目录

一、 初等概率论

1.1 随机变量

定义 1.1.1 设 Ω 是一个非空集合，$\mathcal{F}$ 是 Ω 中的子集类，如果满足：

(1) $\varnothing, \Omega \in \mathcal{F}$;

(2) (关于取补封闭)：若 $A \in \mathcal{F}$，则 $\overline{A} \in \mathcal{F}$;

(3) (关于可列并封闭)：若 $A_1, A_2, \cdots \in \mathcal{F}$，则 $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$,

则称 $\mathcal{F}$ 是 Ω 上 **σ-代数**，称 $(\Omega, \mathcal{F})$ 为**可测空间**.

定义 1.1.2 设 $P: \mathcal{F} \to [0,1]$，如果满足：

(1) $P(\Omega) = 1$;

(2) (可列可加性)：若 $A_1, A_2, \cdots \in \mathcal{F}$ 且两两互斥，则 $P\Big(\bigcup\limits_{i=1}^{\infty} A_i\Big) = \sum\limits_{i=1}^{\infty} P(A_i)$,

则称 P 是 $(\Omega, \mathcal{F})$ 上的**概率测度**，此时称 $(\Omega, \mathcal{F}, P)$ 为**概率空间**.

例 1.1.1 [古典概率模型]

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}, \quad N < \infty; \qquad \mathcal{F} = 2^{\Omega}.$$

定义概率测度

$$\mathrm{P}(A) = \frac{|A|}{N},$$

那么，$(\Omega, \mathcal{F}, \mathrm{P})$ 是一个古典概率空间.

例 1.1.2 [几何概率模型] 设 Ω 是 $\mathbb{R}^d$, $d \geq 1$ 的一个可测区域，$|\Omega| < \infty$，$\mathcal{F} = \mathcal{B}$. 定义概率测度

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|}, \quad A \in \mathcal{F}.$$

那么，$(\Omega, \mathcal{F}, \mathrm{P})$ 是一个几何概率空间.

定义 1.1.3 设 $(\Omega, \mathcal{F})$ 和 $(\Omega', \mathcal{F}')$ 是给定的两个可测空间，ξ 是 Ω 到 Ω' 上的映射. 如果对任何 $A \in \mathcal{F}'$，有 $\xi^{-1}(A) \in \mathcal{F}$，则称 ξ 为**可测映射**. 特别地，当 $(\Omega', \mathcal{F}') = (\mathbb{R}, \mathcal{B})$ 时，称 ξ 为**随机变量** (可测函数)，这里 $\mathcal{B}$ 是 $\mathbb{R}$ 上 Borel σ-代数.

1.2 数字特征

(i) 一些定理

定理 1.2.1 (单调收敛定理) 设 $\{X_n\}$ 是随机变量序列，若 X_n 非负且递增收敛于 X，则

$$\lim_{n\to\infty} E(X_n) = E(X).$$

定理 1.2.2 (Fatou 引理) 设 $\{X_n\}$ 是随机变量序列，若 X_n 非负，那么

$$E\big(\liminf_{n\to\infty} X_n\big) = \liminf_{n\to\infty} E(X_n).$$

定理 1.2.3 (Lebesgue 控制收敛定理) 设 $\{X_n\}$ 是随机变量序列, 若 $\lim\limits_{n\to\infty} X_n = X$, 且存在非负随机变量 η, 使得 $E(\eta) < \infty$, 且对所有 n, $|X_n| \le \eta$, 则有

$$\lim_{n\to\infty} E(X_n) = E(X).$$

定理 1.2.4 (Fubini 定理) 设对 $a \le t \le b$, $X(t)$ 是随机变量, 若 $\int_a^b E|X(t)|\mathrm{d}t < \infty$ 或对所有 $a \le t \le b$, $X(t)$ 都非负, 则

$$E\Big[\int_a^b X(t)\mathrm{d}t\Big] = \int_a^b E[X(t)]\mathrm{d}t.$$

(ii) 两个不等式

定理 1.2.5 (Chebyschev 不等式) 对任意 $\varepsilon > 0$, 有

$$P(|X - E(X)| \ge \varepsilon) \le \frac{Var(X)}{\varepsilon^2}.$$

定理 1.2.6 (Markov 不等式) 假设 f 是一个单调非负函数, 那么对 $\forall x > 0$, 有

$$P(X \ge x) \le \frac{E[f(X)]}{f(x)}.$$

(iii) 特征函数

定义 1.2.1 假设 X 是一个随机变量, 具有分布函数 $F(x)$. 定义

$$\phi(t) = E(\mathrm{e}^{\mathrm{i}tX}) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}tx}\mathrm{d}F(x), \quad t \in \mathbb{R}.$$

称 $\phi(t)$ 为 X 的特征函数.

任何随机变量的特征函数总是存在的.

命题 1.2.1 特征函数具有下列性质:

(1) $\phi(0) = 1$;

(2) $|\phi(t)| \le 1$;

(3) $\phi(t)$ 在 $\mathbb{R}$ 上一致连续;

(4) $\phi(t)$ 是非负定的, 即对任意实数 $t_1, \ldots, t_m$ 和复数 $z_1, \ldots, z_m$, $\sum\limits_{i,j=1}^{n} z_i\overline{z_j}\phi(t_i - t_j) \ge 0$;

(5) 如果 $E|X|^k < \infty$, 那么 $\phi(t)$ k 次可微, 且在 0 处可进行 *Taylor* 展开:

$$\phi(t) = 1 + \sum_{m=1}^{k} \frac{i^m}{m!}\alpha_m t^m + \theta\frac{\beta_k}{(k+1)!}t^{k+1},$$

其中 α_m 为 m-*th* 半不变累积量, β_k 为 k-*th* 阶绝对矩, $|\theta| \le 1$;

(6) 如果 X, Y 相互独立, 那么 $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$.

定理 1.2.7 (唯一性定理) 随机变量的分布函数和特征函数相互唯一确定. 即随机变量

$$X \stackrel{d}{=} Y \iff \phi_X = \phi_Y.$$

特别地, 如果随机变量 X 的特征函数 $\phi(t)$ 绝对可积, 那么 X 具有密度函数

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx}\phi(t)\mathrm{d}t.$$

如果随机变量 X 的特征函数 $\phi(t)$ 形如 $\phi(t) = \sum\limits_{n=1}^{\infty} a_n \mathrm{e}^{\mathrm{i}nt}$, 并且 $a_n \geq 0$, $\sum\limits_{n=1}^{\infty} a_n = 1$, 那么 X 是离散型随机变量, 具有分布 $\mathsf{P}(X = n) = a_n$.

1.3 极限定理

(i) 依概率收敛

定义 1.3.1 假设 $(\Omega, \mathcal{F}, P)$ 是概率空间, $X, X_n, n \geq 1$ 是一列随机变量. 如果对任意 $\varepsilon > 0$,

$$P(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) \to 0,$$

则称 X_n 依概率收敛到 X, 记作 $X_n \xrightarrow{P} X$.

命题 1.3.1 依概率收敛具有以下性质:

(1) 若 $X_n \xrightarrow{P} X$, $X_n \xrightarrow{P} Y$, 那么 $P(X = Y) = 1$.

(2) 若 $X_n \xrightarrow{P} X$, $X_n - Y_n \xrightarrow{P} 0$, 那么 $Y_n \xrightarrow{P} X$.

(3) 若 $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, 那么 $X_n \pm Y_n \xrightarrow{P} X \pm Y$, $X_nY_n \xrightarrow{P} XY$, $X_n/Y_n \xrightarrow{P} X/Y$.

(4) 若 $X_n \xrightarrow{P} X$, f 是连续函数, 那么 $f(X_n) \xrightarrow{P} f(X)$.

下面是依概率收敛的三个重要的大数律.

定理 1.3.1 (Bernoulli) 设 $S_n \sim B(n,p), n \geq 1, 0 < p < 1$, 则 $\dfrac{S_n}{n} \xrightarrow{P} p$.

定理 1.3.2 (Chebyschev) 设 $\xi_n, n \geq 1$ 是一列随机变量, $E(\xi_n^2) < \infty$, 令 $S_n = \sum\limits_{k=1}^{n} \xi_k$, 若 $\dfrac{Var(S_n)}{n^2} \to 0$, 则 $\dfrac{S_n - E(S_n)}{n} \xrightarrow{P} 0$.

定理 1.3.3 (Khinchine) 设 $\xi_n, n \geq 1$ 是一列独立同分布随机变量, $E(\xi_1) = \mu$, 令 $S_n = \sum\limits_{k=1}^{n} \xi_k$, 则 $\dfrac{S_n}{n} \xrightarrow{P} \mu$.

(ii) 依分布收敛

定义 1.3.2 假设 $F(x), F_n(x), n \geq 1$ 是一列分布函数, 如果对每一个 F 的连续点 x, 都有 $F_n(x) \to F(x)$, 那么称 F_n 弱收敛到 F, 记作 $F_n \xrightarrow{w} F$.

定义 1.3.3 假设随机变量 X, X_n 分别定义在概率空间 $(\Omega, \mathcal{F}, P)$, $(\Omega_n, \mathcal{F}_n, P_n)$ 上, $n \geq 1$. 如果相应的分布函数弱收敛, 那么称 X_n 依分布收敛到 X, 记作 $X_n \xrightarrow{d} X$.

依分布收敛最经典的判别法如下定理.

定理 1.3.4 (Lévy 连续性定理) 假设 X, X_n, $n \geq 1$ 是一列随机变量, 特征函数分别为 $\phi_X(t)$, $\phi_n(t)$,

(1) $X_n \xrightarrow{d} X$ 当且仅当相应的特征函数收敛, 即

$$\lim_{n\to\infty} \phi_n(t) = \phi_X(t),\ t \in \mathbb{R}.$$

(2) 如果存在一个函数 $\phi(t)$, 使得

$$\lim_{n\to\infty} \phi_n(t) = \phi(t),\ t \in \mathbb{R}.$$

并且 $\phi(t)$ 在 *0* 点处连续, 那么存在一个随机变量 X 使得 $\phi_X = \phi$, 并且 $X_n \xrightarrow{d} X$.

命题 1.3.2 依分布收敛具有以下性质:

(1) 若 $X_n \xrightarrow{P} X$, 那么 $X_n \xrightarrow{d} X$.

(2) 如果 $X_n \xrightarrow{d} c$, 那么 $X_n \xrightarrow{P} c$.

(3) 如果 $X_n - Y_n \xrightarrow{P} 0$, 并且 $X_n \xrightarrow{d} X$, 那么 $Y_n \xrightarrow{d} X$.

(4) 如果 $X_n \xrightarrow{d} X$, f 是连续函数, 那么 $f(X_n) \xrightarrow{d} f(X)$.

(5) 如果 $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$, 那么 $X_nY_n \xrightarrow{d} cX$.

下面是依分布收敛的两个重要的极限定理.

定理 1.3.5 (De Moivre-Laplace) 假设 $S_n \sim B(n,p), n \geq 1, 0 < p < 1$, 那么

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1).$$

定理 1.3.6 (Lévy-Feller) 设 $\xi_n, n \geq 1$ 是一列独立同分布随机变量, $E(\xi_1) = \mu$, $Var(\xi_1) = \sigma^2 < \infty$, 那么

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1).$$

(iii) 几乎处处收敛

定义 1.3.4 假设 $X, X_n, n \geq 1$ 是定义在概率空间 $(\Omega, \mathcal{F}, P)$ 上的一列随机变量, 如果存在 $\Omega_0 \in \mathcal{F}$, 使得 $P(\Omega_0) = 0$, 并且对任意 $\omega \in \Omega\backslash\Omega_0$,

$$X_n(\omega) \longrightarrow X(\omega),$$

那么称 X_n 几乎处处收敛到 X, 记作 $X_n \longrightarrow X$ *a.s.*.

命题 1.3.3 几乎处处收敛具有以下性质:

(1) 如果 $X_n \longrightarrow X$ *a.s.*, 那么 $X_n \xrightarrow{P} X$.

(2) 如果对任意 $\varepsilon > 0$, $\sum_{n=1}^{\infty} P(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) < \infty$, 那么 $X_n \longrightarrow X$ *a.s.*.

下面是几乎处处收敛的两个重要的大数律.

定理 1.3.7 (Borel) 设 ξ_n, $n \geq 1$ 是一列独立同分布于 $B(1,p)$ 的随机变量, 定义 $S_n = \sum_{k=1}^{n} \xi_k$, 那么 $\frac{S_n}{n} \longrightarrow p$ *a.s.*.

定理 1.3.8 (Kolmogorov) 令 $\xi_n, n \geq 1$ 是一列独立同分布随机变量, $E(\xi_1) = \mu$. 定义 $S_n = \sum_{k=1}^{n} \xi_k$, 那么 $\frac{S_n}{n} \longrightarrow \mu$ *a.s.*.

(iv) 均方收敛

定义 1.3.5 假设 $X, X_n, n \geq 1$ 是一列定义在概率空间 $(\Omega, \mathcal{F}, P)$ 上的随机变量, 具有有限 r 阶矩 $(r > 0)$. 如果

$$E|X_n - X|^r \to 0,$$

则称 X_n r 阶均方收敛到 X, 记作 $X_n \xrightarrow{L^r} X$.

命题 1.3.4 如果 $X_n \xrightarrow{L^r} X$, 那么 $X_n \xrightarrow{P} X$.

1.4 条件概率和条件期望

定义 1.4.1 设 $(\Omega, \mathcal{F}, P)$ 是一个概率空间, $A \in \mathcal{F}$, 且 $P(A) > 0$. 当 A 发生时事件 B 发生的条件概率定义为

$$P(B \mid A) = \frac{P(AB)}{P(A)}.$$

此时 $(\Omega, \mathcal{F}, P(\cdot \mid A))$ 也是一个概率空间.

条件概率具有乘法公式、链式法则、全概率公式和贝叶斯公式. 这些都是初等概率论的重要内容.

定义 1.4.2 若 X 是 $(\Omega, \mathcal{F}, P)$ 上可积的随机变量, 定义 X 在 A 发生时的条件期望为 X 关于条件概率 $P(\cdot \mid A)$ 的期望, 记为 $E(X \mid A)$.

注意到 $\mathrm{E}(X) = \int_\Omega X(\omega)\mathrm{dP}(\omega)$ 是 X 在 Ω 上的加权平均, 因此

$$\mathrm{E}(X \mid A) = \int_\Omega X(\omega)\mathrm{dP}(\omega \mid A) = \frac{1}{\mathrm{P}(A)} \int_A X(\omega)\mathrm{dP}(\omega)$$

是 X 在 A 上的加权平均.

命题 1.4.1 *条件期望具有以下性质:*

(1) $E(X\mathbf{1}_A) = E(X \mid A)P(A)$;

(2) 若 $A_1, A_2, \ldots$ *是* Ω *的划分, 且* $P(A_i) > 0$, *则以下全期望公式成立:*

$$E(X) = \sum_i E(X \mid A_i)P(A_i).$$

例 1.4.1 独立重复地抛一枚硬币直至有 k 次相继出现正面, 计算所抛次数的均值. (设每次出现正面的概率为 p).

解 令 N_k 表示所抛次数, T 表示首次出现反面时所抛的次数, 则

$$\mathsf{E}(N_k \mid T=i)=\begin{cases} i+\mathsf{E}(N_k), & i\le k,\\ k, & i>k.\end{cases}$$

于是由全期望公式

$$\mathsf{E}(N_k)=\sum_{i\ge 1}\mathsf{E}(N_k\mid T=i)\mathsf{P}(T=i)=\sum_{k=1}^{k}[i+\mathsf{E}(N_k)]p^{i-1}(1-p)+kp^k.$$

解得 $\mathsf{E}(N_k)=\dfrac{1}{p}+\dfrac{1}{p^2}+\cdots+\dfrac{1}{p^k}$.

定义 1.4.3 设 Y 是一取值为 $y_1,y_2,\ldots$ 的离散型随机变量, 定义给定 Y 的条件下 X 的条件期望 $E(X\mid Y)$ 为 Y 的一个函数, 它在 $Y=y$ 时的取值是 $E(X\mid Y=y)$, 即

$$h(Y)=E(X\mid Y)=\sum_i E(X\mid Y=y_i)\mathbf{1}_{\{Y=y_i\}}.$$

它是关于 $\sigma(Y)$ 可测的随机变量.

命题 1.4.2 全期望公式: $E(X)=E[E(X\mid Y)]$.

例 1.4.2 设 $Y_1,Y_2,\ldots$ 独立同分布, 均值为 μ, 方差为 σ^2. 设 N 是取非负整数值的随机变量, 与 $Y_1,Y_2,\ldots$ 独立, 且 $\mathsf{E}(N^2)<\infty$. 令 $X=Y_1+\cdots+Y_N$, 计算 $\mathsf{E}(X)$ 和 $\mathsf{Var}(X)$.

解 对非负整数 n, 由于 N 与 $Y_1,Y_2,\ldots,Y_n$ 独立, 所以

$$\begin{aligned}\mathsf{E}(X|N=n) &= \mathsf{E}(Y_1+\cdots+Y_N|N=n)\\ &= \mathsf{E}(Y_1+\cdots+Y_n|N=n)=\mathsf{E}(Y_1+\cdots+Y_n)=n\mu.\end{aligned}$$

因此 $\mathsf{E}(X|N)=N\mu$, 故 $\mathsf{E}(X)=\mathsf{E}[\mathsf{E}(X|N)]=\mathsf{E}(N\mu)=\mu\mathsf{E}(N)$.

进一步,

$$\begin{aligned}\mathsf{E}(X^2|N=n) &= \mathsf{E}[(Y_1+\cdots+Y_N)^2|N=n]\\ &= \mathsf{E}[(Y_1+\cdots+Y_n)^2|N=n]\\ &= \mathsf{E}[(Y_1+\cdots+Y_n)^2]\\ &= \mathsf{Var}(Y_1+\cdots+Y_n)+[\mathsf{E}(Y_1+\cdots+Y_n)]^2=n\sigma^2+n^2\mu^2.\end{aligned}$$

因此 $\mathsf{E}(X^2|N)=N\sigma^2+N^2\mu^2$, 故 $\mathsf{E}(X^2)=\mathsf{E}[\mathsf{E}(X^2|N)]=\sigma^2\mathsf{E}(N)+\mu^2\mathsf{E}(N^2)$. 所以

$$\mathsf{Var}(X)=\mathsf{E}(X^2)-[\mathsf{E}(X)]^2=\sigma^2\mathsf{E}(N)+\mu^2\mathsf{Var}(N).$$

例 1.4.3 设 $\xi\sim U(0,1)$, 在 $\xi=x$ 的条件下 X 服从 $B(n,x)$, 求 X 的分布律.

解 (解法一) 对整数 $0\le k\le n$, 由全概率公式

$$\mathsf{P}(X=k)=\int_{-\infty}^{\infty}\mathsf{P}(X=k|\xi=x)f(x)\mathrm{d}x=\int_0^1\binom{n}{k}x^k(1-x)^{n-k}\mathrm{d}x=\binom{n}{k}\frac{k!(n-k)!}{(n+1)!}=\frac{1}{n+1}.$$

(解法二) 令 $\phi(t)=\mathsf{E}(\mathrm{e}^{\mathrm{i}tX})$, 注意到 $\mathsf{E}(\mathrm{e}^{\mathrm{i}tX}\mid\xi=x)=(1+x(\mathrm{e}^{\mathrm{i}t}-1))^n$. 于是

$$\phi(t)=\mathsf{E}[1+\xi(\mathrm{e}^{\mathrm{i}t}-1)]^n=\int_0^1(1+x(\mathrm{e}^{\mathrm{i}t}-1))^n\mathrm{d}x=\frac{1}{n+1}(1+\mathrm{e}^{\mathrm{i}t}+\mathrm{e}^{2\mathrm{i}t}+\cdots+\mathrm{e}^{n\mathrm{i}t}).$$

所以 $\mathsf{P}(X=k)=\dfrac{1}{n+1}$, $0\le k\le n$.

二、 随机过程基本概念

2.1 随机过程的概念

随机过程是一族随机变量, 用于描述与时间相关的随机现象.

定义 2.1.1 假设 $(\Omega,\mathcal{F},P)$ 是概率空间, $(E,\mathcal{E})$ 是可测空间, T 是指标集. 如果对任何 $t\in T$, X_t 是 $(\Omega,\mathcal{F})$ 到 $(E,\mathcal{E})$ 上的可测映射, 则称 $\boldsymbol{X}=(X_t,t\in T)$ 是 $(\Omega,\mathcal{F},P)$ 上的以 $(E,\mathcal{E})$ 为状态空间的**随机过程**.

随机过程 $\boldsymbol{X}=(X_t,t\in T)$ 的随机性通过 ω 体现出来, 随机过程用映射来表示就是 $X_t(\omega):T\times\Omega\to E$. 固定 t, $X_t(\cdot)$ 是随机变量; 固定 ω, 称 $X_\cdot(\omega)$ 是 T 的函数, 称为随机过程的**样本曲线**或**样本轨迹**.

对过程的一次具体观察结果就是一条样本轨迹.

2.2 随机过程的分布刻画

定义 2.2.1 (1-维分布) 给定 $t\in T$, X_t 是定义在概率空间 $(\Omega,\mathcal{F},P)$ 的随机变量. 对每个 E 的子集 A, 定义

$$\nu_t(A)=P(\omega\in\Omega:X_t(\omega)\in A),$$

称 ν_t 为随机过程 $\boldsymbol{X}$ 在 t 时刻的 *1-维分布*.

特别地, 对 $E=\mathbb{R},T=(-\infty,+\infty)$ 的情况, 我们可以给出 X_t 的分布函数

$$F_t(x)=P(X_t\le x).$$

定义 2.2.2 (k-维分布) 给定 $t_1,\ldots,t_k\in T$, $(X_{t_1},\ldots,X_{t_k})$ 是定义在概率空间 $(\Omega,\mathcal{F},P)$ 的 k-维随机向量. 对 E 的子集 $A_1,\ldots,A_k$, 定义

$$\nu_{t_1,\ldots,t_k}(A_1,\ldots,A_k)=P(\omega\in\Omega:X_{t_1}(\omega)\in A_1,\ldots,X_{t_k}(\omega)\in A_k),$$

称 $\nu_{t_1,\ldots,t_k}$ 为随机过程 $\boldsymbol{X}$ 在 $t_1,\ldots,t_k$ 时刻的 *k-维分布*.

特别地, 对 $E=\mathbb{R},T=(-\infty,+\infty)$ 的情况, 我们可以给出 $(X_{t_1},\ldots,X_{t_k})$ 的联合分布函数

$$F_{t_1,\ldots,t_k}(x_1,\ldots,x_k)=P(X_{t_1}\le x_1,\ldots,X_{t_k}\le x_k).$$

命题 2.2.1 随机过程 $\boldsymbol{X}$ 的有限维分布函数族 $\{F_{t_1,\ldots,t_k}(x_1,\ldots,x_k),k=1,2,\ldots,t_i\in T\}$ 满足:

(1) 横向相容: 对 $(1,2,\ldots,k)$ 的任何置换 τ, $F_{t_{\tau(1)},\ldots,t_{\tau(k)}}(x_{\tau(1)},\ldots,x_{\tau(k)})=F_{t_1,\ldots,t_k}(x_1,\ldots,x_k)$.

(2) 纵向相容: $\lim_{x_{k+1}\to\infty}F_{t_1,\ldots,t_k,t_{k+1}}(x_1,\ldots,x_k,x_{k+1})=F_{t_1,\ldots,t_k}(x_1,\ldots,x_k)$.

例 2.2.1 有 10 把步枪, 其中 2 把已校正, 命中率为 p_1, 其余 8 把未校正, 命中率为 p_2, 这里 $p_1 > p_2$. 某人任取一把开始打靶, 令 X_n 为第 n 次命中的次数, 即

$$X_n = \begin{cases} 1, & \text{第 } n \text{ 次命中}, \\ 0, & \text{第 } n \text{ 次未命中}. \end{cases}$$

(1) 对 $n \neq m$, 求 (X_n, X_m) 的联合分布律和边缘分布律.

(2) 以 S_n 表示前 n 次命中的次数, 求 S_n 的分布律.

(3) 若 $p_1 = 1, p_2 = 0$, 写出所有样本函数和 S_n 的分布律, 并说明此时对 $n \neq m$, X_n 和 X_m 独立吗?

解 (1) 令事件 A 表示取到已校正的枪, 由全概率公式得

$$\begin{aligned} p_{11} &= \mathsf{P}(X_n = 1, X_m = 1) \\ &= \mathsf{P}(X_n = 1, X_m = 1 \mid A)\mathsf{P}(A) + \mathsf{P}(X_n = 1, X_m = 1 \mid \overline{A})\mathsf{P}(\overline{A}) \\ &= 0.2p_1^2 + 0.8p_2^2. \end{aligned}$$

同理可求

$$p_{01} = p_{10} = 0.2p_1(1-p_1) + 0.8p_2(1-p_2), \quad p_{00} = 0.2(1-_1)^2 + 0.8(1-p_2)^2.$$

进而, 边缘分布

$$\mathsf{P}(X_n = 0) = 0.2(1-p_1) + 0.8(1-p_2), \quad \mathsf{P}(X_n = 1) = 0.2p_1 + 0.8p_2.$$

(2) 同样利用全概率公式

$$\begin{aligned} \mathsf{P}(S_n = k) &= \mathsf{P}(S_n = k \mid A)\mathsf{P}(A) + \mathsf{P}(S_n = k \mid \overline{A})\mathsf{P}(\overline{A}) \\ &= 0.2\mathrm{C}_n^k p_1^k (1-p_1)^{n-k} + 0.8\mathrm{C}_n^k p_2^k (1-p_2)^{n-k}, \quad k = 0, 1, \ldots, n. \end{aligned}$$

(3) 若 A 发生, 则百发百中; 若 A 不发生, 则永不命中. 因此 $\{X_n\}$ 只有两条样本函数 $(1,1,1,\ldots)$ 和 $(0,0,0,\ldots)$; $\{S_n\}$ 只有两条样本函数 $(1,2,3,\ldots)$ 和 $(0,0,0,\ldots)$. 于是 S_n 的分布律为

$$\mathsf{P}(S_n = 0) = \mathsf{P}(\overline{A}) = 0.8, \quad \mathsf{P}(S_n = n) = 0.2.$$

而 X_n 与 X_m 不独立, 因为 $\mathsf{P}(X_n = 1 \mid X_m = 1) = 1 \neq 0.2 = \mathsf{P}(X_n = 1)$.

例 2.2.2 [简单随机游走] 甲乙两人游戏, 第 i 次甲赢的钱数为 X_i 元. 设 $X_1, \ldots, X_n, \ldots$ 独立同分布, 满足

$$\mathsf{P}(X_i = 1) = p, \quad \mathsf{P}(X_i = -1) = q = 1 - p.$$

记前 n 次甲赢的总钱数为 S_n, 计算

(1) S_n 的分布律;

(2) $\mathsf{P}(S_1 = 1, S_3 = 1, S_8 = 4)$.

(3) 若 $p = 0.36$, 游戏一直到甲恰好赢 50 次为止, 问游戏需进行 100 次以上的概率约为多少?

解 (1) 以 V_n 记前 n 次甲赢的总次数, 则 $V_n \sim B(n, p)$, 且 $S_n = V_n - (n - V_n) = 2V_n - n$. 所以

$$\mathsf{P}(S_n = k) = \mathsf{P}\left(V_n = \frac{k+n}{2}\right) = \binom{n}{\frac{k+n}{2}} p^{\frac{k+n}{2}} q^{\frac{n-k}{2}}, \quad k \text{ 与 } n \text{ 奇偶性相同, 且} -n \leq k \leq n.$$

(2) 由题意得

$$\begin{aligned}\mathsf{P}(S_1=1,S_3=1,S_8=4) &= \mathsf{P}(S_1=1,S_3-S_1=0,S_8-S_3=3)\\ &= \mathsf{P}(S_1=1)\mathsf{P}(S_3-S_1=0)\mathsf{P}(S_8-S_3=3)\\ &= p(2pq)(\mathrm{C}_5^4p^4q)=10p^6q^2.\end{aligned}$$

(3) 以 W_{50} 记甲恰好赢 50 次时游戏进行的次数, 则 $\{W_{50}>100\}=\{V_{100}<50\}$. 由中心极限定理, V_{100} 近似服从 $N(100p,100pq)$. 因此

$$\mathsf{P}(W_{50}>100)=\mathsf{P}(V_{100}<50)=\mathsf{P}(V_{100}\le 49)\approx\Phi\left(\frac{49-100p}{10\sqrt{pq}}\right)=\Phi(2.71)=0.9966.$$

由此可见概率非常大.

一般情况下, 很难写出所有有限维分布. 下面介绍两种特殊情况:

(1) 利用条件概率的链式法则计算 k-维分布

给定 $t_1,\ldots,t_k\in T$, 我们有

$$\begin{aligned}\nu_{t_1,\ldots,t_k}(A_1,\ldots,A_k) &= \mathsf{P}(\omega\in\Omega: X_{t_1}(\omega)\in A_1,\ldots,X_{t_k}(\omega)\in A_k)\\ &= \mathsf{P}(X_{t_1}\in A_1)\mathsf{P}(X_{t_2}\in A_2|X_{t_1}\in A_1)\cdots\\ &\quad\times\mathsf{P}(X_{t_k}\in A_k|X_{t_1}\in A_1,\ldots,X_{t_{k-1}}\in A_{k-1})\end{aligned}$$

由此, 如果知道随机过程 $\boldsymbol{X}$ 在各个不同时刻的条件概率, 就可以写出整个过程的分布.

(2) 考虑增量的分布

假设 $T\subseteq(-\infty,+\infty), E=\mathbb{R}$. 给定 $t_1,\ldots,t_k\in T$, 不妨设 $t_1<t_2<\cdots<t_k$, 并考虑增量

$$X_{t_1},X_{t_2}-X_{t_1},X_{t_3}-X_{t_2},\ldots,X_{t_k}-X_{t_{k-1}},$$

如果知道上述增量的分布, 那么 $X_{t_1},\ldots,X_{t_k}$ 的分布便可以通过线性变换得到. 特别地, 当所有增量互相独立时, 只要知道任意两个时刻的增量的分布, 就可以完全确定整个过程的分布.

综上所述, 为了揭示随机过程的分布规律, 需要知道随机现象如何随时间推移而变化的内在机制.

2.3 随机过程的数字特征

(i) 数字特征

假设 $\boldsymbol{X}=(X_t,t\in T)$ 是随机过程, $E=\mathbb{R}$.

如果对 $\forall t\in T$, $\mathsf{E}|X_t|<\infty$, 记 $\mu_t=\mathsf{E}(X_t)$, 称 $\mu_t,t\in T$ 为**均值函数**.

如果对 $\forall t\in T$, $\mathsf{E}(X_t^2)<\infty$, 记 $\sigma_t^2=\mathsf{Var}(X_t)$, 称 $\sigma_t^2,t\in T$ 为**方差函数**, $r(s,t)=\mathsf{Cov}(X_s,X_t)$ 为**协方差函数**, $\rho(s,t)=\mathsf{E}(X_sX_t)$ 为 **(自) 相关函数**.

显然我们有 $r(s,t)=\rho(s,t)-\mu_s\mu_t$.

例 2.3.1 [余弦波过程] 求随机相位余弦波 $X(t)=a\cos(\omega t+\Theta)$, $-\infty<t<+\infty$, $\Theta\sim U(0,2\pi)$ 的均值函数、方差函数和自相关函数.

$\boxed{\text{解}}$ 由定义,

$$\mu_t=\mathsf{E}(X_t)=\int_0^{2\pi}a\cos(\omega t+\theta)\frac{1}{2\pi}\mathrm{d}\theta=0.$$

$$\rho(s,t)=\mathsf{E}(X_sX_t)=a^2\int_0^{2\pi}\cos(\omega s+\theta)\cos(\omega t+\theta)\frac{1}{2\pi}\mathrm{d}\theta=\frac{a^2}{2}\cos\omega(t-s).$$

因此

$$\sigma_t^2=\mathsf{Var}(X_t)=\mathsf{E}(X_t^2)-[\mathsf{E}(X_t)]^2=\rho(t,t)-\mu_t^2=\frac{a^2}{2}.$$

(ii) 两个随机过程之间的关系

设 $\boldsymbol{X}=(X_t,t\in T_1)$ 和 $\boldsymbol{Y}=(Y_s,s\in T_2)$ 都是概率空间 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机过程. 如果对 $\forall m,n\geq 1$, $t_1,\ldots,t_m\in T_1$ 和 $s_1,\ldots,s_n\in T_2$ 有 $(X_{t_1},\ldots,X_{t_m})$ 和 $(Y_{s_1},\ldots,Y_{s_n})$ 相互独立, 则称 $\boldsymbol{X}=(X_t,t\in T_1)$ 和 $\boldsymbol{Y}=(Y_s,s\in T_2)$ 相互独立.

对两个随机过程 $\boldsymbol{X}=(X_t,t\in T_1)$ 和 $\boldsymbol{Y}=(Y_s,s\in T_2)$, 称 $\rho_{XY}(t,s)=\mathsf{E}(X_tY_s)$, $t\in T_1$, $s\in T_2$ 为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的**互相关函数**, $r_{XY}(t,s)=\mathsf{Cov}(X_t,Y_s)=\rho_{XY}(t,s)-\mu_X(t)\mu_Y(s)$, $t\in T_1$, $s\in T_2$ 为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的**互协方差函数**.

如果对 $\forall t\in T_1$, $s\in T_2$, 恒有 $r_{XY}(t,s)=0$, 则称随机过程 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 是不相关的.

2.4 几类常见的随机过程

定义 2.4.1 (弱平稳过程) *假设 $\boldsymbol{X}=(X_t,t\in T)$ 是实值随机过程, T 为实数集, 并且对 $\forall t\in T$, 二阶矩 $E(X_t^2)<\infty$. 如果*

(1) 均值函数为常数, 即 $\mu_t=\mu$,

(2) 相关函数 $\rho(s,t)$ 仅依赖于时间间隔长度 $t-s$, 而与时刻 s 和 t 无关,

则称 $\boldsymbol{X}=(X_t,t\in T)$ 是弱平稳过程.

定义 2.4.2 (强平稳过程) *假设 $\boldsymbol{X}=(X_t,t\in T)$ 是状态空间为 E 的随机过程. 如果对 $\forall t,t_1,\ldots,t_k\in T$, 随机向量 $(X_{t_1},X_{t_2},\ldots,X_{t_k})$ 和 $(X_{t_1+t},X_{t_2+t},\ldots,X_{t_k+t})$ 具有相同分布, 即*

$$\nu_{t_1,\ldots,t_k}=\nu_{t_1+t,\ldots,t_k+t},$$

则称 $\boldsymbol{X}=(X_t,t\in T)$ 为强平稳过程.

一般情况下, 强平稳过程并不一定是弱平稳过程, 需要二阶矩存在的条件.

例 2.4.1 设 $X(t)=A\cos t+B\sin t$, $t\in\mathbb{R}$, A 与 B 独立, 满足 $\mathsf{E}(A)=\mathsf{E}(B)=0$, $\mathsf{Var}(A)=\mathsf{Var}(B)=1$.

(1) 计算均值函数和自相关函数, 问 $\{X(t)\}$ 是弱平稳过程吗?

(2) 若 $\mathsf{P}(A=\pm1)=\mathsf{P}(B=\pm1)=0.5$, 求 $X(0)$ 和 $X\left(\frac{\pi}{4}\right)$ 的分布律, 问 $\{X(t)\}$ 是强平稳过程吗?

(3) 若 A, B 均服从 $N(0,1)$, 求 $X(0)$, $X\left(\frac{\pi}{4}\right)$ 和 $X(0)+X\left(\frac{\pi}{4}\right)$ 的分布, 问 $\{X(t)\}$ 是强平稳过程吗?

解 (1) 由定义

$$\mu_X(t)=\mathsf{E}[A\cos t+B\sin t]=\mathsf{E}(A)\cos t+\mathsf{E}(B)\sin t=0$$

为常数, 而

$$\rho_X(t_1,t_2)=\mathsf{E}[(A\cos t_1+B\sin t_1)(A\cos t_2+B\sin t_2)]=\cos(t_2-t_1)$$

只与 t_2-t_1 有关, 因此 $\{X(t)\}$ 是弱平稳过程.

(2) 注意到 $X(0)=A$, $X\left(\frac{\pi}{4}\right)=\frac{\sqrt{2}}{2}(A+B)$. 因此 $X(0)$ 的分布律为

$$\mathsf{P}(X(0)=\pm 1)=\mathsf{P}(A=\pm 1)=0.5.$$

$X\left(\frac{\pi}{4}\right)$ 的分布律为

$$\mathsf{P}\left(X\left(\frac{\pi}{4}\right)=\sqrt{2}\right)=\mathsf{P}(A=1,B=1)=\frac{1}{4},\quad \mathsf{P}\left(X\left(\frac{\pi}{4}\right)=-\sqrt{2}\right)=\mathsf{P}(A=-1,B=-1)=\frac{1}{4},$$

$$\mathsf{P}\left(X\left(\frac{\pi}{4}\right)=0\right)=\mathsf{P}(A=1,B=-1)+\mathsf{P}(A=-1,B=1)=\frac{1}{2}.$$

同样由 (1) 知 $\{X(t)\}$ 是弱平稳过程但不是强平稳过程.

(3) 由正态分布的线性变换不变性可知, $(X_{t_1},X_{t_2},\ldots,X_{t_n})$ 服从 n 维正态分布, 所以 $\{X(t)\}$ 是正态过程, 因此 $\{X(t)\}$ 是弱平稳过程也是强平稳过程. 下面进行计算

$$X(0)=A\sim N(0,1),\quad X\left(\frac{\pi}{4}\right)=\frac{\sqrt{2}}{2}(A+B)\sim N(0,1),\quad X(0)+X\left(\frac{\pi}{4}\right)=\frac{2+\sqrt{2}}{2}A+\frac{\sqrt{2}}{2}B\sim N(0,2+\sqrt{2}).$$

定义 2.4.3 (增量平稳过程) 假设 $\boldsymbol{X}=(X_t,t\in T)$ 是实值随机过程, T 是实数集. 如果对 $\forall s<t$, X_t-X_s 的分布仅与时间间隔长度 $t-s$ 有关, 而与 s 和 t 无关, 那么称 $\boldsymbol{X}=(X_t,t\in T)$ 是增量平稳过程.

定义 2.4.4 (增量独立过程) 假设 $\boldsymbol{X}=(X_t,t\in T)$ 是实值随机过程, T 为实数集. 如果对 $\forall t_1<t_2<\cdots<t_k$, 随机变量

$$X_{t_1},X_{t_2}-X_{t_1},\ldots,X_{t_k}-X_{t_{k-1}}.$$

相互独立, 那么称 $\boldsymbol{X}=(X_t,t\in T)$ 是增量独立过程.

命题 2.4.1 设 $\boldsymbol{X}=(X_t,t\in T)$ 是增量独立过程, $X_0=0$, 则
(1) 有限维分布由所有增量 X_t-X_s 分布确定;
(2) $\mathrm{Cov}(X_t,X_s)=\mathrm{Var}(X_{\min\{s,t\}})$.

例 2.4.2 设 $X_1,\ldots,X_n,\ldots$ 独立同分布, $S_0=0$. 对 $n\geq 1$, 定义 $S_n=\sum_{i=1}^{n}X_i$.
(1) 证明: $\{S_n:n=0,1,\ldots\}$ 是平稳独立增量过程.
(2) 若 $\mathsf{E}(X_1)=\mu$, $\mathsf{Var}(X_1)=\sigma^2$. 计算 $\{S_n\}$ 的均值函数和自相关函数.

解 (1) 证明: 对任意 $n>m>0$, $S_n-S_m=\sum_{i=m+1}^{n}X_i$, 因为 $X_1,X_2,\ldots,X_n$ 独立同分布, 因此

$$(X_{m+1},X_{m+2},\ldots,X_n)\overset{d}{=\!=}(X_1,X_2,\ldots,X_{n-m}).$$

所以 S_n-S_m 与 S_{n-m} 同分布, 即 $\{S_n\}$ 是增量平稳过程.

另一方面, 对任意 $n\geq 3$, $0\leq m_1<m_2<\cdots<m_n$, 由于 $X_1,X_2,\ldots,X_{m_n}$ 独立, 所以

$$\sum_{i=m_1+1}^{m_2}X_i,\ \sum_{i=m_2+1}^{m_3},\ldots,\ \sum_{i=m_{n-1}+1}^{m_n}X_i$$

独立, 这说明 $S_{m_2}-S_{m_1},\ldots,S_{m_n}-S_{m_{n-1}}$ 独立, 因此 $\{S_n\}$ 是增量独立过程. 综上所述, $\{S_n\}$ 是平稳独立增量过程.

(2) 由定义

$$\mu_S(n)=\mathsf{E}\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n \mathsf{E}(X_i)=n\mu.$$

对 $0\le m\le n$,

$$\mathsf{Cov}(S_m,S_n)=\mathsf{Cov}(S_m,S_m)+\mathsf{Cov}(S_m,S_n-S_m)=\mathsf{Var}(S_m)=m\sigma^2.$$

于是

$$\mathsf{E}(S_mS_n)=\mathsf{Cov}(S_mS_n)+\mathsf{E}(S_m)\mathsf{E}(S_n)=m\sigma^2+mn\mu^2.$$

因此对 $m\ge 0$, $n\ge 0$, 有 $\rho_S(m,n)=mn\mu^2+\min(m,n)\sigma^2$.

定义 2.4.5 (正态过程) 假设 $\boldsymbol{X}=(X_t,t\in T)$ 是实值随机过程, T 为实数集. 如果任意 k-维分布都是联合正态分布, 则称 $\boldsymbol{X}=(X_t,t\in T)$ 为正态过程.

命题 2.4.2 如果 $\boldsymbol{X}=(X_t,t\in T)$ 是弱平稳正态过程, 那么它一定是强平稳正态过程.

由定义知, $\boldsymbol{X}=(X_t,t\in T)$ 是正态过程当且仅当它对任意有限线性组合是正态随机变量.

定义 2.4.6 (白噪声) 假设 $\boldsymbol{X}=(X_t,t\in T)$ 是状态空间为 E 的随机过程, 如果 X_s 和 X_t 互不相关, 则称 $\boldsymbol{X}=(X_t,t\in T)$ 为白噪声.

例 2.4.3 设 $\boldsymbol{X}=(X_t,t\ge 0)$ 是正态过程, $\mu_X(t)=t$, $r_X(s,t)=st+1$, 求 X_1, X_2, X_1+X_2 的分布.

$\boxed{\text{解}}$ 计算方差

$$\sigma_X^2(t)=r_X(t,t)=t^2+1.$$

因此 $X_t\sim N(t,t^2+1)$. 即 $X_1\sim N(1,2)$, $X_2\sim N(2,5)$.

又因为 X_1+X_2 服从正态分布, 而

$$\mathsf{E}(X_1+X_2)=1+2=3,\quad \mathsf{Var}(X_1+X_2)=\mathsf{Var}(X_1)+\mathsf{Var}(X_2)+2r_X(1,2)=13.$$

所以 $X_1+X_2\sim N(3,13)$.

三、 Poisson 过程

3.1 Poisson 流

Poisson 流用于描述随机服务系统, 令 $N(0)=0$, 当 $t>0$ 时, $N(t)$ 表示 $(0,t]$ 时间内某服务系统接受顾客服务的次数. 因此 $\boldsymbol{N}=(N(t),t\ge 0)$ 是连续时间、取非负整数值的随机过程, 亦称作计数过程, 其具有下列属性:

(1) 平稳性——前往某服务设施寻求服务的顾客流是平稳的

$$N(t)-N(s)\overset{d}{=\!=}N(t-s),\quad s<t.$$

(2) 独立性——各个顾客在某时刻是否前往某服务设施寻求服务相互独立

$$N(t)-N(s) \text{ 与 } N(s) \text{ 独立}.$$

(3) 稀有性——短时间内, 最多只有一名顾客寻求服务

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

其中, 参数 λ 称为 Poisson 流的强度, 它的大小表示该服务系统的繁忙程度.

令 $N(0)=0$, 当 $t>0$ 时, $N(t)$ 表示强度为 λ 的 Poisson 流在 $(0,t]$ 时间内的顾客数. 那么 $N(t)$ 服从参数为 λt 的 Poisson 分布:

$$\mathsf{P}(N(t)=k)=\frac{(\lambda t)^k\mathrm{e}^{-\lambda t}}{k!},\quad k=0,1,2,\ldots.$$

3.2 齐次 Poisson 过程

定义 3.2.1 *假设 $\boldsymbol{N}=(N(t),t\geq 0)$ 是非负整数值随机过程, 并且具有下列性质:*

(1) 初始值为 0: $P(N(0)=0)=1$.

(2) 平稳性: 假设 $s<t$, $N(t)-N(s)$ 与 $N(t-s)$ 具有相同分布:

$$N(t)-N(s)\overset{d}{=\!=}N(t-s),\quad s<t.$$

(3) 独立性: 假设 $s<t$, $N(t)-N(s)$ 与 $N(t-s)$ 相互独立.

(4) Poisson 分布:

$$P(N(t)=k)=\frac{(\lambda t)^k e^{-\lambda t}}{k!},\quad k=0,1,2,\ldots.$$

则称 $\boldsymbol{N}=(N(t),t\geq 0)$ 是强度为 λ 的 Poisson 过程.

(i) Poisson 过程的数字特征

(1) 均值函数

$$\mu_t=\mathsf{E}[N(t)]=\lambda t.$$

于是 $\lambda=\mathsf{E}[N(1)]=\dfrac{\mathsf{E}[N(t)]}{t}$ 表示单位时间内顾客寻求服务的个数, 同时是平均强度.

(2) 方差函数

$$\mathsf{Var}(N(t))=\lambda t.$$

(3) 相关函数

假设 $s<t$, 则

$$\begin{aligned}\mathsf{E}[N(s)N(t)]&=\mathsf{E}\big[N(s)\big(N(t)-N(s)+N(s)\big)\big]\\&=\mathsf{E}[N(s)(N(t)-N(s))]+\mathsf{E}[N^2(s)]\\&=\mathsf{E}[N(s)]\mathsf{E}[N(t)-N(s)]+\mathsf{E}[N^2(s)]\\&=\mathsf{E}[N(s)]\mathsf{E}[N(t-s)]+\mathsf{E}[N^2(s)]\\&=\lambda^2st+\lambda s.\end{aligned}$$

因此 $\mathsf{E}[N(s)N(t)]=\lambda^2st+\lambda\min(s,t)$.

(4) 协方差函数

假设 $s<t$, 利用 (1) 和 (3) 的结论可得

$$\mathsf{Cov}(N(s),N(t))=\mathsf{E}[N(s)N(t)]-\mathsf{E}[N(s)]\mathsf{E}[N(t)]=\lambda s.$$

因此 $\mathsf{Cov}(N(s),N(t))=\lambda\min(s,t)$.

(ii) Poisson 过程的分布规律

(1) 1-维分布

$$N(0)=0,\quad N(t)\sim\mathcal{P}(\lambda t),\quad t>0.$$

(2) k-维分布

假设 $t_1<t_2<\cdots<t_k$, 则

$$\begin{aligned}
&\mathsf{P}(N(t_1)=m_1,N(t_2)=m_2,\ldots,N(t_k)=m_k)\\
=&\mathsf{P}(N(t_1)=m_1,N(t_2)-N(t_1)=m_2-m_1,\ldots,N(t_k)-N(t_{k-1})=m_k-m_{k-1})\\
=&\mathsf{P}(N(t_1)=m_1)\mathsf{P}(N(t_2)-N(t_1)=m_2-m_1)\cdots\mathsf{P}(N(t_k)-N(t_{k-1})=m_k-m_{k-1})\\
=&\frac{(\lambda t_1)^{m_1}\mathrm{e}^{-\lambda t_1}}{m_1!}\cdots\frac{(\lambda(t_k-t_{k-1}))^{m_k-m_{k-1}}\mathrm{e}^{-\lambda(t_k-t_{k-1})}}{(m_k-m_{k-1})!}.
\end{aligned}$$

(3) 条件分布

对 $s<t$, $m\le n$, 有

$$\begin{aligned}
\mathsf{P}(N(t)=n\mid N(s)=m)&=\mathrm{e}^{-\lambda(t-s)}\frac{[\lambda(t-s)]^{n-m}}{(n-m)!},\\
\mathsf{P}(N(s)=m\mid N(t)=n)&=\binom{n}{m}\left(\frac{s}{t}\right)^m\left(1-\frac{s}{t}\right)^{n-m}.
\end{aligned}$$

也就是说, 在 $N(t)=n$ 的条件下, $N(s)|_{N(t)=n}\sim B\Big(n,\dfrac{s}{t}\Big)$.

例 3.2.1 顾客依 Poisson 过程到达某商店, 速率为 4 人/小时. 已知商店上午 9:00 开门.

(1) 求到 9:30 时仅到 1 位顾客, 而到 11:30 时已到 5 位顾客的概率.

(2) 求第 2 位顾客在 10 点前到达的概率.

(3) 求第 1 位顾客在 9:30 前到达且第 2 位顾客在 10:00 前到达的概率.

$\boxed{\text{解}}$ 以上午 9:00 作为 0 时刻, 以 1 小时作为单位时间. 设 $N(t)$ 表示 $(0,t]$ 内来到的顾客数, 则 $\boldsymbol{N}=\{N(t),t>0\}$ 是 $\lambda=4$ 的 Poisson 过程.

(1) 由题意

$$\begin{aligned}
\mathsf{P}(N(0.5)=1,N(2.5)=5)&=\mathsf{P}(N(0.5)=1,N(2.5)-N(0.5)=4)\\
&=\mathsf{P}(N(0.5)=1)\mathsf{P}(N(2.5)-N(0.5)=4)\\
&=\mathsf{P}(N(0.5)=1)\mathsf{P}(N(2)=4)\\
&=\frac{2^1}{1!}\mathrm{e}^{-2}\cdot\frac{8^4}{4!}\mathrm{e}^{-8}=0.0155.
\end{aligned}$$

(2) 令 S_n 表示第 n 个顾客到达的时刻, 则

$$\mathsf{P}(S_2\le 1)=\mathsf{P}(N(1)\ge 2)=1-\mathrm{e}^{-4}-4\mathrm{e}^{-4}=1-5\mathrm{e}^{-4}.$$

(3) 类似地,

$$\begin{aligned}\mathsf{P}(S_1 \le 0.5, S_2 \le 1) &= \mathsf{P}(N(0.5) \ge 1, N(1) \ge 2)\\ &= \mathsf{P}(N(0.5) = 1, N(1) - N(0.5) \ge 1) + \mathsf{P}(N(0.5) \ge 2)\\ &= \mathsf{P}(N(0.5) = 1)\mathsf{P}(N(0.5) \ge 1) + \mathsf{P}(N(0.5) \ge 2)\\ &= 2\mathrm{e}^{-2}(1 - \mathrm{e}^{-2}) + 1 - \mathrm{e}^{-2} - 2\mathrm{e}^{-2} = 1 - \mathrm{e}^{-2} - 2\mathrm{e}^{-4}.\end{aligned}$$

(iii) Poisson 过程的样本曲线

根据定义, $\boldsymbol{N} = (N(t), t \ge 0)$ 是连续时间、取非负整数值的随机过程, 因此样本曲线一定是阶梯型曲线.

由于 $N(0) = 0$, 所以在第一个顾客到达之前, $N(t) = 0$. 令 S_1 表示第一个顾客到达时刻, 则有 $N(t) = 0,\ t < S_1;\quad N(S_1) = 1$. 一般地, 令 S_n 表示第 n 个顾客到达时刻, 那么有

$$N(t) = n - 1,\ S_{n-1} \le t < S_n;\quad N(S_n) = n.$$

这样, $S_1, S_2, \dots, S_n, \dots$ 是 Poisson 过程的跳跃点, 跳跃高度为 1.

记 $S_0 = 0$, 于是实直线上点列 $S_0, S_1, S_2, \cdots, S_n, \cdots$, 可以看作是 Poisson 流, 从而

$$N(t) = \sharp\{n \ge 1; S_n \le t\} = \max_n\{n \ge 1; S_n \le t\}.$$

(iv) 到达时刻的分布

由于 $S_n > t \Leftrightarrow N(t) \le n - 1,\ t \ge 0$, 所以

$$\mathsf{P}(S_n > t) = \mathsf{P}(N(t) \le n - 1) = \sum_{k=0}^{n-1}\mathsf{P}(N(t) = k) = \sum_{k=0}^{n-1}\frac{(\lambda t)^k\mathrm{e}^{-\lambda t}}{k!},\quad t \ge 0,$$

求导得, $S_n \sim \dfrac{\lambda^n t^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t}$, 即 $S_n \sim \Gamma(n, \lambda)$.

(v) 等候时间的分布

令 $X_0 = 0, X_1 = S_1 - S_0, \dots, X_n = S_n - S_{n-1}$, 于是 $X_1 \sim \mathrm{Exp}(\lambda)$. 假设 $X_{n-1} \sim \mathrm{Exp}(\lambda)$, 则

$$\begin{aligned}\mathsf{P}(X_n > t) &= \mathsf{P}(S_n - S_{n-1} > t)\\ &= \int_0^\infty \mathsf{P}(S_n - S_{n-1} > t | S_{n-1} = s)\mathsf{P}(S_{n-1} = s)\mathrm{d}s\\ &= \int_0^\infty \mathsf{P}(N(t+s) - N(s) = 0)\mathsf{P}(S_{n-1} = s)\mathrm{d}s\\ &= \int_0^\infty \mathrm{e}^{-\lambda t}\lambda\mathrm{e}^{-\lambda s}\mathrm{d}s = \mathrm{e}^{-\lambda t}.\end{aligned}$$

所以, $X_n \sim \mathrm{Exp}(\lambda)$, 亦有结论 $X_1, X_2, \dots, X_n, \dots$ 相互独立.

(vi) Poisson 过程的另一种解释

假设 $X_1, X_2, \cdots, X_n, \cdots$ 是一列独立同分布的指数随机变量, 令 $S_0 = 0, S_n = \sum_{k=1}^{n} X_k$.

定义

$$N(0) = 0, N(t) = \max_n \{n \geq 1, S_n \leq t < S_{n+1}\}, \quad t > 0.$$

于是 $\boldsymbol{N} = (N(t), t \geq 0)$ 是强度为 λ 的 Poisson 过程.

定理 3.2.1 *$\boldsymbol{N} = (N(t), t \geq 0)$ 是强度为 λ 的 Poisson 过程当且仅当其时间间隔 $X_1, X_2, \ldots, X_n, \ldots$ 独立同分布于参数为 λ 的指数分布.*

(vii) 到达时刻的条件分布

定理 3.2.2 *假设 $\boldsymbol{N} = (N(t), t \geq 0)$ 是参数为 λ 的 Poisson 过程, 令 $S_1, S_2, \ldots$ 为顾客依次到达时刻. 给定 $t > 0$, 那么*

$$(S_1, S_2, \ldots, S_n | N(t) = n) \stackrel{d}{=} (U_{(1)}, U_{(2)}, \ldots, U_{(n)}),$$

其中 $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ 为 $(0, t)$ 上 n 个独立同分布均匀随机变量 $U_1, U_2, \ldots, U_n$ 的次序统计量.

推论 3.2.1 *假设 $\boldsymbol{N} = (N(t), t \geq 0)$ 是参数为 λ 的 Poisson 过程. 对 $t > s \geq 0$, 令 $W_1, W_2, \ldots$ 分别为 $(s, t]$ 内发生的第 1 个事件, 第 2 个事件, $\ldots$ 的到达时刻. 任意给定正整数 n, 则*

$$(W_1, W_2, \ldots, W_n | N(t) - N(s) = n) \stackrel{d}{=} (U_{(1)}, U_{(2)}, \ldots, U_{(n)}),$$

其中 $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ 为 (s, t) 上 n 个独立同分布均匀随机变量 $U_1, U_2, \ldots, U_n$ 的次序统计量.

例 3.2.2 保险理赔按速率 λ 的 Poisson 过程 $\{N(t)\}$ 到达. 设各人理赔金额独立同分布 (且独立于此 Poisson 过程), 具有均值为 μ 的分布 G. 以 S_i 和 C_i 分别表示第 i 次理赔的时间和金额. 采用贴现算法, 即 t 时刻的 1 元相当于 0 时刻的 $\mathrm{e}^{-\alpha t}$ 元, 则到 t 时刻为止总理赔的贴现价值为

$$D(t) = \sum_{i=1}^{N(t)} \mathrm{e}^{-\alpha S_i} C_i,$$

计算 $\mathsf{E}(D(t))$.

$\boxed{\text{解}}$ 由到达时刻的条件分布知,

$$(S_1, S_2, \ldots, S_n | N(t) = n) \stackrel{d}{=} (U_{(1)}, U_{(2)}, \cdots, U_{(n)}),$$

其中 $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ 为 $(0, t)$ 上 n 个独立同分布均匀随机变量 $U_1, U_2, \ldots, U_n$ 的次序统计量. 因此

$$\mathsf{E}[D(t) \mid N(t) = n] = \mathsf{E}\left[\sum_{i=1}^{n} C_i \mathrm{e}^{-\alpha U_{(i)}}\right] = \sum_{i=1}^{n} \mathsf{E}(C_i)\mathsf{E}(\mathrm{e}^{-\alpha U_{(i)}}) = \mu \mathsf{E}\left[\sum_{i=1}^{n} \mathrm{e}^{-\alpha U_i}\right] = n\mu \frac{1 - \mathrm{e}^{-\alpha t}}{\alpha t}.$$

所以

$$\mathsf{E}[D(t)] = \mathsf{E}[\mathsf{E}(D(t) \mid N(t))] = \frac{\lambda \mu}{\alpha}(1 - \mathrm{e}^{-\alpha t}).$$

3.3 Poisson 过程的合并与分解

定理 3.3.1 (Poisson 过程的合并) *假设 $\boldsymbol{N}_1 = (N_1(t), t \ge 0)$ 和 $\boldsymbol{N}_2 = (N_2(t), t \ge 0)$ 是两个独立 Poisson 过程, 参数分别为 λ_1 和 λ_2. 令 $N(t) = N_1(t) + N_2(t)$, 那么 $\boldsymbol{N} = (N(t), t \ge 0)$ 是 Poisson 过程, 参数为 $\lambda_1 + \lambda_2$.*

上述结论对任意有限个 Poisson 过程也成立.

例 3.3.1 某银行有两个窗口可以接受服务. 上午九点钟小王到达这个银行, 此时两个窗口分别有一个顾客在接受服务, 另外有 2 个顾客排在小王的前面等待接受服务, 一会儿又来了很多顾客. 假设服务的规则是先来先服务, 也就是说一旦有一个窗口顾客接受完服务, 那么排在队伍中的第一个顾客就马上在此窗口接受服务. 假设各个顾客接受服务的时间独立同分布, 而且服从均值为 20 分钟的指数分布. 问: 小王在十点钟之前能够接受服务的概率?

$\boxed{\text{解}}$ 以上午九点钟为 0 时刻, 以 1 小时作为单位时间. 对 $i = 1, 2$, 令 $N_i(t)$ 表示 $(0, t]$ 内第 i 个窗口完成服务的顾客数, 则 $\{N_i(t); t \ge 0\}$ 是强度为 3 的 Poisson 过程, 且 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 相互独立.

令 $N(t)$ 表示 $(0, t]$ 内这两个窗口完成服务的顾客总数, 则由 Poisson 过程的合并可知 $N(t) = N_1(t) + N_2(t)$, 且 {N(t)} 是强度为 6 的 Poisson 过程. 以 S_i 表示第 i 个顾客服务完成的时刻, 则所求概率为

$$\mathsf{P}(S_3 \le 1) = \mathsf{P}(N(1) \ge 3) = 1 - \mathrm{e}^{-6} - 6\mathrm{e}^{-6} - 18\mathrm{e}^{-6} = 0.938.$$

定理 3.3.2 (Poisson 过程的分解) *假设寻求某种服务的顾客分为两类, 来自类型 I 的可能性为 p, 来自类型 II 的可能性为 $1 - p$. 令 $t > 0$, 用 $N(t)$ 表示 $(0, t]$ 内寻求服务的顾客总数, $N_1(t)$ 表示 $(0, t]$ 内寻求服务的 I 型顾客个数, $N_2(t)$ 表示 $(0, t]$ 内寻求服务的 II 型顾客个数. 假设 $\boldsymbol{N} = (N(t), t \ge 0)$ 是参数为 λ 的 Poisson 过程, 那么 $\boldsymbol{N}_1 = (N_1(t), t \ge 0)$ 和 $\boldsymbol{N}_2 = (N_2(t), t \ge 0)$ 分别是参数为 λp 和 $\lambda(1 - p)$ 的 Poisson 过程, 且二者相互独立.*

上述结论对多个类型服务也成立.

例 3.3.2 设 $N(t)$ 表示手机在 $(0, t]$ 天内收到的短信数. 假设 $\{N(t); t \ge 0\}$ 是强度为 10 条的 Poisson 过程, 其中每条短信独立地以概率 0.2 是垃圾短信. 求:

(1) 一天内没有收到垃圾短信的概率;

(2) 第一天内收到 3 条有用短信, 1 条垃圾短信, 第二天没有收到垃圾短信的概率.

$\boxed{\text{解}}$ 以 $X(t)$, $Y(t)$ 分别表示手机在 $(0, t]$ 天内收到的垃圾短信数和有用短信数, 则 $\{X(t); t \ge 0\}$ 和 $\{Y(t); t \ge 0\}$ 分别是强度为 2 和 8 的 Poisson 过程, 且相互独立.

(1) $\mathsf{P}(X(1) = 0) = \mathrm{e}^{-2} = 0.135$.

(2)

$$\begin{aligned}
& \mathsf{P}(Y(1) = 3, X(1) = 1, X(2) - X(1) = 0) \\
= \ & \mathsf{P}(Y(1) = 3)\mathsf{P}(X(1) = 1, X(2) - X(1) = 0) \\
= \ & \mathsf{P}(Y(1) = 3)\mathsf{P}(X(1) = 1)\mathsf{P}(X(2) - X(1) = 0) \\
= \ & \frac{8^3}{3!}\mathrm{e}^{-8} \cdot 2\mathrm{e}^{-2} \cdot \mathrm{e}^{-2} = \frac{512}{3}\mathrm{e}^{-12}.
\end{aligned}$$

3.4 复合 Poisson 过程

定义 3.4.1 假设 $\boldsymbol{N}=(N(t),t\ge 0)$ 是强度为 λ 的 *Poisson* 过程，$\xi_1,\xi_2,\ldots,\xi_n,\ldots$ 是一列随机变量. 令

$$Z(t)=\sum_{i=1}^{N(t)}\xi_i,$$

于是，$\boldsymbol{Z}=(Z(t),t\ge 0)$ 是连续时间参数、取实值的随机过程. 一般情况下，$\boldsymbol{Z}=(Z(t),t\ge 0)$ 不再是 *Poisson* 过程，称 $\boldsymbol{Z}=(Z(t),t\ge 0)$ 为复合 *Poisson* 过程.

定理 3.4.1 在如上定义的基础上，进一步假设 $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ 是一列独立同分布随机变量，均值为 μ，方差为 σ^2，并且与 $\boldsymbol{N}=(N(t),t\ge 0)$ 独立，则

(1) $EZ(t)=\mu\lambda t$, $Var(Z(t))=(\mu^2+\sigma^2)\lambda t$;

(2) $\boldsymbol{Z}=(Z(t),t\ge 0)$ 增量独立且增量平稳.

例 3.4.1 某零件在运行中会受到撞击. 记在 $(0,t]$ 内受到的撞击次数为 $N(t)$，设 $\{N(t)\}$ 是参数为 λ 的 Poisson 过程. 各次撞击带来的磨损量分别为 $\xi_1,\xi_2,\ldots$，假设它们独立同服从参数为 β 的指数分布，且与 $\{N(t)\}$ 独立. 如果磨损量大于 $\alpha>0$，那么更换零件. 计算零件的平均寿命.

解 (解法一) 令 η 为此零件的寿命，$Z(t)=\sum_{k=1}^{N(t)}\xi_k$，则

$$\eta=\inf\{t>0:Z(t)>\alpha\}.$$

于是 $\mathsf{E}(\eta)=\int_0^{\infty}\mathsf{P}(\eta>t)\mathrm{d}t=\int_0^{\infty}\mathsf{P}(Z(t)\le\alpha)\mathrm{d}t$. 而由全概率公式

$$\mathsf{P}(Z(t)\le\alpha)=\sum_{n=0}^{\infty}\mathsf{P}(Z(t)\le\alpha\mid N(t)=n)\mathsf{P}(N(t)=n)=\sum_{n=0}^{\infty}\mathsf{P}\left(\sum_{k=1}^{n}\xi_k\le\alpha\right)\mathsf{P}(N(t)=n).$$

令 $M(t)=\sup\left\{n\ge 0:\sum_{k=1}^{n}\xi_k\le t\right\}$，则 $\{M(t)\}$ 是参数为 β 的 Poisson 过程. 因此

$$\mathsf{P}\left(\sum_{k=1}^{n}\xi_k\le\alpha\right)=\mathsf{P}(M(\alpha)\ge n).$$

综合上述，我们有

$$\mathsf{E}(\eta)=\int_0^{\infty}\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\mathsf{P}(M(\alpha)\ge n)\mathrm{d}t=\sum_{n=0}^{\infty}\int_0^{\infty}\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\mathsf{P}(M(\alpha)\ge n)\mathrm{d}t.$$

注意到

$$\int_0^{\infty}\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\mathrm{d}t=\frac{1}{\lambda},\quad \mathsf{E}(M(\alpha))=\sum_{n=0}^{\infty}\mathsf{P}(M(\alpha)>n),$$

所以 $\mathsf{E}(\eta)=\frac{1}{\lambda}\sum_{n=0}^{\infty}\mathsf{P}(M(\alpha)\ge n)=\frac{1+\mathsf{E}(M(\alpha))}{\lambda}=\frac{1+\alpha\beta}{\lambda}$.

(解法二) 令 $S_0=0$, S_n 为第 n 次撞击发生的时刻，$T_n=S_n-S_{n-1}$, $n\ge 1$. 令 X 为磨损量首次大于 α 时受到的撞击数目，η 为此零件的寿命，于是

$$\eta=S_X,\quad X=\min\{n\ge 1:\xi_1+\xi_2+\cdots+\xi_n>\alpha\}.$$

由于 $\{\xi_1,\xi_2,\dots\}$ 与 $\{N(t),t\geq 0\}$ 独立, 所以

$$\mathsf{E}(S_X|X=n)=\mathsf{E}\left(\sum_{i=1}^n T_i\Big|X=n\right)=\mathsf{E}\left(\sum_{i=1}^n T_i\right)=\frac{n}{\lambda}.$$

于是 $\mathsf{E}(S_X)=\mathsf{E}[\mathsf{E}(S_X|X)]=\frac{1}{\lambda}\mathsf{E}(X)$. 同样地, 令 $M(t)=\sup\left\{n\geq 0:\sum_{k=1}^n\xi_k\leq t\right\}$, 则 $\{M(t)\}$ 是参数为 β 的 Poisson 过程. 而 X 是此过程中时刻 α 后首次发生的事件, 因此 $X=M(\alpha)+1$. 所以

$$\mathsf{E}(\eta)=\frac{\mathsf{E}(X)}{\lambda}=\frac{1+\mathsf{E}(M(\alpha))}{\lambda}=\frac{1+\alpha\beta}{\lambda}.$$

(解法三) 令 η 为此零件的寿命, 记 $f(\alpha)=\mathsf{E}(\eta)$. 令 $S_0=0$, S_n 为第 n 次撞击时刻, $T_n=S_n-S_{n-1}$, $n\geq 1$.

(1) 如果 $\alpha=0$, 则 $\eta=S_1=T_1$, 所以 $f(0)=\mathsf{E}(\eta)=\frac{1}{\lambda}$.

(2) 设 $\alpha>0$, 则零件至少可经过 1 次撞击, 之后的撞击时间间隔 $T_2,T_3,\dots$ i.i.d. $\sim\mathrm{Exp}(\lambda)$ 且与 T_1 独立. 因此之后形成的计数过程 $\{N_1(t);t\geq 0\}$ 仍是参数为 λ 的 Poisson 过程且与 T_1 独立. 其中 $N_1(t)=N(T_1+t)-N(T_1)$.

第 1 次撞击的磨损量 $\xi_1\sim\mathrm{Exp}(\beta)$, 且 $\{\xi_1,T_1\}$ 与 $\{\xi_n,N_1(t);n\geq 1,t\geq 0\}$ 独立.

如果 $\xi_1=x>\alpha$, 则 $\eta=T_1$, 所以 $\mathsf{E}(\eta|\xi_1=x)=\mathsf{E}(T_1)=\frac{1}{\lambda}$.

如果 $\xi_1=x\leq\alpha$, 则还需磨损量大于 $\alpha-x$, 所以

$$\mathsf{E}(\eta|\xi_1=x)=\mathsf{E}(T_1)+f(\alpha-x)=\frac{1}{\lambda}+f(\alpha-x).$$

由全期望公式:

$$f(\alpha)=\mathsf{E}(\eta)=\int_0^\infty\mathsf{E}(\eta|\xi_1=x)f_{\xi_1}(x)\mathrm{d}x=\frac{1}{\lambda}+\beta\mathrm{e}^{-\alpha\beta}\int_0^\alpha\mathrm{e}^{\beta u}f(u)\mathrm{d}u.$$

从而可求得 $f(\alpha)=\frac{1+\alpha\beta}{\lambda}$.

3.5 非齐次 Poisson 过程

令 $t>0$, $N(t)$ 表示 $(0,t]$ 内顾客寻求服务的个数. 假设

(1) 初始条件: $N(0)=0,N(t)\geq 0$;

(2) 独立性: 假设 $s<t$, 则 $N(t)-N(s)$ 与 $N(s)$ 独立;

(3) 稀有性: 存在一个非负函数 $\lambda(t)$ 使得

$$\mathsf{P}(N(t+\Delta t)-N(t)=1)=\lambda(t)\Delta t+o(\Delta t),$$
$$\mathsf{P}(N(t+\Delta t)-N(t)\geq 2)=o(\Delta t).$$

其中参数 $\lambda(t)$ 被称为 Poisson 流的强度函数, 它的大小表示该服务系统的繁忙程度.

定义 3.5.1 *在上述假设下, $N(t)$ 服从参数为 $m(t)=\int_0^t\lambda(u)\mathrm{d}u$ 的 Poisson 分布. 特别地, 对 $\forall s<t$, 有*

$$P(N(t)-N(s)=k)=\frac{[m(t)-m(s)]^k}{k!}\mathrm{e}^{-[m(t)-m(s)]},\ \forall k\geq 0,$$

称 $\boldsymbol{N}=(N(t),t\geq 0)$ 是强度为 $\lambda(\cdot)$ 的非齐次 Poisson 过程.

(i) 基本性质

(1) $\mathsf{E}N(t)=m(t)$, $\quad \mathsf{Var}N(t)=m(t)$.

(2) 假设 $s<t$, 由增量独立与增量平稳的性质, 得

$$\begin{aligned}\mathsf{E}N(s)N(t)&=\mathsf{E}[N(s)(N(t)-N(s)+N(s))]\\&=\mathsf{E}N(s)\mathsf{E}(N(t)-N(s))+\mathsf{E}N(s)^2\\&=m(s)m(t-s)+m(s)+[m(s)]^2\end{aligned}$$

(3) 顾客到达时刻与等待时间的分布

给定 $t>0$, 由于 $S_1>t\Leftrightarrow N(t)=0$, 所以

$$\mathsf{P}(S_1>t)=\mathsf{P}(N(t)=0)=\mathrm{e}^{-m(t)},$$

即 $S_1\sim\lambda(t)\mathrm{e}^{-m(t)}$, 所以 S_1 不再服从指数分布.

类似地, 可以证明 $X_1,X_2,\cdots,X_n,\cdots$ 不再独立同分布, 更不服从指数分布.

例 3.5.1 设 $\{N(t),t\geq 0\}$ 是强度为 $\lambda(t)=t^2$ 的非齐次 Poisson 过程. 计算

(1) $\mathsf{E}(N(2))$.

(2) $\mathsf{P}(N(1)=1,N(2)=2)$.

(3) $\mathsf{P}(N(2)=2\mid N(1)=1)$.

(4) $\mathsf{P}(N(1)=1\mid N(2)=2)$.

$\boxed{\text{解}}$ (1) $\mathsf{E}(N(2))=\displaystyle\int_0^2\lambda(t)\mathrm{d}t=\int_0^2 t^2\mathrm{d}t=\frac{8}{3}$.

(2) 计算得 $m(1)=\displaystyle\int_0^1 t^2\mathrm{d}t=\frac{1}{3}$, $m(2)=\displaystyle\int_0^2 t^2\mathrm{d}t=\frac{8}{3}$, 因此

$$\mathsf{P}(N(1)=1,N(2)=2)=\mathsf{P}(N(1)=1)\mathsf{P}(N(2)-N(1)=1)=\frac{1}{3}\mathrm{e}^{-\frac{1}{3}}\cdot\frac{7}{3}\mathrm{e}^{-\frac{7}{3}}=\frac{7}{9}\mathrm{e}^{-\frac{8}{3}}.$$

(3) $\mathsf{P}(N(2)=2\mid N(1)=1)=\mathsf{P}(N(2)-N(1)=1)=\dfrac{7}{3}\mathrm{e}^{-\frac{7}{3}}$.

(4) $\mathsf{P}(N(1)=1\mid N(2)=2)=\dfrac{\mathsf{P}(N(1)=1,N(2)=2)}{\mathsf{P}(N(2)=2)}=\dfrac{\frac{7}{9}\mathrm{e}^{-\frac{8}{3}}}{(\frac{8}{3})^2\mathrm{e}^{-\frac{8}{3}}/2}=\dfrac{7}{32}$.

下面定理给出了非齐次 Poisson 过程到达时刻的条件分布.

定理 3.5.1 令 $t>0,n\geq 1$. 假设 $V_1,V_2,\ldots,V_n$ 独立同分布, 密度函数为 $\dfrac{\lambda(u)}{m(t)}$,$0\leq u\leq t$. 那么

$$(S_1,S_2,\ldots,S_n|N(t)=n)\overset{d}{=}(V_{(1)},V_{(2)},\ldots,V_{(n)}),$$

其中 $V_{(1)},V_{(2)},\ldots,V_{(n)}$ 是 $V_1,V_2,\ldots,V_n$ 的次序统计量.

下面定理给出了非齐次 Poisson 过程的合并与分解.

定理 3.5.2 (非齐次 Poisson 过程的合成) 设 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 分别是强度为 $\lambda_1(t)$ 和 $\lambda_2(t)$ 的 *Poisson* 过程, 且相互独立, 则 $\{N_1(t)+N_2(t)\}$ 是强度为 $\lambda_1(t)+\lambda_2(t)$ 的 *Poisson* 过程.

定理 3.5.3 (非齐次 Poisson 过程的分解) 设 $\{N(t)\}$ 是强度为 $\lambda(t)$ 的 *Poisson* 过程, 在 t 时刻发生的事件独立地 *(也独立于过程* $\{N(t)\}$*) 以概率* $p(t)$ *为类型 1, 以* $1-p(t)$ *为类型 2. 令* $N_1(t)$ 和 $N_2(t)$ 分别表示到 t 为止类型 *1* 和类型 *2* 发生的个数, 则 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 分别是强度为 $\lambda(t)p(t)$ 和 $\lambda(t)(1-p(t))$ 的 *Poisson* 过程, 且相互独立.

例 3.5.2 [无穷条服务线的排队问题 $M/G/\infty$] 顾客按速率 λ 的 Poisson 过程到达服务站, 到达后马上接受服务, 服务时间独立同服从分布 G. 令 $X(t)$ 表示到 t 为止完成服务的顾客数, $Y(t)$ 表示 t 时在接受服务的顾客数. 问: $X(t)$ 和 $Y(t)$ 分别服从何分布?

解 固定 $t>0$, 令 $N(t)$ 表示 $(0,t]$ 内到达的顾客数, 则 $N(t)\sim\mathcal{P}(\lambda t)$. 对于在 $s\leq t$ 时到达的顾客, 独立地以概率 $p(s)=G(t-s)$ 到 t 为止完成服务 (称为类型 1), 以概率 $1-p(s)$ 到 t 时还在接受服务 (称为类型 2).

在 $N(t)=n$ 的条件下, 在 $(0,t]$ 到达的这 n 个顾客 (不考虑到达先后顺序) 到达时刻独立同服从 $U(0,t)$, 所以每个事件独立地以概率

$$q=\int_0^t p(s)\frac{1}{t}\mathrm{d}s=\frac{1}{t}\int_0^t G(s)\mathrm{d}s$$

为类型 1, 以概率 $1-q$ 为类型 2. 因此

$$X(t)\sim\mathcal{P}\left(\lambda\int_0^t G(s)\mathrm{d}s\right),\quad Y(t)\sim\mathcal{P}\left(\lambda\int_0^t(1-G(s))\mathrm{d}s\right)$$

且二者相互独立.

(ii) 时间变换

定理 3.5.4 设 $\boldsymbol{N}=(N(t),t\geq 0)$ 是强度为 *1* 的 *Poisson* 过程, 令 $\lambda(u)$ 是 $[0,\infty)$ 上非负的在任何有界区间上可积的函数, 令

$$m(t)=\int_0^t\lambda(u)\mathrm{d}u,\quad M(t)=N(m(t)),$$

则 $\boldsymbol{M}=(M(t),t\geq 0)$ 是强度为 $\lambda(t)$ 的非齐次 *Poisson* 过程.

定理 3.5.5 假设 $\widetilde{\boldsymbol{N}}=(\widetilde{N}(t),t\geq 0)$ 是强度函数为 $\lambda(t)$ 的非齐次 *Poisson* 过程, 并且 $\lambda(t)$ 是严格正的函数, $m(t)$ 是严格单调增的正函数, 其逆函数存在, 记为 $m^{-1}(t)$. 定义一个新过程

$$N(0)=0,\quad N(t)=\widetilde{N}(m^{-1}(t)),\ t>0,$$

则 $\boldsymbol{N}=(N(t),t\geq 0)$ 是强度为 *1* 的齐次 *Poisson* 过程.

四、 Markov 链

4.1 Markov 链及例子

定义 4.1.1 (Markov 链) 如果 $\{X_n,n=0,1,\dots\}$ 是离散状态的随机过程, 并且具有 *Markov* 性, 即对任何 $k\geq 1$, 任何状态 $i_0,\dots,i_{k-1}$, i, j, 有

$$P(X_{k+1}=j\mid X_0=i_0,\dots,X_{k-1}=i_{k-1},X_k=i)=P(X_{k+1}=j\mid X_k=i),$$

则称 $\{X_n,n=0,1,\dots\}$ 是 ***Markov* 链**.

令 $A=\{X_0=i_0,\ldots,X_{k-1}=i_{k-1}\}$ 表示过去, $B=\{X_k=i\}$ 表示现在, $C=\{X_{k+1}=j\}$ 表示未来, 则 Markov 性意味着

$$\mathsf{P}(C\mid AB)=\mathsf{P}(C\mid B).$$

也就是说, 在已知现在状态的条件下, 过去与将来相互独立.

对 $n\geq m\geq 0$ 和状态 i, j, 记

$$p_{ij}(m,n)=\mathsf{P}(X_n=j\mid X_m=i),$$

表示在 m 时处于状态 i 的条件下, 到 n 时转移到状态 j 的概率. 显然, 转移概率具有如下性质:

$$p_{ij}(m,n)\geq 0,\quad \sum_{j\in I}p_{ij}(m,n)=1,$$

其中 I 是 Markov 链的状态空间.

定义 4.1.2 *如果对任何状态 i, j, $P(X_{n+1}=j\mid X_n=i)$ 不依赖于 n, 则称 $\{X_n,n=0,1,\ldots\}$ 是**时齐的 Markov 链**. 在时齐 Markov 链中, 称 $p_{ij}=P(X_{n+1}=j\mid X_n=i)$ 为从 i 到 j 的一步转移概率, $\boldsymbol{P}=(p_{ij})_{I\times I}$ 为一步转移矩阵.*

例 4.1.1 [0-1 传输系统] 在只传输 0 和 1 的串联系统中, 设每一级的传真率为 p, 误码率为 $1-p$. 以 X_0 表示第 1 级的输入, X_n 表示第 n 级的输出 $(n\geq 1)$. 于是 $\{X_n\}$ 是一时齐的 Markov 链, 状态空间 $I=\{0,1\}$,

$$p_{ij}=\mathsf{P}(X_{n+1}=j\mid X_n=i)=\begin{cases}p, & j=i,\\ 1-p, & j\neq i,\end{cases}\quad i,j=0,1.$$

故一步转移矩阵为 $\boldsymbol{P}=\begin{pmatrix}p & 1-p\\ 1-p & p\end{pmatrix}$.

例 4.1.2 [排队模型] 有一修理店, 每天只能修好一个机器, 并且不修当天送来的机器. 假定第 n 天有 ξ_n 个机器损坏, 则第 $n+1$ 天把这 ξ_n 个机器送往此店维修. 令 X_n 表示第 n 天结束时此店中机器的个数. 于是

$$X_{n+1}=\max(X_n-1,0)+\xi_n.$$

如果 $\xi_1,\xi_2,\ldots$ 独立同分布, 且与 X_0 独立, 则 $\{X_n;n\geq 0\}$ 是时齐的 Markov 链, 状态空间 $I=\{0,1,2,\ldots\}$. 令 $\mathsf{P}(\xi_1=k)=a_k$, $k\geq 0$, 则一步转移概率

$$p_{ij}=\begin{cases}a_j, & i=0,\\ a_{j-i+1}, & i>0\text{ 且 }j\geq i-1,\\ 0, & \text{其他}.\end{cases}$$

例 4.1.3 [爬梯子模型] 考虑某人患某种病的情况. 对 $n\geq 0$, 令 $\xi_n=\begin{cases}0, & \text{第 } n \text{ 天没患病},\\ 1, & \text{第 } n \text{ 天患病}.\end{cases}$

假设 $\{\xi_n; n\geq 0\}$ 是时齐 Markov 链, 状态空间 $I=\{0,1\}$, 一步转移矩阵 $\boldsymbol{Q}=\begin{pmatrix}1-p & p\\ q & 1-q\end{pmatrix}$.

令 $X_n=\begin{cases}\max\{1\leq k\leq n+1:\xi_n\xi_{n-1}\cdots\xi_{n-k+1}=1\}, & \xi_n=1,\\ 0, & \xi_n=0\end{cases}$ 表示第 n 天时连续患病的天数, 则

$$X_{n+1}=\begin{cases}X_n+1, & \xi_{n+1}=1,\\ 0, & \xi_{n+1}=0.\end{cases}$$

于是 $\{X_n; n\geq 0\}$ 是时齐 Markov 链, 状态空间 $I=\{0,1,2,\ldots\}$, 转移概率为

$$p_{0,0}=1-p,\quad p_{0,1}=p,\quad p_{i,i+1}=1-q,\quad p_{i,0}=q,\quad \forall i\geq 1.$$

例 4.1.4 [Polya 罐子模型] 设一罐子装有 r 个红球, b 个黑球. 现随机从罐中取出一球, 记录其颜色后将其放回, 并加入 a 个相同颜色的球. 持续进行这一过程, 令 X_n 表示第 n 次试验结束时罐中的红球数, 则 $\{X_n; n\geq 0\}$ 是一 Markov 链, 状态空间 $I=\{r,r+a,r+2a,\ldots\}$, 它的一步转移概率为

$$p_{ij}(n,n+1)=\begin{cases}\dfrac{i}{r+b+na}, & j=i+a,\\ 1-\dfrac{i}{r+b+na}, & j=i,\\ 0, & \text{其他}.\end{cases}$$

所以该 Markov 链不是时齐的.

例 4.1.5 独立重复地掷骰子, 以 X_n 记第 n 次掷出的点数, 令 $Y_n=X_{n+1}+X_{n+2}$, $n\geq 0$. 计算 $\mathsf{P}(Y_2=12\mid Y_0=2,Y_1=7)$, $\mathsf{P}(Y_2=12\mid Y_1=7)$ 并判断 $\{Y_n\}$ 是否是 Markov 链.

解 由定义

$$\mathsf{P}(Y_2=12\mid Y_0=2,Y_1=7)=\mathsf{P}(X_3=6,X_4=6\mid X_1=1,X_2=1,X_3=6)=\frac{1}{6}.$$

$$\mathsf{P}(Y_2=12\mid Y_1=7)=\frac{\mathsf{P}(X_2=1,X_3=X_4=6)}{\mathsf{P}(X_2+X_3=7)}=\frac{(\frac{1}{6})^3}{\frac{6}{36}}=\frac{1}{36}.$$

二者不想等, 因此 $\{Y_n\}$ 不是 Markov 链.

4.2 Markov 链的有限维分布

定义 4.2.1 (初始分布) 令 $p_0(i)=P(X_0=i)$, 并记 $\boldsymbol{p}_0=(p_0(1),p_0(2),\ldots,p_0(N))$, 称 $\boldsymbol{p}_0$ 为 *Markov* 链的**初始分布**.

引理 4.2.1 (Chapman-Kolmogorov(C-K) 方程) 对 $\forall n\geq 0$, $m,l\geq 1$, $i,j\in I$,

$$p_{ij}(n,n+m+l)=\sum_k p_{ik}(n,n+m)p_{kj}(n+m,n+m+l).$$

事实上, 该方程也可以写成矩阵形式:

$$\boldsymbol{P}(n,n+m+l)=\boldsymbol{P}(n,n+m)\boldsymbol{P}(n+m,n+m+l).$$

因此对时齐 Markov 链 $\{X_n\}$, 由 C-K 方程可知 $\boldsymbol{P}^{(m)}\triangleq\boldsymbol{P}(n,n+m)=\boldsymbol{P}^m$.

$\boldsymbol{p}_0$ 描述着随机系统的初始状态分布规律, 它和转移概率矩阵 $\boldsymbol{P}$ 共同决定 Markov 的分布. 由 C-K 方程, Markov 链在第 n 时刻处于各个状态的概率分布

$$\boldsymbol{p}_n = (p_n(1), p_n(2), \ldots, p_n(N)) = \boldsymbol{p}_0\boldsymbol{P}^n, \quad \forall n \geq 1.$$

进一步, 利用条件概率的链式法则可知, 对 $\forall n_1 < n_2 < \cdots < n_k$,

$$\mathsf{P}(X_{n_1} = i_1, \ldots, X_{n_k} = i_k) = \mathsf{P}(X_{n_1} = i_1)p_{i_1i_2}^{(n_2-n_1)} \cdots p_{i_{k-1}i_k}^{(n_k-n_{k-1})},$$

其中 $p_{ij}^{(m)} = p_{ij}(n, n+m)$.

例 4.2.1 设 $\{X_n, n \geq 0\}$ 是具有三个状态 0,1,2 的时齐 Markov 链, 一步转移概率矩阵为

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

且 $\mathsf{P}(X_0 = 0) = \mathsf{P}(X_0 = 1) = \dfrac{1}{2}$. 试求:

(1) $\mathsf{P}(X_0 = 0, X_1 = 1, X_3 = 1)$;

(2) $\mathsf{P}(X_3 = 1, X_1 = 1 \mid X_0 = 0)$;

(3) $\mathsf{P}(X_3 = 1)$;

(4) $\mathsf{P}(X_0 = 0 \mid X_3 = 1)$.

解 经计算得

$$\boldsymbol{P}^2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{7}{8} & \frac{1}{8} \\ \frac{3}{8} & \frac{3}{16} & \frac{7}{16} \end{pmatrix}, \quad \boldsymbol{P}^3 = \begin{pmatrix} 0 & \frac{7}{8} & \frac{1}{8} \\ \frac{7}{16} & \frac{3}{32} & \frac{15}{32} \\ \frac{3}{32} & \frac{45}{64} & \frac{13}{64} \end{pmatrix}.$$

所以 (1) $\mathsf{P}(X_0 = 0, X_1 = 1, X_3 = 1) = \mathsf{P}(X_0 = 0)p_{01}p_{11}^{(2)} = \dfrac{1}{2} \times 1 \times \dfrac{7}{8} = \dfrac{7}{16}$.

(2) $\mathsf{P}(X_3 = 1, X_1 = 1 \mid X_0 = 0) = p_{01}p_{11}^{(2)} = \dfrac{7}{8}$.

(3) $\mathsf{P}(X_3 = 1) = \mathsf{P}(X_0 = 0)p_{01}^{(3)} + \mathsf{P}(X_0 = 1)p_{11}^{(3)} = \dfrac{1}{2} \times \dfrac{7}{8} + \dfrac{1}{2} \times \dfrac{3}{32} = \dfrac{31}{64}$.

(4) $\mathsf{P}(X_0 = 0 \mid X_3 = 1) = \dfrac{\mathsf{P}(X_3 = 1 \mid X_0 = 0)\mathsf{P}(X_0 = 0)}{\mathsf{P}(X_3 = 1)} = \dfrac{\frac{1}{2} \times \frac{7}{8}}{\frac{31}{64}} = \dfrac{28}{31}$.

例 4.2.2 淘宝网上有 5 家店卖同一种产品. 设每位购买此种产品的顾客独立地任选一家网店购买. 问经过 5 名顾客购买后, 恰有 3 个网店被购买过的概率?

解 以 X_n 表示第 $n+1$ 个顾客购买后被购买过的网店数目, 则 $\{X_n\}$ 是以 1,2,3,4,5 为状态的 Markov 链, 转移概率

$$p_{i,i} = \frac{i}{5} = 1 - p_{i,i+1}.$$

所求概率即为 $p_{13}^{(4)}$.

由一步转移概率矩阵可得两步转移概率矩阵，即

$$\boldsymbol{P}^2 = \begin{pmatrix} 1/5 & 4/5 & 0 & 0 & 0 \\ 0 & 2/5 & 3/5 & 0 & 0 \\ 0 & 0 & 3/5 & 2/5 & 0 \\ 0 & 0 & 0 & 4/5 & 1/5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 0.04 & 0.48 & 0.48 & 0 & 0 \\ 0 & 0.16 & 0.60 & 0.24 & 0 \\ 0 & 0 & 0.36 & 0.56 & 0.08 \\ 0 & 0 & 0 & 0.64 & 0.36 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

因此

$$p_{13}^{(4)} = \sum_{i=1}^{5} p_{1i}^{(2)} p_{i3}^{(2)} = 0.04 \times 0.48 + 0.48 \times 0.60 + 0.48 \times 0.36 = 0.48.$$

4.3 常返性和瞬时性

定义 4.3.1 (首中时) 称

$$\tau_i = \min\{n \geq 1 : X_n = i\}$$

为状态 i 的**首中时**，并约定 $\min \varnothing = \infty$.

定义 4.3.2 (常返态与瞬时态) 若状态 i 满足 $P(\tau_i < \infty \mid X_0 = i) = 1$，则称状态 i 为**常返态**，否则 $P(\tau_i < \infty \mid X_0 = i) < 1$，称状态 i 为**瞬时态**.

从定义中可以看出，常返的意义是从状态 i 出发以概率 1 在有限时间内能返回状态 i，而瞬时的意义是从状态 i 出发以正概率不再返回状态 i.

定义 4.3.3 (平均回转时) 若状态 i 是常返态，定义

$$\mu_i = E(\tau_i \mid X_0 = i)$$

为状态 i 的**平均回转时**.

定义 4.3.4 (正常返与零常返) 若常返态 i 满足 $\mu_i < \infty$，则称状态 i **正常返**，否则 $\mu_i = \infty$，称状态 i **零常返**.

从定义中可以看出，正常返的意义是从状态 i 出发不但以概率 1 在有限时间内返回状态 i，而且平均回转时有限，而零常返的意义是从状态 i 出发虽然以概率 1 在有限时间内返回状态 i，但平均回转时无限. 因此，正常返态返回速度比零常返态快.

为判定 Markov 链中状态的常返性与瞬时性，令

$$f_{ij}^{(n)} = \mathsf{P}(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i)$$

表示从状态 i 出发首次到达状态 j 的概率，而 $f_{ij} = \mathsf{P}(\tau_j < \infty \mid X_0 = i)$ 表示从状态 i 出发能到达状态 j 的概率，因此

$$f_{ij} = \lim_{N\to\infty} \sum_{n=1}^{N} f_{ij}^{(n)} = \sum_{n=1}^{\infty} f_{ij}^{(n)}.$$

定理 4.3.1 状态 i 是常返态当且仅当 $f_{ii} = 1$，且有平均回转时 $\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$；状态 i 是瞬时态当且仅当 $f_{ii} < 1$.

下面给出当 $n \geq 1$ 时, $p_{ij}^{(n)}$ 与 $f_{ij}^{(n)}$ 之间的关系: 由全概率公式

$$\begin{aligned}
p_{ij}^{(n)} = \mathsf{P}(X_n = j \mid X_0 = i) &= \sum_{k=1}^{n} \mathsf{P}(\tau_j = k, X_n = j | X_0 = i) \\
&= \sum_{k=1}^{n} \mathsf{P}(\tau_j = k \mid X_0 = i)\mathsf{P}(X_n = j \mid \tau_j = k, X_0 = i) \\
&= \sum_{k=1}^{n} f_{ij}^{(k)} \mathsf{P}(X_n = j \mid X_k = j, X_{k-1} \neq j, \ldots, X_1 \neq j, X_0 = i) \\
&= \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}.
\end{aligned}$$

下面定理给出了常返性和瞬时性的等价概率含义.

定理 4.3.2 *以 $N_i = \sharp\{n \geq 0; X_n = i\}$ 表示 $X_0, X_1, X_2, \ldots$ 中访问状态 i 的次数, 于是状态 i 是常返态 $\Leftrightarrow P(N_i = \infty \mid X_0 = i) = 1$; 状态 i 是瞬时态 $\Leftrightarrow P(N_i < \infty \mid X_0 = i) = 1$.*

这一定理由 Markov 性即得. 更重要的是, 这一定理蕴含着“常返”和“瞬时”的直观含义. 更进一步, 有

$$\mathsf{P}(N_i = 0 \mid X_0 = i) = 1 - f_{ii}, \quad \mathsf{P}(N_i = n \mid X_0 = i) = (1 - f_{ii}) f_{ii}^n, \ n \geq 1.$$

即 N_i 是几何随机变量. 由此得到一个重要结论, 即

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathsf{E}(N_i \mid X_0 = i) = \begin{cases} \infty, & \text{状态 } i \text{ 是常返态}, \\ \dfrac{1}{1 - f_{ii}} < \infty, & \text{状态 } i \text{ 是瞬时态}. \end{cases}$$

事实上, 由于 $N_i = \sum\limits_{n=0}^{\infty} \mathbf{1}_{\{X_n = i\}}$, 所以

$$\mathsf{E}(N_i \mid X_0 = i) = \sum_{n=0}^{\infty} \mathsf{E}(\mathbf{1}_{\{X_n = i\}} \mid X_0 = i) = \sum_{n=0}^{\infty} \mathsf{P}(X_n = i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)}.$$

归纳成如下定理.

定理 4.3.3 *状态 i 是常返态 $\Leftrightarrow \sum\limits_{n=0}^{\infty} p_{ii}^{(n)} = \infty$; 状态 i 是瞬时态 $\Leftrightarrow \sum\limits_{n=0}^{\infty} p_{ii}^{(n)} < \infty$, 进一步可推出 $\lim\limits_{n\to\infty} p_{ii}^{(n)} = 0$.*

定义 4.3.5 (吸收态) *如果 $p_{ii} = 1$, 则称状态 i 为*吸收态.

Markov 链一旦到达吸收态 i, 就永远待在状态 i, 从而 $f_{ii} = 1$, $\mu_i = 1$, 状态 i 是正常返态.

例 4.3.1 对状态空间 $I = \{0, 1\}$ 上, 状态转移矩阵 $\boldsymbol{P} = \begin{pmatrix} 0.2 & 0.8 \\ 0 & 1 \end{pmatrix}$ 的时齐 Markov 链, 判断状态 0 的常返性.

解 注意到 $f_{00}^{(1)} = p_{00} = 0.2$, 而对 $n \geq 2$ 都有 $f_{00}^{(n)} = 0$. 因此 $f_{00} = 0.2 < 1$, 故状态 0 是瞬时态.

例 4.3.2 对状态空间 $I=\{0,1,2,3\}$ 上, 状态转移矩阵 $\boldsymbol{P}=\begin{pmatrix}0 & 0.5 & 0 & 0.5\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0.5 & 0 & 0 & 0.5\end{pmatrix}$ 的时齐 Markov 链, 判断状态 0 和状态 3 的常返性.

解 首先讨论状态 0 的常返性. 计算得

$$f_{00}^{(1)}=0,\quad f_{00}^{(2)}=p_{03}p_{30}=\frac{1}{4},\quad f_{00}^{(3)}=p_{03}p_{33}p_{30}=\frac{1}{8},$$

而当 $n\geq 4$ 时,

$$f_{00}^{(n)}=p_{03}p_{33}^{n-2}p_{30}+p_{01}p_{12}p_{23}p_{33}^{n-4}p_{30}=\frac{5}{2^n}.$$

因此

$$f_{00}=\sum_{n=1}^{\infty}f_{00}^{(n)}=\frac{1}{4}+\frac{1}{8}+\sum_{n=4}^{\infty}\frac{5}{2^n}=1,\quad \mu_0=\frac{1}{2}+\frac{3}{8}+\sum_{n=4}^{\infty}\frac{5n}{2^n}=4<\infty,$$

所以状态 0 是正常返态.

下面讨论状态 3 的常返性. 同样地, 计算得

$$f_{33}^{(1)}=\frac{1}{2},\quad f_{33}^{(2)}=p_{30}p_{03}=\frac{1}{4},\quad f_{33}^{(3)}=0,\quad f_{33}^{(4)}=p_{30}p_{01}p_{12}p_{23}=\frac{1}{4},$$

而当 $n\geq 5$ 时, $f_{33}^{(n)}=0$, 因此

$$f_{33}=\sum_{n=1}^{\infty}f_{33}^{(n)}=1,\quad \mu_3=\frac{1}{2}+\frac{1}{2}+1=2<\infty,$$

所以状态 3 是正常返态.

例 4.3.3 [爬梯子模型] 设 $\{X_n\}$ 是时齐 Markov 链, $I=\{0,1,2,\dots\}$, 转移概率

$$p_{i,i+1}=p_i,\quad p_{i,0}=1-p_i,\qquad 0<p_i<1, i\geq 0.$$

讨论状态 0 的常返性.

解 因为 $f_{00}^{(1)}=p_{00}=1-p_0$, 而对 $n\geq 2$,

$$f_{00}^{(n)}=p_{01}p_{12}\cdots p_{n-2,n-1}p_{n-1,0}=p_0p_1\cdots p_{n-2}(1-p_{n-1}).$$

令 $u_0=1$, $u_n=p_0p_1\cdots p_{n-1}$, $\forall n\geq 1$, 则

$$f_{00}=(1-u_1)+(u_1-u_2)+(u_2-u_3)+\cdots+(u_{n-1}-u_n)+\cdots=1-\lim_{n\to\infty}u_n.$$

所以状态 0 是常返态 $\Leftrightarrow \lim\limits_{n\to\infty}u_n=0$.

进一步, 当状态 0 是常返态时,

$$\mu_0=(1-u_1)+2(u_1-u_2)+3(u_2-u_3)+\cdots+n(u_{n-1}-u_n)+\cdots=\sum_{n=0}^{\infty}u_n.$$

所以状态 0 是正常返态 $\Leftrightarrow \sum\limits_{n=0}^{\infty}u_n<\infty$.

下面考虑正常返与零常返的等价刻画.

定理 4.3.4 状态 i 是正常返态 $\Leftrightarrow \lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{k=1}^{n}p_{ii}^{(k)}=\dfrac{1}{\mu_i}>0 \Leftrightarrow \limsup\limits_{n\to\infty}p_{ii}^{(n)}>0$; 状态 i 是零常返态 $\Leftrightarrow \sum\limits_{n=0}^{\infty}p_{ii}^{(n)}=\infty$ 但 $\lim\limits_{n\to\infty}p_{ii}^{(n)}=0$.

例 4.3.4 [对称随机游动] 设 d 是正整数, $\boldsymbol{X}=(X_n,n\geq 0)$ 是 $\mathbb{Z}^d$ 上的对称随机游动, 讨论状态 0 (原点) 的常返性.

解 若 $d=1$, 则对任意整数 i 有 $p_{i,i-1}=p_{i,i+1}=\dfrac{1}{2}$. 于是有

$$p_{00}^{(2n-1)}=0,\quad p_{00}^{(2n)}=\binom{2n}{n}\left(\frac{1}{2}\right)^{2n},\ n=1,2,\ldots.$$

由 Strling 公式 $n!\sim\sqrt{2\pi n}e^{-n}n^n$ 得 $p_{00}^{(2n)}\sim\dfrac{1}{\sqrt{\pi n}}$. 因此 $\sum\limits_{n=0}^{\infty}p_{00}^{(n)}=\infty$ 但 $\lim\limits_{n\to\infty}p_{00}^{(n)}=0$. 所以状态 0 零常返.

若 $d=2$, 则对每一对整数 (i,j), 有

$$p_{(i,j),(i+1,j)}=p_{(i,j),(i-1,j)}=p_{(i,j),(i,j+1)}=p_{(i,j),(i,j-1)}=\frac{1}{4}.$$

于是有

$$p_{00}^{(2n)}=\frac{1}{4^{2n}}\sum_{k=0}^{n}\binom{2n}{n}\binom{n}{k}^2=\frac{1}{4^{2n}}\binom{2n}{n}^2\sim\frac{1}{\pi n}.$$

因此 $\sum\limits_{n=0}^{\infty}p_{00}^{(n)}=\infty$ 但 $\lim\limits_{n\to\infty}p_{00}^{(n)}=0$. 所以状态 0 零常返.

事实上, 对于 $d\geq 3$, 有 $p_{00}^{(2n-1)}=0$, $p_{00}^{(2n)}=O(n^{-d/2})$, 此时状态 0 是瞬时态.

4.4 状态空间分解

定义 4.4.1 (可达与互达) 设 $i,j\in I$,

(1) 若存在 $n\geq 0$, 使得 $p_{ij}^{(n)}>0$, 则称 i **可达** *j, 记为 $i\to j$;*

(2) 若 $i\to j$, 且 $j\to i$, 则称 i,j **互达***, 记为 $i\leftrightarrow j$.*

互达是等价关系 (满足对称性、传递性和自反性), 按等价关系可以将状态空间分成若干个 (有限或无穷) 等价类, 即将状态空间分成不交的互达等价类的并.

定义 4.4.2 (不可约) 如果 Markov 链的任意两个状态都是互达的, 则称之**不可约**.

定义 4.4.3 (周期) 假设 $i\in I$, 令

$$d_i=\gcd\{n\geq 1;p_{ii}^{(n)}>0\},$$

其中 gcd 表示最大公约数, 称 d_i 为状态 i 的**周期**, 并规定 $\gcd\varnothing=0$.

如果 $d_i=1$, 则称状态 i 是**非周期**的; 如果状态 i 是非周期正常返的, 则称其**遍历**.

如果 $\{X_n\}$ 的所有状态常返 (瞬时、正常返、零常返、非周期), 则称 $\{X_n\}$ 常返 (瞬时、正常返、零常返、非周期); 如果 $\{X_n\}$ 是不可约非周期正常返的, 则称其**遍历**.

下面定理表明互达等价类中各状态具有相同的周期和常返性.

定理 4.4.1 (互达等价类的同一性质) 如果 $i \leftrightarrow j$, 则

(1) $d_i = d_j$;

(2) i 是常返态 $\Leftrightarrow$ j 是常返态;

(3) i 是正常返态 $\Leftrightarrow$ j 是正常返态.

证明 设 $i \leftrightarrow j$, $i \neq j$, 则存在正整数 m, n, 使得 $p_{ij}^{(m)} > 0$, $p_{ji}^{(n)} > 0$.

(1) 如果 $p_{ii}^{(k)} > 0$, 则 $p_{jj}^{(k+m+n)} \geq p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)} > 0$. 所以

$$d_j \mid k + m + n.$$

特别地, 有 $d_j \mid m + n$. 从而 $d_j \mid k$, 故 $d_j \mid d_i$. 同理有 $d_i \mid d_j$, 从而 $d_i = d_j$.

(2) 如果状态 i 常返, 则 $\sum\limits_{k=0}^{\infty} p_{ii}^{(k)} = \infty$. 所以

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

即状态 j 也常返.

(3) 如果 $\lim\limits_{n\to\infty} p_{jj}^{(n)} = 0$, 则

$$p_{ii}^{(k)} \leq p_{jj}^{(k+m+n)} \div (p_{ji}^{(n)} p_{ij}^{(m)}) \to 0.$$

从而结论成立. □

有了上述定理, 我们可以更快地判断一个 Markov 链中各状态的周期和常返性. 在例 4.3.2 中, 我们已经判断了状态 0 和状态 3 是正常返态, 又注意到 $p_{33} > 0$, 因此 $d_3 = 1$. 结合所有状态互达可知, 这是一个遍历的 Markov 链. 在例 4.3.3 中, 我们已经判断了状态 0 的常返性, 又 $p_{00} > 0$, 因此 $d_0 = 1$. 所以所有状态非周期, 并且与状态 0 具有相同的常返性.

4.5 极限分布与平稳分布

(i) 极限分布

定义 4.5.1 (极限分布) 假设 $\{X_n\}$ 是 *Markov* 链, 转移概率矩阵为 $\boldsymbol{P}$, 初始分布为 $\boldsymbol{p}_0$. 如果存在 I 上的一个概率分布 $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)$ 使得对任意 $j \in I$,

$$\lim_{n\to\infty} p_n(j) = \mu_j,$$

则称 $\boldsymbol{\mu}$ 是该 *Markov* 链的**极限分布**.

注意到

$$p_n(j) = \sum_{i=1}^{N} p_0(i) p_{ij}^{(n)},$$

所以如果对每个 $i \in I$, 存在 I 上一个概率分布 $\nu_i = (\nu_{i1}, \nu_{i2}, \ldots, \nu_{iN})$ 使得 $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \nu_{ij}, j \in \mathcal{E}$. 令 $\mu_j = \sum\limits_{i=1}^{N} p_0(i)\nu_{ij}$, $j \in I$, 则 $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)$ 是该 Markov 链的极限分布.

根据上述分析, 要求极限分布, 实际上只要计算 $\lim\limits_{n\to\infty} \boldsymbol{P}^{(n)}$, 也就是计算 $\lim\limits_{n\to\infty} \boldsymbol{P}^n$.

注意, 不是所有 Markov 链都一定有极限分布.

(ii) 平稳分布

假设 $\{X_n\}$ 是 Markov 链, 初始分布为 $\boldsymbol{p}_0$, 转移概率矩阵为 $\boldsymbol{P}$. 则 $\boldsymbol{X}$ 是强平稳过程当且仅当对任意 $n \geq 0, m \geq 1$,

$$(X_m, X_{m+1}, \ldots, X_{m+n}) \stackrel{d}{=} (X_0, X_1, \ldots, X_n).$$

由于转移概率不依赖于时间, 所以上式成立 $\Leftrightarrow \boldsymbol{p}_0 = \boldsymbol{p}_1 = \boldsymbol{p}_2 = \cdots \Leftrightarrow \boldsymbol{p}_0 = \boldsymbol{p}_0\boldsymbol{P}$. 一般情况下, 初始分布 $\boldsymbol{p}_0$ 和转移概率矩阵 $\boldsymbol{P}$ 不一定满足该式.

定义 4.5.2 (平稳分布) *如果存在 I 上的一个概率分布 $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_N)$ 使得,*

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

*称 $\boldsymbol{\pi}$ 是该 Markov 链的***平稳分布***.*

如果确实存在平稳分布 $\boldsymbol{\pi}$, 取 $\boldsymbol{\pi}$ 为初始分布, 就可以得到平稳 Markov 链. 为求平稳分布, 只需要解方程组

$$\begin{cases} \sum\limits_{i=1}^{N} \pi_i p_{ij} = \pi_j, \ j = 1, 2, \cdots, N, \\ \sum\limits_{j=1}^{N} \pi_j = 1, \\ \pi_j \geq 0, \ j = 1, 2, \cdots, N. \end{cases}$$

注意, 有时候 Markov 链的平稳分布不一定唯一.

例 4.5.1 求例 4.3.2 中 Markov 链的平稳分布.

$\boxed{\text{解}}$ 设平稳分布 $\boldsymbol{\pi} = (\pi_0, \pi_1, \pi_2, \pi_3)$, 则由平稳方程得

$$\begin{cases} \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1, \\ \pi_0 = \dfrac{1}{2}\pi_3, \\ \pi_1 = \dfrac{1}{2}\pi_0, \\ \pi_2 = \pi_1. \end{cases} \implies \boldsymbol{\pi} = \left(\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{2}\right).$$

例 4.5.2 求例 4.3.3 中 Markov 链的平稳分布.

$\boxed{\text{解}}$ 设平稳分布为 $\boldsymbol{\pi}$, 则

$$\pi_1 = p_0\pi_1, \quad \pi_2 = p_1\pi_1, \quad \ldots, \quad \pi_n = p_{n-1}\pi_{n-1}, \quad \ldots.$$

因此 $\pi_n = u_n\pi_0$. 又 $\sum\limits_{n=0}^{\infty} \pi_n = 1$, 所以平稳分布存在当且仅当 $\sum\limits_{n=0}^{\infty} u_n < \infty$, 即当且仅当 $\{X_n\}$ 正常返.

命题 4.5.1 *设状态 j 瞬时或零常返, 则*

(a) 对所有状态 i, $\lim\limits_{n\to\infty} p_{ij}^{(n)} = 0$;

(b) 不管初始分布如何, 恒有 $\lim\limits_{n\to\infty} P(X_n = j) = 0$;

(c) 设 $\boldsymbol{\pi}$ 是 $\{X_n\}$ 的平稳分布, 则有 $\pi_j = 0$.

定理 4.5.1 假设 $\{X_n\}$ 是非周期不可约 *Markov* 链, 则 $\{X_n\}$ 存在平稳分布当且仅当 $\{X_n\}$ 正常返的, 并且此时平稳分布唯一, 同时

$$\pi_j = \frac{1}{\mu_j}, \quad j \in I,$$

其中 μ_j 为状态 j 的平均回转时. 进一步, 对任何状态 i, j, 有 $\lim\limits_{n\to} \frac{1}{n}\sum\limits_{k=1}^{n} p_{ij}^{(k)} = \pi_j$.

特别地, 若 $\{X_n\}$ 遍历, 则对任何状态 i, j, 有 $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \pi_j$.

推论 4.5.1 (极限分布与平稳分布的关系) 假设 $\{X_n\}$ 是非周期不可约 *Markov* 链, 状态空间为 I, 转移概率矩阵为 $\boldsymbol{P}$. 那么该 *Markov* 链存在极限分布当且仅当该 *Markov* 链存在平稳分布, 并且二者相等.

注意, 对周期 Markov 链, 推论不一定成立; 不可约零常返或瞬时 Markov 链没有平稳分布.

命题 4.5.2 若 $\{X_n\}$ 是有限 *Markov* 链, 则至少存在一个正常返态.

命题 4.5.3 若 $\{X_n\}$ 是不可约的有限 *Markov* 链, 则 $\{X_n\}$ 正常返.

定义 4.5.3 (闭集) 设 C 为状态空间 I 的子集, 若对于任意状态 $i \in C$ 和任意状态 $j \notin C$, 都有 $p_{ij} = 0$, 则称之为**闭集**.

闭集的定义告诉我们, 闭集是封闭的, 从 C 中的状态出发将永远不会跑出 C 以外.

定理 4.5.2 *(1)* 如果状态 i 常返且 $i \to j$, 则 $i \leftrightarrow j$ 且 $f_{ij} = f_{ji} = 1$.
(2) 如果状态 i 常返, 则状态 i 的互达等价类是闭的.
(3) 如果状态 i 的互达等价类是有限闭集, 则状态 i 正常返.

例 4.5.3 以 X_n 表示 n 时刻系统里的顾客数. 设 $\{X_n\}$ 是 Markov 链, $I = \{0, 1, 2, 3\}$, 一步转移概率矩阵

$$\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.6 & 0.2 & 0.2 & 0 \end{pmatrix}.$$

(1) 讨论各状态的周期和常返性;
(2) 计算正常返态的平均回转时;
(3) 计算相邻两次变空的平均时间间隔.

$\boxed{\text{解}}$ (1) 注意到 $\{0, 1, 2\}$ 是有限闭的互达等价类, $\{3\}$ 是不闭的. 因此状态 0,1,2 正常返, 而状态 3 瞬时. 另一方面, 注意到 $p_{00}^{(2)} > 0$, $p_{00}^{(3)} > 0$, 因此 $d_0 = d_1 = d_2 = 1$; 而从状态 3 出发永远不返回状态 3, 所以 $d_3 = 0$.

(2) 将 Markov 链限制在 $\{0, 1, 2\}$ 上, 则该 Markov 链为非周期不可约的 Markov 链. 记其平稳分布为 $\boldsymbol{\pi} = (\pi_0, \pi_1, \pi_2)$, 由平稳方程可知

$$\begin{cases} \pi_0 + \pi_1 + \pi_2 = 1, \\ \pi_1 = 0.7\pi_2, \\ \pi_2 = \pi_0, \end{cases} \implies \boldsymbol{\pi} = \left(\frac{10}{27}, \frac{7}{27}, \frac{10}{27}\right), \implies (\mu_0, \mu_1, \mu_2) = \left(\frac{27}{10}, \frac{27}{7}, \frac{27}{10}\right).$$

(3) 相邻两次变空的平均时间间隔为 $\mu_0 = 2.7$.

4.6 吸收概率与平均吸收时间

由定理 4.5.2, 我们可以对有限 Markov 链进行状态分解:

$$I = T \cup C_1 \cup C_2 \cup \cdots \cup C_k,$$

其中 $C_1, C_2, \ldots, C_k$ 是所有闭的互达等价类, T 是余下的状态. 完成分解后, $C_1, C_2, \ldots, C_k$ 中各状态正常返, T 中各状态瞬时. 也就是说, 如果 $X_0 \in T$, 则 Markov 链最终会进入某个 C_i 并将不再离开.

对状态 i, 令

$$T_i = \min\{n \geq 0 : X_n = i\}$$

表示首次访问状态 i 的时刻; 对 I 的子集 A, 令

$$T_A = \min\{n \geq 0 : X_n \in A\}$$

表示首次访问子集 A 的时刻; 规定 $\min \varnothing = \infty$.

例 4.6.1 [赌徒输光问题] 甲、乙两人独立重复玩游戏, 每一局甲赢 1 元和输 1 元的概率都是 1/2. 游戏一直到某人输光结束. 一开始甲带有 i 元钱, 乙带有 $m-i$ 元钱. 这里 i 与 $m-i$ 都是非负整数. 计算:

(1) 游戏在有限时间内结束的概率;

(2) 甲输光的概率;

(3) 游戏平均持续时间.

$\boxed{\text{解}}$ (1) 以 S_n 表示 n 局游戏后甲所拥有的钱数, 则 $\{S_n; n \geq 0\}$ 是时齐 Markov 链, 状态空间 $I = \{0, 1, \ldots, m\}$, 一步转移概率为

$$p_{00} = p_{mm} = 1, \quad p_{i,i+1} = p_{i,i-1} = \frac{1}{2}, \quad \forall 0 < i < m.$$

于是 $C_1 = \{0\}$, $C_2 = \{m\}$, $T = \{1, 2, \ldots, m-1\}$ 为状态空间的划分, 则 T 中状态均为瞬时态, 所以最终会进入 $C_1 \cup C_2$, 此时游戏结束. 因此游戏在有限时间内结束的概率为 1.

(2) 令 $h_i = \mathsf{P}(\text{甲输光} \mid S_0 = i) = \mathsf{P}(T_0 < \infty \mid S_0 = i)$, 则 $h_0 = 1$, $h_m = 1$, 且

$$\begin{aligned} h_i &= \sum_j \mathsf{P}(\text{甲输光} \mid S_1 = j, S_0 = i)\mathsf{P}(S_1 = j \mid S_0 = i) \\ &= \sum_j \mathsf{P}(\text{甲输光} \mid S_0 = j)p_{ij} = \frac{1}{2}(h_{i+1} + h_{i-1}), \quad 0 < i < m. \end{aligned}$$

即 $h_{i+1} - h_i = h_i - h_{i-1}$, $0 < i < m$. 所以 $h_i = \dfrac{m-i}{m}$.

(3) 令 $C = \{0, m\}$, 则 T_C 为游戏结束时间. 令 $a_i = \mathsf{E}(T_C \mid S_0 = i)$, 则 $a_0 = 0$, $a_m = 0$, 且

$$\begin{aligned} a_i &= \sum_j \mathsf{E}(T_C \mid S_1 = j, S_0 = i)\mathsf{P}(S_1 = j \mid S_0 = i) \\ &= \sum_j [1 + \mathsf{E}(T_C \mid S_0 = j)]p_{ij} = 1 + \frac{1}{2}(a_{i+1} + a_{i-1}), \quad 0 < i < m. \end{aligned}$$

所以 $a_i = i(m-i)$.

上例告诉我们, 当 Markov 链有多个闭集时, 可以利用 Markov 性和全概率公式, 利用一步分析法建立方程, 计算被某个特定闭集吸收的概率或平均吸收时间.

例 4.6.2 以 X_n (单位: 元) 表示 n 时刻某股票的价格. 设 $\{X_n\}$ 是 Markov 链, 状态空间 $I=\{1,2,3,4\}$, 一步转移矩阵

$$\boldsymbol{P}=\begin{pmatrix}1/2 & 1/2 & 0 & 0\\ 1/3 & 1/3 & 1/3 & 0\\ 0 & 1/4 & 1/4 & 1/2\\ 0 & 0 & 1/2 & 1/2\end{pmatrix}.$$

已知 $\mathsf{P}(X_0=2)=\mathsf{P}(X_0=3)=1/2$, 计算:

(1) 股票价格在涨到 4 元前不曾跌到 1 元的概率;

(2) 股票价格到达 4 元的平均时间.

解 (1) 所求概率为 $\mathsf{P}(T_4<T_1)$, 这个值与到达状态 1 或状态 4 之后的过程没有关系, 所以可以将状态 1 和状态 4 看成吸收态. 令 $h_i=\mathsf{P}(T_4<T_1\mid X_0=i)$, 则 $h_1=0$, $h_4=1$, 且

$$h_2=\frac{1}{3}h_1+\frac{1}{3}h_2+\frac{1}{3}h_3,\quad h_3=\frac{1}{4}h_2+\frac{1}{4}h_3+\frac{1}{2}h_4.$$

可求得 $h_2=\dfrac{2}{5}$, $h_3=\dfrac{4}{5}$, 因此 $\mathsf{P}(T_4<T_1)=\sum\limits_{i=1}^{4}\mathsf{P}(X_0=i)h_i=\dfrac{3}{5}$.

(2) 所求概率为 $\mathsf{E}(T_4)$, 这个值与到达状态 4 之后的过程没有关系, 所以可以将状态 4 看成吸收态. 令 $a_i=\mathsf{E}(T_4\mid X_0=i)$, 则 $a_4=0$, 且

$$a_1=1+\frac{1}{2}a_1+\frac{1}{2}a_2,\quad a_2=1+\frac{1}{3}a_1+\frac{1}{3}a_2+\frac{1}{3}a_3,\quad a_3=1+\frac{1}{4}a_2+\frac{1}{4}a_3+\frac{1}{2}a_4.$$

可求得 $a_1=\dfrac{23}{2}$, $a_2=\dfrac{19}{2}$, $a_3=\dfrac{9}{2}$, 因此 $\mathsf{E}(T_4)=\sum\limits_{i=1}^{4}\mathsf{P}(X_0=i)a_i=7$.

4.7 可逆 Markov 链

定义 4.7.1 (可逆 Markov 链) 假设 $\boldsymbol{X}=(X_n,n\geq 0)$ 是 *Markov* 链, 状态空间为 $\mathcal{E}$, 转移概率矩阵为 $\boldsymbol{P}$, 初始分布为 $\boldsymbol{p}_0$. 假设 $\boldsymbol{X}$ 是不可约 *Markov* 链, 对任意 $N\geq 0$, 令

$$Y_n=X_{N-n},\quad 0\leq n\leq N,$$

如果对任意 $N\geq 1$, 过程 $\boldsymbol{Y}=(Y_n,0\leq n\leq N)$ 是 *Markov* 链, 其中转移概率矩阵为 $\boldsymbol{P}$, 初始分布为 $\boldsymbol{p}_0$, 那么称 $\boldsymbol{X}$ 是**可逆 *Markov* 链**.

显然, 如果 $\boldsymbol{X}$ 是可逆 Markov 链, 那么 $\forall N\geq 1$, $(Y_n,0\leq n\leq N)\overset{d}{=}(X_n,0\leq n\leq N)$.

另外, 如果 $\boldsymbol{X}$ 是可逆 Markov 链, 那么 $X_N\overset{d}{=}X_0$, $\forall N\geq 1$, 所以该 Markov 链一定是平稳的, 初始分布 $\boldsymbol{p}_0$ 为平稳分布.

定理 4.7.1 假设平稳分布 $\boldsymbol{\pi}$ 存在, 选择 $\boldsymbol{\pi}$ 作为初始分布 $\boldsymbol{p}_0$, 则

$$\boldsymbol{X}\ \text{可逆}\iff (X_1,X_0)\overset{d}{=}(X_0,X_1),$$

即 $\pi_i p_{ij}=\pi_j p_{ji}$, $\forall i,j\in\mathcal{E}$.

定义 4.7.2 (可逆分布) 一个概率分布 $\boldsymbol{\pi}$ 称为**可逆分布**如果对所有 i, j 有 $\pi_i p_{ij} = \pi_j p_{ji}$.

定理 4.7.2 可逆分布一定是平稳分布.

另外, 如果对所有状态 i, j, 满足 $\pi_i p_{ij} = \pi_j p_{ji}$, 则 $\boldsymbol{X}$ 就是可逆的 Markov 链.

例 4.7.1 有 A, B 两只容器, 中间有一细管相连. 有 m 只跳蚤, 每次有一只随机地从一个容器跳到另一个容器. 以 X_n 表示 n 次后 A 中跳蚤数, 则 $\{X_n\}$ 是时齐 Markov 链, 转移概率为

$$p_{i,i+1} = \frac{m-i}{m}, \quad p_{i,i-1} = \frac{i}{m}, \quad i = 0, 1, \ldots, m.$$

试问其平稳分布.

解 先尝试求其可逆分布 $\boldsymbol{\pi}$, 则

$$\pi_i p_{i,i-1} = \pi_{i-1} p_{i-1,i}, \quad i = 1, 2, \ldots, m.$$

于是

$$\pi_i = \frac{m-i+1}{i}\pi_{i-1} = \cdots = \frac{(m-i+1)(m-i+2)\cdots m}{i!}\pi_0 = \binom{m}{i}\pi_0, \quad i = 0, 1, \ldots, m.$$

因此 $\pi_i = \dfrac{\binom{m}{i}}{2^m}$, $i = 0, 1, \ldots, m$. 由于 $\{X_n\}$ 不可约, 因此这个可逆分布 $\boldsymbol{\pi}$ 也是唯一的平稳分布.

例 4.7.2 设 $\{X_n\}$ 是平稳的时齐 Markov 链, 状态空间 $I = \{1, 2\}$, 一步转移概率矩阵

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix},$$

问 $\{X_n\}$ 可逆吗?

解 设 $\boldsymbol{\pi} = (\pi_1, \pi_2)$ 是可逆分布, 则

$$\pi_1 p_{12} = \pi_2 p_{21}, \quad 即\ \pi_1 = 0.5\pi_2.$$

又 $\pi_1 + \pi_2 = 1$, 所以 $\boldsymbol{\pi} = \left(\dfrac{1}{3}, \dfrac{2}{3}\right)$, 因此 $\{X_n\}$ 可逆.

例 4.7.3 设 $\{X_n\}$ 是平稳的时齐 Markov 链, 状态空间 $I = \{1, 2, 3\}$, 一步转移概率矩阵

$$\boldsymbol{P} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.1 & 0.7 & 0.2 \\ 0.3 & 0 & 0.7 \end{pmatrix},$$

问 $\{X_n\}$ 可逆吗?

解 设 $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3)$ 是可逆分布, 则

$$\begin{aligned} \pi_2 p_{23} &= \pi_3 p_{32}, \quad 即\ \pi_2 = 0, \\ \pi_1 p_{12} &= \pi_2 p_{21}, \quad 得\ \pi_1 = 0, \\ \pi_1 p_{13} &= \pi_3 p_{31}, \quad 得\ \pi_3 = 0. \end{aligned}$$

所以不存在可逆分布, 因此 $\{X_n\}$ 不可逆.

定理 4.7.3 (Kolmogorov 准则) 假设 $\boldsymbol{X}=(X_n,n\geq 0)$ 是不可约平稳 *Markov* 链, 转移概率矩阵为 $\boldsymbol{P}$. 该 *Markov* 链是可逆的, 当且仅当对任意闭路径 $i_0,i_1,\cdots,i_{N-1},i_N=i_0$, 有

$$p_{i_0i_1}p_{i_1i_2}\cdots p_{i_{N-1}i_0}=p_{i_0i_{N-1}}\cdots p_{i_2i_1}p_{i_1i_0}.$$

根据上述定理也可以判断前面的例子中的 Markov 链是不可逆的.

五、 Galton-Watson 分枝过程

5.1 模型简介

令 ξ 是一个取非负整数值随机变量, 其分布如下:

$$\mathsf{P}(\xi=k)=p_k,\quad k\geq 0,$$

其中 $p_0<1$. 如不加说明, 本章内容中出现的 ξ 均假设服从这一分布.

考虑某物种自然繁殖过程, 设祖先为 Z_0, Z_1 为第一代个体个数, 其分布与 ξ 相同. 第一代 Z_1 个个体独立进行繁殖各自的后代, 并且与祖先繁殖后代的能力一样, 即每个个体繁殖后代的个数为随机变量, 与 ξ 同分布.

令 $(\xi_{1j},j\geq 1)$ 为一列独立同分布随机变量, 与 ξ 同分布且与 Z_1 独立, 则第二代个体总数为

$$Z_2=\sum_{j=1}^{Z_1}\xi_{1j}.$$

类似地, 令 Z_n 为第 n 代个体总数, 并假设 $(\xi_{nj},j\geq 1)$ 为一列独立同分布随机变量, 与 ξ 同分布且与 $\{\xi_{ij};1\leq i<n\}$ 相互独立, 则第 $n+1$ 代个体总数为

$$Z_{n+1}=\sum_{j=1}^{Z_n}\xi_{nj}.$$

定理 5.1.1 $\boldsymbol{Z}=(Z_n,n\geq 0)$ 是 *Markov* 链, 状态空间为 $\mathbb{Z}_+$, 转移概率

$$p_{ij}=P\left(\sum_{l=1}^{i}\xi_l=j\right),\quad i,j\geq 0,$$

其中 $\xi_1,\xi_2,\ldots$ 是一列独立同分布随机变量, 与 ξ 同分布.

例 5.1.1 设 $\mathsf{P}(\xi=1)=p$, $\mathsf{P}(\xi=0)=1-p$, $0<p<1$, 则

$$\mathsf{P}(Z_n=1)=p^n,\quad \mathsf{P}(Z_n=0)=1-p^n.$$

例 5.1.2 设 $\xi\sim B(n,p)$, 则在 $Z_n=i$ 的条件下, $Z_{n+1}=\sum\limits_{l=1}^{i}\xi_{nl}\sim B(ni,p)$, 所以

$$p_{ij}=\binom{ni}{j}p^j(1-p)^{ni-j},\quad j=0,1,\ldots,ni.$$

例 5.1.3 设 $\xi \sim P(\lambda)$, 则在 $Z_n = i$ 的条件下, $Z_{n+1} = \sum_{l=1}^{i} \xi_{nl} \sim P(i\lambda)$, 所以

$$p_{ij} = \frac{(i\lambda)^j}{j!} \mathrm{e}^{-i\lambda}, \quad j = 0, 1, \ldots.$$

例 5.1.4 设 $\xi \sim G(p)$, 即 $\mathsf{P}(\xi = k) = (1-p)^k p$, $k = 0, 1, \ldots$, 则

$$p_{ij} = \mathsf{P}\left(\sum_{l=1}^{i} \xi_{nl} = j\right) = \binom{i+j-1}{i-1} (1-p)^j p^i.$$

一般来说, 求转移概率并不是一件容易的事情, Z_n 的分布律也很难算.

命题 5.1.1 (数字特征) 假设随机变量 ξ 满足 $E(\xi) = \mu$, $Var(\xi) = \sigma^2$, 那么对每个 $n \geq 1$, 有

$$E(Z_n) = \mu^n, \quad Var(Z_n) = \sigma^2 \mu^{n-1} (1 + \mu + \cdots + \mu^{n-1}).$$

这一结果告诉我们, 总体上 Z_n 以几何级数的方式增长 ($\mu > 1$) 或衰减 ($\mu < 1$).

5.2 生成函数

定义 5.2.1 (生成函数) 假设 ξ 是非负整数值随机变量, 定义

$$\phi(s) = E(s^\xi) = \sum_{k=0}^{\infty} p_k s^k, \quad 0 \leq s \leq 1.$$

可以知道 $\phi(s)$ 在 $[0, 1]$ 上收敛, 称 $\phi(s)$ 为随机变量 ξ 的生成函数.

正如特征函数一样, 非负整数值随机变量 ξ 的分布和其生成函数相互唯一确定.

命题 5.2.1 生成函数有许多良好的性质, 如

(1) $\phi_0 = p_0$, $\phi(1) = 1$, $0 \leq \phi(s) \leq 1$;

(2) $\phi(s)$ 在 $[0, 1]$ 上单调递增且一致连续;

(3) 如果 $E(\xi^k) < \infty$, 则

$$E[\xi(\xi-1)\cdots(\xi-k+1)] = \phi^{(k)}(1),$$

特别地, 当 $E(\xi^2) < \infty$ 时, $\phi'(1) = E(\xi)$, $\phi''(1) = E(\xi^2) - E(\xi)$;

(4) $\phi(s)$ 在 $s = 0$ 处无穷次可微, 并且

$$p_k = \frac{\phi^{(k)}(0)}{k!}, \quad \forall k \geq 0,$$

其中约定 $\phi^{(0)}(0) = \phi(0)$;

(5) 假设 ξ, η 为两个独立非负整数值随机变量, 那么 $\xi + \eta$ 的生成函数等于各自生成函数的乘积, 即

$$\phi_{\xi+\eta}(s) = \phi_\xi(s) \phi_\eta(s).$$

该性质可以推广到任意有限个独立随机变量和.

5.3 分枝过程的生成函数

记 ξ 的生成函数为 $\phi(s)$, Z_n 的生成函数为 $\phi_n(s)$, 显然有 $\phi_1(s)=\phi(s)$, 而对 $n\geq 1$, 有

$$\begin{aligned}\phi_{n+1}(s)=\mathsf{E}(s^{Z_{n+1}})=\mathsf{E}\left(s^{\sum\limits_{j=1}^{Z_n}\xi_{nj}}\right)&=\sum_{N=0}^{\infty}\mathsf{E}\left(s^{\sum\limits_{j=1}^{N}\xi_{nj}}\Big|Z_n=N\right)\mathsf{P}(Z_n=N)\\&=\sum_{N=0}^{\infty}\mathsf{E}\left(s^{\sum\limits_{j=1}^{N}\xi_{nj}}\right)\mathsf{P}(Z_n=N)\\&=\sum_{N=0}^{\infty}(\phi_n(s))^N\mathsf{P}(Z_n=N)=\phi(\phi_n(s)).\end{aligned}$$

所以分枝过程的生成函数满足

$$\phi_{n+1}(s)=\phi_n(\phi(s))=\phi(\phi_n(s)),\quad n\geq 1.$$

根据前面的讨论, 可以由 $\phi_n(s)$ 来计算 Z_n 的分布. 当然, 通常情况下 Z_n 的分布并不能详细给出, 因为即使容易算出 $\phi(s)$, 也很难得到 n 次复合函数 $\phi_n(s)$ 的解析表达式.

例 5.3.1 设 $\mathsf{P}(\xi=k)=\dfrac{1}{2^{k+1}}$, $k=0,1,\ldots$. 对 $n\geq 1$, 计算 $\phi_n(s)$ 及 Z_n 的分布律.

$\boxed{\text{解}}$ ξ 的生成函数

$$\phi(s)=\sum_{k=0}^{\infty}\frac{s^k}{2^{k+1}}=\frac{1}{2-s},\quad 0\leq s\leq 1.$$

所以

$$\phi_1(s)=\phi(s)=\frac{1}{2-s},\quad \phi_2(s)=\phi(\phi_1(s))=\frac{1}{2-\frac{1}{2-s}}=\frac{2-s}{3-2s}.$$

若已算得 $\phi_n(s)=\dfrac{n-(n-1)s}{n+1-ns}$, 则

$$\phi_{n+1}(s)=\phi(\phi_n(s))=\frac{1}{2-\frac{n-(n-1)s}{n+1-ns}}=\frac{(n+1)-ns}{(n+2)-(n+1)s}.$$

由归纳法知, 对 $n\geq 1$, $\phi_n(s)=\dfrac{n-(n-1)s}{n+1-ns}$.

注意到

$$\begin{aligned}\phi_n(s)&=\frac{n-(n-1)s}{n+1}\cdot\frac{1}{1-\frac{n}{n+1}s}\\&=\frac{n-(n-1)s}{n+1}\sum_{k=0}^{\infty}\left(\frac{n}{n+1}s\right)^k\\&=\frac{n}{n+1}+\sum_{k=0}^{\infty}\left(\frac{n}{n+1}\frac{n^{k+1}s^{k+1}}{(n+1)^{k+1}}-\frac{n-1}{n+1}\frac{n^ks^{k+1}}{(n+1)^k}\right)\\&=\frac{n}{n+1}+\sum_{k=1}^{\infty}\frac{n^{k-1}}{(n+1)^{k+1}}s^k.\end{aligned}$$

所以 $\mathsf{P}(Z_n=0)=\dfrac{n}{n+1}$, $\mathsf{P}(Z_n=k)=\dfrac{n^{k-1}}{(n+1)^{k+1}}$, $k\geq 1$.

5.4 生存与灭绝概率

Galton-Watson 分枝过程理论研究中, 一个基本问题是: 该物种是否一直生存下去? 最终灭绝的概率是多少? 显然, 如果存在 $n \geq 1$, 使得 $Z_n = 0$, 那么对所有 $m > n$ 都有 $Z_m = 0$, 即该物种完全灭绝.

根据前面有关生成函数的讨论, 物种在第 n 代灭绝的概率为 $\mathsf{P}(Z_n = 0) = \phi_n(0)$. 注意到 $\phi_n(0)$ 是一个有界单调不减数列, 所以极限存在, 记

$$\tau = \lim_{n\to\infty} \phi_n(0).$$

称 τ 为**灭绝概率**. 特别地, 如果 $\tau = 1$, 则该物种最终一定会灭绝.

利用生成函数的性质, 有

$$\phi_n(0) = \phi(\phi_{n-1}(0)),$$

两边取极限得到最终灭绝概率 τ 满足方程

$$\tau = \phi(\tau).$$

定理 5.4.1 *假设 $E(\xi) = \mu$, $\boldsymbol{Z} = (Z_n, n \geq 0)$ 是 Galton-Watson 分枝过程.*
(1) 如果 $\mu \leq 1$, 那么 $\tau = 1$, 即该物种最终一定会灭绝;
(2) 如果 $\mu > 1, 0 < p_0 < 1$, 那么灭绝概率 τ 为方程 $\tau = \phi(\tau)$ 的最小正解, 且 $0 < \tau < 1$;
(3) 如果 $p_0 = 0$, 那么最终灭绝的概率为 0.

例 5.4.1 设 $\mathsf{P}(\xi = k) = \dfrac{1}{2^{k+1}}$, $k = 0, 1, \ldots$, 则 $\phi(s) = \dfrac{1}{2-s}$, $\mu = \phi'(1) = 1$, 因此 $\tau = 1$. 进一步, 令 $T_0 = \min\{n \geq 1 : Z_n = 0\}$ 为首次灭绝的时刻, 计算 T_0 的分布律.

解 已算得 $\alpha_n = \mathsf{P}(Z_n = 0) = \dfrac{n}{n+1}$, 因此对 $n \geq 1$,

$$\mathsf{P}(T_0 = n) = \mathsf{P}(Z_n = 0, Z_{n-1} \neq 0) = \mathsf{P}(Z_n = 0) - \mathsf{P}(Z_{n-1} = 0) = \frac{1}{n(n+1)}.$$

此即首次灭绝时刻的分布律.

例 5.4.2 设 $\mathsf{P}(\xi = k) = (1-p)^k p$, $k = 0, 1, \ldots$, 则 $\phi(s) = \dfrac{p}{1-(1-p)s}$, $\mu = \phi'(1) = \dfrac{1-p}{p}$. 讨论灭绝概率.

解 由方程

$$s = \phi(s) \iff (1-p)s^2 - s + p = 0 \iff (s-1)((1-p)s - p) = 0.$$

所以当 $p \geq \dfrac{1}{2}$ 时, $\tau = 1$; 当 $p < \dfrac{1}{2}$ 时, $\tau = \dfrac{p}{1-p}$.

六、 平稳随机过程遍历性

6.1 时间平均

随机变量的数学期望实际上是一种加权平均: 按照其概率大小进行加权平均.

假设 $X, X_1, X_2, \ldots, X_n, \ldots$ 是一列独立同分布随机变量, $\mathsf{E}|X| < \infty$ 并且 $\mathsf{E}(X) = \mu$, 由 Khinchine 大数律得

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{\mathsf{P}} \mu, \quad n \to \infty.$$

进而由 Kolmogorov 大数律, 上式可加强为几乎处处收敛. 基于这一事实, 有时称数学期望 μ 为**样本平均**或**统计平均**.

下面考虑随机过程的时间平均.

假设 $\boldsymbol{X} = (X_n, n \geq 0)$ 是平稳随机过程, 前 n 个时刻观测值的平均值为

$$\overline{X}_n = \frac{X_0 + X_1 + \cdots + X_{n-1}}{n}.$$

如果存在一个随机变量 τ, 使得 $(\overline{X}_n, n \geq 1)$ 在均方意义下收敛于 τ, 即

$$\lim_{n \to \infty} \mathsf{E}(\overline{X}_n - \tau)^2 = 0,$$

则称 τ 为该随机过程的**时间平均**, 简记为 $\lim\limits_{n \to \infty} \overline{X}_n = \tau$.

假设 $\boldsymbol{X} = (X(t), t \geq 0)$ 是平稳随机过程, 定义 $[0, T]$ 内的过程平均值为

$$\overline{X}_T = \frac{1}{T} \int_0^T X(t) \mathrm{d}t.$$

假设对任何 $T > 0$, 上式积分存在且有限, 则定义其为 $[0, T]$ 上的时间平均. 如果存在一个随机变量 τ, 使得 $(\overline{X}_T, T > 0)$ 在均方意义下收敛于 τ, 即

$$\lim_{T \to \infty} \mathsf{E}(\overline{X}_T - \tau)^2 = 0,$$

则称 τ 为该随机过程的**时间平均**, 简记为 $\lim\limits_{T \to \infty} \overline{X}_T = \tau$.

事实上, 任意给定样本点 ω, $X(\omega, t)$ 是一条样本曲线, 但作为 t 的函数, 它不一定在 $[0, T]$ 内 Riemann 可积. 这里我们采用均方可积的概念: 将 $[0, T]$ 进行划分, 分点为 $0 = t_0 < t_1 < \cdots < t_n = T$, 在每个小区间 $[t_{k-1}, t_k]$ 内任取一点 t_k^*, 作和 $S_n = \sum\limits_{k=1}^{n} X(t_k^*)(t_k - t_{k-1})$.

如果存在一个随机变量 ξ_T(不依赖于上述划分和取点), 使得

$$\lim_{\max\limits_k (t_k - t_{k-1}) \to 0} \mathsf{E}(S_n - \xi_T)^2 = 0,$$

则称 $X(t)$ 在 $[0, T]$ 内均方可积, 积分为 ξ_T, 记为 $\int_0^T X(t) \mathrm{d}t = \xi_T$.

下面定理给出均方可积的充分条件.

定理 6.1.1 假设 $\boldsymbol{X} = (X(t), t \geq 0)$ 是二阶矩过程, $E(X(t))^2 < \infty$. 给定 $T > 0$, 如果

$$\int_0^T \int_0^T E(X(s)X(t)) \mathrm{d}s \mathrm{d}t < \infty,$$

那么 $X(t)$ 在 $[0, T]$ 内均方可积.

推论 6.1.1 如果 $\boldsymbol{X} = (X(t), t \geq 0)$ 是弱平稳过程，那么对任何 $[a, b] \subseteq T$, $X(t)$ 在 $[a, b]$ 上均方可积.

6.2 均值遍历性

定义 6.2.1 令 $\boldsymbol{X} = (X(t), t \in T)$ 是平稳随机过程, $\mathsf{E}[X(t)] = \mu$, 如果时间平均等于样本平均，即

$$\tau = \mu, \quad \text{a.s.},$$

则称 $\boldsymbol{X} = (X(t), t \in T)$ 满足**均值遍历性**.

定理 6.2.1 假设 $\boldsymbol{X} = (X_n, n \geq 0)$ 是离散时间平稳随机过程, $\mathsf{E}(X_n) = \mu$, 那么 $\boldsymbol{X} = (X_n, n \geq 0)$ 满足均值遍历性当且仅当

$$\frac{1}{n^2}\sum_{k=1}^{n}(n-k)(\mathsf{E}(X_0X_k) - \mu^2) \to 0, \quad n \to \infty.$$

推论 6.2.1 在上述定理的假设下, $\boldsymbol{X} = (X_n, n \geq 0)$ 满足均值遍历性当且仅当

$$\frac{1}{n}\sum_{k=1}^{n}(\mathsf{E}(X_0X_k) - \mu^2) \to 0, \quad n \to \infty.$$

推论 6.2.2 在上述定理的假设下，如果

$$\mathsf{E}(X_0X_k) \to \mu^2, \quad k \to \infty,$$

那么 $\boldsymbol{X} = (X_n, n \geq 0)$ 满足均值遍历性.

该推论意味着随机过程渐近不相关. 特别地，如果 $\boldsymbol{X} = (X_n, n \geq 0)$ 是平稳白噪声序列，即 $\mathsf{E}(X_n) = 0$ 且对 $\forall k \geq 1, \mathsf{E}(X_0X_k) = 0$, 那么 $\boldsymbol{X} = (X_n, n \geq 0)$ 满足均值遍历性.

下面考虑连续时间随机过程.

定理 6.2.2 假设 $\boldsymbol{X} = (X(t), t \geq 0)$ 是连续时间平稳随机过程, $\mathsf{E}X(t) = \mu$, 那么 $\boldsymbol{X} = (X(t), t \geq 0)$ 满足均值遍历性当且仅当

$$\frac{1}{T^2}\int_0^T (T-t)(\mathsf{E}(X(0)X(t)) - \mu^2)\mathrm{d}t \to 0, \quad T \to \infty.$$

推论 6.2.3 在上述定理的假设下, $\boldsymbol{X} = (X(t), t \geq 0)$ 满足均值遍历性当且仅当

$$\frac{1}{T}\int_0^T (\mathsf{E}(X(0)X(t)) - \mu^2)\mathrm{d}t \to 0, \quad T \to \infty.$$

推论 6.2.4 在上述定理的假设下，如果

$$\mathsf{E}(X(0)X(t)) \to \mu^2, \quad t \to \infty,$$

那么 $\boldsymbol{X} = (X(t), t \geq 0)$ 满足均值遍历性.

例 6.2.1 判断随机相位余弦波过程 $X(t)=a\cos(\omega t+\Theta)$, $\Theta\sim U(0,2\pi)$ 是否具有均值遍历性.

解 时间平均

$$\tau_X=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}a\cos(\omega t+\Theta)\mathrm{d}t=\lim_{T\to\infty}\frac{a[\sin(\omega T+\Theta)-\sin(-\omega T+\Theta)]}{2T\omega}=0.$$

样本平均

$$\mu_X=\int_0^{2\pi}a\cos(\omega t+\theta)\frac{1}{2\pi}\mathrm{d}\theta=\frac{a[\sin(\omega t+2\pi)-\sin(\omega t)]}{2\pi}=0.$$

因此 $\tau_X=\mu_X$, 具有均值遍历性.

例 6.2.2 考虑随机电报信号 $\mathsf{P}(X(t)=\pm I)=\dfrac{1}{2}$, 而正负号在区间 $(t,t+\tau]$ 内变化的次数服从参数为 $\lambda\tau$ 的 Poisson 分布. 证明: $X(t)$ 是弱平稳过程, 并判断是否有均值遍历性.

解 首先 $\mu_X=\mathsf{E}[X(t)]=0$ 为常数; 其次, 设 $\tau>0$, 则

$$\begin{aligned}\rho(t,t+\tau)&=\mathsf{E}[X(t)X(t+\tau)]\\&=I^2\mathsf{P}(X(t)X(t+\tau)=I^2)-I^2\mathsf{P}(X(t)X(t+\tau)=-I^2)\\&=I^2\left\{\sum_{k=0}^{\infty}\frac{(\lambda\tau)^{2k}}{(2k)!}\mathrm{e}^{-\lambda\tau}-\sum_{k=0}^{\infty}\frac{(\lambda\tau)^{2k+1}}{(2k+1)!}\mathrm{e}^{-\lambda\tau}\right\}\\&=I^2\mathrm{e}^{-\lambda\tau}\sum_{k=0}^{\infty}\frac{(-\lambda\tau)^k}{k!}=I^2\mathrm{e}^{-2\lambda\tau}.\end{aligned}$$

若 $\tau<0$, 则 $\rho(t,t+\tau)=\rho(t+\tau,t)=I^2\mathrm{e}^{2\lambda\tau}$; 另外 $\rho(t,t)=\mathsf{E}[X^2(t)]=I^2$.

综合上述, $\rho(t,t+\tau)=I^2\mathrm{e}^{-2\lambda|\tau|}$ 仅与 τ 有关, 因此 $X(t)$ 是弱平稳过程. 又因为

$$\lim_{\tau\to\infty}\rho(t,t+\tau)=0=\mu_X^2.$$

所以具有均值遍历性.

例 6.2.3 设 $X(t)=\mathrm{e}^{-\frac{\alpha t}{2}}B(\mathrm{e}^{\alpha t})$, 其中 $\alpha>0$, $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 Brown 运动.

(1) 证明: $\boldsymbol{X}=(X(t),t\geq 0)$ 是强平稳过程;

(2) $\boldsymbol{X}$ 具有均值遍历性吗? 为什么?

解 (1) 首先 $\mu_X=\mathsf{E}[X(t)]=\mathrm{e}^{-\frac{\alpha t}{2}}\mathsf{E}[B(\mathrm{e}^{\alpha t})]=0$ 为常数; 其次, 设 $t,\tau\geq 0$, 则

$$\begin{aligned}\rho(t,t+\tau)&=\mathsf{E}[X(t)X(t+\tau)]\\&=\mathrm{e}^{-\frac{\alpha t}{2}}\mathrm{e}^{-\frac{\alpha(t+\tau)}{2}}\mathsf{E}[B(\mathrm{e}^{\alpha t})B(\mathrm{e}^{\alpha(t+\tau)})]\\&=\mathrm{e}^{-\frac{\alpha t}{2}}\mathrm{e}^{-\frac{\alpha(t+\tau)}{2}}\mathrm{e}^{\alpha t}=\mathrm{e}^{-\frac{\alpha\tau}{2}}.\end{aligned}$$

所以对所有 τ, $\rho(t,t+\tau)=\mathrm{e}^{-\frac{\alpha|\tau|}{2}}$ 仅与 τ 有关, 因此 $\boldsymbol{X}$ 是弱平稳过程. 又由于它是正态过程, 所以它也是强平稳过程.

(2) 因为

$$\lim_{\tau\to\infty}\rho(t,t+\tau)=0=\mu_X^2.$$

所以具有均值遍历性.

例 6.2.4 设 X_0 具有密度函数

$$f(x)=\begin{cases}2x, & 0\le x\le 1,\\ 0, & \text{其他}.\end{cases}$$

对 $n\ge 0$, 在已知 $X_0,\dots,X_n$ 的条件下, X_{n+1} 服从 $(1-X_n,1]$ 上的均匀分布.

(1) 证明: $\{X_n;n=0,1,\dots\}$ 是弱平稳过程;

(2) 判断当 $N\to\infty$ 时, $\frac{1}{N}\sum_{i=1}^{N}X_i$ 是否依概率收敛? 如果收敛, 收敛到何值?

$\boxed{解}$ (1) 首先有 $\mathsf{E}(X_0)=\int_0^1 x(2x)\mathrm{d}x=\frac{2}{3}$, $\mathsf{E}(X_0^2)=\int_0^1 (x^2)(2x)\mathrm{d}x=\frac{1}{2}$. 对 $n\ge 0$,

$$\mathsf{E}(X_{n+1}\mid X_0,X_1,\dots,X_n)=\frac{1}{2}(1-X_n+1)=1-\frac{X_n}{2},$$
$$\mathsf{E}(X_{n+1}^2\mid X_0,X_1,\dots,X_n)=\frac{1}{12}X_n^2+\left(1-\frac{X_n}{2}\right)^2=\frac{X_n^2}{3}-X_n+1.$$

于是

$$\mathsf{E}(X_{n+1})=\mathsf{E}[\mathsf{E}(X_{n+1}\mid X_0,X_1,\dots,X_n)]=1-\frac{\mathsf{E}(X_n)}{2},$$
$$\mathsf{E}(X_{n+1}^2)=\mathsf{E}[\mathsf{E}(X_{n+1}^2\mid X_0,X_1,\dots,X_n)]=\frac{\mathsf{E}(X_n^2)}{3}-\mathsf{E}(X_n)+\frac{1}{3}.$$

由归纳法可知, $\mathsf{E}(X_n)=\frac{2}{3}$, $\mathsf{E}(X_n^2)=\frac{1}{2}$ 对所有 n 成立.

对 $n,m\ge 0$, 因为

$$\mathsf{E}(X_nX_{n+m+1}\mid X_0,X_1,\dots,X_{n+m})=X_n\left(1-\frac{X_{n+m}}{2}\right),$$

所以

$$\mathsf{E}(X_nX_{n+m+1})=\frac{2}{3}-\frac{1}{2}\mathsf{E}(X_nX_{n+m}).$$

由归纳法可知,

$$\rho(n,n+m)=\mathsf{E}(X_nX_{n+m})=\frac{4}{9}+\frac{(-1)^m}{18\times 2^m},\quad \forall m\ge 0.$$

所以 $\{X_n;n=0,1,\dots\}$ 是弱平稳过程;

(2) 当 $m\to\infty$ 时,

$$\lim_{m\to\infty}\rho(n,n+m)=\frac{4}{9}=\mu_X^2,$$

因此具有均值遍历性. 从而当 $N\to\infty$ 时, $\frac{1}{N}\sum_{i=1}^{N}X_i\xrightarrow{\mathsf{P}}\mu_X=\frac{2}{3}$.

七、 Brown 运动

7.1 Brown 运动及基本性质

定义 7.1.1 (Brown 运动) 令 $\boldsymbol{B}=(B(t),t\ge 0)$ 是实数值随机过程, 如果满足

1° 初始值: $B(0)=0$;

2° 增量独立: 假设 $0<t_1<t_2<\cdots<t_k$, 那么 $B(t_1),B(t_2)-B(t_1),\dots,B(t_k)-B(t_{k-1})$ 相互独立;

3° 增量平稳: 假设 $s<t$, 那么 $B(t)-B(s)$ 与 $B(t-s)$ 同分布, 即分布仅依赖于 $t-s$;

4° 正态分布: 对任何 $t>0$, 有 $B(t)\sim N(0,\sigma^2t)$,

则称 $\boldsymbol{B}=(B(t),t\ge 0)$ 是参数为 σ^2 的 ***Brown* 运动**. 当 $\sigma^2=1$ 时, 称为**标准 *Brown* 运动**.

以下总假定 $\boldsymbol{B}=(B(t),t\geq 0)$ 为标准 Brown 运动.

命题 7.1.1 (数字特征) *设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 Brown 运动, 则*
(1) 均值函数 $\mu(t)=\mathsf{E}[B(t)]=0,\ t\geq 0$;
(2) 方差函数 $\sigma^2(t)=\mathsf{Var}[B(t)]=t,\ t\geq 0$;
(3) $m(\geq 1)$ 阶矩 $\mathsf{E}[B(t)^m]=\begin{cases} t^k(2k-1)!!, & m=2k,\\ 0, & m=2k+1;\end{cases}$
(4) 协方差函数 $r(s,t)=\mathsf{Cov}(B(s),B(t))=\min\{s,t\}\triangleq s\wedge t$.

例 7.1.1 设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 Brown 运动, 求:
(1) $B(1)+3B(2)$ 的分布;
(2) $\mathsf{Cov}(B(1)+B(3),B(3)-B(2))$;
(3) $\mathsf{P}(B(7)\leq 3\mid B(1)=1,B(3)=2)$.

解 (1) 注意到 $B(1)+3B(2)=4B(1)+3[B(2)-B(1)]$, 而 $\boldsymbol{B}$ 增量独立, 因此

$$B(1)+3B(2)\sim N(0,25).$$

或直接计算其方差为

$$\mathsf{Var}(B(1)+3B(2))=\mathsf{Var}(B(1))+9\mathsf{Var}(B(2))+6\mathsf{Cov}(B(1),B(2))=25.$$

(2)

$$\begin{aligned}\mathsf{Cov}(B(1)+B(3),B(3)-B(2)) &= \mathsf{Cov}(B(1)+B(2)+[B(3)-B(2)],B(3)-B(2))\\ &= \mathsf{Var}[B(3)-B(2)]=1.\end{aligned}$$

或直接计算得

$$\mathsf{Cov}(B(1)+B(3),B(3)-B(2))=\mathsf{Cov}(B(1),B(3))-\mathsf{Cov}(B(1),B(2))+\mathsf{Var}(B(3))-\mathsf{Cov}(B(3),B(2))=1.$$

(3)

$$\begin{aligned}\mathsf{P}(B(7)\leq 3\mid B(1)=1,B(3)=2) &= \mathsf{P}(B(7)-B(3)\leq 1\mid B(1)=1,B(3)=2)\\ &= \mathsf{P}(B(7)-B(3)\leq 1)=\Phi\left(\frac{1}{2}\right).\end{aligned}$$

这里用到了 Brown 运动的 Markov 性.

Brown 运动的样本轨道极其特殊, 具有以下性质.

命题 7.1.2 *设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 Brown 运动, 则*
(1) 样本轨道几乎处处连续

$$P\big(\omega\in\Omega: B(\omega,t)\ \text{在}\ [0,+\infty)\ \text{上连续}\big)=1.$$

(2) 样本轨道几乎无处可导

$$|B(t)-B(s)|\asymp|t-s|^{\frac{1}{2}},$$

其中 $\asymp$ 表示近似. 所以 Brown 运动样本轨道极其不正则.

(3) *样本轨道的刻画：通常采用简单随机游动的折线去近似 Brown 运动的样本轨道.*

考虑 $T=[0,1]$, $\boldsymbol{B}=(B(t),0\le t\le 1)$ *是连续时间取实值的 Brown 运动. 令* $n\ge 1$, $S_0,S_1,\cdots,S_n$ *为简单随机游动，构造* $[0,1]$ *上连续函数* $X_n(t)$ *如下：*

$$X_n(t)=\begin{cases}\dfrac{S_k}{\sqrt{n}}, & t=\dfrac{k}{n},\\ \text{线性插值}, & \dfrac{k}{n}<t<\dfrac{k+1}{n}.\end{cases}$$

则 $\boldsymbol{X}_n=(X_n(t),0\le t\le 1)$ *是* $[0,1]$ *区间上的折线，称为部分和过程. 可以证明，当* $n\to\infty$ *时，有*

$$(X_n(t),0\le t\le 1)\Rightarrow(B(t),0\le t\le 1),$$

其中 "⇒" 表示过程依分布收敛.

下面命题给出了 Brown 运动的分布规律.

命题 7.1.3 *设* $\boldsymbol{B}=(B(t),t\ge 0)$ *是标准 Brown 运动，则*

(1) ***1-维分布***: $B(t)\sim N(0,t)$, *密度函数为* $p(t,x)=\dfrac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{x^2}{2t}}$, $x\in\mathbb{R}$.

(2) ***2-维分布***: *假设* $s<t$, *有* $(B(s),B(t))\sim N\left(0,s;0,t;\left(\frac{s}{t}\right)^{\frac{1}{2}}\right)$, *矩阵形式为*

$$(B(s),B(t))\sim N\left((0,0),\begin{pmatrix} s & s\wedge t\\ s\wedge t & t\end{pmatrix}\right).$$

在给定 $B(s)=x$ *的条件下，*$B(t)$ *的条件密度函数为*

$$p_{t|s}(y|x)=P(B(t)=y\mid B(s)=x)=\frac{1}{\sqrt{2\pi(t-s)}}\mathrm{e}^{-\frac{(y-x)^2}{2(t-s)}},\quad y\in\mathbb{R}.$$

在给定 $B(t)=y$ *的条件下，*$B(s)$ *的条件密度函数为*

$$p_{s|t}(x|y)=P(B(s)=x\mid B(t)=y)=\frac{\sqrt{t}}{\sqrt{2\pi s(t-s)}}\mathrm{e}^{-\frac{(tx-sy)^2}{2st(t-s)}},\quad x\in\mathbb{R}.$$

(3) ***n-维分布***

$$(B(t_1),B(t_2),\cdots,B(t_n))\sim N\left((0,0,\cdots,0),\begin{pmatrix} t_1 & t_1\wedge t_2 & \cdots & t_1\wedge t_n\\ t_1\wedge t_2 & t_2 & \cdots & t_2\wedge t_n\\ \vdots & \vdots & \ddots & \vdots\\ t_n\wedge t_1 & t_2\wedge t_n & \cdots & t_n\end{pmatrix}\right).$$

这表明 Brown 运动是正态过程，均值为 0, 任意两个时刻的协方差为 $s\wedge t$.

定理 7.1.1 *样本轨道连续的随机过程* $\boldsymbol{B}=(B(t),t\ge 0)$ *是 Brown 运动当且仅当它是正态过程，且满足* $E(B(t))=0$, $E(B(t)B(s))=s\wedge t$.

推论 7.1.1 假设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 Brown 运动, 则下列三个过程也是标准 Brown 运动:

(1) 给定 $t_0\geq 0$, $\widetilde{B}(t)=B(t+t_0)-B(t_0)$, $t\geq 0$;

(2) 给定常数 $a\neq 0$, $\widetilde{B}(t)=\dfrac{1}{a}B(a^2t)$, $t\geq 0$;

(3) $\widetilde{B}(t)=\begin{cases} tB(t^{-1}), & t>0,\\ 0, & t=0.\end{cases}$

例 7.1.2 设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 Brown 运动, 求 $\mathsf{P}(B(0.5)\leq 1\mid B(1)=1,B(2)=2)$.

解 由 Brown 运动的性质, $\widetilde{\boldsymbol{B}}=(\widetilde{B}(t),t\geq 0)$ 是标准 Brown 运动, 其中 $B(t)=t\widetilde{B}(t^{-1})$. 所以

$$\begin{aligned}\mathsf{P}(B(0.5)\leq 1\mid B(1)=1,B(2)=2)&=\mathsf{P}(0.5\widetilde{B}(2)\leq 1\mid \widetilde{B}(1)=1,2\widetilde{B}(0.5)=2)\\&=\mathsf{P}(\widetilde{B}(2)\leq 2\mid \widetilde{B}(1)=1,\widetilde{B}(0.5)=1)\\&=\mathsf{P}(\widetilde{B}(2)-\widetilde{B}(1)\leq 1\mid \widetilde{B}(1)=1,\widetilde{B}(0.5)=1)\\&=\mathsf{P}(\widetilde{B}(2)-\widetilde{B}(1)\leq 1)=\Phi(1)=0.8413.\end{aligned}$$

这里用到了 0 与 ∞ 的对称性.

7.2 与 Brown 运动相关的过程

(i) Brown 桥

定义 7.2.1 设 $B^0(t)=B(t)-tB(1)$, 称 $\boldsymbol{B}^0=(B^0(t),0\leq t\leq 1)$ 为 ***Brown* 桥过程**.

边界 $B^0(0)=B^0(1)=0$. 可以证明 $\boldsymbol{B}^0$ 是均值为 0 的正态过程, 自相关函数为

$$\mathsf{E}(B^0(s)B^0(t))=s\wedge t(1-s\vee t),\quad 0\leq s,t\leq 1.$$

于是 1-维分布 $B^0(t)\sim N(0,t(1-t))$.

条件分布: 在 $B(1)=0$ 的条件下, $B(t)$ 的分布密度函数为

$$p_{t|1}(x|0)=\frac{1}{\sqrt{2\pi t(1-t)}}\mathrm{e}^{-\frac{x^2}{2t(1-t)}}.$$

这表明 $B^0(t)\overset{d}{=}(B(t)\mid B(1)=0)$.

(ii) 反射 Brown 运动

定义 7.2.2 设 $X(t)=|B(t)|$, 称 $\boldsymbol{X}=(X(t),t\geq 0)$ 为**反射 *Brown* 运动**.

反射 Brown 运动不再是一个正态过程, 其数字特征

$$\mathsf{E}[X(t)]=\sqrt{\frac{2t}{\pi}},\quad \mathsf{E}[X^2(t)]=t,\quad t\geq 0.$$

给定 $t>0$, 其分布

$$\mathsf{P}(X(t)\leq x)=2\Phi\Big(\frac{x}{\sqrt{t}}\Big)-1,\quad p(t,x)=\sqrt{\frac{2}{\pi t}}\mathrm{e}^{-\frac{x^2}{2t}},\quad x\geq 0.$$

(iii) 几何 Brown 运动

定义 7.2.3 设 $\alpha,\beta\in\mathbb{R}$, 记 $X(t)=\mathrm{e}^{\alpha t+\beta B(t)}$, 称 $\boldsymbol{X}=(X(t),t\geq 0)$ 是**几何 Brown 运动**.

几何 Brown 运动不再是一个正态过程, 其数字特征

$$\mathsf{E}[X(t)]=\mathrm{e}^{\alpha t+\beta^2 t/2},\quad \mathsf{E}[X^2(t)]=\mathrm{e}^{2\alpha t+2\beta^2 t},\quad t\geq 0.$$

给定 $t>0$, $X(t)$ 的密度函数

$$p(t,x)=\frac{1}{\beta x\sqrt{2\pi t}}\mathrm{e}^{-\frac{(\ln x-\alpha t)^2}{2\beta^2 t}},\quad x>0.$$

(iv) 积分过程

对 Brown 运动而言, 几乎处处每一条样本曲线都连续, 即存在一个零概率事件 Ω_0, 对每一个 $\omega\in\Omega\backslash\Omega_0$, $B(\omega,t)$ 作为 t 的函数在 $[0,+\infty)$ 连续. 定义

$$X(\omega,t)=\begin{cases}\displaystyle\int_0^t B(\omega,s)\mathrm{d}s, & \omega\in\Omega\backslash\Omega_0,\\ 0, & \omega\in\Omega_0.\end{cases}$$

这个积分是一个 Riemann 积分, 它存在并有限.

定义 7.2.4 设 $X(t)=\displaystyle\int_0^t B(s)\mathrm{d}s$, 称 $\boldsymbol{X}=(X(t),t\geq 0)$ 为**积分过程**.

积分过程是一个正态过程, 这是因为

$$\int_0^t B(s)\mathrm{d}s=\lim_{n\to\infty}\sum_{k=1}^n (s_i-s_{i-1})B(s_i),$$

而 $\boldsymbol{B}$ 是一个正太过程. 积分过程 $\boldsymbol{X}$ 的数字特征

$$\mathsf{E}[X(t)]=0,\ t\geq 0;\quad \mathsf{Cov}(X(s),X(t))=\frac{s^2 t}{2}-\frac{s^3}{6},\ s\leq t.$$

7.3 最大值与首中时分布

定义 7.3.1 (首中时) 设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 *Brown* 运动, $a\neq 0$, 令

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

表示首次击中 a 的时刻, 称为 a 的**首中时**.

定理 7.3.1 (反射原理) 固定实数 a, 令

$$\widetilde{B}(t)=\begin{cases}B(t), & t<T_a,\\ 2a-B(t), & t\geq T_a,\end{cases}$$

则 $\widetilde{\boldsymbol{B}}=(\widetilde{B}(t),t\geq 0)$ 也是 *Brown* 运动.

证明 对 $t \geq 0$, 令

$$Y(t) = B(t)\mathbf{1}_{\{t \leq T_a\}}, \quad Z(t) = B(t+T_a) - a.$$

则 $\boldsymbol{Z} = (Z(t), t \geq 0)$ 是独立于 $\boldsymbol{Y} = (Y(t), t \geq 0)$ 的 Brown 运动. 因而 $-\boldsymbol{Z}$ 也是独立于 $\boldsymbol{Y}$ 的 Brown 运动, 所以 $(\boldsymbol{Y}, \boldsymbol{Z})$ 与 $(\boldsymbol{Y}, -\boldsymbol{Z})$ 具有相同的有限维分布.

$$\varphi : (\boldsymbol{Y}, \boldsymbol{Z}) \to \{Y(t)\mathbf{1}_{t \leq T_a} + (a + Z(t - T_a))\mathbf{1}_{t > T_a}, t \geq 0\}$$

生成一个连续过程, $\varphi(\boldsymbol{Y}, -\boldsymbol{Z})$ 也是一个连续过程, 且两者具有相同的有限维分布. 又 $\varphi(\boldsymbol{Y}, \boldsymbol{Z}) = \boldsymbol{B}$, $\varphi(\boldsymbol{Y}, -\boldsymbol{Z}) = \widetilde{\boldsymbol{B}}$, 所以 $\widetilde{\boldsymbol{B}}$ 也是 Brown 运动. □

命题 7.3.1 (最大值分布) 设 $\boldsymbol{B} = (B(t), t \geq 0)$ 是标准 *Brown* 运动, 定义

$$M(t) = \sup_{0 \leq s \leq t} B(t), \quad t \geq 0,$$

那么 $M(t) \stackrel{d}{=} |B(t)|$.

证明 当 $x < 0$ 时, $\mathsf{P}(M(t) \leq x) = 0$. 首先由反射定理知, 对 $\forall a, y, t \geq 0$,

$$\mathsf{P}(M(t) \geq a, B(t) \leq a - y) = \mathsf{P}(B(t) \geq a + y).$$

当 $x \geq 0$ 时, 考虑

$$\begin{aligned} \mathsf{P}(M(t) > x) &= \mathsf{P}(M(t) > x, B(t) > x) + \mathsf{P}(M(t) > x, B(t) < x) + \mathsf{P}(M(t) > x, B(t) = x) \\ &= \mathsf{P}(B(t) > x) + \mathsf{P}(B(t) > x) + 0 = 2\mathsf{P}(B(t) > x) = \mathsf{P}(|B(t)| > x). \end{aligned}$$

因此 $M(t) \stackrel{d}{=} |B(t)|$. □

注意到 $T_a \leq t \Leftrightarrow M(t) \geq a$, 因此首中时 T_a 的密度函数为

$$p_{T_a}(t) = \frac{|a|}{\sqrt{2\pi t^3}} \mathrm{e}^{-\frac{a^2}{2t}}, \quad t > 0.$$

更进一步, 给定 $a \in \mathbb{R}$, 有 $\mathsf{P}(T_a < \infty) = 1$, $\mathsf{E}(T_a) = \infty$. 这一结论告诉我们, $|a|$ 无论多大, 从 0 点出发的 Brown 运动总会在有限时间内到达 a; 另一方面, $|a|$ 无论多么靠近 0, Brown 运动到达 a 所需要的平均时间为 ∞.

定理 7.3.2 令 $a < 0 < b$, 那么

$$\mathsf{P}(T_a < T_b) = \frac{b}{b-a}, \quad \mathsf{P}(T_a > T_b) = \frac{-a}{b-a}.$$

例 7.3.1 设 $\boldsymbol{B} = (B(t), t \geq 0)$ 是标准 Brown 运动. 设 $X(t) = \left|\min\limits_{0 \leq s \leq t} B(s)\right|$, 求 $X(t)$ 的分布函数.

解 注意到

$$X(t) = \left|\min_{0 \leq s \leq t} B(s)\right| = -\min_{0 \leq s \leq t} B(s) = \max_{0 \leq s \leq t} B_1(s),$$

其中 $B_1(s) = -B(s)$. 因此, 当 $y > 0$ 时,

$$\mathsf{P}(X(t) \leq y) = 1 - \mathsf{P}\big(\max_{0 \leq s \leq t} B_1(s) > y\big) = 1 - 2\mathsf{P}(B_1(t) > y) = 2\Phi\Big(\frac{y}{\sqrt{t}}\Big) - 1.$$

所以 $F_{X(t)}(y)=\begin{cases}2\Phi\left(\dfrac{y}{\sqrt{t}}\right)-1, & y>0,\\ 0, & y\leqslant 0.\end{cases}$

例 7.3.2 以 $X(t)$ 表示 t 时刻的股票价格 (单位: 元). 设 $X(t)=2^{B(t)}$, 其中 $\{B(t);t\geq 0\}$ 是标准 Brown 运动. 求 $[0,4]$ 内股票价格不曾达到 8 元的概率.

解 由题意

$$\mathsf{P}(\max_{0\leq t\leq 4} X(t)<8)=\mathsf{P}(\max_{0\leq t\leq 4} B(t)<3)=1-2\mathsf{P}(B(4)\geq 3)=2\Phi(1.5)-1=0.8664.$$

注意这里利用了几何 Brown 运动与 Brown 运动的单调性.

7.4 随机积分 (Itô 积分)

(i) 有界变差与无界变差

定义 7.4.1 (有界变差) 假设 G 是 $[a,b]$ 上的实值函数. 对 $[a,b]$ 进行分割 $\Delta: a=x_0<x_1<\cdots<x_n=b$, 如果

$$\sup_{\Delta}\sum_{i=1}^{n}|G(x_i)-G(x_{i-1})|<\infty,$$

则称 G 在 $[a,b]$ 上具有**有界变差**.

假设 G 是 $[0,+\infty)$ 上的实值函数. 如果

$$\sup_{t>0}\sup_{\Delta}\sum_{i=1}^{n}|G(x_i)-G(x_{i-1})|<\infty,$$

其中 Δ 是 $[0,t]$ 上的分割, 则称 G 在 $[0,+\infty)$ 上具有**有界变差**.

可以证明, 任何有界变差函数都可以写成两个单调增函数的差, 都几乎处处可微; 另外, $[a,b]$ 上具有有界导函数的函数是有界变差函数, $[a,b]$ 上的单调函数是有界变差函数.

命题 7.4.1 假设 f 是 $[a,b]$ 上的连续函数, G 是 $[a,b]$ 上的有界变差函数, 那么 *Riemann-Stieltjes* 积分 $\displaystyle\int_a^b f(x)\mathrm{d}G(x)$ 存在.

如果对所有连续函数 f, 积分 $\displaystyle\int_a^b f(x)\mathrm{d}G(x)$ 都存在, 则 G 一定是 $[a,b]$ 上的有界变差函数.

下面定理表明标准 Brown 运动 $\boldsymbol{B}=(B(t),t\geq 0)$ 是无界变差函数.

定理 7.4.1 仅考虑 $t\in[0,1]$. 将该区间进行 2^n 等分, 则和式

$$S_n=\sum_{i=1}^{2^n}\left|B\left(\frac{i}{2^n}\right)-B\left(\frac{i-1}{2^n}\right)\right|\to\infty,\quad \text{a.s.},\quad n\to\infty.$$

证明 注意到

$$\mathsf{E}(S_n)=\sum_{i=1}^{2^n}\mathsf{E}\left|B\left(\frac{i}{2^n}\right)-B\left(\frac{i-1}{2^n}\right)\right|=2^n\mathsf{E}\left|N\left(0,\frac{1}{2^n}\right)\right|=2^n\left(\frac{2}{\pi 2^n}\right)^{1/2}.$$

$$\mathsf{Var}(S_n)=\sum_{i=1}^{2^n}\mathsf{Var}\left|B\left(\frac{i}{2^n}\right)-B\left(\frac{i-1}{2^n}\right)\right|=2^n\mathsf{Var}\left|N\left(0,\frac{1}{2^n}\right)\right|=2^n\frac{t}{2^n}\left(1-\frac{2}{\pi}\right)=\left(1-\frac{2}{\pi}\right)t.$$

所以
$$\mathsf{P}\left(S_n < \frac{\mathsf{E}(S_n)}{2}\right) \le \mathsf{P}\left(|S_n - \mathsf{E}(S_n)| \ge \frac{\mathsf{E}(S_n)}{2}\right) \le \frac{4\mathsf{Var}(S_n)}{[\mathsf{E}(S_n)]^2} \to 0.$$

所以 $S_n \xrightarrow{\mathsf{P}} \infty$, 又 S_n 单调递增, 所以 $S_n \to \infty$, a.s.. □

(ii) 随机积分

假设 $\boldsymbol{f} = (f(t), t \ge 0)$ 是非随机、有界变差函数, 定义

$$\int_0^t f(s)\mathrm{d}B(s) = f(t)B(t) - \int_0^t B(s)\mathrm{d}f(s),$$

其中 $\int_0^t B(s)\mathrm{d}f(s)$ 是 Brown 运动关于 f 的 Riemann-Stieltjes 积分.

假设 $\boldsymbol{f} = (f(t), t \ge 0)$ 是随机过程, 如何定义 $\int_0^t f(s)\mathrm{d}B(s)$?

由于 Brown 运动具有无界变差, 按 Riemann-Stieltjes 积分方式定义可能会引起麻烦. 以 $f(s) = B(s)$ 为例, 可以证明, 将 $[0, t]$ 进行划分, 分点记为 $0 = t_0 < t_1 < \cdots < t_n = t$, 在均方收敛的意义下, 取左端点计算积分值为 $\big(B(t)^2 - t\big)/2$, 取右端点计算积分值为 $\big(B(t)^2 + t\big)/2$.

(iii) Itô 积分

假设 $\boldsymbol{B} = (B(t), t \ge 0)$ 是标准 Brown 运动, $\boldsymbol{f} = (f(t), t \ge 0)$ 是一个关于 $\boldsymbol{B}$ 适应的随机过程, 即 $f(t)$ 仅依赖于 $\{B(s), 0 \le s \le t\}$. 进一步假设

$$\int_0^\infty \mathsf{E}[f(s)^2]\mathrm{d}s < \infty.$$

对任意 $t > 0$, 将 $[0, t]$ 进行划分, 分点记为 $0 = s_0 < s_1 < \cdots < s_n = t$, 使得 $\Delta = \max\limits_{1 \le i \le n} |s_i - s_{i-1}| \to 0$, 选择 $\xi = s_{i-1}$(区间的左端点), 作 Riemann-Stieltjes 和

$$S_n(t) = \sum_{i=1}^n f(s_{i-1})\big(B(s_i) - B(s_{i-1})\big).$$

一定存在一个随机变量 $X(t)$, 使得 $\mathsf{E}\big(S_n(t) - X(t)\big)^2 \to 0,\ n \to \infty$.

定义 $\int_0^t f(s)\mathrm{d}B(s) = X(t)$, 称 $X(t)$ 是 f 关于 $\boldsymbol{B}$ 的 Itô 积分.

下面命题给出了 Itô 积分运算性质.

命题 7.4.2 *设 $X(t) = \int_0^t f(s)\mathrm{d}B(s)$ 是 f 关于 $\boldsymbol{B}$ 的 Itô 积分, 则*

(1) 数学期望: $\mathsf{E}[X(t)] = 0$;

(2) 方差: $\mathsf{Var}[X(t)] = \mathsf{E}[X^2(t)] = \int_0^t \mathsf{E}[f^2(s)]\mathrm{d}s$;

(3) 线性性: $\int_0^t \big(af(s) + bg(s)\big)\mathrm{d}B(s) = a\int_0^t f(s)\mathrm{d}B(s) + b\int_0^t g(s)\mathrm{d}B(s)$;

(4) 可加性: $\int_0^T f(s)\mathrm{d}B(s) = \int_0^t f(s)\mathrm{d}B(s) + \int_t^T f(s)\mathrm{d}B(s)$.

(iv) Itô 公式

前面给出了 Itô 积分的定义计算, 但通常需要一个基本公式帮助计算.

引理 7.4.1 假设 $f:\mathbb{R}\mapsto\mathbb{R}$ 是一个二次可微分函数, 那么

$$f\big(B(t)\big)=f\big(B(0)\big)+\int_0^t f'\big(B(s)\big)\mathrm{d}B(s)+\frac{1}{2}\int_0^t f''\big(B(s)\big)\mathrm{d}s.$$

需要指出的是, 该公式与 Newton-Lebinitz 公式相比多了最后一个二阶导数项. 这是由于 Brown 运动无界变差引起的, 称作变差项.

由上面的引理, 可以得到 Itô 积分的一个计算公式:

$$\int_0^t f'\big(B(s)\big)\mathrm{d}B(s)=f\big(B(t)\big)-f\big(B(0)\big)-\frac{1}{2}\int_0^t f''\big(B(s)\big)\mathrm{d}s.$$

例 7.4.1 计算 $\displaystyle\int_0^t B(s)\mathrm{d}B(s)$.

解 由上述计算公式, $\displaystyle\int_0^t B(s)\mathrm{d}B(s)=\frac{1}{2}B^2(t)-\frac{1}{2}\int_0^t 1\mathrm{d}s=\frac{B^2(t)-t}{2}$.

例 7.4.2 计算 $\displaystyle\int_0^t \cos B(s)\mathrm{d}B(s)$.

解 由上述计算公式, $\displaystyle\int_0^t \cos B(s)\mathrm{d}B(s)=\sin B(t)+\frac{1}{2}\int_0^t \sin B(s)\mathrm{d}s$.

例 7.4.3 计算 $\displaystyle\int_0^t \mathrm{e}^{B(s)}\mathrm{d}B(s)$.

解 由上述计算公式, $\displaystyle\int_0^t \mathrm{e}^{B(s)}\mathrm{d}B(s)=\mathrm{e}^{B(t)}-1-\frac{1}{2}\int_0^t \mathrm{e}^{B(s)}\mathrm{d}s$.

引理 7.4.2 假设 $f(t,x)$ 是二元函数, 关于 t 和 x 的二阶偏导数存在并且连续, 那么

$$f\big(t,B(t)\big)=f(0,0)+\int_0^t\frac{\partial}{\partial s}f\big(s,B(s)\big)\mathrm{d}s+\int_0^t\frac{\partial}{\partial x}f\big(s,B(s)\big)\mathrm{d}B(s)+\frac{1}{2}\int_0^t\frac{\partial^2}{\partial x^2}f\big(s,B(s)\big)\mathrm{d}s.$$

例如, 对几何 Brown 运动有

$$f\big(t,B(t)\big)=\mathrm{e}^{(\mu-\frac{\sigma^2}{2})t+\sigma B(t)}=1+\mu\int_0^t f\big(s,B(s)\big)\mathrm{d}s+\sigma\int_0^t f\big(s,B(s)\big)\mathrm{d}B(s).$$

(v) Itô 过程

定义 7.4.2 假设 $\boldsymbol{B}=(B(t),t\geq 0)$ 是标准 *Brown* 运动, $a(t,x)$ 和 $b(t,x)$ 是两个二元函数, 称满足下列方程

$$X(t)=X(0)+\int_0^t a\big(s,X(s)\big)\mathrm{d}s+\int_0^t b\big(s,X(s)\big)\mathrm{d}B(s)$$

的随机过程 $\boldsymbol{X}=(X(t),t\geq 0)$ 为 *Itô* 扩散过程, 称 $a(t,x)$ 为漂移系数, $b(t,x)$ 为扩散系数.

Itô 扩散过程 $\boldsymbol{X}$ 通常简记为如下的随机微分方程

$$\mathrm{d}X(t) = a\big(t, X(t)\big)\mathrm{d}t + b\big(t, X(t)\big)\mathrm{d}B(t).$$

引理 7.4.3 假设 $\boldsymbol{X} = (X(t), t \geq 0)$ 是 *Itô* 扩散过程，具有漂移系数 $a(t,x)$ 和扩散系数 $b(t,x)$; $f(t,x)$ 是二元函数，关于 t 和 x 的二阶偏导数存在并且连续，那么

$$\begin{aligned} f\big(t, X(t)\big) = f\big(0, X(0)\big) &+ \int_0^t \frac{\partial}{\partial s} f\big(s, X(s)\big)\mathrm{d}s + \int_0^t \frac{\partial}{\partial x} f\big(s, X(s)\big)\mathrm{d}X(s) \\ &+ \frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2} f\big(s, X(s)\big) b^2\big(s, X(s)\big)\mathrm{d}s \end{aligned}$$

命题 7.4.3 (分部积分公式) 假设 $\boldsymbol{X} = (X(t), t \geq 0), \boldsymbol{Y} = (Y(t), t \geq 0)$ 是两个 *Itô* 扩散过程：

$$\begin{aligned} \mathrm{d}X(t) &= a_1\big(t, X(t)\big)\mathrm{d}t + b_1\big(t, X(t)\big)\mathrm{d}B(t) \\ \mathrm{d}Y(t) &= a_2\big(t, Y(t)\big)\mathrm{d}t + b_2\big(t, Y(t)\big)\mathrm{d}B(t) \end{aligned}$$

那么 $\boldsymbol{Z} = (Z(t) = X(t)Y(t), t \geq 0)$ 也是 *Itô* 扩散过程，并且

$$\mathrm{d}\big(X(t)Y(t)\big) = X(t)\mathrm{d}Y(t) + Y(t)\mathrm{d}X(t) + b_1\big(t, X(t)\big)b_2\big(t, Y(t)\big)\mathrm{d}t.$$

7.5 Black-Scholes 公式

现有欧式买入期权，规定持有人在 T 时刻以指定价格 K 购买股票，但不具有义务. 假定在 $(0, T)$ 内任意时刻可以买卖该期权，Black-Scholes 公式给出了 t 时刻欧式买入期权的合理价格为

$$V(t) = S(t)\Phi\Big(d_1\big(t, S(t)\big)\Big) - Ke^{-r(T-t)}\Phi\Big(d_2\big(t, S(t)\big)\Big),$$

其中 $\Phi(x)$ 为标准正态分布函数，$S(t) = \mathrm{e}^{(\mu - \sigma^2/2)t + \sigma B(t)}$ 为 t 时刻 1 单位股票的价格，并且

$$d_1(t, x) = \frac{\ln x - \ln K + (r + \sigma^2/2)(T - t)}{\sigma(T - t)^{1/2}}, \quad d_2(t, x) = \frac{\ln x - \ln K + (r - \sigma^2/2)(T - t)}{\sigma(T - t)^{1/2}}.$$