leftrightarrow j$; 且对任意的 $k \in C_2$, 有 $i \leftrightarrow k$, 从而 $j \leftrightarrow k$, 此即说明 $C_1 = C_2$.

(2) 这可以由定理4.7与定理4.8得到.

(3) 设 C 是常返等价类, $i \in B$. 若存在 $j \in I - B$, 使得 $i \to j$, 则由 i 是常返的知 $i \leftrightarrow j$, 从而 $j \in B$, 矛盾.

(4) 设 B 是零常返等价类, 则 B 是闭集, 从而对 $i \in B$, 有

$$\sum_{k \in B} p_{ik} = 1 \implies \sum_{k \in B} p_{ik}^{(n)} = 1,$$

且根据 i 是零常返的, 知 $\lim\limits_{n \to \infty} p_{ik}^{(n)} = 0$. 若 $|B| < \infty$, 则

$$\lim_{n \to \infty} \sum_{k \in B} p_{ik}^{(n)} = \sum_{k \in B} \lim_{n \to \infty} p_{ik}^{(n)} = 0 \neq 1,$$

矛盾, 因此 $|B| = \infty$. 类似地, 可以证明 (5).

另外, 假设状态空间 I 有限, 所得到的 Markov 链 $\{X_n\}$ 称为有限 *Markov* 链. 根据定理4.11得到, 其具有以下的性质.

定理 4.12

设 I 是有限的状态空间.

(1) 非常返等价类不可能为闭集;

(2) 没有零常返状态;

(3) 必有正常返状态;

(4) 设 I 不可约, 则 I 中的状态是正常返状态;

(5) 设 I 不可约, 且非周期, 则 I 中的状态是遍历状态;

(6) 状态空间

$$I = T \cup C_1 \cup \cdots \cup C_k = T \cup \bigcup_{j=1}^{m} C_j,$$

其中 T 为非常返等价类, C_i 为常返等价类. 另外, 可以将转移矩阵写成分块矩阵

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_1 & 0 & \cdots & 0 & 0 \\ 0 & \boldsymbol{P}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{P}_m & 0 \\ \boldsymbol{R}_1 & \boldsymbol{R}_2 & \cdots & \boldsymbol{R}_m & \boldsymbol{Q}_T \end{bmatrix},$$

则

$$\boldsymbol{P}^{(n)} = \begin{bmatrix} \boldsymbol{P}_1^n & 0 & \cdots & 0 & 0 \\ 0 & \boldsymbol{P}_2^n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{P}_m^n & 0 \\ \boldsymbol{R}_1^{(n)} & \boldsymbol{R}_2^{(n)} & \cdots & \boldsymbol{R}_m^{(n)} & \boldsymbol{Q}_T^{(n)} \end{bmatrix}.$$

其中 $\boldsymbol{Q}_T^{(n)} = \boldsymbol{Q}_T^n$.

4.4 Markov 链的不变分布

首先, 设 $\{X_n\}$ 是 Markov 链, 转移矩阵为 $\boldsymbol{P}$, $I = \{1, 2, \cdots\}$, 记初始分布 $\boldsymbol{\pi} = [\pi_1, \pi_2, \cdots]$, 并设 n 步转移后的分布 $\boldsymbol{\pi}^{(n)} = [\pi_1^{(n)}, \pi_2^{(n)}, \cdots]$. 若 $\boldsymbol{\pi}^{(1)} = \boldsymbol{\pi}$, 则可以推出 $\boldsymbol{\pi}^{(n)} = \boldsymbol{\pi}$: 当 $n = 1$ 时, $\boldsymbol{\pi}^{(1)} = \boldsymbol{\pi}$; 假设 $\boldsymbol{\pi}^{(n-1)} = \boldsymbol{\pi}$, 则

$$\boldsymbol{\pi}^{(n)} = \boldsymbol{\pi}\boldsymbol{P}^{(n)} = \boldsymbol{\pi}^{(n-k)}\boldsymbol{P}^{(k)} = \boldsymbol{\pi}^{(n-1)}\boldsymbol{P} = \boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}^{(1)} = \boldsymbol{\pi},$$

因此根据数学归纳法知上面的命题成立. 在此基础上, 给出不变分布的定义.

定义 4.10 (不变分布)

若 $\boldsymbol{\pi} = [\pi_1, \pi_2, \cdots]$ 满足

$$\sum_{j \in I} \pi_j = 1, \quad \boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

则称 $\boldsymbol{\pi}$ 为 Markov 链 $\{X_n\}$ 的不变分布.

定理 4.13

设 C^+ 为 Markov 链 $\{X_n\}$ 的所有正常返状态, $i \in C^+$.

(1) 若 C^+ 是遍历等价类, 则

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j}, \quad j \in I$$

是唯一的不变分布;

(2) 若 C^+ 是周期为 d 的等价类, 则

$$\pi_j = \frac{1}{d}\lim_{n\to\infty} p_{jj}^{(nd)} = \lim_{n\to\infty}\frac{1}{d}\sum_{s=1}^{d} p_{ij}^{(nd+s)} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} = \frac{1}{\mu_j}, \quad j \in I$$

是唯一的不变分布;

(3) $\{X_n\}$ 有唯一不变分布的充要条件是 C^+ 是等价类;

(4) $\{X_n\}$ 有不变分布的充要条件是 C^+ 非空;

(5) 状态有限的 Markov 链必有不变分布.

例 4.13 设 $I = \{1, 2\}$, 转移矩阵

$$\boldsymbol{P} = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{5}{8} & \frac{3}{8} \end{bmatrix},$$

(1) 求不变分布 $\boldsymbol{\pi}$;

(2) 计算 μ_1, μ_2.

解答 (1) 根据 $1 \leftrightarrow 2$, 以及 I 有限, 知存在唯一不变分布. 由

$$[\pi_1, \pi_2]\begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{5}{8} & \frac{3}{8} \end{bmatrix} = [\pi_1, \pi_2]$$

解得 $\pi_1 = \frac{5}{7}, \pi_2 = \frac{2}{7}$. 从而不变分布 $\boldsymbol{\pi} = \left[\frac{5}{7}, \frac{2}{7}\right]$.

(2) 在上面的基础上, 得到

$$\lim_{n\to\infty} \boldsymbol{P}^{(n)} = \begin{bmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{bmatrix} = \begin{bmatrix} \frac{5}{7} & \frac{2}{7} \\ \frac{5}{7} & \frac{2}{7} \end{bmatrix}.$$

从而 $\mu_1 = \frac{1}{\pi_1} = \frac{7}{5}, \mu_2 = \frac{1}{\pi_2} = \frac{7}{2}$.

例 4.14Ehrenfest 模型 容器内有 $2a$ 个例子, 一张薄膜将该容器分为对称的 A, B 两部分. 将粒子穿过薄膜时占用的时间忽略不计. 用 X_0 表示初始时 A 中的粒子数, X_n 表示有 n 个粒子穿过薄膜后 A 中的粒子数. 当所有的粒子以相同的规律独立行动时, $\{X_n\}$ 是

Markov 链, 有状态空间 $I=\{0,1,2,\cdots,2a\}$, 设 Markov 链 $\{X_n\}$ 有转移概率

$$p_{ij}=\begin{cases}\dfrac{2a-i}{2a}, & 0\le i\le 2a-1, j=i+1,\\ \dfrac{i}{2a}, & 1\le i\le 2a, j=i-1,\\ 0, & \text{其他情况},\end{cases}$$

计算该 Markov 链的不变分布.

解答 从问题的背景知这是一个正常返 Markov 链, 周期等于 2, 不变分布唯一存在. 补充定义 $\pi_{-1}=\pi_{2a+1}=0$, 可将方程组 $\boldsymbol{\pi}=\boldsymbol{\pi P}$ 写成

$$\pi_i=\pi_{i-1}p_{i-1,i}+\pi_{i+1}p_{i+1,i},\quad 0\le i\le 2a.$$

于是有

$$\pi_{i+1}=\frac{\pi_i-\pi_{i-1}p_{i-1,i}}{p_{i+1,i}}.$$

经过计算可依次得到

$$\begin{aligned}\pi_1&=\frac{\pi_0}{p_{10}}=2a\pi_0=\binom{2a}{1}\pi_0,\\ \pi_2&=\frac{\pi_1-\pi_0p_{01}}{p_{21}}=(\pi_1-\pi_0)\frac{2a}{2}=\binom{2a}{2}\pi_0,\\ &\cdots,\\ \pi_{2a}&=\binom{2a}{2a}\pi_0.\end{aligned}$$

利用 $\pi_0+\pi_1+\cdots+\pi_{2a}=2^{2a}\pi_0=1$ 得到 $\pi_0=2^{-2a}$. 最后得到不变分布

$$\pi_i=\binom{2a}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{2a-i},\quad 0\le i\le 2a,$$

也即二项分布 $\mathcal{B}\left(2a,\dfrac{1}{2}\right)$.

4.5 Markov 链的平稳可逆分布

4.5.1 平稳性

设 $\boldsymbol{\pi}$ 是 $\{X_n\}$ 的不变分布, 将其作为初始分布, 则

$$\mathbb{P}(X_n=i)=\mathbb{P}(X_0=i),$$

从而随机向量 $\boldsymbol{\xi}_n = (X_n, X_{1+n}, \cdots, X_{m+n})$ 与 $\boldsymbol{\xi}_0 = (X_0, X_1, \cdots, X_n)$ 同分布. 若记 $A_k = (X_{k+n} = i_k), B_k = (X_k = i_k)$, 则可以计算得

$$\begin{aligned}\mathbb{P}(A_0 A_1 \cdots A_m) &= \mathbb{P}(X_n = i_0)\mathbb{P}(A_1|A_0)\cdots\mathbb{P}(A_m|A_{m-1})\\ &= \mathbb{P}(X_0 = i_0) p_{i_0,i_1} \cdots p_{i_{m-1},i_m}\\ &= \mathbb{P}(B_0 B_1 \cdots B_m),\end{aligned}$$

这便说明了这一点.

定义 4.11 (平稳序列)

设 $\{X_n\}$ 是随机序列. 如果对任意的 $m, n \geq 1$, 随机向量

$$\boldsymbol{\xi}_n = (X_n, X_{1+n}, \cdots, X_{m+n}) \quad 和 \quad \boldsymbol{\xi}_0 = (X_0, X_1, \cdots, X_n)$$

同分布, 则称 $\{X_n\}$ 为严平稳序列, 简称为平稳序列. 如果 Markov 链 $\{X_n\}$ 是平稳序列, 则称 $\{X_n\}$ 处于平稳状态. ♣

若 $\{X_n\}$ 是平稳序列, 则 (X_0, X_1) 与 (X_n, X_{n+1}) 同分布, 从而

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

此即说明 $\mathbb{P}(X_0 = i) = \mathbb{P}(X_n = i)$, 从而对任意的 n, X_0 和 X_n 同分布. 特别地, 对于 $n = 1$, X_0 和 X_1 同分布.

定理 4.14

设 $\{X_n\}$ 是遍历 Markov 链, 初始分布为平稳不变分布 $\boldsymbol{\pi}$, $\{Y_n\}$ 的转移概率和 $\{X_n\}$ 相同.

(1) $\lim\limits_{n\to\infty} \mathbb{P}(Y_n = i_0, Y_{1+n} = i_1, \cdots, Y_{m+n} = i_m) = \mathbb{P}(X_0 = i_0, X_1 = i, \cdots, X_m = i_m)$;

(2) 对于充分大的 n, $(Y_n, Y_{1+n}, \cdots, Y_{m+n})$ 和 $(X_n, X_{1+n}, \cdots, X_{m+n})$ 同分布. ♡

证明 (1) 由 $\{X_n\}$ 是遍历 Markov 链知其周期为 1, 且每个状态都是正常返的. 计算得

$$\begin{aligned}&\lim_{n\to\infty} \mathbb{P}(Y_n = i_0, Y_{1+n} = i_1, \cdots, Y_{m+n} = i_m)\\ &= \lim_{n\to\infty} \sum_{i\in I} \mathbb{P}(Y_0 = i, Y_n = i_0, Y_{1+n} = i_1, \cdots, Y_{m+n} = i_m)\\ &= \lim_{n\to\infty} \sum_{i\in I} \mathbb{P}(Y_0 = i) p_{ii_0}^{(n)} p_{i_0 i_1} \cdots p_{i_{m-1} i_m}\\ &= \sum_{i\in I} \mathbb{P}(Y_0 = i) \cdot \lim_{n\to\infty} p_{ii_0}^{(n)} \cdot p_{i_0 i_1} \cdots p_{i_{m-1} i_m}\\ &= \pi_{i_0} \cdot p_{i_0 i_1} \cdots p_{i_{m-1} i_m}\\ &= \mathbb{P}(X_0 = i_0, X_1 = i, \cdots, X_m = i_m),\end{aligned}$$

其中 $\lim\limits_{n\to\infty} p_{ii_0}^{(n)} = \pi_{i_0}$.

(2) 根据 (1) 得, 对于充分大的 n, $(Y_n, Y_{1+n}, \cdots, Y_{m+n})$ 和 $(X_0, X_1, \cdots, X_m)$ 同分布. 又

根据 $\{X_n\}$ 的平稳性知 $(X_0, X_1, \cdots, X_m)$ 和 $(X_n, X_{1+n}, \cdots, X_{m+n})$ 同分布, 因此对于充分大的 n, $(Y_n, Y_{1+n}, \cdots, Y_{m+n})$ 和 $(X_n, X_{1+n}, \cdots, X_{m+n})$ 同分布.

4.5.2 平稳可逆性

设 $\boldsymbol{\pi}$ 是 $\{X_n\}$ 的不变分布, 则

$$\begin{aligned}\mathbb{P}(X_{n-1}=j|X_n=i, X_{n+1}=i_1, \cdots, X_{n+m}=i_m) &= \mathbb{P}(X_{n-1}=j|X_n=i)\\ &= \frac{\mathbb{P}(X_{n-1}=j, X_n=i)}{\mathbb{P}(X_n=i)}\\ &= \frac{\mathbb{P}(X_{n-1}=j)\mathbb{P}(X_n=i|X_{n-1}=j)}{\mathbb{P}(X_n=i)}\\ &= \frac{\pi_j p_{ji}}{\pi_i},\end{aligned}$$

右边是一个和 m 与 n 无关的常数.

定义 4.12 (平稳可逆性)

设 $\{X_n\}$ 是 Markov 链, 转移矩阵 $\boldsymbol{P} = (p_{ij})_{i,j\in I}$.

(1) 如果有不全为 0 的非负序列 $\boldsymbol{\eta} = \{\eta_j\}$, 使得

$$\eta_i p_{ij} = \eta_j p_{ji}, \quad \forall i, j \in I,$$

则称 $\{X_n\}$ 是对称 *Markov* 链, $\boldsymbol{\eta}$ 为 $\{X_n\}$ 的对称化序列. 特别地, 若概率分布 $\boldsymbol{\pi}$ 满足

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad \forall i, j \in I,$$

称其为 $\{X_n\}$ 的可逆分布或平稳可逆分布;

(2) 若 $\{Y_n\}$ 是平稳序列, 且对任意的 $n > m \geq 0$,

$$(Y_m, Y_{m+1}, \cdots, Y_n) \quad 和 \quad (Y_n, Y_{n-1}, \cdots, Y_m)$$

同分布, 则称 $\{Y_n\}$ 是时间可逆的平稳序列或平稳可逆序列.

(3) 若 $\{X_n\}$ 是平稳可逆序列, 则称 $\{X_n\}$ 为可逆 *Markov* 链.

♣

对于可逆 Markov 链和平稳可逆分布, 有如下的性质.

(1) 若 $\{X_n\}$ 是可逆 Markov 链, 设其转移概率为 p_{ij}, “倒向”的转移概率为 p_{ij}^*, 则

$$\mathbb{P}(X_n=i, X_{n+1}=j) = \mathbb{P}(X_n=i)p_{ij} = \mathbb{P}(X_{n+1}=i, X_n=j) = \mathbb{P}(X_{n+1}=i)p_{ij}^*,$$

又根据平稳性得 $\mathbb{P}(X_n=i) = \mathbb{P}(X_{n+1}=i)$, 从而 $p_{ij} = p_{ij}^*$;

(2) 若 $\boldsymbol{\eta}$ 是对称化序列, 且 $\sum\limits_{i\in I} \eta_i < \infty$, 令

$$\pi_j = \frac{\eta_j}{\sum\limits_{i\in I}\eta_i},$$

则 $\boldsymbol{\pi}$ 为平稳可逆分布;

(3) 若 $\{X_n\}$ 是可逆 Markov 链, 则其初始分布 $\boldsymbol{\pi}$ 为平稳可逆分布, 这是因为 (X_0, X_1) 与 (X_1, X_0) 同分布, 从而

$$\mathbb{P}(X_0 = i, X_1 = j) = \pi_i p_{ij} = \mathbb{P}(X_0 = j, X_1 = i) = \pi_j p_{ji}, \quad \forall i, j \in I.$$

定理 4.15

设 $\{X_n\}$ 是 Markov 链, 转移矩阵 $\boldsymbol{P} = (p_{ij})_{i,j\in I}$, $\boldsymbol{\pi}$ 为平稳可逆分布.

(1) $\boldsymbol{\pi}$ 是 $\{X_n\}$ 是平稳不变分布;

(2) $\{X_n\}$ 的初始分布为 $\boldsymbol{\pi}$ 时, $\{X_n\}$ 为可逆 Markov 链.

证明 (1) 此时 $\pi_i p_{ij} = \pi_j p_{ji}$, 因此

$$\sum_{i\in I} \pi_i p_{ij} = \sum_{i\in I} \pi_j p_{ji} = \pi_j \cdot \sum_{i\in I} p_{ji} = \pi_j,$$

上式说明了 $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}$.

(2) 设 $m < n$, 则

$$\begin{aligned}
&\mathbb{P}(X_m = i_m, X_{m+1} = i_{m+1}, \cdots, X_n = i_n) \\
&= \pi_{i_m} p_{i_m i_{m+1}} \cdots p_{i_{n-1} i_n} \\
&= \pi_{i_{m+1}} p_{i_{m+1} i_m} \cdots p_{i_{n-1} i_n} \\
&= \cdots \\
&= \pi_{i_n} p_{i_n i_{m-1}} \cdots p_{i_{m+1} i_m} \\
&= \mathbb{P}(X_n = i_m, X_{n-1} = i_{m+1}, \cdots, X_m = i_n),
\end{aligned}$$

此即说明 $\{X_n\}$ 为可逆 Markov 链.

4.5.3 平稳可逆分布的计算

定理 4.16

设互通 Markov 链 $\{X_n\}$ 以平稳不变分布 $\boldsymbol{\pi}$ 为初始分布.

(1) $\{X_n\}$ 可逆 $\iff$ 对任意的 $i, i_1, \cdots, i_k \in I$, 有

$$p_{ii_1} p_{i_1 i_2} \cdots p_{i_k i} = p_{ii_k} p_{i_k i_{k-1}} \cdots p_{i_1 i},$$

称该条件为 *Kolmogorov* 条件;

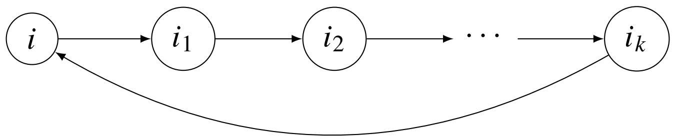
(2) 若 $\{X_n\}$ 是平稳可逆序列, 对于选定的 i 及从 i 到 j 的通路

$$i \to i_1 \to i_2 \to \cdots \to i_k \to j,$$

定义

$$\eta_i = 1, \quad \eta_j = \frac{p_{ii_1} p_{i_1 i_2} \cdots p_{i_k j}}{p_{ji_k} p_{i_k i_{k-1}} \cdots p_{i_1 i}}, \quad j \neq i,$$

则 $\{\eta_j\}$ 是 $\{X_n\}$ 的对称化序列.



证明 (1) 若 $\{X_n\}$ 可逆, 则对任意的 $i,j\in I$, $\pi_j p_{ji}=\pi_i p_{ij}$, 从而

$$\begin{aligned}\pi_i p_{ii_1}p_{i_1i_2}\cdots p_{i_ki}&=p_{i_1i}\pi_{i_1}p_{i_1i_2}\cdots p_{i_ki}\\&=\cdots\\&=p_{ii_k}p_{i_ki_{k-1}}\cdots p_{i_1i}\pi_i,\end{aligned}$$

从而 $\{X_n\}$ 满足 Kolmogorov 条件; 若 $\{X_n\}$ 满足 Kolmogorov 条件, 则

$$\begin{aligned}&\mathbb{P}(X_0=i,X_1=i_1,\cdots,X_k=i_k,X_{k+1}=i)\\&=\pi_i p_{ii_1}p_{i_1i_2}\cdots p_{i_ki}\\&=\mathbb{P}(X_0=i,X_1=i_k,\cdots,X_k=i_1,X_{k+1}=i)\\&=p_{ii_k}p_{i_ki_{k-1}}\cdots p_{i_1i},\end{aligned}$$

在上式中对 $i_1,i_2,\cdots,i_{k-1}$ 求和得

$$\mathbb{P}(X_0=i,X_k=i_k,X_{k+1}=i)=\mathbb{P}(X_{k+1}=i,X_1=i_k,X_0=i),$$

令 $i_k=j$ 得

$$\pi_i p_{ij}^{(k)}p_{ji}=\pi_i p_{ij}p_{ji}^{(k)}\implies\frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}\cdot p_{ji}=\frac{1}{n}\sum_{k=1}^{n}p_{ji}^{(k)}\cdot p_{ij},$$

在上式中令 $n\to\infty$, 可得

$$\pi_j p_{ji}=\pi_i p_{ij},\quad\forall i,j\in I,$$

从而 $\{X_n\}$ 可逆.

(2) 首先根据 Kolmogorov 条件知 $\{\eta_j\}$ 的选取与路径无关. 另外, 考虑 i 到 j 的通路和 i 到 l 的通路

$$\begin{cases}i\to i_1\to i_2\to\cdots\to i_k\to j,\\i\to j_1\to j_2\to\cdots\to j_s\to l,\end{cases}$$

同样根据 Kolmogorov 条件, 计算得

$$\begin{aligned}\eta_j p_{jl}&=\frac{p_{ii_1}p_{i_1i_2}\cdots p_{i_kj}p_{jl}}{p_{ji_k}p_{i_ki_{k-1}}\cdots p_{i_1i}}\\&=\frac{p_{ij_1}p_{j_1j_2}\cdots p_{j_sl}p_{lj}}{p_{lj_s}p_{j_sj_{s-1}}\cdots p_{j_1i}}\\&=\eta_l p_{lj},\end{aligned}$$

从而 $\{\eta_j\}$ 是 $\{X_n\}$ 的对称化序列.

根据定理4.16的证明过程, 可以得到如下定理.

定理 4.17

互通 Markov 链 $\{X_n\}$ 存在对称化序列的充要条件是其满足 Kolmogorov 条件. 同时, 由定理4.16(2) 所定义的序列 $\{\eta_j\}$ 是 $\{X_n\}$ 的对称化序列.

证明 一方面, 由定理4.16知, 若 $\{X_n\}$ 满足 Kolmogorov 条件, 则由定理4.16(2) 所定义的序列 $\{\eta_j\}$ 是 $\{X_n\}$ 的对称化序列.

另外一方面, 若存在对称化序列 $\{\eta_j\}$, 则

$$\begin{aligned}\eta_i p_{ii_1} p_{i_1 i_2} \cdots p_{i_k i} &= p_{i_1 i} \eta_{i_1} p_{i_1 i_2} \cdots p_{i_k i} \\ &= \cdots \\ &= p_{i_1 i} p_{i_2 i_1} \cdots p_{i i_k} \eta_i,\end{aligned}$$

从而 $\{X_n\}$ 满足 Kolmogorov 条件.

综合本节中的定理, 对于互通的 Markov 链 $\{X_n\}$, 给出如下计算可逆分布的步骤.

(1) 检验 Kolmogorov 条件;

(2) 若 Kolmogorov 条件成立, 则根据定理4.16(2) 计算对称化序列 $\{\eta_j\}$;

(3) 若 $\sum\limits_{j\in I} \eta_i < \infty$, 令

$$\pi_i = \frac{\eta_i}{\sum\limits_{j\in I} \eta_j},$$

由此得到可逆分布 $\boldsymbol{\pi}$.

定理 4.18

对于互通的 Markov 链 $\{X_n\}$, 若 Kolmogorov 条件成立, 且 $\sum\limits_{j\in I} \eta_i < \infty$, 则 $\{X_n\}$ 正常返.

证明 只需证明若 $\sum\limits_{j\in I} \eta_i = \infty$, 则 $\{X_n\}$ 非正常返. 假设当 $\sum\limits_{j\in I} \eta_i = \infty$ 时, $\{X_n\}$ 正常返, 则 $\{X_n\}$ 存在不变分布 $\{\pi_i\}$, 又根据 Kolmogorov 条件成立, 知 $\{\pi_i\}$ 是可逆分布, 从而

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad \forall i, j \in I.$$

考虑通路 $i \to i_1 \to i_2 \to \cdots \to i_k \to j$, 利用定理4.16(2) 中所构造的 $\{\eta_j\}$, 则有

$$\pi_i p_{ii_1} p_{i_1 i_2} \cdots p_{i_k j} = \pi_j p_{i_1 i} p_{i_2 i_1} \cdots p_{j i_k} \implies \pi_i \eta_j = \pi_j,$$

在上式左右端对 j 求和得

$$\infty = \pi_i \sum_{j\in I} \eta_j = \sum_{j\in I} \pi_j = 1,$$

这便推出了矛盾.

例 4.15 在例4.3的基础上, 考虑两端为反射壁的简单随机游走, 证明其是可逆 Markov 链, 并计算对称化序列和平稳可逆分布.

解答 首先计算 $\{\eta_j\}$. 考虑从 0 出发, 令

$$\eta_0 = 1, \quad \eta_1 = \frac{p_{01}}{p_{10}} = \frac{1}{q}, \quad \eta_2 = \frac{p_{01}p_{12}}{p_{10}p_{21}} = \frac{p}{q^2}, \quad \cdots,$$
$$\eta_k = \frac{p^{k-1}}{q^k}, \quad \cdots, \quad \eta_n = \frac{p^{n-1}}{q^{n-1}}.$$

接下来, 验证 $\eta_i p_{ij} = \eta_j p_{ji}$. 计算得

$$\eta_0 p_{01} = 1 = \frac{1}{q} \cdot q = \eta_1 p_{10},$$
$$\eta_i p_{i,i+1} = \frac{p^{i-1}}{q^i} p = \frac{p^i}{q^{i+1}} q = \eta_{i+1} p_{i+1,i}, \quad 1 \le i \le n-1,$$
$$\eta_n p_{n,n-1} = \frac{p^{n-2}}{q^{n-1}} p = \frac{p^{n-1}}{q^{n-1}} = \eta_n p_{n,n-1},$$

从而 $\{\eta_i\}$ 是对称化序列, 另外 $\sum\limits_{j=0}^{n} \eta_j < \infty$, 因此令

$$\pi_i = \frac{\eta_i}{\sum\limits_{j=0}^{n} \eta_j},$$

由此得到可逆分布 $\boldsymbol{\pi}$.

例 4.16 质点在 $I = \{0, 1, \cdots\}$ 中作随机游动, 有转移概率

$$p_{ij} = \begin{cases} p_i, & j = i+1, i \ge 0, \\ 1 - p_i, & j = i-1, i \ge 1, \end{cases}$$

其中 $p_{01} = p_0 = 1$, 当 $i > 1$ 时, $p_i \in (0, 1)$.

(1) 证明转移概率 $\{p_{ij}\}$ 存在对称化序列 $\boldsymbol{\eta}$;

(2) 求转移概率 $\{p_{ij}\}$ 有平稳可逆分布的充分必要条件;

(3) 给出 $\{p_{ij}\}$ 有平稳不变分布的充分必要条件.

解答 (1) 质点每次只能向左或右走一步, 设 $q_i = 1 - p_i$, 引入

$$\eta_0 = 1, \quad \eta_i = \frac{p_{01}p_{12}\cdots p_{i-1,i}}{p_{10}p_{21}\cdots p_{i,i-1}} = \frac{p_0 p_1 \cdots p_{i-1}}{q_1 q_2 \cdots q_i}, \quad i \ge 1.$$

以下验证 $\{\eta_i\}$ 是对称化序列. 首先

$$\eta_0 p_{01} = 1 = \frac{p_0}{q_1} q_1 = \eta_1 p_{10},$$

而对于一般情形,

$$\eta_i p_{i,i+1} = \frac{p_0 p_1 \cdots p_{i-1}}{q_1 q_2 \cdots q_i} p_i = \frac{p_0 p_1 \cdots p_i}{q_1 q_2 \cdots q_{i+1}} q_{i+1} = \eta_{i+1} q_{i+1}, \quad i \ge 1.$$

(2) 如果 $\sum_{i=0}^{\infty} \eta_i < \infty$, 则该过程存在平稳可逆分布

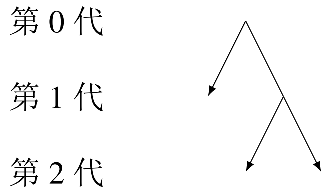
$$\pi_i = \frac{\eta_i}{\sum_{i=0}^{\infty} \eta_i}, \quad i = 0, 1, \cdots.$$

否则该过程不是正常返的, 平稳可逆分布也不存在. 因此该过程有平稳可逆分布的充要条件是 $\sum_{i=0}^{\infty} \eta_i < \infty$.

(3) 由 (2) 知该过程正常返的充要条件是 $\sum_{i=0}^{\infty} \eta_i < \infty$, 因此该过程有平稳不变分布的充要条件是 $\sum_{i=0}^{\infty} \eta_i < \infty$.

4.6 离散分支过程

设每一代的粒子都独立进行分裂, 产生后代, 并用 $\{X_n\}$ 表示第 n 代生物的总数, 则称随机序列 $\{X_n\}$ 是离散时间分支过程, 也被称为 *Galton-Watson* 分支过程.



设第 0 代仅有一个粒子, $\xi_{n,k}$ 表示第 n 代的第 k 个个体分裂成的后代数, 其是来自总体 ξ 的随机变量, 则

$$X_1 = \eta_{01}, \quad X_n = \sum_{k=1}^{X_{n-1}} \xi_{n-1,k}.$$

由上面定义的随机序列 $\{X_n\}$ 是 Markov 链.

接下来, 设 $\mu = \mathbb{E}\xi$, 尝试计算 $\{X_n\}$ 的期望. 计算得

$$\mathbb{E}X_n = \mathbb{E}\left(\mathbb{E}\left(\sum_{k=1}^{X_{n-1}} \xi_{n-1,k} \middle| X_{n-1}\right)\right) = \mu\mathbb{E}X_{n-1},$$

且 $\mathbb{E}X_0 = 1$, 由此递推得到 $\mathbb{E}X_n = \mu^n$.

再设 $\sigma^2 = \mathrm{Var}\xi$, 首先设 $X_{n-1} = l$, 计算得

$$\mathbb{E}(X_n^2 | X_{n-1} = l) = \mathbb{E}\left(\sum_{k=1}^{l} \xi_{n-1,k}\right)^2 = l \cdot \mathbb{E}\xi^2 + (l^2 - l) \cdot (\mathbb{E}\xi)^2 = l \cdot \sigma^2 + l^2\mu^2,$$

从而 $\mathbb{E}(X_n^2|X_{n-1}) = \sigma^2 \cdot X_{n-1} + \mu^2 \cdot X_{n-1}^2$, 进而

$$\mathbb{E}X_n^2 = \mathbb{E}(\mathbb{E}(X_n^2|X_{n-1})) = \sigma^2 \cdot \mathbb{E}X_{n-1} + \mu^2 \cdot \mathbb{E}X_{n-1}^2,$$

根据 $\mathrm{Var}X_n = \mathbb{E}X_n^2 - (\mathbb{E}X_n)^2 = \mathbb{E}X_n^2 - \mu^{2n}$ 得

$$\mathrm{Var}X_n = \sigma^2\mu^{n-1} + \mu^2 \cdot \mathbb{E}X_{n-1}^2 - \mu^{2n} = \sigma^2\mu^{n-1} + \mu^2 \cdot \mathrm{Var}X_{n-1},$$

并且 $\mathrm{Var}X_0 = 0$, 根据上式递推得

$$\mathrm{Var}X_n = \begin{cases} \sigma^2\mu^{n-1} \cdot \dfrac{\mu^n - 1}{\mu - 1}, & \mu \neq 1, \\ n\sigma^2, & \mu = 1. \end{cases}$$

设 T_0 为首达 0 的时刻, 接下来对 $\mathbb{P}(T_0 < \infty)$ 感兴趣, 也即灭绝概率. 此时

$$\mathbb{P}(T_0 < \infty | X_0 = i) = (\mathbb{P}(T_0 < \infty | X_0 = 1))^i,$$

从而我们只需要计算在 $X_0 = 1$ 的条件下, $T_0 < \infty$ 的概率. 计算得

$$\begin{aligned} \rho &= \mathbb{P}(T_0 < \infty, X_0 = 1) \\ &= \sum_{i=0}^{\infty} \mathbb{P}(T_0 < \infty, X_1 = i | X_0 = 1) \\ &= \sum_{i=0}^{\infty} \mathbb{P}(T_0 < \infty | X_1 = i, X_0 = 1)\mathbb{P}(X_1 = i | X_0 = 1) \\ &= \sum_{i=0}^{\infty} \mathbb{P}(T_0 < \infty | X_1 = i)\mathbb{P}(X_1 = i | X_0 = 1) \\ &= \sum_{i=0}^{\infty} \rho^i p_i, \end{aligned}$$

其中 $p_i = \mathbb{P}(X_1 = i | X_0 = 1)$, 上式为关于 ρ 的方程, 且 $\rho = 1$ 是一个平凡解. 记

$$f(\rho) = \sum_{i=0}^{\infty} p_i \rho^i, \quad 0 \leq \rho \leq 1,$$

则

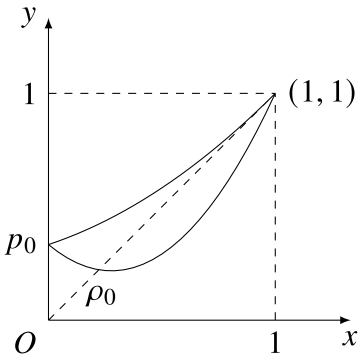
$$f'(\rho) = \sum_{i=1}^{\infty} i p_i \rho^{i-1}, \quad f''(\rho) = \sum_{i=2}^{\infty} i(i-1) p_i \rho^{i-2} > 0,$$

从而 f 是严格凸函数, 又 $f'(1) = \sum_{i=1}^{\infty} i p_i = \mu$, 以下画图进行讨论.

(1) 若 $\mu = f'(1) \leq 1$, 则 $\rho = 1$ 为 $f(\rho) = \rho$ 的唯一解;

(2) 若 $\mu = f'(1) > 1$, 则 $f(\rho) = \rho$ 在 $(0, 1)$ 内还有一解 ρ_0.

对上面的过程进行整理, 即可得到以下定理.



定理 4.19

对于一个简单分支过程, 设开始有 i 个粒子, 一个粒子的裂变分布为

$$\mathbb{P}(\xi = k) = p_k, \quad k = 0, 1, 2, \cdots.$$

令 $\mu = \mathbb{E}\xi$, 则灭绝概率 $\rho_{i,0} = \rho^i$, 其中

$$\rho = \rho_{1,0} = \begin{cases} 1, & \mu \le 1, \\ \rho_0, & \mu > 1, \end{cases}$$

其中 ρ_0 为方程 $\rho = \sum_{k=0}^{\infty} p_k \rho^k$ 在 $(0, 1)$ 内的唯一交点.

4.7 课后习题

问题 4.1 设 I 是状态空间, $A, A_j \subset I, j = 0, 1, 2, \cdots$, 用 (1) 和 (2) 推导 (3) 和 (4).

(1) 已知 $X_n = i$ 的情况下, 将来 $(X_m : m \ge n + 1)$ 与过去 $(X_j : j \le n - 1)$ 独立;

(2) $\mathbb{P}(X_{n+k} = j | X_n = i) = \mathbb{P}(X_k = j, X_0 = i)$;

(3) $\mathbb{P}(X_{n+k} = j | X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) = \mathbb{P}(X_k = j | X_0 = i)$;

(4) $\mathbb{P}(X_{n+k} \in A | X_n = i, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) = \mathbb{P}(X_k \in A | X_0 = i)$;

问题 4.2 对 $n \ge 0, k \ge 1$, $\xi_{n,k}$ 是取非负整数值的独立同分布随机变量. X_0 是取正整数值的随机变量, 与 $\{\xi_{n,k}\}$ 独立, 定义

$$X_{n+1} = \sum_{k=1}^{X_n} \xi_{n,k}, \quad n \ge 0,$$

证明 $\{X_n\}$ 是 Markov 链, 并且当 $X_0 = 1$ 时, 计算 $\mathbb{E}X_n$.

问题 4.3 对于固定的 j, 证明 $M(n) = \max\{p_{ij}^{(n)} | i \in I\}$ 关于 n 单调不增.

问题 4.4 设 $I = \{1, 2, 3\}$, 转移矩阵

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

对每个状态进行分类.

问题 4.5 探讨 Markov 链 $\{X_n\}$ 的转移矩阵 $\boldsymbol{P}$ 各行向量相同的充分必要条件.

问题 4.6 一只蜘蛛在座钟的 12 个数字上做随机游动, 每次以概率 p 顺时针走一步, 以概率 q 逆时针走一步, 用 X_n 表示 n 时蜘蛛的位置.

(1) 说明 $\{X_n\}$ 是 Markov 链, 写出转移矩阵, 计算平稳不变分布;

(2) 给出 $\{X_n\}$ 存在可逆分布的条件, 求可逆分布.

问题 4.7 对于分支过程计算群体灭绝概率 ρ_0, 其中 $p_0 = 0.2, p_1 = 0.3, p_2 = 0.5$.

问题 4.8 设 $\xi \sim \mathcal{B}(2, p)$. 分支过程中每个粒子分裂成的后代数是来自总体 ξ 的随机变量, 当 $X_0 = 1$ 时, 计算

(1) 群体灭绝的概率;

(2) 群体恰在第 2 代灭绝的概率;

(3) 如果 $X_0 \sim P(\mu), p > 0.5$, 计算群体最终灭绝的概率.

第 5 章 连续时间 Markov 链

在本节中, Markov 链指的是连续时间 Markov 链.

5.1 Markov 链与 Poisson 过程

5.1.1 Markov 链的定义

上一节中所讨论的是离散时间 Markov 链, 在这里给出连续时间 Markov 链的定义.

定义 5.1 (连续时间 Markov 链)

设 I 是状态空间, $\{X(t), t \geq 0\}$ 是以 I 为状态空间的一个随机过程, 对任意的正整数 $n, t_0 < t_1 < \cdots < t_{n+1}$ 以及状态 $i, j, i_{n-1}, \cdots, i_0 \in I$, 有

$$\mathbb{P}(X(t_{n+1}) = j | X(t_n) = i, X(t_{n-1}) = i_{n-1}, \cdots, X(t_0) = t_0) = \mathbb{P}(X(t_{n+1}) = j | X(t_n) = i),$$

则称 $\{X(t)\}$ 是连续时间离散状态的 *Markov* 链, 简称连续时间 *Markov* 链. ♣

上述定义和离散时间 Markov 链是类似的. 若 $\{X(t)\}$ 还具有时齐性, 也即

$$\mathbb{P}(X(t+s) = j | X(s) = i) = \mathbb{P}(X(t) = j | X(0) = i),$$

则可记上述概率为 $p_{ij}(t)$, 其中 t 表示间隔时间. 以下我们讨论的 Markov 链都是时齐的. 另外, Markov 性还可以用

$$\mathbb{P}(X(t_{n+1}) = j | X(t_n) = i, X(t') = i', t' \in [0, t_n)) = \mathbb{P}(X(t_{n+1}) = j | X(t_n) = i)$$

来刻画.

对于 $\{X(t)\}$ 和 $p_{ij}(t)$, 有如下的性质.

(1) $0 \leq p_{ij}(t) \leq 1, \forall i, j \in I, t \geq 0$, 且 $\sum_{j \in I} p_{ij}(t) = 1, t \geq 0$;

(2) 已知 $\{X(t) = i\}$ 的条件下, $\{X(u), 0 \leq u < t\}$ 与 $\{X(v), v > t\}$ 独立;

(3) *Kolmogorov-Chapman* 方程, 也即

$$p_{ij}(t+s) = \sum_{k \in I} p_{ik}(t) p_{kj}(s), \quad t, s \geq 0.$$

为了方便, 记转移概率矩阵 $\boldsymbol{P}(t) = (p_{ij}(t))_{i,j \in I}$, 则 $\boldsymbol{P}(t+s) = \boldsymbol{P}(t)\boldsymbol{P}(s)$;

(4) 考虑 $\{X(t)\}$ 的有限维分布, 在给定 $X(0) = i_0$ 的条件下,

$$\begin{aligned} &\mathbb{P}(X(t_1) = i_1, X(t_2) = i_2, \cdots, X(t_n) = i_n | X(0) = i_0) \\ &= p_{i_0 i_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1}, i_n}(t_n - t_{n-1}), \end{aligned}$$

从而 $\{X(t)\}$ 的有限维分布

$$\begin{aligned}&\mathbb{P}(X(t_1)=i_1,X(t_2)=i_2,\cdots,X(t_n)=i_n)\\&=\sum_{k\in I}\mathbb{P}(X(0)=k)p_{i_0i_1}(t_1)p_{i_1i_2}(t_2-t_1)\cdots p_{i_{n-1},i_n}(t_n-t_{n-1}).\end{aligned}$$

特别地, 当 $t_{i+1}-t_i=a$ 时, 有

$$\mathbb{P}(X(a)=i,X(2a)=i,\cdots,X(na)=i|X(0)=i)=[p_{ii}(a)]^n;$$

(5) 考虑 $X(t)$ 的分布, 设 $X(0)$ 的分布与 $X(t)$ 的分布

$$\boldsymbol{\pi}(0)=(\mathbb{P}(X(0)=i))_{i\in I},\quad \boldsymbol{\pi}(t)=(\mathbb{P}(X(t)=i))_{i\in I},$$

转移概率矩阵为 $\boldsymbol{P}(t)$, 则 $X(t)$ 的分布 $\boldsymbol{\pi}(t)=\boldsymbol{\pi}(0)\boldsymbol{P}(t)$, 此即说明 $X(t)$ 的分布由 $X(0)$ 的分布和 $\boldsymbol{P}(t)$ 唯一决定;

(6) 对于 $\boldsymbol{P}(t)$ 中的元素, 有

$$p_{ij}(t+s)\geq p_{ik}(t)p_{kj}(s),$$

特别地当 $i=j=k$ 时, 有 $p_{jj}(t+s)\geq p_{jj}(t)p_{jj}(s)$, 并且 $p_jj(t)\geq p_{jj}^n\left(\frac{t}{n}\right)$.

5.1.2 Poisson 过程是连续时间 Markov 链

设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, 则对 $s,t\geq 0$, 有 $\mathbb{P}(N(s,s+t]=k)=\frac{(\lambda t)^k}{k!}\cdot \mathrm{e}^{-\lambda}$. 为了验证 Markov 性与时齐性, 计算得

$$\begin{aligned}&\mathbb{P}(N(t_{n+1})=j|N(t_n)=i,N(t_{n-1})=i_{n-1},\cdots,N(t_0)=i_0)\\&=\mathbb{P}(N(t_{n-1},t_n]=j-i|N(t_n)=i,N(t_{n-1})=i_{n-1},\cdots,N(t_0)=i_0)\\&=\mathbb{P}(N(t_{n-1},t_n]=j-i)\\&=\frac{\lambda(t_{n+1}-t_n)^{j-i}}{(j-i)!}\cdot \mathrm{e}^{-\lambda(t_{n+1}-t_n)}\\&=\mathbb{P}(N(t_{n+1})=j|N(t_n)=i),\end{aligned}$$

从而 Poisson 过程是连续时间 Markov 链. 以下, Poisson 过程也用 $\{X(t),t\geq 0\}$ 表示. 对于 Poisson 过程 $\{X(t)\}$, 其具有以下特殊的性质.

(1) 初始分布 $\mathbb{P}(X(0)=0)=1,\mathbb{P}(X(0)=i)=0(i\neq 0)$;

(2) 概率转移矩阵中的元素

$$p_{ij}(t)=\mathbb{P}(X(t+s)-X(s)=j-i)=\begin{cases}\dfrac{(\lambda t)^{j-i}}{(j-i)!}\cdot \mathrm{e}^{-\lambda t}, & j\geq i,\\ 0, & i<j,\end{cases}\quad t\geq 0.$$

特别地,

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

接下来, 我们尝试求出 $p_{ij}(t)$ 在 0 处的导数. 若 $j < i$, 则 $p_{ij}(t) = 0$, 从而 $p'_{ij}(0) = 0$; 若 $j = i$, 则 $p_{ij}(0) = 1$, 从而

$$p'_{ij}(0) = \lim_{t \to 0^+} \frac{\mathrm{e}^{-\lambda t} - 1}{t} = -\lambda;$$

若 $j > i$, 则

$$p'_{ij}(0) = \lim_{t \to 0^+} \frac{\dfrac{(\lambda t)^{j-i}}{(j-i)!} \cdot \mathrm{e}^{-\lambda t}}{t} = \begin{cases} \lambda, & j = i + 1, \\ 0, & j > i + 1. \end{cases}$$

记 $q_{ij} = p'_{ij}(0)$, $\boldsymbol{Q} = (q_{ij})_{i,j \in I}$, 则

$$\boldsymbol{Q} = -\lambda \begin{bmatrix} 1 & -1 & 0 & \cdots \\ 0 & 1 & -1 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \Longrightarrow \boldsymbol{Q}^k = (-\lambda)^k \begin{bmatrix} k & -k & \frac{k(k-1)}{2} & \cdots \\ 0 & k & -k & \cdots \\ 0 & 0 & k & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

其中 $\boldsymbol{Q}^k = (q_{ij}^{(k)})_{i,j \in I}$, 且 $q_{ij}^{(k)} = (-\lambda)^k (-1)^{j-i} \dbinom{k}{j-i}$, 为了方便, 定义当 $j - i < 0$ 或 $j - i > k$ 时, $\dbinom{k}{j-i} = 0$.

$q_{ij}^{(k)}$ 的表达式可以用数学归纳法证明. 假设 $q_{ij}^{(k)}$ 由上式给出, 则

$$q_{ij}^{(k+1)} = (-\lambda)^k \left[0, 0, \cdots, \binom{k}{0}, -\binom{k}{1}, \binom{k}{2}, \cdots \right] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1 \\ 1 \\ 0 \\ \vdots \end{bmatrix}.$$

若 $i > j$, 则 $q_{ij}^{(k+1)} = 0$; 若 $i \leq j$, 则

$$\begin{aligned} q_{ij}^{(k+1)} &= (-\lambda)^k (-1) \cdot \binom{k}{j-1-i} (-1)^{j-1-i} + (-\lambda)^k \binom{k}{j-i} (-1)^{j-i} \\ &= (-\lambda)^k (-1)^{j-1} \binom{k+1}{j-i}. \end{aligned}$$

综合以上过程即可说明 $q_{ij}^{(k)} = (-\lambda)^k (-1)^{j-i} \dbinom{k}{j-i}$, $k \geq 1$. 以下, 为了进一步建立 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$

的关系, 考虑级数

$$\begin{aligned}\sum_{k=0}^{\infty} q_{ij}^{(k)} \cdot \frac{t^k}{k!} &= \sum_{k=j-1}^{\infty} \binom{k}{j-i} (-1)^{j-1} \cdot \frac{(-\lambda t)^k}{k!} \\ &= \frac{(\lambda t)^{j-i}}{(j-i)!} \cdot \sum_{k=j-i}^{\infty} \frac{1}{(k-(j-i))!} \cdot (-\lambda t)^{k-(j-i)} \\ &= \frac{(\lambda t)^{j-i}}{(j-i)!} \cdot \mathrm{e}^{-\lambda t} \\ &= p_{ij}(t),\end{aligned}$$

右边正是我们的转移概率, 将上式写成矩阵形式得

$$\boldsymbol{P}(t) = \sum_{k=0}^{\infty} \boldsymbol{Q}^k \cdot \frac{t^k}{k!} = \mathrm{e}^{\boldsymbol{Q}t}.$$

上面的矩阵 $\boldsymbol{Q}$ 被称为转移概率矩阵, 这将在下一节中介绍.

5.2 Markov 链的转移概率矩阵

5.2.1 规则 Markov 链与保守 Markov 链

定义 5.2 (规则 Markov 链)

在概率 1 的意义下, 若 Markov 链 $\{X(t)\}$ 在任何有限时间内只能转移有限次, 则称其为规则 *Markov* 链. ♣

可以看出, 规则 Markov 链的轨迹是右连续的阶梯函数. 以下我们讨论的 Markov 链都是规则 Markov 链.

定理 5.1

设 I 是状态空间, $\{X_n\}$ 是 Markov 链.

(1) $p_{ij}(t)$ 在 $t=0$ 处连续, 也即 $\lim\limits_{t\to 0^+} p_{ij}(t) = p_{ij}(0) = \delta_{ij}$;

(2) $p_{ij}(t)$ 在 $[0,+\infty)$ 上一致连续, 且

$$\sum_{j\in I} |p_{ij}(t+h) - p_{ij}(t)| \le 2(1 - p_{ii}(h));$$

(3) 对于 $t \ge 0$, $p_{ii}(t) > 0$;

(4) $p_{ij}(t)$ 在 $t=0$ 处有右导数, 也即

$$\lim_{t\to 0^+} \frac{p_{ij}(t) - p_{ij}(0)}{t} = q_{ij};$$

(5) 对于 $i \in I$, $\sum\limits_{j\neq i} q_{ij} \le |q_{ii}|$. ♡

定理 5.2

设 I 是状态空间, $\{X_n\}$ 是 Markov 链, 定义 $q_i = -q_{ii}$.

(1) 若 $q_i = 0$, 则对所有的 $t \geq 0$, $p_{ii} = 1$;

(2) $q_i = \sup\limits_{t>0} \dfrac{1 - p_{ii}(t)}{t}$.

在以上的结果的基础上, 称 $\boldsymbol{Q} = (q_{ij})_{i,j\in I}$ 为转移速率矩阵或强度矩阵. 考虑到 $\boldsymbol{Q}$ 其实是 $\boldsymbol{P}(t)$ 在 0 处的导数, 因此 $\boldsymbol{Q}$ 可以称为无穷小矩阵. 在 $\boldsymbol{Q}$ 的定义下, 给出保守 Markov 链的定义.

定义 5.3 (保守 Markov 链)

如果 Markov 链 $\{X(t)\}$ 的转移速率矩阵满足

$$\sum_{j\neq i} q_{ij} = |q_{ii}| < \infty, \quad \forall i \in I,$$

则称其为保守 *Markov* 链.

可以说, 至今在实际应用中见到的 Markov 链都是保守的. 事实上, 根据 $q_{ii} = p'_{ii}(0) \leq 0$ 得

$$\sum_{j\neq i} q_{ij} = |q_{ii}| \iff \sum_{j\in I} q_{ij} = 0.$$

对于有限状态 Markov 链, 若 $|q_{ii}| < \infty$, 则其一定是保守的, 这是因为

$$\sum_{j\in I} q_{ij} = \sum_{j\in I} \lim_{h\to 0^+} \frac{p_{ij}(h) - p_{ij}(0)}{h} = \lim_{h\to 0^+} \sum_{j\in I} \frac{p_{ij}(h) - p_{ij}(0)}{h} = 0.$$

例 5.1 设 $X(t) = 1$ 表示 t 时刻占用, $X(t) = 0$ 表示 t 时刻未占用, 转移矩阵

$$\boldsymbol{P}(t) = \begin{bmatrix} \dfrac{1+7\mathrm{e}^{-8t}}{8} & \dfrac{7-7\mathrm{e}^{-8t}}{8} \\ \dfrac{1-\mathrm{e}^{-8t}}{8} & \dfrac{7+\mathrm{e}^{-8t}}{8} \end{bmatrix}$$

且初始分布 $\mathbb{P}(X(0) = 0) = 0.1, \mathbb{P}(X(0) = 1) = 0.9$. 求:

(1) $\boldsymbol{P}(0)$;

(2) $\mathbb{P}(X(0.2) = 0)$, $\mathbb{P}(X(0.2) = 0|X(0) = 0)$, $\mathbb{P}(X(0.1) = 0, X(0.6) = 1, X(1.1) = 0)$, $\mathbb{P}(X(0.1) = 0, X(0.6) = 1, X(1.1) = 0|X(0) = 0)$;

(3) t 时刻的一维分布;

(4) 转移速率矩阵 $\boldsymbol{Q}$ 是否保守?

解答 (1) 代入 $t = 0$ 得

$$\boldsymbol{P}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}.$$

(2) 先设 $X(0)=0$, 则

$$\mathbb{P}(X(0.2)=0|X(0)=0)=p_{00}(0.2)=\frac{1+7\mathrm{e}^{-1.6}}{8}\approx 0.3017,$$

同理可以求出

$$\mathbb{P}(X(0.2)=0|X(0)=1)=p_{10}(0.2)=\frac{1-\mathrm{e}^{-1.6}}{8}\approx 0.0998,$$

从而根据全概率公式得

$$\begin{aligned}\mathbb{P}(X(0.2)=0)&=\sum_{k\in\{0,1\}}\mathbb{P}(X(0.2)=0|X(0)=k)\mathbb{P}(X(0)=k)\\&=\frac{5-\mathrm{e}^{-1.6}}{40}\\&\approx 0.1200.\end{aligned}$$

另外, 先设 $X(0)=0$, 则

$$\begin{aligned}&\mathbb{P}(X(0.1)=0,X(0.6)=1,X(1.1)=0|X(0)=0)\\&=p_{00}(0.1)p_{01}(0.6-0.1)p_{10}(1.1-0.6)\\&=\frac{1+7\mathrm{e}^{-0.8}}{8}\cdot\frac{7-7\mathrm{e}^{-4}}{8}\cdot\frac{1-\mathrm{e}^{-4}}{8}\\&\approx 0.0546,\end{aligned}$$

同理可以求出

$$\begin{aligned}&\mathbb{P}(X(0.1)=0,X(0.6)=1,X(1.1)=0|X(0)=1)\\&=p_{10}(0.1)p_{01}(0.6-0.1)p_{10}(1.1-0.6)\\&=\frac{1-\mathrm{e}^{-0.8}}{8}\cdot\frac{7-7\mathrm{e}^{-4}}{8}\cdot\frac{1-\mathrm{e}^{-4}}{8}\\&\approx 0.0073,\end{aligned}$$

从而根据全概率公式得

$$\begin{aligned}\mathbb{P}(X(0.2)=0)&=\sum_{k\in\{0,1\}}\mathbb{P}(X(0.1)=0,X(0.6)=1,X(1.1)=0|X(0)=k)\mathbb{P}(X(0)=k)\\&\approx 0.0120.\end{aligned}$$

(3) 计算得

$$\begin{aligned}\mathbb{P}(X(t)=0)&=\sum_{k\in\{0,1\}}\mathbb{P}(X(t)=0|X(0)=k)\mathbb{P}(X(0)=k)\\&=0.1p_{00}(t)+0.9p_{10}(t)\\&=\frac{5-\mathrm{e}^{-8t}}{40},\end{aligned}$$

同理

$$\mathbb{P}(X(t)=1)=0.1p_{01}(t)+0.9p_{11}(t)=\frac{35+2\mathrm{e}^{-8t}}{40},$$

因此 $X(t)$ 的分布为 $\left(\dfrac{5-\mathrm{e}^{-8t}}{40}, \dfrac{35+2\mathrm{e}^{-8t}}{40}\right)$.

(4) 此时

$$\boldsymbol{Q} = \boldsymbol{P}'(0) = \begin{bmatrix} -7 & 7 \\ 1 & -1 \end{bmatrix},$$

注意到 $\boldsymbol{Q}$ 的行和都是 0, 因此 $\{X(t)\}$ 是保守 Markov 链.

5.2.2 Kolmogorov 方程

接下来, 我们尝试探索 $\boldsymbol{P}(t)$ 与 $\boldsymbol{Q}$ 之间的关系. 在上一节中, 对于 Poisson 过程, 我们发现 $\boldsymbol{P}(t) = \exp(\boldsymbol{Q}t)$. 对于一般的 Markov 链, 该性质是否仍然成立呢? 定理5.3建立了 $\boldsymbol{P}(t)$ 与 $\boldsymbol{Q}$ 之间的关系.

定理 5.3 (Kolmogorov 方程)

设 $\{X(t)\}$ 是 Markov 链, $\boldsymbol{P}(t)$ 是其转移矩阵, $\boldsymbol{Q}$ 是其转移速率矩阵, 则有

(1) Kolmogorov 向前方程:

$$\begin{cases} p'_{ij}(t) = \sum\limits_{k\in I} q_{ik}p_{kj}(t), \\ \boldsymbol{P}'(t) = \boldsymbol{P}(t)\boldsymbol{Q}; \end{cases}$$

(2) Kolmogorov 向后方程:

$$\begin{cases} p'_{ij}(t) = \sum\limits_{k\in I} p_{kj}(t)q_{ik}, \\ \boldsymbol{P}'(t) = \boldsymbol{Q}\boldsymbol{P}(t). \end{cases}$$

♡

证明 根据 Kolmogorov-Chapman 方程, 我们知道

$$p_{ij}(t+s) = \sum_{k\in I} p_{ik}(t)p_{kj}(s) = p_{ij}(t)p_{jj}(s) + \sum_{k\neq j} p_{ik}(t)p_{kj}(s).$$

对于上式中的 $p_{jj}(s)$, 当 $s \to 0$ 时, 有

$$p_{jj}(s) = p_{jj}(0) + p'_{jj}(0)\cdot s + o(s) = 1 + q_{jj}(0)\cdot s + o(s);$$

而当 $k \neq j$ 时, 当 $s \to 0$ 时, 有

$$p_{kj}(s) = p_{kj}(0) + p'_{kj}(0)\cdot s + o(s) = q_{kj}\cdot s + o(s).$$

综合以上结果, 代入 Kolmogorov-Chapman 方程得

$$p_{ij}(t+s) = p_{ij}(t) + \sum_{k\in I} p_{ik}(t)q_{kj}\cdot s + o(s),$$

对上式整理得

$$\frac{p_{ij}(t+s)-p_{ij}(t)}{s}=\sum_{k\in I}p_{ik}(t)q_{kj}+\frac{o(s)}{s}$$

$$p'_{ij}(t)=\lim_{s\to 0}\frac{p_{ij}(t+s)-p_{ij}(t)}{s}=\sum_{k\in I}p_{ik}(t)q_{kj},$$

将上式改写为矩阵形式得 $\boldsymbol{P}'(t)=\boldsymbol{P}(t)\boldsymbol{Q}$, 此即 Kolmogorov 向前方程; 另外, 还可以证明

$$p'_{ij}(t)=\sum_{k\in I}q_{ik}p_{kj}(t),$$

将上式改写为矩阵形式得 $\boldsymbol{P}'(t)=\boldsymbol{Q}\boldsymbol{P}(t)$, 此即 Kolmogorov 向后方程.

另外, 对于分布而言, 还可以得到如下的方程.

定理 5.4 (Fokker-Plank 方程)

设 $\{X(t)\}$ 的初始分布为 $\boldsymbol{\pi}(0)$, t 时刻的分布为 $\boldsymbol{\pi}(t)$, 则 $\boldsymbol{\pi}'(t)=\boldsymbol{\pi}(t)\boldsymbol{Q}$.

证明 根据 $\boldsymbol{\pi}(t)=\boldsymbol{\pi}(0)\boldsymbol{P}(t)$, 结合定理5.3得

$$\boldsymbol{\pi}'(t)=\boldsymbol{\pi}(0)\boldsymbol{P}'(t)=\boldsymbol{\pi}(0)\boldsymbol{Q}\boldsymbol{P}(t)=\boldsymbol{\pi}(0)\boldsymbol{P}(t)\boldsymbol{Q}=\boldsymbol{\pi}(t)\boldsymbol{Q},$$

从而 $\boldsymbol{\pi}'(t)=\boldsymbol{\pi}(t)\boldsymbol{Q}$.

对于 Poisson 过程而言, 我们知道 $\boldsymbol{P}(t)$ 和 $\boldsymbol{Q}$ 可以通过指数函数联系在一起. 然而, 一般的 Markov 过程并不能得到这样的关系. 如果我们限制 Markov 链是有限状态的 Markov 链的话, 应用定理5.3, 可以得到如下结果.

定理 5.5

有限状态 Markov 链 $\{X(t)\}$ 的转移概率矩阵 $\boldsymbol{P}(t)$ 由转移速率矩阵 $\boldsymbol{Q}$ 唯一决定, 且

$$P(t)=\exp(\boldsymbol{Q}t)=\sum_{k=0}^{\infty}\frac{(t\boldsymbol{Q})^k}{k!}.$$

以下是一个有限状态 Markov 链的例子.

例 5.2 设 $\{X(t), t\geq 0\}$ 是 Markov 链, 状态空间 $I=\{1,2,\cdots,m\}$, 且

$$q_{ij}=\begin{cases}-(m-1), & i=j,\\ 1, & i\neq j.\end{cases}$$

求 $p_{ij}(t)$.

解答 根据 Kolmogorov 方程得

$$\begin{aligned} p'_{ij}(t) &= \sum_{k\in I} p_{ik}(t)q_{kj} \\ &= \sum_{k\neq j} p_{ik}(t) - (m-1)p_{ij}(t) \\ &= 1 - p_{ij}(t) - (m-1)p_{ij}(t) \\ &= 1 - mp_{ij}(t), \end{aligned}$$

解微分方程 $p'_{ij}(t) = 1 - mp_{ij}(t)$ 得

$$p_{ij}(t) = \begin{cases} \dfrac{1+(m-1)\mathrm{e}^{-mt}}{m}, & i = j, \\ \dfrac{1-\mathrm{e}^{-mt}}{m}, & i \neq j. \end{cases}$$

另外, 也可以尝试根据定理5.5求解.

5.3 Markov 链的结构

在连续时间 Markov 链的转移矩阵 $\boldsymbol{P}(t)$ 中, 有

$$p_{ij}(t) = \mathbb{P}(X(s+t) = j|X(s) = i) = \mathbb{P}(X(t) = j|X(0) = i).$$

类比离散时间 Markov 链, 我们在此研究连续时间 Markov 链的结构.

定理 5.6

设 $\{X(t)\}$ 是 Markov 链, $q_i = |q_{ii}|, t, h > 0$.

(1) $\mathbb{P}(X(t+h) = j|X(u) = i, u \in [0,t]) = p_{ij}(h)$;

(2) $\mathbb{P}(X(u) = i, u \in [0,t]|X(0) = i) = \mathrm{e}^{-q_i t}$.

证明 (1) 根据 Markov 链的定义得

$$\mathbb{P}(X(t+h) = j|X(u) = i, u \in [0,t]) = \mathbb{P}(X(t+h) = j|X(t) = i) = p_{ij}(h).$$

(2) 若 $q_i = 0$, 则以上概率为 1; 若 $q_i > 0$, 考虑将区间 $[0,t]$ 离散化. 令

$$\begin{cases} B_n = \left\{\dfrac{jt}{2^n}, 1 \leq j \leq 2^n\right\}, \\ A_n = (X(u) = i, u \in B_n), \end{cases}$$

则 $\{B_n\}$ 单调递增, $\{A_n\}$ 单调递减, $\{B_n\}$ 在 $[0,t]$ 上稠密. 且 $X(t)$ 的轨迹是右连续的阶梯函数, 从而

$$\bigcap_{n=1}^{\infty} A_n = (X(u) = i, u \in [0,t]),$$

因此

$$\begin{aligned}\mathbb{P}(X(u)=i,u\in[0,t]|X(0)=i)&=\mathbb{P}\left(\bigcap_{n=1}^{\infty}A_n\middle|X(0)=i\right)\\&=\lim_{n\to\infty}\mathbb{P}(A_n|X(0)=i)\\&=\lim_{n\to\infty}\left(p_{ii}\left(\frac{t}{2^n}\right)\right)^{2^n},\end{aligned}$$

其中 $p_{ii}\left(\frac{t}{2^n}\right)=1-q_i\cdot\frac{t}{2^n}+o\left(\frac{t}{2^n}\right)$, 因此 $\mathbb{P}(X(u)=i,u\in[0,t]|X(0)=i)=\mathrm{e}^{-q_i t}$.

定理 5.7

设 $\{X(t)\}$ 是 Markov 链, $q_i=|q_{ii}|$, τ 表示质点在状态 i 的停留时间.

(1) $\mathbb{P}(\tau>t|X(0)=i)=\mathrm{e}^{-q_i t}$;

(2) $\mathbb{E}(\tau|X(0)=i)=\frac{1}{q_i}$;

(3) 当 $j\neq i$ 时, $\mathbb{P}(X(\tau)=j,\tau\leq t|X(0)=i)=\frac{q_{ij}}{q_i}(1-\mathrm{e}^{-q_i t})$;

(4) 当 $j\neq i$ 时, $\mathbb{P}(X(\tau)=j|X(0)=i)=\frac{q_{ij}}{q_i}$;

(5) 在 $X(0)=i$ 的条件下, τ 和 $X(\tau)$ 独立;

(6) 当所有的 $q_i<\infty$ 时, $\{X(t)\}$ 是保守的.

证明 (1) 计算得

$$\mathbb{P}(\tau>t|X(0)=i)=\mathbb{P}(X(u)=i,u\in[0,t]|X(0)=i)=\mathrm{e}^{-q_i t}.$$

(2) 对 (1) 中的结果取期望得

$$\mathbb{E}(\tau|X(0)=i)=\int_0^{\infty}\mathbb{P}(\tau>t|X(0)=i)\mathrm{d}t=\frac{1}{q_i}.$$

(3) 为了计算出 $\mathbb{P}(X(\tau)=j,\tau\leq t|X(0)=i)$, 首先考虑 $\mathbb{P}(X(\tau)=j,\tau=s|X(0)=i)$, 其表示在 $[0,s)$ 上都停留在 i 状态, 但是在 s 之后转移到了 j 状态. 采取类似之前的方法对 $[0,s)$ 进行划分, 令

$$\begin{cases}B_n=\left\{\frac{j}{2^n},1\leq j\leq 2^n-1\right\},\\ A_n=(X(u)=i,u\in B_n),\end{cases}$$

则 $\{B_n\}$ 单调递增, $\{A_n\}$ 单调递减, $\{B_n\}$ 在 $[0,s)$ 上稠密, 并且

$$\bigcap_{n=1}^{\infty}A_n=(X(u)=i,u\in[0,s)).$$

因此

$$\begin{aligned}
\mathbb{P}(X(\tau)=j,\tau=s|X(0)=i) &= \mathbb{P}(X(t)=i,t\in[0,s),X(s)=j|X(0)=i)\\
&= \mathbb{P}\left(\bigcap_{n=1}^{\infty}A_n, X(s)=j\middle|X(0)=i\right)\\
&= \mathbb{P}\left(X(s)=j\middle|\bigcap_{n=1}^{\infty}A_n, X(0)=i\right)\cdot\mathbb{P}\left(\bigcap_{n=1}^{\infty}A_n\middle|X(0)=i\right)\\
&= \lim_{n\to\infty}\mathbb{P}\left(X(s)=j\middle|X\left(\frac{2^n-1}{2^n}\cdot s\right)=i\right)\cdot\mathbb{P}(A_n|X(0)=i)\\
&= q_{ij}\cdot \mathrm{e}^{-q_i s}\cdot\mathrm{d}s,
\end{aligned}$$

其中

$$\begin{aligned}
\mathbb{P}\left(X(s)=j\middle|X\left(\frac{2^n-1}{2^n}\right)\right) &= p_{ij}\left(\frac{s}{2^n}\right)\\
&= p_{ij}(0)+p'_{ij}(0)\cdot\frac{s}{2^n}+o\left(\frac{s}{2^n}\right)\\
&= q_{ij}\cdot\mathrm{d}s,
\end{aligned}$$

最后得到

$$\mathbb{P}(X(\tau)=j,\tau\le t|X(0)=i)=\int_0^t q_{ij}\mathrm{e}^{-q_i s}\mathrm{d}s=\frac{q_{ij}}{q_i}\left(1-\mathrm{e}^{-q_i t}\right).$$

(4) 在 (3) 的结果中, 令 $t\to\infty$, 则

$$\mathbb{P}(X(\tau)=j|X(0)=i)=\frac{q_{ij}}{q_i}.$$

(5) 根据

$$\begin{aligned}
\mathbb{P}(X(\tau)=j,\tau\le t|X(0)=i) &= \frac{q_{ij}}{q_i}\cdot\left(1-\mathrm{e}^{-q_i t}\right)\\
&= \mathbb{P}(X(\tau)=j|X(0)=i)\cdot\mathbb{P}(\tau\le t|X(0)=i),
\end{aligned}$$

因此 $X(\tau)$ 与 τ 独立.

(6) 根据 $q_i<\infty$, 有

$$\sum_{j\ne i}\frac{q_{ij}}{q_i}=1 \implies \sum_{j\ne i}q_{ij}=q_i=|q_{ii}|,$$

因此 $\{X(t)\}$ 是保守的.

考虑从初始状态开始的停留时间 τ.

- $\tau_0=0$, 对应的状态为 $X(\tau_0)$;
- $\tau_1=\inf\{X(t)\ne X(\tau_0),\tau_1\ge\tau_0\}$ 表示首次从状态 $X(\tau_0)$ 转出的时间, 对应的状态为 $X(\tau_1)$;
- $\tau_2=\inf\{X(t)\ne X(\tau_1),\tau_2\ge\tau_1\}$ 表示首次从状态 $X(\tau_1)$ 转出的时间, 对应的状态为 $X(\tau_2)$;
- $\cdots$.

对于上面所定义的 $\tau_0, \tau_1, \tau_2, \cdots$, 令

$$X_n = X(\tau_n),$$

则 $\{X_n\}$ 是离散时间 Markov 链. 考虑 $\{X_n\}$ 的转移矩阵 $\boldsymbol{K} = (k_{ij})_{i,j\in I}$, 则有

$$k_{ij} = \begin{cases} \dfrac{q_{ij}}{q_i}, & q_i \neq 0, j \neq i, \\ 0, & q_i \neq 0, j = i, \\ \delta_{ij}, & q_i = 0. \end{cases}$$

现在在上面的基础上, 对 Markov 链的结构作如下的整理.

(1) $X_n = X(\tau_n)$ 是以 $\boldsymbol{K} = (k_{ij})_{i,j\in I}$ 为转移矩阵的离散 Markov 链, 称为 $\{X(t), t \geq 0\}$ 的嵌入链或跳跃链;

(2) 沿着嵌入链轨迹 $i_0 \to i_1 \to i_2 \to \cdots$, 并设在每个状态的停留时间为 $T_0, T_1, \cdots$, 则 $T_1, T_2, \cdots$ 相互独立, 且 $T_i \sim \mathrm{Exp}(q_i)$.

为了加深对 Markov 链的结构的理解, 在此考虑一个特殊的 Markov 链的结构, 也即 Poisson 过程.

例 5.3 对于 Poisson 过程 $\{N(t)\}$ 而言,

$$\boldsymbol{Q} = (-\lambda)\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \Longrightarrow \boldsymbol{K} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

另外, 根据之前的结论知, Poisson 过程的等待时间间隔 $X_n = S_n - S_{n-1} \sim \mathrm{Exp}(\lambda)$. 从另外一个角度来看, 根据 Poisson 过程是一个计数过程, 可以得到矩阵 $\boldsymbol{K}$ 的形式如上式所示, 再借助 $\boldsymbol{Q}$ 和 $\boldsymbol{K}$ 的关系也可以求出 $\boldsymbol{Q}$.

例 5.4 设 Markov 链 $\{X(t)\}$ 有转移概率矩阵

$$\boldsymbol{P}(t) = \frac{1}{5}\begin{bmatrix} 2+3\mathrm{e}^{-3t} & 1-\mathrm{e}^{-3t} & 2-2\mathrm{e}^{-3t} \\ 2-2\mathrm{e}^{-3t} & 1+4\mathrm{e}^{-3t} & 2-2\mathrm{e}^{-3t} \\ 2-2\mathrm{e}^{-3t} & 1-\mathrm{e}^{-3t} & 2+3\mathrm{e}^{-3t} \end{bmatrix}.$$

(1) 计算转移速率矩阵 $\boldsymbol{Q}$;

(2) 计算质点在各状态的平均停留时间;

(3) 计算嵌入链的一步转移概率矩阵.

解答 (1) 计算得

$$\boldsymbol{Q} = \boldsymbol{P}'(0) = \frac{1}{5}\begin{bmatrix} -9 & 3 & 6 \\ 6 & -12 & 6 \\ 6 & 3 & -9 \end{bmatrix} = \frac{3}{5}\begin{bmatrix} -3 & 1 & 2 \\ 2 & -4 & 2 \\ 2 & 1 & -3 \end{bmatrix}.$$

(2) 根据 $\mathbb{E}(\tau|X(0)=i)=\dfrac{1}{q_i}=\dfrac{1}{|q_{ii}|}$, 得知平均停留时间分别为 $\dfrac{5}{9},\dfrac{5}{12},\dfrac{5}{9}$.

(3) 此时 $q_i\neq 0$, 当 $i\neq j$ 时, $k_{ij}=\dfrac{q_{ij}}{q_i}$, 而当 $i=j$ 时, $k_{ij}=0$, 据此写出

$$\boldsymbol{K}=(k_{ij})_{i,j\in I}=\begin{bmatrix}0 & \frac{1}{3} & \frac{2}{3}\\ \frac{1}{2} & 0 & \frac{1}{2}\\ \frac{2}{3} & \frac{1}{3} & 0\end{bmatrix}.$$

5.4 生灭过程

5.4.1 指数分布的性质

在此首先复习指数分布的性质, 设 $\mathrm{Exp}(\lambda)$ 是参数为 λ 的指数分布, $t\geq 0$.

首先, 若 t 时刻存活的细胞在 $(t,t+h]$ 分裂的概率为 $\lambda h+o(h)$, 则细胞寿命 $T\sim\mathrm{Exp}(\lambda)$. 条件也即

$$\mathbb{P}(t<T\leq t+h|T>t)=\lambda h+o(h),$$

因此

$$\begin{aligned}\mathbb{P}(t<T\leq t+h|T>t)&=\frac{\mathbb{P}(t<T\leq t+h,T>t)}{\mathbb{P}(T>t)}\\&=\frac{\mathbb{P}(T>t)-\mathbb{P}(T>t+h)}{\mathbb{P}(T>t)}\\&=\lambda h+o(h),\end{aligned}$$

令 $\overline{F}(t)=\mathbb{P}(T>t)$, 对上式整理得

$$\frac{\overline{F}(t+h)-\overline{F}(t)}{h}=-\lambda\overline{F}(t)+\frac{o(h)}{h}\implies\overline{F}'(t)=-\lambda\overline{F}(t),$$

又 $\overline{F}(0)=1$, 解此微分方程得 $\overline{F}(t)=\mathrm{e}^{-\lambda t}$, 从而 $T\sim\mathrm{Exp}(\lambda)$.

接下来, 若 t 时刻有 m 个细胞, 对于第 i 个细胞, 其在 $(t,t+h]$ 分裂的概率为 $\lambda_i h+o(h)$, 则从 t 时刻开始, 等待第一次分裂的时间 $T\sim\mathrm{Exp}(\lambda_1+\lambda_2+\cdots+\lambda_m)$. 为了验证这一点, 设 T_i 是第 i 个细胞的分裂, 则 $T>t\iff T_i>t$ 对任意的 i 成立, 从而

$$\mathbb{P}(T>t)=\prod_{i=1}^{n}\mathbb{P}(T_i>t)=\mathrm{e}^{-(\lambda_1+\lambda_2+\cdots+\lambda_m)t},$$

从而 $T\sim\mathrm{Exp}(\lambda_1+\lambda_2+\cdots+\lambda_m)$.

5.4.2 线性生灭过程

定义 5.4 (线性生灭过程)

设一个 t 时存活的生物个体在 $(t,t+h]$ 内的分裂情况与其在 t 时的年龄无关, 并且

(1) 在 $(t,t+h]$ 内死亡的概率为 $\mu h+o(h)$;

(2) 在 $(t,t+h]$ 内不死亡也不分裂的概率是 $1-(\lambda+\mu)h+o(h)$;

(3) 在 $(t,t+h]$ 内分裂一次成为两个个体的概率为 $\lambda h+o(h)$.

用 $X(t)$ 表示 t 时刻生物的总数, 则称 $\{X(t)\}$ 为线性生灭过程, 并称 μ 和 λ 为生物个体的死亡强度和出生强度.

♣

另外, 考虑个体在 $(t,t+h]$ 内分裂为超过两个个体的概率

$$1-(\mu h+o(h))-(1-(\lambda+\mu)h+o(h))-(\lambda h+o(h))=o(h),$$

从而分裂为超过两个个体的概率趋于 0. 接下来探究线性生灭过程 $\{X(t)\}$ 的性质.

(1) 首先, 根据指数分布的无记忆性, 以及单个个体分裂和死亡的过程与年龄无关, 得知 $\{X(t)\}$ 是 Markov 链.

(2) 接下来, 考虑该 Markov 链的转移矩阵 $\boldsymbol{P}(h)$ 的元素 $p_{ij}(h)$. 先设 $i=0$, 则 $p_{00}(h)=1, p_{0j}=0, j\neq 0$; 再设 $i\neq 0$, 则

$$\begin{aligned} p_{ii}(h) &= (1-(\lambda+\mu)h+o(h))^i+o(h) \\ &= 1-i(\lambda+\mu)h+o(h); \end{aligned}$$

最后设 $i,j\neq 0$, 若 $|j-i|\geq 2$, 则 $p_{ij}(h)=o(h)$, 只需计算 $p_{i,i-1}(h)$ 和 $p_{i,i+1}(h)$, 根据题目的条件, 容易得到

$$\begin{aligned} p_{i,i-1}(h) &= i(\mu h+o(h))(1-(\lambda+\mu)h+o(h))^{i-1} \\ &= i\mu h+o(h), \end{aligned}$$

同理 $p_{i,i+1}(h)=i\lambda h+o(h)$.

(3) 在 $\boldsymbol{P}(h)$ 的基础上, 考虑转移速率矩阵 $\boldsymbol{Q}$, 则

$$q_{ij}=\begin{cases} i\mu, & j=i-1, i\neq 0, \\ -i(\lambda+\mu), & j=i, i\neq 0, \\ i\lambda, & j=i+1, i\neq 0, \\ 0, & \text{其他情况}. \end{cases}$$

容易验证 $\boldsymbol{Q}$ 的行和为 0, 从而 $\{X(t)\}$ 是保守的.

(4) 在 $\boldsymbol{Q}$ 的基础上, 记 $p=\dfrac{\lambda}{\lambda+\mu}$, $q=1-p=\dfrac{\mu}{\lambda+\mu}$, 可以得到嵌入链 $\{X_n\}$ 的转移矩阵

$$\boldsymbol{K}=\begin{bmatrix}1 & 0 & 0 & \cdots\\ q & 0 & p & \cdots\\ 0 & q & 0 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{bmatrix},$$

这说明了嵌入链 $\{X_n\}$ 是以 0 为吸收壁的随机游走, 并且我们知道这样的 Markov 链的常返等价类是 $\{0\}$, 非常返等价类是 $\{1,2,\cdots\}$, 从而个体的数量应该会趋于无穷大.

(5) 根据上一节的结论, 我们知道每一个个体的寿命服从 $\mathrm{Exp}(\lambda+\mu)$. 设 $\{X_n\}$ 在 i 状态停留的时间为 T_i, 并记 τ_j 为第 j 个个体的剩余寿命, 则

$$T_i=\min\{\tau_1,\tau_2,\cdots,\tau_i\}\sim \mathrm{Exp}(i(\lambda+\mu)),$$

在此基础上, 还可以得到平均停留时间 $\mathbb{E}T_i=\dfrac{1}{i(\lambda+\mu)}$.

(6) 考虑 $\{X_n\}$ 的一个轨迹 $j_0\to j_1\to j_2\to\cdots$, 根据随机游走的性质得 $j_i\leq j_0+i$, 故

$$\sum_{i=0}^{\infty}\mathbb{E}T_{j_i}=\sum_{i=0}^{\infty}\frac{1}{j_i(\lambda+\mu)}\geq\frac{1}{\lambda+\mu}\cdot\sum_{i=0}^{\infty}\frac{1}{j_0+i}=\infty,$$

根据指数分布的性质[1] 知 $\mathbb{P}\left(\sum_{i=0}^{\infty}T_{j_i}=\infty\right)\iff\sum_{i=0}^{\infty}\mathbb{E}T_{j_i}=\infty$, 从而嵌入链 $\{X_n\}$ 是规则的.

5.4.3 线性纯生过程

线性纯生过程是线性生灭过程的特例, 此时 $\mu=0$. 该过程也被叫做 *Yule* 过程. 在上一节的基础上, 我们得到转移速率矩阵和嵌入链的转移矩阵分别为

$$\boldsymbol{Q}=\begin{bmatrix}0 & 0 & 0 & \cdots\\ 0 & -\lambda & \lambda & \cdots\\ 0 & 0 & -2\lambda & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{bmatrix},\quad \boldsymbol{K}=\begin{bmatrix}1 & 0 & 0 & \cdots\\ 0 & 0 & 1 & \cdots\\ 0 & 0 & 0 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{bmatrix}.$$

沿用上一节中的记号, 此时 T_i 为已知有 i 个生物时, 等待下一次分裂的时间. 记 $S_0=0$, S_k 为等待第 k 次分裂的时间, 则

$$S_k=T_1+T_2+\cdots+T_k.$$

并且在这里指出, 等待第 k 次分裂的时候, 共有 k 个个体, 也即 $X_n=k$.

[1]这是 [**1**] 第 21 页例 8.1 的内容.

定理 5.8

对于等待时间 S_k, 有

$$\mathbb{P}(S_k \le t|X(0)=1) = \left(1-\mathrm{e}^{-\lambda t}\right)^k, \quad k \ge 1.$$

证明 记 $\mathbb{P}_1(\cdot) = \mathbb{P}(\cdot|X(0)=1)$, 以下使用数学归纳法完成证明.

当 $k=1$ 时, $\mathbb{P}_1(S_1 \le t) = \mathbb{P}_1(T_1 \le t) = 1-\mathrm{e}^{-\lambda t}$. 假设

$$\mathbb{P}(S_{k-1} \le t) = \left(1-\mathrm{e}^{-\lambda t}\right)^{k-1},$$

并且注意到 $S_k = S_{k-1}+T_k$, 其中 S_{k-1} 与 T_k 独立, 且 $T_k \sim \mathrm{Exp}(k\lambda)$, 因此

$$\begin{aligned}
\mathbb{P}_1(S_k \le t) &= \mathbb{P}_1(S_{k-1}+T_k \le t) \\
&= \int_0^\infty \mathbb{P}_1(S_{k-1}+T_k \le t|T_k = s)\mathrm{d}\mathbb{P}_1(T_k \le s) \\
&= \int_0^\infty \mathbb{P}_1(S_{k-1} \le t-s)\mathrm{d}\mathbb{P}_1(T_k \le s) \\
&= \int_0^\infty \left(1-\mathrm{e}^{-\lambda(t-s)}\right)^{k-1} k\lambda \mathrm{e}^{-k\lambda s}\mathrm{d}s \\
&= \int_0^\infty \sum_{l=0}^{k-1}\binom{k-1}{l}(-1)^l \mathrm{e}^{-\lambda l(t-s)}k\lambda\mathrm{e}^{-k\lambda s}\mathrm{d}s \\
&= \sum_{l=0}^{k-1}\int_0^\infty \binom{k-1}{l}(-1)^l \mathrm{e}^{-\lambda l(t-s)}k\lambda\mathrm{e}^{-k\lambda s}\mathrm{d}s \\
&= \left(1-\mathrm{e}^{-\lambda t}\right)^k,
\end{aligned}$$

结合上式, 根据数学归纳法知命题成立.

在命题5.8的基础上, 在 $X(0)=1$ 的条件下, 考虑转移矩阵 $\boldsymbol{P}(t)$, 则有

$$\begin{aligned}
p_{1j}(t) &= \mathbb{P}(X(t)=j|X(0)=1) \\
&= \mathbb{P}(S_{j-1} < t \le S_j|X(0)=1) \\
&= \left(1-\mathrm{e}^{-\lambda}\right)^{k-1} - \left(1-\mathrm{e}^{-\lambda}\right)^k \\
&= \left(1-\mathrm{e}^{-\lambda}\right)^{k-1} \cdot \mathrm{e}^{-\lambda t},
\end{aligned}$$

若记 $p=\mathrm{e}^{-\lambda t}$, $q=1-p$, 则 $p_{1j}(t) \sim \mathcal{G}(p)$. 以下考虑 $X(0)=i$ 的情况, 对于 $1 \le i \le k$, 设 $Y_k(t)$ 为第 k 个个体 t 时刻的后代数, 则 $Y_k(t) \sim \mathcal{G}(\mathrm{e}^{-\lambda t})$, 此时 t 时刻生物的总数

$$X(t) = \sum_{k=1}^{i} Y_k(t)$$

则根据几何分布与 Pascal 分布的关系得

$$p_{ij}(t) = \mathbb{P}(X(t)=j|X(0)=i) = \binom{j-1}{i-1}(1-\mathrm{e}^{-\lambda t})^{j-i}(\mathrm{e}^{-\lambda t})^i.$$

5.4.4 一般生灭过程

以下考虑一般的生灭过程. 设 $\{\lambda_i, i \geq 0\}$ 和 $\{\mu_i, i \geq 0\}$ 是非负数列, $\lambda_i + \mu_i > 0$, $\{X(t)\}$ 是 Markov 链, 且有转移速率矩阵

$$\boldsymbol{Q} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

设 $\{X_n\}$ 是 $\{X(t)\}$ 的嵌入链, 则其转移矩阵

$$\boldsymbol{K} = \begin{bmatrix} q_0 & p_0 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots \\ 0 & q_2 & 0 & p_2 & \cdots \\ 0 & 0 & q_3 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

其中 $p_0 = 1 - q_0$, 且对于 $\boldsymbol{K}$ 的第一行, 有

$$q_0 = \begin{cases} 1, & \lambda_0 = 0, \\ 0, & \lambda_0 > 0; \end{cases}$$

而对于第 i 行, $i \geq 1$, 有 $p_i = \dfrac{\lambda_i}{\lambda_i + \mu_i}$, $q_i = \dfrac{\mu_i}{\lambda_i + \mu_i}$.

作为生灭过程的例子, 我们考虑现实生活中排队的情形, 此时包含了三个因素:

(1) 输入过程;

(2) 服务时间;

(3) 服务窗口的个数.

通常用“输入分布/服务时间/窗口个数”来表示一个排队系统.

例 5.5 设某排队系统的顾客按强度为 λ 的 Poisson 流到达, 每个顾客的服务时间服从参数为 β 的指数分布, 共有 m 个服务的窗口. $X(t)$ 表示 t 时刻系统中的顾客数 (正在服务和排队中的顾客数), 求 $\{X(t)\}$ 的转移速率矩阵 $\boldsymbol{Q}$.

解答 首先考虑进的过程. 设 $\{N(t)\}$ 是强度为 λ 的 Poisson 过程, 则

$$\begin{cases} \mathbb{P}(N(t,t+h] = 0) = 1 - \lambda h + o(h), \\ \mathbb{P}(N(t,t+h] = 1) = \lambda h + o(h), \\ \mathbb{P}(N(t,t+h] \geq 2) = o(h); \end{cases}$$

其次考虑出的过程. 设 $T \sim \mathrm{Exp}(\beta)$, 则

$$\begin{cases} \mathbb{P}(T \leq h) = 1 - \mathrm{e}^{-\beta h} = \beta h + o(h), \\ \mathbb{P}(T > h) = 1 - \beta h + o(h). \end{cases}$$

设在 $(t, t+h]$ 内, 人数从 i 变化到 j, 令 $l = \min\{m, i\}$. 此时进与出超过两个人的概率都为 $o(h)$, 从而只需考虑每次进和出不超过一个人. 计算得

$$\begin{aligned} p_{ii}(h) &= \mathbb{P}(\text{进 } 1 \text{ 出 } 1) + \mathbb{P}(\text{进 } 0 \text{ 出 } 0) + o(h) \\ &= (\lambda h + o(h))\binom{l}{1}(\beta h + o(h))(1 - \beta h + o(h))^{l-1} \\ &+ (1 - \lambda h + o(h))(1 - \beta h + o(h))^{l} + o(h) \\ &= 1 - (\lambda + l\beta)h + o(h), \end{aligned}$$

从而 $q_{ii} = -(\lambda + l\beta)$, 同样地, 还可以得到 $q_{i,i+1} = \lambda, q_{i,i-1} = l\beta$. 而对于其他情况, 有 $q_{ij} = 0$. 从而

$$\boldsymbol{Q} = \begin{bmatrix} -l\beta & l\beta & 0 & 0 & \cdots \\ \lambda & -(\lambda + l\beta) & l\beta & 0 & \cdots \\ 0 & \lambda & -(\lambda + l\beta) & l\beta & \cdots \\ 0 & 0 & \lambda & -(\lambda + l\beta) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

在这个基础上, 还可以求出嵌入链的转移矩阵 $\boldsymbol{K}$.

5.5 课后习题

问题 5.1 一个 t 时刻存活的生物在 $(t, t+h]$ 内寿终的概率是 $\lambda h + o(h)$, 计算这个生物寿命的生存函数.

问题 5.2 t 时刻有 m 个生物独立存活, 每个生物在长为 h 的时间内寿终的概率是 $\lambda h + o(h)$. 从 t 时刻开始, 用 T_m 表示最早寿终的生物的寿终时间, 求 T_m 的分布.

参考文献

[1] 何书元. 随机过程 [M]. 北京: 北京大学出版社,2008.

[2] 何书元. 概率论 [M]. 北京: 北京大学出版社,2015.

[3] 程士宏. 测度论与概率论基础 [M]. 北京: 北京大学出版社,2004.

[4] Sheldon M. Ross. *Stochastic Processes*. Wiley,1995.

[5] 王梓坤. 随机过程论 [M]. 北京: 科学出版社,1965.