随机过程
Stochastic Process

龚舒凯
中国人民大学
Renmin University of China
School of Applied Economics/ School of Statistics
shukai_gong@ruc.edu.cn

目录

1　预备知识

1.1　概率公式

矩母函数

矩母函数定义为:

$$\phi(t) = \mathbb{E}[e^{tX}]$$

矩母函数有如下重要性质

1. 分布函数 $F_X(x)$ 与矩母函数 $\phi(t)$ 一一对应
2. 独立随机变量和的矩母函数是他们各自矩母函数的乘积
3. 当随机变量的矩有限时，成立

$$\mathbb{E}[X^k] = \phi^{(k)}(0)$$

正态分布的矩母函数非常重要：$Z = \mathcal{N}(0,1)$ 的矩母函数如下：

$$\mathbb{E}[e^{tZ}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} \cdot e^{-\frac{1}{2}z^2} \mathrm{d}z = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t)^2}{2}} \mathrm{d}z = e^{\frac{t^2}{2}}$$

从而 $\mathcal{N}(\mu,\sigma)$ 的矩母函数如下

$$\mathbb{E}[e^{t(\mu+\sigma Z)}] = e^{t\mu}\mathbb{E}[e^{t\sigma Z}] = \exp(t\mu + \frac{t^2\sigma^2}{2})$$

除了矩母函数之外，还有类似的、可以与分布函数一一对应的变换

1. 母函数: $\mathbb{E}[t^X], |t| \leq 1$
2. Laplace 变换: $\mathbb{E}[e^{-tX}], t \geq 0$，适用于非负随机变量
3. 特征函数: $\mathbb{E}[e^{\mathrm{i}tX}]$，对任意随机变量都成立

Fubini 定理 (求和换序/积分换序)

1. 对于两重求和 $\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} a_{ij}$，若 $\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} |a_{ij}| < \infty$，则求和可换序.
2. 对于二重积分 $\int_0^{\infty}\int_0^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y$，若 $\int_0^{\infty}\int_0^{\infty} |f(x,y)|\mathrm{d}x\mathrm{d}y < \infty$，则积分可换序

由 Fubini 定理衍生出 Tonelli 定理

Tonelli 定理

1. 若 $a_{ij} \geq 0$，则 $\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} a_{ij}$ 可换序.

2. 若 $f(x,y)\geq 0$，则 $\int_0^\infty\int_0^\infty f(x,y)\mathrm{d}x\mathrm{d}y<\infty$ 可换序

以下定理说明，对于非负随机变量，**可以将数学期望表达为尾部概率的积分/级数和.**

数学期望为尾部概率的积分/级数和

若 X 为一个非负连续型随机变量，分布函数为 $F(x)$，则

$$\mathbb{E}[X]=\int_0^\infty(1-F(x))\mathrm{d}x=\int_0^\infty P(X>x)\mathrm{d}x$$

若 X 是一个非负离散型随机变量，则

$$\mathbb{E}[X]=\sum_{k=1}^{\infty}P(X\geq k)$$

证明. 对于连续情况

$$\begin{aligned}\mathbb{E}[X]&=\int_0^\infty t\mathrm{d}F(t)=\int_{t=0}^\infty(\int_{x=0}^t 1\mathrm{d}x)\mathrm{d}F(t)=\int_{x=0}^\infty(\int_{t=x}^\infty\mathrm{d}F(t))\mathrm{d}x\\&=\int_{x=0}^\infty(1-F(x))\mathrm{d}x=\int_{x=0}^\infty P(X>x)\mathrm{d}x\end{aligned}$$

对于离散情况

$$\begin{aligned}\mathbb{E}[X]&=\sum_{k=1}^\infty kP(X=k)=\sum_{k=1}^\infty(\sum_{j=1}^k 1)P(X=k)=\sum_{j=1}^\infty\sum_{k=j}^\infty P(X=k)\\&=\sum_{j=1}^\infty P(X\geq j)\end{aligned}$$

□

全期望公式与条件方差公式

全期望公式为：$\mathbb{E}[X]=\mathbb{E}[\mathbb{E}[X|Y]]$

1. 离散：$\mathbb{E}[X]=\sum_y\mathbb{E}[X|Y=y]P(Y=y)$
2. 连续：$\mathbb{E}[X]=\int_y\mathbb{E}[X|Y=y]f_Y(y)\mathrm{d}y$

条件方差公式为：$\mathrm{Var}(X)=\mathbb{E}[\mathrm{Var}(X|Y)]+\mathrm{Var}(\mathbb{E}[X|Y])$

全概率公式

记 A 为任意随机事件，Y 为任意随机变量，X 为 A 的示性 $X=\mathbf{1}(A\text{发生})$ 则有

1. 若 Y 离散，则 $P(A)=\sum_y P(A|Y=y)P(Y=y)=\sum_y\mathbb{E}(X|Y=y)P(Y=y)$
2. 若 Y 连续，则 $P(A)=\int_{-\infty}^\infty P(A=Y|y)f_Y(y)\mathrm{d}y=\int_{-\infty}^\infty E(X=Y|y)f_Y(y)\mathrm{d}y$

1.2 概率空间与测度论

我们把样本点 ω 看作抽象的点，它们的全体构成样本空间 Ω。

σ 域

若把事件的全体记为 $\mathcal{F}$(由 Ω 的一些子集构成的集类)，它满足限制：

1. $\Omega \in \mathcal{F}$
2. 若 $A \in \mathcal{F}$，则 $\bar{A} \in \mathcal{F}$
3. 若 $A_n(n = 1, 2, ...) \in \mathcal{F}$，则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$

则称空间 Ω 上满足上述三条要求的集类 ($\mathcal{F}$) 为 σ 域或 σ 代数。

σ 域对交、并、差、补是封闭的，若 $\mathcal{F}$ 为 σ 域，$A_n \in \mathcal{F}, n = 1, 2, ...$，则

1. $\varnothing \in \mathcal{F}$
2. 可列个事件的交：$\bigcap\limits_{n=1}^{\infty} A_n = \overline{\bigcup\limits_{n=1}^{\infty} \overline{A_n}} \in \mathcal{F}$
3. 有限个个事件的并：$\bigcup\limits_{n=1}^{k} A_n = A_1 \cup A_2 \cup ... \cup A_k \cup \varnothing \cup ... \in \mathcal{F}$
4. 有限个个事件的交：$\bigcap\limits_{n=1}^{k} A_n = A_1 \cap A_2 \cap ... \cap A_k \cap \Omega \cap ... \in \mathcal{F}$

$\mathcal{F}$ 上的随机变量

设 X 为定义在 Ω 上取值为实数的函数，若对任意 $x \in \mathbb{R}$，有 $\{w : X(w) \in \mathcal{B}\} \in \mathcal{F}$ (简写为 $\{X \in \mathcal{B}\} \in \mathcal{F}$，$\mathcal{B}$ 为 Borel 点集)，则称 X 是 $\mathcal{F}$ 上的随机变量。

- 随机变量 X 可以将基本概率空间 $(\Omega, \mathcal{F}, P)$ 诱导到一个定义在实数集合上的概率空间 $(\mathbb{R}, \mathcal{B}, P_X)$

$\mathcal{G}$ 可测

设 X 为取值为实数的随机变量，子 σ 代数 $\mathcal{G} \subset \mathcal{F}$，若对任意 $x \in \mathbb{R}$，有 $\{w : X(w) \in \mathcal{B}\} \in \mathcal{G}$，则称 X 是 $\mathcal{G}$ 可测的，记为 $X \in \mathcal{G}$

最小生成 σ 代数

最小 σ 域

若给定 Ω 的一个非空集类 $\mathcal{G}$，必存在唯一的一个 Ω 中的 σ 域 $\mathfrak{M}(\mathcal{G})$，具有如下两个性质:

1. 包含 $\mathcal{G}$;
2. 若有其他 σ 域包含 $\mathcal{G}$，则必包含 $\mathfrak{M}(\mathcal{G})$，称 $\mathfrak{M}(\mathcal{G})$ 为由 $\mathcal{G}$ 生成的 σ 域。

称 $\mathfrak{M}(\mathcal{G})$ 为包含 $\mathcal{G}$ 的最小的 σ 域

设 X 为取值为实数的随机变量，记包含所有形如 $\{X \in \mathcal{B}\}$ 的最小 σ 代数为 $\sigma(X)$.

设 $Y_1, \cdots, Y_n$ 为一列取值为实数的随机变量，记包含所有形如 $\{Y_1 \in \mathcal{B}_1, \cdots, Y_n \in \mathcal{B}_n\}$ 的最小 σ 代数为 $\mathcal{G}_n = \sigma(Y_1, \cdots, Y_n)$.

- $\sigma(X)$ 可以看作所有与 X 有关的信息或历史
- $\mathcal{G}_n = \sigma(Y_1, \cdots, Y_n)$ 可以看作与 $\{Y_1, \cdots, Y_n\}$ 有关的信息或历史

定义在最小生成 σ 代数上的随机变量

X 是 $\sigma(Y_1, \cdots, Y_n)$ 可测的 $\iff$ X 是 $(Y_1, \cdots, Y_n)$ 的函数。

σ 代数流

若 $\forall n$，有 $X_n \in \sigma(Y_1, \cdots, Y_n)$，则称 $\{X_n\}$ 为 $\sigma(Y_1, \cdots, Y_n)$ 适应的，称 $\sigma(Y_1, \cdots, Y_n)$ 为一个 σ 代数流。

[注]：σ 代数流指的是：$X_1 \in \sigma(Y_1), X_2 \in \sigma(Y_1, Y_2), \cdots$

2 随机过程的基本概念

2.1 随机过程的定义

随机过程 (Stochastic Process)

随机过程 $\{X(t,\omega), t\in T, \omega\in\Omega\}$ 是一个关于时间 t 和样本点 ω 的二元函数。对于固定的样本点 $\omega\in\Omega$，$X(t,\omega)$ 为 t 的函数，是一条样本路径 (或一个样本函数)。

- 应用中，通常不论述随机过程的样本空间，即简记 $X(t,\omega)=X(t)$，将随机过程定义为随机变量的集合 $\{X(t), t\in T\}$。
- 参数 T 可以是自然数集 $\mathbb{N}$，整数集 $\mathbb{Z}$ 或者实数区间 $[a,b]$（这里 a 可以是 $-\infty$，b 可以是 $+\infty$）

假设抛掷 n 枚硬币，正面用 H 表示，反面用 T 表示。若第 t 次是正面则获得 t 个单位的报酬，反之没有报酬。在这个“报酬过程”中，其样本空间 Ω 为

$$\Omega=\{(H,H,\cdots,H),\cdots,(T,T,\cdots,T)\}$$

共包含 2^n 个样本点，每个样本点是一个 n 维向量，报酬过程 $X(t,\omega)$ 可以描述为

$$X(t,\omega)=t\mathbf{1}(\omega_t=H), t=1,2,\cdots,n, \omega\in\Omega$$

这里 ω_t 是样本点 ω 的第 t 个分量。

2.2 随机过程的有限维分布和数字特征

有限维分布族

设 $t_i\in T$，$1\le i\le n$，记

$$F(t_1,t_2,\cdots,t_n;x_1,x_2,\cdots,x_n)=P(X(t_1)\le x_1, X(t_2)\le x_2,\cdots,X(t_n)\le x_n)$$

其全体 $\{F(t_1,t_2,\cdots,t_n;x_1,x_2,\cdots,x_n), t_1,\cdots,t_n\in T\}$ 称为随机过程的有限维分布族，具有

1. 对称性:
$$F(t_1,t_2,\cdots,t_n;x_1,x_2,\cdots,x_n)=F(t_{j_1},t_{j_2},\cdots,t_{j_n};x_{j_1},x_{j_2},\cdots,x_{j_n})$$
2. 相容性:
$$F(t_1,t_2,\cdots,t_m,t_{m+1},\cdots,t_n;x_1,x_2,\cdots,x_m,\infty,\cdots,\infty)=F(t_1,t_2,\cdots,t_m;x_1,x_2,\cdots,x_m)$$

随机过程通常用以下数字特征描述

1. 均值函数: $m(t)=\mathbb{E}[X(t)]$
2. 方差函数: $\mathrm{Var}(X(t))=\mathbb{E}[(X(t)-m(t))^2]$
3. 协方差函数: $R(s,t)=\mathrm{Cov}(X(s),X(t))$
4. 相关函数: $\rho(s,t)=\dfrac{\mathrm{Cov}(X(s),X(t))}{\sqrt{\mathrm{Var}(X(t))\mathrm{Var}(X(s))}}$

2.3 经典随机过程

计数过程 (Counting Process)

对于非负且取值为整数的随机过程 $N(t)$，如果 $N(t)$ 表示时间间隔 $[0,t]$ 内事件发生的总数，并满足如下两条特性:

1. 若 $t_1 < t_2$，则 $N(t_1) \le N(t_2)$；
2. 若 $t_1 < t_2$，则 $N(t_2) - N(t_1)$ 为时间间隔 $[t_1, t_2]$ 间事件发生的总数

则 $N(t)$ 为计数过程。

- 如果一个计数过程在不相交的时间区间内发生的事件数是独立的，则称它具有**独立增量性**。即 $N(s) \perp N(t) - N(s)$。
- 如果一个计数过程在任一时间区间内事件发生次数的分布仅取决于时间区间的长度，则称它具有**平稳增量性**。即 $\forall s > 0, t_1 < t_2$，$N(t_2+s) - N(t_1+s)$ 与 $N(t_2) - N(t_1)$ 有着相同的分布。

独立增量 (Independent Increments) 过程

对 $t_1 < \cdots < t_n, t_i \in T, 1 \le i \le n$，设 $\Delta X(t_1) = X(t_1)$，$\Delta X(t_i) = X(t_i) - X(t_{i-1})$ $(2 \le i \le n)$，若增量

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

相互独立，则称 $\{X(t), t \in T\}$ 为**独立增量过程**.

- $\forall 0 \le s < t$，$X(t) - X(s)$ 只依赖 $t - s$，则称 $\{X(t), t \in T\}$ 具有**平稳增量**.
- 有平稳增量的独立增量过程称为平稳独立增量过程 (Stationary Process with Independent Increments).
- Poisson 过程和 Brownian Motion 就是典型的平稳独立增量过程。

Markov 过程

若对任意 $t_1 < \cdots < t_n < t$，以及 $x_1, \cdots, x_n$，以及 $A \subset \mathbb{R}$，总成立

$$P(X(t) \in A | X(t_1) = x_1, \cdots, X(t_n) = x_n) = P(X(t) \in A | X(t_n) = x_n)$$

则称 $\{X(t), t \in T\}$ 为 Markov 过程。也就是说，将来状态 $X(t)$ 与过去状态 $X(s), s < t_n$ 无关 (条件独立)，只与现在状态 $X(t_n)$ 有关.

- 状态空间 S: $X(t)$ 的取值全体构成的集合.
- Markov 链: 对于 Markov 过程 $\{X(t), t \ge 0\}$，当其状态空间 S（不加说明时默认非负整数集 $\{0, 1, 2, \cdots\}$）为可列无限集或有限集时，称为 Markov 链.

平稳过程 (Stationary Process) 与二阶矩过程

1. 宽平稳过程 (Wide-sense Stationary Process): 若 $\forall \tau, t \in T$，$\mathbb{E}[X(t)] < \infty$ 且

$$\mathbb{E}[X(t)^2] = m, \text{Cov}(X_t, X_{t+\tau}) = R(\tau)$$

   仅依赖 τ，则称 $\{X(t), t \in T\}$ 为宽平稳过程.

2. 严平稳过程 (Strict Stationary Process): 若 $\forall t_1, \cdots, t_n \in T$ 及 $\forall h > 0$，$(X(t_1), \cdots, X(t_n))$ 与 $(X(t_1 + h), \cdots, X(t_n + h))$ 有相同的联合分布，则称 $\{X(t), t \in T\}$ 为严平稳过程.

3. 二阶矩过程 (Finite Second Moment Process): 若 $\text{Var}[X(t)] < \infty$，则 $\{X(t), t \in T\}$ 为二阶矩过程.

鞅 (Martingales)

若对 $\forall t \in T, \mathbb{E}[X(t)] < \infty$，且对 $\forall t_1 < \cdots < t_n < t_{n+1}$，有

$$\mathbb{E}[X(t_{n+1}) | X(t_1), \cdots, X(t_n)] = X(t_n) \quad \text{a.s.}$$

，则称 $\{X(t), t \in T\}$ 为鞅

更新过程 (Renewal Process)

设 $\{X_k, k \geq 1\}$ 为独立同分布的正的随机变量序列，对 $\forall t > 0$，令 $T_0 = 0, T_n = \sum_{k=1}^{n} X_k$，并定义

$$N(t) = \sup\{n : n \geq 0, T_n \leq t\}$$

称 $\{N(t), t \geq 0\}$ 为更新过程.

- $N(t)$ 可以解释 $[0, t]$ 内更换零件的个数或系统来的信号/粒子数，或服务站来的顾客数等等。

3 Poisson 过程

3.1 Poisson 过程的定义

Poisson 过程有两个等价的定义

Poisson 过程 (定义 1)

如果计数过程 $\{N(t), t \geq 0\}$ 满足以下条件：

1. $N(0) = 0$
2. $N(t)$ 具有独立增量性
3. 在长度 t 的时间内事件发生的次数变量服从参数为 λt 的 Poisson 分布，即 $\forall s, t \geq 0$，

$$P(N(t+s) - N(s) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

(平稳增量性 & 增量 $\sim$ Poi(λt))

则称 $\{N(t), t \geq 0\}$ 为参数为 λ 的 Poisson 过程。

- 期望：$m(t) = \mathbb{E}[N(t)] = \lambda t$
- 方差：Var$[N(t)] = \lambda t$
- 协方差：Cov$(N(s), N(t)) = \lambda \min\{s, t\}$
- $\lambda = \dfrac{\mathbb{E}[N(t)]}{t}$，相当于单位时间事件的平均发生次数，所以称参数 λ 为过程 $N(t)$ 的强度。

证明. 这里只对协方差进行说明：不妨设 $s \leq t$，注意到

$$\begin{aligned}
\text{Cov}(N(s), N(t)) &= \mathbb{E}[N(s)N(t)] - \mathbb{E}[N(s)]\mathbb{E}[N(t)] \\
&= \mathbb{E}[N(s)(N(t) - N(s)) + N^2(s)] - \lambda s \lambda t \\
&= \mathbb{E}[N(s)]\mathbb{E}[N(t) - N(s)] + \mathbb{E}[N^2(s)] - \lambda s \lambda t \\
&= \lambda s \lambda (t - s) + \mathbb{E}[N^2(s)] - \lambda s \lambda t
\end{aligned}$$

其中

$$\begin{aligned}
\mathbb{E}[N^2(s)] &= \sum_{n=0}^{\infty} n^2 P(N(s) = n) = \sum_{n=0}^{\infty} n^2 e^{-\lambda s}\frac{(\lambda s)^n}{n!} \\
&= \sum_{n=0}^{\infty} n e^{-\lambda s}\frac{(\lambda s)^n}{(n-1)!} = \lambda s \sum_{n=1}^{\infty} n e^{-\lambda s}\frac{(\lambda s)^{n-1}}{(n-1)!} \\
&= \lambda s\left(\sum_{n=1}^{\infty}(n-1)e^{-\lambda s}\frac{(\lambda s)^{n-1}}{(n-1)!} + \sum_{n=1}^{\infty} e^{-\lambda s}\frac{(\lambda s)^{n-1}}{(n-1)!}\right) \\
&= \lambda s\,(\lambda s + 1) = (\lambda s)^2 + \lambda s
\end{aligned}$$

从而 Cov$(N(s), N(t)) = \lambda s \lambda (t - s) + (\lambda s)^2 + \lambda s - \lambda s \lambda t = \lambda s$，同理 $t < s$ 时有 Cov$(N(s), N(t)) = \lambda t$. □

Poisson 过程 (定义 2)(不考)

如果计数过程 $\{N(t), t \geq 0\}$ 满足以下条件：

1. $N(0) = 0$
2. $N(t)$ 具有独立增量性和平稳增量性
3. $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$
4. $P(N(t+h) - N(t) \geq 2) = o(h)$ (事件在同一瞬间同时发生多个事件的可能性极小)

则称 $\{N(t), t \geq 0\}$ 为参数为 λ 的 Poisson 过程。

要证明定义 1 $\Longleftrightarrow$ 定义 2，只需证明 (3) $\Longleftrightarrow$ (c)(d)，必要性的证明常采用在充分小的时间内建立一个微分方程。

证明. **必要性:** 由增量平稳性，记

$$P_n(t) = P(N(t) = n) = P(N(s+t) - N(s) = n)$$

先看 $n = 0$ 的情形。因

$$(N(t+h) = 0) = (N(t) = 0, N(t+h) - N(t) = 0), \quad h > 0,$$

故

$$\begin{aligned} P_0(t+h) &= P(N(t+h) = 0) = P(N(t) = 0, N(t+h) - N(t) = 0) \\ &= P(N(t) = 0)P(N(t+h) - N(t) = 0) \quad \text{(增量独立)} \\ &= P_0(t)P_0(h) \end{aligned}$$

另一方面

$$P_0(h) = P(N(t+h) - N(t) = 0) = 1 - (\lambda h + o(h)),$$

代入上式，有

$$\frac{P_0(t+h) - P_0(t)}{h} = -\left(\lambda P_0(t) + \frac{o(h)}{h}\right).$$

令 $h \to 0$，两边取极限，得

$$P_0'(t) = -\lambda P_0(t).$$

这是一阶线性常微分方程。由初始条件 $P_0(0) = P(N(0) = 0) = 1$，可得

$$P_0(t) = e^{-\lambda t}.$$

再看 $n > 0$ 的情形，因

$$\begin{aligned} \{N(t+h) = n\} &= \{N(t) = n, N(t+h) - N(t) = 0\} \cup \{N(t) = n-1, N(t+h) - N(t) = 1\} \\ &\cup \left(\bigcup_{l=2}^{n} \{N(t) = n-l, N(t+h) - N(t) = l\}\right) \end{aligned}$$

故

$$P_n(t+h) = P_n(t)(1 - \lambda h - o(h)) + P_{n-1}(t)(\lambda h + o(h)) + o(h).$$

化简可得

$$\frac{P_n(t+h)-P_n(t)}{h}=-\lambda P_n(t)+\lambda P_{n-1}(t)+\frac{o(h)}{h}.$$

令 $h\to 0$，两边取极限，有

$$P_n'(t)=-\lambda P_n(t)+\lambda P_{n-1}(t).$$

将上述两边乘以 $e^{-\lambda t}$，移项后可得

$$\frac{d}{dt}\left[e^{\lambda t}P_n(t)\right]=\lambda e^{\lambda t}P_{n-1}(t),$$

且满足初始条件

$$P_n(0)=P(N(0)=n)=0.$$

当 $n=1$ 时

$$\frac{d}{dt}\left[e^{\lambda t}P_1(t)\right]=\lambda.$$

注意到初始条件 $P_1(0)=0$，可得

$$P_1(t)=(\lambda t)e^{-\lambda t}.$$

再用归纳法即有

$$P_n(t)=\frac{(\lambda t)^n}{n!}e^{-\lambda t}.$$

充分性：

$$\begin{aligned}
P\{N(t+h)-N(t)=1\}&=P\{N(h)-N(0)=1\}=e^{-\lambda h}\frac{\lambda h}{1!}=\lambda h\sum_{n=0}^{\infty}\frac{(\lambda h)^n}{n!}\\
&=\lambda h[1-\lambda h+o(h)]=\lambda h+o(h)\\
P\{N(t+h)-N(t)\geq 2\}&=P\{N(h)-N(0)\geq 2\}=\sum_{n=2}^{\infty}e^{-\lambda h}\frac{(\lambda h)^n}{n!}\\
&=o(h)
\end{aligned}$$

□

必要性的证明也可以采用 Laplace 变换。由于此处 $N(t)\geq 0$，其 Laplace 变换必存在。回顾 X∼Poisson 分布的矩母函数

$$\mathbb{E}[e^{-uX}]=\exp(\lambda(e^{-u}-1))$$

如果我们能证明

$$\mathbb{E}[e^{-uN(t)}]=\exp(\lambda t(e^{-u}-1))$$

由 Laplace 变换的一一对应性，我们就能说明 $N(t)\sim$ Poi(λt)

证明. 令 $u\geq 0$，并记 $g(t)=\mathbb{E}[e^{-uN(t)}]$，则

$$\begin{aligned}
g(t+h)&=\mathbb{E}[e^{-uN(t)}\cdot e^{-u[N(t+h)-N(t)]}]\\
&=\mathbb{E}[e^{-uN(t)}]\cdot\mathbb{E}[e^{-u[N(t+h)-N(t)]}] && \text{(独立增量性)}\\
&=g(t)\cdot\mathbb{E}[e^{-uN(h)}] && \text{(平稳增量性)}
\end{aligned}$$

其中

$$\begin{aligned}\mathbb{E}[e^{-uN(h)}] &= \sum_{k=0}^{\infty} e^{-uk} P(N(h)=k) = \sum_{k=0}^{\infty} e^{-uk} \cdot e^{-\lambda h}\frac{(\lambda h)^k}{k!} \\ &\approx 1\cdot(1-\lambda h+o(h)) + e^{-u}(\lambda h+o(h)) + o(h) \\ &= 1-\lambda h + e^{-u}\lambda h + o(h)\end{aligned}$$

从而

$$\begin{aligned} g(t+h) &= g(t)(1-\lambda h+e^{-u}\lambda h)+o(h) \\ \Rightarrow \frac{g(t+h)-g(t)}{h} &= g(t)\lambda(e^{-u}-1)+\frac{o(h)}{h} \\ \Rightarrow g'(t) &= g(t)\lambda(e^{-u}-1)\end{aligned}$$

解上述微分方程可得 $g(t)=e^{\lambda t(e^{-u}-1)}$，这说明了 $N(t)\sim \text{Poi}(\lambda t)$。 □

还有一种通过相邻时间间隔的分布来定义 Poisson 过程的定义.

Poisson 过程 (定义 3)

如果每次事件发生的时间间隔 $X_1, X_2, \cdots$ 相互独立，且服从同一参数 λ 的指数分布，则计数过程 $\{N(t), t\geq 0\}$ 是参数为 λ 的 Poisson 过程.

3.2 Poisson 过程相联系的若干分布

统一标记： 设第 n 次事件发生的时刻为 T_n，则第 n 次与第 $n-1$ 次事件发生的时间间隔定义为 $X_n = S_n - S_{n-1}$.

时间间隔的分布

Poisson 过程 $\{N(t), t\geq 0\}$ 第 $n-1$ 次与第 n 次事件发生的时间间隔 $X_n \sim \text{Exp}(\lambda)$

证明. 首先考虑 X_1 的分布，注意到 $\{X_1 > t\} \iff \{N(t)=0\}$，因此

$$P(X_1>t) = P(N(t)=0) = e^{-\lambda t} \Rightarrow P(X_1\leq t) = 1-e^{-\lambda t}$$

再考虑 X_2 的分布，根据 Poisson 过程的独立增量性

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

所以 X_2 与 X_1 独立，且都服从 $\text{Exp}(\lambda)$。以此类推，$X_n\sim\text{Exp}(\lambda)$ □

到达时刻的分布

Poisson 过程 $\{N(t), t\geq 0\}$ 第 n 次事件发生的时刻 $T_n \sim \Gamma(n,\lambda)$

证明. 注意到 $T_n = \sum_{i=1}^{n} X_i$，$\{X_n\}$ 相互独立且服从 $\text{Exp}(\lambda) = \Gamma(1,\lambda)$，所以

$$T_n = \sum_{i=1}^{n} X_i \sim \Gamma(n,\lambda)$$

S_n 还有另一种导出方式：因为 $\{T_n \le t\} \iff \{N(t) \ge n\}$，所以

$$
\begin{aligned}
F_{T_n}(t) = P(T_n \le t) = P(N(t) \ge n) &= \sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!} \\
\Rightarrow f_{T_n}(t) = \sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!}\cdot(-\lambda) &+ \sum_{j=n}^{\infty} e^{-\lambda t}\frac{j(\lambda t)^{j-1}\cdot\lambda}{j!} \\
&= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{j=n+1}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^j}{j!} \\
&= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}
\end{aligned}
$$

所以 $T_n \sim \Gamma(n,\lambda)$ □

[注]：在给定 $N(t)=n$ 时，T_n 不服从 $\Gamma(n,\lambda)$ 分布。从下面的定理可以得知，当给定 $N(t)=n$ 时，这 n 个事件的发生时刻 $T_1,\cdots,T_n$ 服从 $(0,t]$ 上的均匀分布。$T_n=\max\{T_1,\cdots,T_n\}$，因此 $T_n|N(t)=n$ 的分布求解就转化为了 n 个 i.i.d. 均匀随机变量的最大值分布求解。简记 $T_n|N(t)=n$ 的分布函数为 $F_{T_n}(x)$，则

$$
\begin{aligned}
F_{T_n}(x) &= P(\max\{T_1,\cdots,T_n\}\le x) = P(T_1\le x,\cdots,T_n\le x) = \left(\frac{x}{t}\right)^n \\
\Rightarrow f_{T_n}(x) &= \frac{n}{t}\left(\frac{x}{t}\right)^{n-1}
\end{aligned}
$$

令 $Y=\frac{T_n}{t}$，则

$$
\begin{aligned}
f_Y(y) = f_{T_n}(yt)\cdot t = ny^{n-1} = \frac{y^{n-1}(1-y)^0}{B(n,1)} &\Rightarrow Y\sim \text{Beta}(n,1) \\
\Rightarrow T_n \sim \text{Beta}(n,1)\cdot t
\end{aligned}
$$

所以 $T_n|N(t)=n$ 服从一个缩放后的 Beta 分布！

到达时刻的条件分布

设 $\{N(t),t\ge 0\}$ 是 Poisson 过程，$\forall s\le t$，有

$$P(T_1\le s|N(t)=1)=\frac{s}{t}$$

证明.

$$
\begin{aligned}
P(T_1\le s|N(t)=1) &= \frac{P(T_1\le s, N(t)=1)}{P(N(t)=1)} \\
&= \frac{P([0,s)\text{ 中 1 次事件，}[s,t]\text{ 中 0 次事件})}{P(N(t)=1)} \\
&= \frac{P(N(s)=1)\cdot P(N(t-s)=0)}{P(N(t)=1)} \\
&= \frac{e^{-\lambda s}\frac{(\lambda s)^1}{1!}\cdot e^{-\lambda(t-s)}\frac{(\lambda(t-s))^0}{0!}}{e^{-\lambda t}\frac{(\lambda t)^1}{1!}} = \frac{s}{t}
\end{aligned}
$$

□

[注]：直观上，在已知 t 时间内只发生一次事件的时候，由于 Poisson 过程是平稳增量和独立增量的，$[0,t]$ 内的每个相等长度的区间都有相同的概率包含这个事件。换言之，事件发生的时刻应该均匀分布在 $[0,t]$ 上。
[注]：下面的结论推广到了 $N(t)=n$ 时，$T_1,\cdots,T_n$ 这 n 个事件发生时刻的条件分布

到达时刻的分布与顺序统计量

若设 $\{N(t),t\geq 0\}$ 是 Poisson 过程，则在 $N(t)=n$ 的条件下，$T_1,\cdots,T_n$ 的条件分布函数与 n 个 $(0,t]$ 上**相互独立同均匀分布的顺序统计量的分布函数相同**。即在已给 $N(t)=n$ 的情况下，事件相继发生的时刻 $(T_1,\cdots,T_n)$ 的联合密度为

$$f(t_1,\cdots,t_n|N(t)=n)=\begin{cases}\dfrac{n!}{t^n} & 0<t_1<\cdots<t_n<t\\ 0 & \text{otherwise}\end{cases}$$

证明. 不妨设 $0<t_1<\cdots<t_n<t_{n+1}=t$，再取充分小的 h_i 使得 $t_i+h_i<t_{i+1},i=1,\cdots,n$，则

$$\begin{aligned}P(t_i<T_i\leq t_i+h_i,i=1,\cdots,n|N(t)=n)=&\frac{P(t_i<T_i\leq t_i+h_i,i=1,\cdots,n,N(t)=n)}{P(N(t)=n)}\\=&\frac{P(N(t_i+h_i)-N(t_i)=1,N(t_{i+1})-N(t_i+h_i)=0,i=1,\cdots,n,N(t_1)=0)}{P(N(t)=n)}\\=&\frac{(\lambda h_1e^{-\lambda h_1})\cdots(\lambda h_ne^{-\lambda h_n})\cdot e^{-\sum\limits_{i=1}^{n}\lambda(t_{i+1}-t_i-h_i)}\cdot e^{-\lambda t_1}}{e^{-\lambda t}\dfrac{(\lambda t)^n}{n!}}\\=&\frac{(\lambda h_1e^{-\lambda h_1})\cdots(\lambda h_ne^{-\lambda h_n})\cdot e^{-\lambda(t-h_1-\cdots-h_n)}}{e^{-\lambda t}\dfrac{(\lambda t)^n}{n!}}\\=&\frac{n!}{t^n}h_1\cdots h_n\end{aligned}$$

那么

$$f(t_1,\cdots,t_n|N(t)=n)=\lim_{h_1,\cdots,h_n\to 0}\frac{P(t_i<T_i\leq t_i+h_i,i=1,\cdots,n|N(t)=n)}{h_1h_2\cdots h_n}=\frac{n!}{t^n},\ 0<t_1<\cdots<t_n<t$$

□

[注]：直观上，在已知 $N(t)=n$ 的条件下，在 $(0,t]$ 内独立发生了 n 个事件，且每个事件的发生时间都服从 $(0,t]$ 上的均匀分布。

3.3 Poisson 过程的变体

Poisson 过程的随机稀疏

假设事件 A 的发生形成强度为 λ 的 Poisson 过程 $\{N(t),t\geq 0\}$，如果每次事件发生时**以概率 p 被记录下来**，以 $M(t)$ 表示 t 时刻被记录下来的时间总数，则 $\{M(t),t\geq 0\}$ 是一个强度为 λp 的 Poisson 过程：

$$P(M(t)=m)=\frac{(\lambda pt)^m}{m!}e^{-\lambda pt}$$

证明. 根据全概率公式有

$$\begin{aligned}
P\{M(t)=m\} &= \sum_{n=0}^{\infty} P\{M(t)=m \mid N(t)=m+n\}P\{N(t)=m+n\} \\
&= \sum_{n=0}^{\infty} C_{m+n}^{m} p^m (1-p)^n \frac{(\lambda t)^{m+n}}{(m+n)!} e^{-\lambda t} \\
&= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda p t)^m [\lambda(1-p)t]^n}{m!n!} \\
&= e^{-\lambda t} \frac{(\lambda p t)^m}{m!} \sum_{n=0}^{\infty} \frac{[\lambda(1-p)t]^n}{n!} \\
&= e^{-\lambda t} \frac{(\lambda p t)^m}{m!} e^{\lambda(1-p)t} \\
&= e^{-\lambda p t} \frac{(\lambda p t)^m}{m!}.
\end{aligned}$$

□

Poisson 过程的叠加

若 $\{N_i(t), t \geq 0\}$ 为强度为 λ_i 的独立 Poisson 过程，则 $\{\sum_i N_i(t), t \geq 0\}$ 为强度为 $\sum_i \lambda_i$ 的 Poisson 过程。

证明. 由于 $\{N_i(t), t \geq 0\}$ 相互独立，所以

$$\begin{aligned}
P\left(\sum_i N_i(t)=n\right) &= \sum_{n_1+\cdots+n_k=n} P(N_1(t)=n_1, \cdots, N_k(t)=n_k) \\
&= \sum_{n_1+\cdots+n_k=n} e^{-\lambda_1 t}\frac{(\lambda_1 t)^{n_1}}{n_1!} \cdots e^{-\lambda_k t}\frac{(\lambda_k t)^{n_k}}{n_k!} \\
&= \frac{e^{-\sum_i \lambda_i t}(\sum_i \lambda_i t)^n}{n!}
\end{aligned}$$

□

[注]：直观上而言，由于 $N_i(t)$ 具有独立增量性，且两个独立的 Poisson 随机变量之和也为 Poisson 变量，Poisson 参数为两个参数之和。因此 $N_1(t)+N_2(t)$ 是参数为 $(\lambda_1+\lambda_2)t$ 的 Poisson 过程。以此类推。

3.4 非时齐 Poisson 过程

非时齐 Poisson 过程

如果计数过程 $\{N(t), t \geq 0\}$ 满足以下条件：

- $N(0)=0$
- $N(t)$ 具有独立增量性和平稳增量性
- $P(N(t+h)-N(t)=1)=\lambda(t)h+o(h)$
- $P(N(t+h)-N(t) \geq 2)=o(h)$ (事件在同一瞬间同时发生多个时间的可能性极小)

则称 $\{N(t), t \geq 0\}$ 为强度为 $\lambda(t)$ 的非时齐 Poisson 过程.

[注]: 记累计强度函数：$m(t) = \int_0^t \lambda(s)\mathrm{d}s,\ \lambda(s) > 0$

非时齐 Poisson 过程事件发生次数的分布

若 $\{N(t), t \geq 0\}$ 是非时齐具有强度函数 $\{\lambda(t), t \geq 0\}$ 的 Poisson 过程，则 $\forall s, t \geq 0$，成立

$$P(N(s+t) - N(s) = n) = \frac{[m(s+t) - m(s)]^n \exp(-[m(s+t) - m(s)])}{n!}$$

[注]: 显然，时齐 Poisson 过程是非时齐的一个特例。时齐时 $m(t) = \lambda t$。

非齐次 Poisson 过程的导出

设 $\{N(t), t \geq 0\}$ 是一个强度为 λ 的 Poisson 过程，在时刻 t 发生的事件以概率 $\frac{\lambda(t)}{\lambda}$ 被计数 ($\lambda(t) \leq \lambda$)，则对 $\forall t \geq 0$，被计数的事件构成的过程是具有强度函数 $\lambda(t)$ 的非齐次 Poisson 过程。

[注]: 直观上，当 $\lambda(t)$ 有上界 λ 时，可将非齐次 Poisson 过程理解为对齐次 Poisson 过程的随机采样 (随机稀疏)。

非齐次 Poisson 过程的逆 (不考)

设 $\{N(t), t \geq 0\}$ 是一个强度为 $\lambda(t)$ 的非齐次 Poisson 过程。令 $N^*(t) = N(m^{-1}(t))$，则 $\{N^*(t), t \geq 0\}$ 是一个强度为 1 的 Poisson 过程。

[注]: $m(t) = \int_0^t \lambda(s)\mathrm{d}s,\ \lambda(s) > 0$ 单调递增，所以 $m^{-1}(t)$ 存在且也单调递增的。

3.5 复合 Poisson 过程

复合 Poisson 过程

称随机过程 $\{X(t), t \geq 0\}$ 为复合 Poisson 过程，如果对于 $t \geq 0$，$X(t)$ 可以表示为

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

其中 $\{N(t), t \geq 0\}$ 时一个 Poisson 过程，Y_i 为 i.i.d 的随机变量，且独立于 $\{N(t), t \geq 0\}$。

[注]: 索赔过程服从一个 Poisson 过程 $\{N(t), t \geq 0\}$，每次索赔数额 Y_i 都相互独立同分布，且与发生时刻无关，因此赔付总金额 $\{X(t) = \sum_{i=1}^{N(t)} Y_i, t \geq 0\}$ 服从复合 Poisson 过程。

- 期望：$m(t) = \lambda t \mathbb{E}[Y_i]$
- 方差：$\mathrm{Var}[X(t)] = \lambda t \mathbb{E}[Y_i^2]$

期望和方差的证明可以使用条件期望/方差公式说明

证明. 期望：根据全期望公式

$$m(t) = \mathbb{E}[X(t)] = \mathbb{E}[\mathbb{E}[X(t)|N(t)]] = \mathbb{E}[N(t)]\mathbb{E}[Y_i] = \lambda t \mathbb{E}[Y_i]$$

方差：根据条件方差公式

$$
\begin{aligned}
\mathsf{Var}[X(t)] &= \mathbb{E}[\mathsf{Var}[X(t)|N(t)]] + \mathsf{Var}[\mathbb{E}[X(t)|N(t)]] \\
&= \mathbb{E}[N(t)]\mathsf{Var}[Y_i] + \mathsf{Var}[N(t)]\mathbb{E}[Y_i]^2 \\
&= \lambda t \mathsf{Var}[Y_i] + \lambda t \mathbb{E}[Y_i]^2 \\
&= \lambda t(\mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2) + \lambda t \mathbb{E}[Y_i]^2 \\
&= \lambda t \mathbb{E}[Y_i^2]
\end{aligned}
$$

□

期望和方差的证明也可以使用矩母函数说明

证明.

$$
\begin{aligned}
\phi_t(u) = \mathbb{E}[\exp\{uX(t)\}] &= \sum_{n=0}^{\infty} P\{N(t)=n\}\mathbb{E}[\exp\{uX(t)\} \mid N(t)=n\} \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \exp(-\lambda t)\mathbb{E}[\exp\{u(Y_1+Y_2+\cdots+Y_n)\} \mid N(t)=n] \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \exp(-\lambda t)\mathbb{E}[\exp\{u(Y_1+Y_2+\cdots+Y_n)\}] \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \exp(-\lambda t)\left(\mathbb{E}[\exp\{uY\}]\right)^n \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \exp(-\lambda t)\left(\phi_Y(u)\right)^n \\
&= \sum_{n=0}^{\infty} \exp(-\lambda t)\frac{(\lambda t \phi_Y(u))^n}{n!} = \exp\{\lambda t[\phi_Y(u)-1]\}
\end{aligned}
$$

对 $\phi_t(u)$ 在 $t=0$ 处分别取一阶导、二阶导得

$$
\begin{aligned}
\mathbb{E}[X(t)] &= \phi_t'(0) = \lambda t \phi_Y'(0) = \lambda t \mathbb{E}[Y_i] \\
\mathbb{E}[X^2(t)] &= \phi_t''(0) = (\lambda t \phi_Y'(0))^2 + \lambda t \phi_Y''(0) = (\lambda t \mathbb{E}[Y_i])^2 + \lambda t \mathbb{E}[Y_i^2] \\
\Rightarrow \mathsf{Var}[X(t)] &= \mathbb{E}[X^2(t)] - (\mathbb{E}[X(t)])^2 = \lambda t \mathbb{E}[Y_i^2]
\end{aligned}
$$

□

3.6 条件 (混合)Poisson 过程

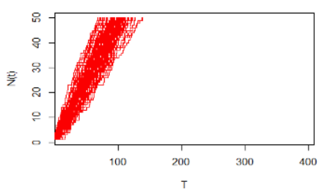
Poisson 过程描述的是一个有“风险”参数 λ 的个体发生某一事件的频率。如果我们认为一个群体中，个体的风险参数 λ 是“因人而异”的 (一个随机变量 Λ)，那么我们就可以描述一个群体的事件发生频率。这就是条件 Poisson 过程。

条件 Poisson 过程

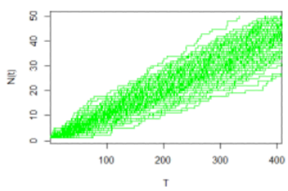
设随机变量 $\Lambda > 0$，在 $\Lambda = \lambda$ 的条件下，计数过程 $\{N(t), t \geq 0\}$ 是参数为 λ 的 Poisson 过程，则称 $\{N(t), t \geq 0\}$ 为条件 Poisson 过程。即若 $\Lambda \sim G(\lambda)$，成立

$$P(N(t+s) - M(s) = n) = \int_0^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \mathrm{d}G(\lambda)$$

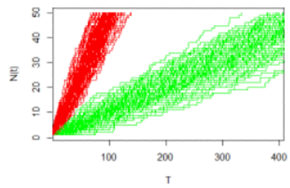
举个最简单的例子，假设将保险公司的新投保客户的事故发生率 λ 看作一个服从 $P(\lambda = 0.5) = P(\lambda = 0.1) = \frac{1}{2}$ 的随机变量 (两点分布)，则条件 Poisson 过程可绘制为：



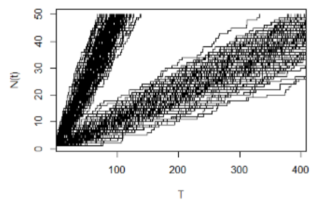
1. 1. 强度为0. 5的泊松过程



1.2. 强度为0.1的泊松过程



1.3. 以1/2概率各取强度为0. 5和0.1的计数过程



1.4. 条件泊松过程

图 1: 条件 Poisson 过程

条件 Poisson 过程的数字特征

1. 期望：$\mathbb{E}[N(t)] = t\mathbb{E}[\Lambda]$
2. 方差：$\mathsf{Var}[N(t)] = t^2\mathsf{Var}[\Lambda] + t\mathbb{E}[\Lambda]$
3. 协方差：$\mathsf{Cov}[N(s), N(t)] = \min\{s,t\}\mathbb{E}[\Lambda] + st\mathsf{Var}[\Lambda]$

证明. 期望：根据全期望公式

$$\mathbb{E}[N(t)] = \mathbb{E}[\mathbb{E}[N(t)|\Lambda = \lambda]] = \mathbb{E}[\lambda t] = t\mathbb{E}[\Lambda]$$

方差：根据条件方差公式

$$\begin{aligned}\mathsf{Var}[N(t)] &= \mathbb{E}[\mathsf{Var}[N(t)|\Lambda = \lambda]] + \mathsf{Var}[\mathbb{E}[N(t)|\Lambda = \lambda]] \\ &= \mathbb{E}[\lambda t] + \mathsf{Var}[\lambda t] = t^2\mathsf{Var}[\Lambda] + t\mathbb{E}[\Lambda]\end{aligned}$$

协方差：不妨设 $s \leq t$，回顾 Poisson 过程的协方差计算中，$\mathbb{E}[N(s)N(t)] = \mathbb{E}[N(s)]\mathbb{E}[N(t) - N(s)] + \mathbb{E}[N^2(s)] =$

$\lambda s\lambda t+\lambda s$，因此

$$\begin{aligned}\mathsf{Cov}[N(s),N(t)]&=\mathbb{E}[\mathbb{E}[N(s)N(t)|\Lambda=\lambda]]-\mathbb{E}[\mathbb{E}[N(s)|\Lambda=\lambda]]\cdot\mathbb{E}[\mathbb{E}[N(t)|\Lambda=\lambda]]\\&=\mathbb{E}[\lambda s\lambda t+\lambda s]-\mathbb{E}[\lambda s]\mathbb{E}[\lambda t]\\&=st\mathbb{E}[\Lambda^2]+s\mathbb{E}[\Lambda]-st\mathbb{E}[\Lambda]^2\\&=s\mathbb{E}[\Lambda]+st\mathsf{Var}[\Lambda]\end{aligned}$$

□

条件 Poisson 过程的性质

条件 Poisson 过程是平稳增量的，但不是独立增量的。

证明. 平稳增量性从条件 Poisson 分布的定义显然得出，关于独立增量性，事实上

$$\begin{aligned}&P(N(s)=j,N(t+s)-N(s)=k)\\=&\int_0^\infty\frac{(\lambda s)^j}{j!}e^{-\lambda s}\frac{(\lambda t)^k}{k!}e^{-\lambda t}\mathrm{d}G(\lambda)\\\neq&\left(\int_0^\infty\frac{(\lambda s)^j}{j!}e^{-\lambda s}\mathrm{d}G(\lambda)\right)\left(\int_0^\infty\frac{(\lambda t)^k}{k!}e^{-\lambda t}\mathrm{d}G(\lambda)\right)\\=&P(N(s)=j)\cdot P(N(t+s)-N(s)=k)\end{aligned}$$

具体到数值，可以参考如下例子：令 $P(\Lambda=1)=P(\Lambda=2)=\frac{1}{2}$，则

$$\begin{aligned}&P(N(1)=0,N(2)-N(1)=0)\\=&\frac{1}{2}P(N(1)=0,N(2)-N(1)=0\mid\Lambda=1)+\frac{1}{2}P(N(1)=0,N(2)-N(1)=0\mid\Lambda=2)\\=&\frac{1}{2}e^{-1}\cdot e^{-1}+\frac{1}{2}e^{-2}\cdot e^{-2}\approx0.0768\end{aligned}$$

假设 $N(t)$ 是独立增量过程，则

$$\begin{aligned}P(N(1)=0,N(2)-N(1)=0)&=P(N(1)=0)\cdot P(N(2)-N(1)=0)\\&=\left[\frac{1}{2}P(N(1)=0\mid\Lambda=1)+\frac{1}{2}P(N(1)=0\mid\Lambda=2)\right]\\&\quad\cdot\left[\frac{1}{2}P(N(2)-N(1)=0\mid\Lambda=1)+\frac{1}{2}P(N(2)-N(1)=0\mid\Lambda=2)\right]\\&=\left[\frac{1}{2}e^{-1}+\frac{1}{2}e^{-2}\right]^2\approx0.0633\end{aligned}$$

矛盾！假设不成立。 □

4 更新过程

4.1 更新过程的基本概念

Poisson 过程可以被定义为：事件发生的时间间隔 $X_1, X_2, \cdots$ 相互独立，且服从 Exp(λ) 的计数过程。本章考虑更一般的情形，去除“指数分布”的限制，考察事件发生的时间间隔 $X_1, X_2, \cdots$ 独立同分布的计数过程。

更新过程

设 $\{X_k, k \geq 1\}$ 为独立同分布的正的随机变量序列，对 $\forall t > 0$，令 $T_0 = 0, T_n = \sum\limits_{k=1}^{n} X_k$，并定义

$$N(t) = \sup\{n : n \geq 0, T_n \leq t\}$$

称 $\{N(t), t \geq 0\}$ 为更新过程。

- 在更新过程中事件发生一次叫做一次更新，X_n 是第 $n-1$ 次和第 n 次更新相距的时间。X_n 的分布函数记为 $F(x) = P(X \in \mathcal{B})$。特别的，$F(0) = P(X \leq 0) = 0$。均值 $\mu = \mathbb{E}[X_n] = \displaystyle\int_0^{\infty} x\mathrm{d}F(x)$。
- T_n 表示第 n 次更新的时刻。T_n 的分布函数记为 $F_n(t) = P(T_n \leq t)$。
- $N(t)$ 是 $[0, t]$ 事件内发生的总更新次数，称 $\{N(t)\}$ 的更新函数为 $M(t) = \mathbb{E}[N(t)]$

更新过程中事件的集合关系

1. $\{N(t) \geq n\} = \{T_n \leq t\}$：$t$ 时刻及之前更新次数大于等于 $n \iff$ 第 n 次更新的发生时刻在 t 时刻及之前。
2. $\{N(t) > n\} \subset \{T_n \leq t\}$：$t$ 时刻及之前更新次数大于 $n \Rightarrow$ 第 n 次更新的发生时刻在 t 时刻及之前。
3. $\{N(t) = n\} = \{T_n \leq t < T_{n+1}\}$：$t$ 时刻及之前更新次数恰等于 $n \iff t$ 时刻时已发生第 n 次更新，但还未发生第 $n+1$ 次更新。

4.2 更新过程的性质

有限时间内至多发生有限次更新

有限时间内至多发生有限次更新，即 $P(N(t) < \infty) = 1$

证明. 直观上说由强大数定理，$\dfrac{T_n}{n} = \dfrac{\sum\limits_{i=1}^{n} X_n}{n} \to \mu$ a.s.，从而 $T_n \to \infty$ a.s.，即无穷次更新只可能发生在无限长时间内发生，从而有限时间内最多只能发生有限次更新，即 $P(N(t) < \infty) = 1$ 以概率 1 成立。 □

$N(t)$ 的分布列

设 T_n 的分布函数为 $F_n(t)$，则 $P(N(t) = n) = F_n(t) - F_{n+1}(t)$

证明.

$$P(N(t)=n)=P(N(t)\geq n)-P(N(t)\geq n+1)=P(T_n\leq t)-P(T_{n+1}\leq t)=F_n(t)-F_{n+1}(t)$$

□

更新函数

更新函数 $M(t)=\mathbb{E}[N(t)]=\sum_{n=1}^{\infty}P(N(t)\geq n)=\sum_{n=1}^{\infty}F_n(t)$

[注]：更新函数 $M(t)$ 是关于 t 的函数，不是随机变量。

证明. 先证明 $M(t)=\mathbb{E}[N(t)]=\sum_{n=1}^{\infty}P(N(t)\geq n)$，注意到

$$\begin{aligned}
M(t)=\mathbb{E}[N(t)]&=\sum_{n=1}^{\infty}nP(N(t)=n)=\sum_{n=1}^{\infty}n\left(P(N(t)\geq n)-P(N(t)\geq n+1)\right)\\
&=P(N(t)\geq 1)-P(N(t)\geq 2)+2P(N(t)\geq 2)-2P(N(t)\geq 3)\\
&+3P(N(t)\geq 3)-3P(N(t)\geq n)+\cdots\\
&=\sum_{n=1}^{\infty}P(N(t)\geq n)
\end{aligned}$$

再证明 $M(t)=\sum_{n=1}^{\infty}F_n(t)$，注意到

$$M(t)=\mathbb{E}[N(t)]=\sum_{n=1}^{\infty}P(N(t)\geq n)=\sum_{n=1}^{\infty}P(T_n\leq t)=\sum_{n=1}^{\infty}F_n(t)$$

□

更新函数的有限性 (证明不考)

$M(t)$ 是 t 的不减函数，且 $\forall t\in[0,\infty)$，$M(t)<\infty$

4.3 更新方程

更新过程的一大特点是：更新后系统恢复如初。因此，新过程的概率性质与原过程相同。具体而言，若设 s 时刻为一更新点，则 $(s,t]$ 时间区间内更新发生的次数与 $(0,t-s]$ 内更新发生的次数**同分布**。由此，我们可以使用**对更新点取条件期望/概率**，利用更新点处系统恢复如新的性质，使条件期望/概率转化为无条件期望/概率。

更新函数的积分方程表达

更新函数 $M(t)$ 满足积分方程

$$M(t)=F(t)+\int_0^t M(t-s)\mathrm{d}F(s)$$

更新密度 $m(t)=M'(t)$ 满足积分方程

$$m(t)=f(t)+\int_0^t m(t-s)f(s)\mathrm{d}s$$

证明. 依重期望公式，对第一次更新的时刻取条件期望：

$$M(t) = \mathbb{E}[N(t)] = \mathbb{E}[\mathbb{E}[N(t)|T_1]] = \int_0^\infty \mathbb{E}(N(t)|T_1 = x)\mathrm{d}F_1(x)$$

注意到

1. $T_1 = x > t$ 时，$N(t) = 0$。
2. $T_1 = x$ 时系统恢复如初，因此 $(x, t]$ 内更新次数的分布和 $(0, t-x]$ 内更新次数的分布相同。
3. T_1 的分布函数 $F_1(t)$ 等于 X_n 的分布函数 $F(t)$。

所以

$$\begin{aligned}
M(t) &= \int_0^\infty \mathbb{E}[N(t)|T_1 = x]\mathrm{d}F_1(x) = \int_0^t \mathbb{E}[N(t)|T_1 = x]\mathrm{d}F(x) + \int_t^\infty \mathbb{E}[N(t)|T_1 = x]\mathrm{d}F(x) \\
&= \int_0^t \mathbb{E}[N(t)|T_1 = x]\mathrm{d}F(x) = \int_0^t \mathbb{E}[1 + N(t-x)]\mathrm{d}F(x) \\
&= F(t) + \int_0^t \mathbb{E}[N(t-x)]\mathrm{d}F(x) \\
&= F(t) + \int_0^t M(t-x)\mathrm{d}F(x)
\end{aligned}$$

□

更新方程

更一般地，对于已知的 $H(t)$ 和 $F(t)$，且满足 $H(t) = 0, F(t) = 0, t < 0$，我们称如下积分方程为更新方程：

$$K(t) = H(t) + \int_0^t K(t-s)\mathrm{d}F(s)$$

我们关心这个方程的解 $K(t)$ 是否存在且唯一，有什么性质。下面这个定理给出了更新方程有解的条件：

更新方程解的存在性定理 (证明不考)

若更新方程 $K(t) = H(t) + \int_0^t K(t-s)\mathrm{d}F(s)$ 中 $H(t)$ 为有界函数，则方程**存在唯一的、在有限区间内有界的解** $K(t)$：

$$K(t) = H(t) + \int_0^t H(t-s)\mathrm{d}M(s)$$

这里 $M(t)$ 是分布函数 $F(t)$ 的更新函数。

根据此定理，$\{N(t)\}$ 的更新函数 $M(t)$ 的积分方程可以转化为

$$M(t) = F(t) + \int_0^t M(t-s)\mathrm{d}F(s) \Rightarrow M(t) = F(t) + \int_0^t F(t-s)\mathrm{d}M(s)$$

Wald 等式

设非负随机序列 $\{X_n\}$ 独立同分布，且 $\mu=\mathbb{E}[X_n]<\infty, n=1,2,\cdots$，则成立

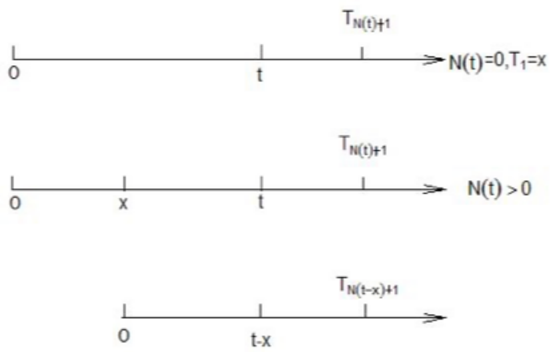
$$\mathbb{E}[T_{N(t)+1}]=\mathbb{E}[\sum_{i=1}^{N(t)+1}X_i]=\mathbb{E}[X_1]\mathbb{E}[N(t)+1]$$

证明. 在证明定理前，直观的理解 $T_{N(t)+1}$ 和 $T_{N(t)}$ 是很有帮助的：

- $T_{N(t)+1}$ 表示第 $N(t)+1$ 次更新的时刻，即 t **时刻后第一次更新的时刻。**
- $T_{N(t)}$ 表示第 $N(t)$ 次更新的时刻，即 t **时刻前最后一次更新的时刻。**

t 时刻时，若还没第一次更新（即 $x>t$），则 t 时刻后第一次更新的时刻 $T_{N(t)+1}=x$，若已经有第一次更新（即 $x\leq t$），将时间轴零点挪至 x 处，相当于求 $t-x$ 时刻后第一次更新的时刻 $T_{N(t-x)+1}$。因此，关于第一次更新的时刻 T_1 取条件期望可得

$$\mathbb{E}[T_{N(t)+1}|T_1=x]=\begin{cases}x & x>t\\ x+\mathbb{E}[T_{N(t-x)+1}] & x\leq t\end{cases}$$



再令 $K(t)=\mathbb{E}[T_{N(t)+1}]$，则

$$\begin{aligned}K(t)&=\mathbb{E}[\mathbb{E}[T_{N(t)+1}|T_1=x]]=\int_0^{\infty}\mathbb{E}[T_{N(t)+1}|T_1=x]\mathrm{d}F(x)\\&=\int_0^t x+\mathbb{E}[T_{N(t-x)+1}]\mathrm{d}F(x)+\int_t^{\infty}x\mathrm{d}F(x)\\&=\int_0^{\infty}x\mathrm{d}F(x)+\int_0^t\mathbb{E}[T_{N(t-x)+1}]\mathrm{d}F(x)\\&=\mathbb{E}[X_1]+\int_0^t K(t-x)\mathrm{d}F(x)\end{aligned}$$

显然上式为一个更新方程，其解 $K(t)$ 满足

$$K(t)=\mathbb{E}[X_1]+\int_0^t\mathbb{E}[X_1]\mathrm{d}M(x)=\mathbb{E}[X_1](1+M(t))=\mu\mathbb{E}[N(t)+1]$$

□

根据 Wald 等式 $\mathbb{E}[T_{N(t)+1}] = \mathbb{E}[X_1]\mathbb{E}[N(t)+1]$，把下标缩进一格，是否能得到 $\mathbb{E}[T_{N(t)}] = \mathbb{E}[X_1]\mathbb{E}[N(t)]$? 我们知道 $T_{N(t)+1} = T_{N(t)} + X_{N(t)+1}$，那么

$$\mathbb{E}[T_{N(t)}] = \mathbb{E}[X_1]\mathbb{E}[N(t)+1] - \mathbb{E}[X_{N(t)+1}]$$

如果 $\mathbb{E}[X_1] = \mathbb{E}[X_{N(t)+1}]$，那么“缩进后”的 Wald 等式依然成立。可惜这是不正确的，有如下结论：

不正确的 Wald 等式

$$\mathbb{E}[X_{N(t)+1}] \neq \mathbb{E}[X_1]$$

从而

$$\mathbb{E}[T_{N(t)}] \neq \mathbb{E}[X_1]\mathbb{E}[N(t)]$$

证明. 使用**检验悖论 (The Inspection Paradox)** 进行说明。检验悖论是指在某些随机过程 (如更新过程) 中，观察到的随机变量 **(如时间间隔)** 往往比其期望值更大。**这是因为较大的间隔更有可能被“采样”到。**具体到这个例子来说，在一个更新过程中，如果你在一个随机时间点 t 进行观察，那么你所在的时间间隔往往比平均时间间隔要长。

$X_{N(t)+1}$ 为紧接在 t 时刻前最后一次更新的一个时间间隔。根据检验悖论，$X_{N(t)+1} > \mathbb{E}[X_1]$，从而 $\mathbb{E}[X_{N(t)+1}] > \mathbb{E}[X_1]$。由于 $\mathbb{E}[X_{N(t)+1}] \neq \mathbb{E}[X_1]$，所以缩进后 Wald 等式不成立。 □

4.4 更新定理

有时我们需要计算**更新方程解 $H(t)$ 的渐近表示**。在更新时间间隔 X_n 的分布为连续分布时，可以使用关键更新定理给出 $H(t)$ 的渐进。

关键更新定理

记 $\mu = \mathbb{E}[X_n]$，设函数 $h(t), t \geq 0$ 满足

1. $h(t)$ 非负不增
2. $\displaystyle\int_0^\infty h(t)\mathrm{d}t < \infty$

$H(t)$ 是更新方程 $H(t) = h(t) + \displaystyle\int_0^t H(t-x)\mathrm{d}F(x)$ 的解，那么若 F 为连续分布，有

$$\lim_{t\to\infty} H(t) = \begin{cases} \dfrac{1}{\mu}\displaystyle\int_0^\infty h(x)\mathrm{d}x & \mu < \infty \\ 0 & \mu = \infty \end{cases}$$

下面看关键更新定理的一个重要应用

剩余寿命与年龄的极限分布

以 $r(t) = T_{N(t)+1} - t$ 表示 t 时刻的剩余寿命（即从 t 开始到下次更新剩余的时间），$s(t) = t - T_{N(t)}$ 为 t 时刻的年龄。设 $F(x)$ 为更新间隔 X_n 的分布函数，则剩余寿命的分布 $\overline{R_y}(t) = P(r(t) > y)$ 满足更新

方程

$$\overline{R_y}(t)=1-F(t+y)+\int_0^t\overline{R_y}(t-x)\mathrm{d}F(x)$$

且 $\overline{R_y}(t)$ 的极限分布为

$$\lim_{t\to\infty}\overline{R_y}(t)=\frac{1}{\mu}\int_y^\infty(1-F(z))\mathrm{d}z,\quad z>0$$

证明. 类似地，仍然关于 $\overline{R}_y(t)=P(r(t)>y)$ 取第一次更新时刻 $X_1=x$ 的条件概率。

1. 当 $x\le t$ 时（即已进行第一次更新），更新过程从 x 时刻起重新计算，于是 t 时刻剩余寿命的分布为 $\overline{R}_y(t-x)$。
2. 当 $x>t$ 时（即未进行第一次更新），分两种情况
   (a) $t<x\le t+y$，那么剩余寿命 $x-t\le y$，从而 $P(r(t)>y)=0$
   (b) $x>t+y$，那么剩余寿命 $x-t>y$，从而 $P(r(t)>y)=1$

由此我们可以写出

$$P(r(t)>y|X_1=x)=\begin{cases}1 & x>t+y\\ 0 & t<x\le t+y\\ \overline{R}_y(t-x) & 0<x\le t\end{cases}$$

进而由全概率公式得

$$\begin{aligned}\overline{R}_y(t)=P(r(t)>y)=&\int_0^\infty P(r(t)>y|X_1=x)\mathrm{d}F(x)\\ =&\int_0^t\overline{R}_y(t-x)\mathrm{d}F(x)+\int_{t+y}^\infty\mathrm{d}F(x)\\ =&1-F(t+y)+\int_0^t\overline{R}_y(t-x)\mathrm{d}F(x)\end{aligned}$$

显然 $1-F(t+y)$ 是一个有界函数，因此上述更新方程的解 $\overline{R}_y(t)$ 为

$$\overline{R}_y(t)=1-F(t+y)+\int_0^t[1-F(t+y-x)]\mathrm{d}M(x)$$

注意到 $1-F(t+y)$ 非负不增，且由于 $\mu=\mathbb{E}[X_1]<\infty$，因此

$$\begin{aligned}\mu=&\int_0^\infty xf(x)\mathrm{d}x=\int_0^\infty F(x)+xf(x)-F(x)\mathrm{d}x=\int_0^\infty\mathrm{d}(xF(x))-\int_0^\infty F(x)\mathrm{d}x\\ =&xF(x)\Big|_0^\infty-\int_0^\infty F(x)\mathrm{d}x=\int_0^\infty(1-F(x))\mathrm{d}x<\infty\end{aligned}$$ (回顾第一章“期望是尾部概率的积分”)

那么 $\int_0^\infty 1-F(t+y)\mathrm{d}t=\int_y^\infty 1-F(z)\mathrm{d}z<\infty$，根据关键更新定理，$\overline{R}_y(t)$ 的极限分布为

$$\lim_{t\to\infty}\overline{R}_y(t)=\frac{1}{\mu}\int_y^\infty(1-F(z))\mathrm{d}z$$

□

类似地，还能导出年龄的极限分布。注意到

- $\{r(t) > x\} \iff$ 过程在 $[t, t+x]$ 没有更新
- $\{s(t) > y\} \iff$ 过程在 $[t-y, t]$ 没有更新
- $\{r(t-y) > x+y\} \iff$ 过程在 $[t-y, t-y+(x+y)]$ 没有更新

所以 $\{r(t) > x, s(t) > y\} \iff \{r(t-y) > x+y\}$，从而

$$\lim_{t\to\infty} P(r(t) > x, s(t) > y) = \lim_{t\to\infty} P(r(t-y) > x+y) = \frac{1}{\mu}\int_{x+y}^{\infty}(1-F(z))\mathrm{d}z$$

那么年龄的极限分布为

$$\lim_{t\to\infty} P(s(t) > y) = \lim_{t\to\infty} P(s(t) > y, r(t) > 0) = \frac{1}{\mu}\int_{y}^{\infty}(1-F(z))\mathrm{d}z$$

4.5 更新过程的拓展

更新回报过程

设 $\{N(t), t \geq 0\}$ 是一个更新过程，允许回报 R_n 依赖于 X_n（回报的多少与等待时间有关），只要求随机向量列 (X_n, R_n) 独立同分布，则

$$R(t) = \sum_{t=1}^{N(t)} R_t$$

称为更新回报过程。

更新回报定理

若 $\{R(t), t \geq 0\}$ 是一个更新回报过程，其更新间隔 $X_1, X_2 \cdots$ 满足 $\mathbb{E}[X_n] < \infty$，每次得到的回报 $\{R_n\}$ 满足 $\mathbb{E}[R_n] < \infty$，则

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{\mathbb{E}[R_1]}{\mathbb{E}[X_1]} \quad \text{a.s.}$$

$$\lim_{t\to\infty} \frac{\mathbb{E}[R(t)]}{t} = \frac{\mathbb{E}[R_1]}{\mathbb{E}[X_1]}$$

- 直观上，长时间后，单位时间的平均报酬等于一次更新的平均报酬除以一次更新所需的平均时间。

产品保修问题

设某公司所售出的商品售价为 c，成本 $c_0 < c$，产品寿命 $X \sim F(x)$ 且 $\mathbb{E}[X] = \mu < \infty$。假定产品一旦损坏，顾客立刻更换或者购买新的产品。现在公司采取如下更换策略

1. 若产品在保修期 $(0, w]$ 内损坏，公司将免费更换新产品，但仍执行原来的更换期。
2. 若产品在 $(w, w+T]$ 内损坏，则按使用时间折价更换新产品。比如产品在 $t \in (w, w+T)$ 时损坏，顾客只需要支付 $c \cdot \dfrac{t-w}{T}$ 购买新产品。
3. 若产品在 $w+T$ 之后损坏，则顾客需要支付原价购买产品。

求长期执行此策略时,
(1) 顾客购买产品后的单位时间期望支出;
(2) 公司单位时间的期望成本;
(3) 公司单位时间的净利润。

解. 不妨设

1. $Y_1, Y_2, \cdots$: 顾客第 1 次、第 2 次、... **付费更换**的时间间隔。$Y_i \sim G(t), i = 1, 2, \cdots$ (因为 $\{Y_n\}$ 独立同分布)。
2. $N(t)$: $(0, t]$ 时间内的更换次数 (包括免费更换、折价更换、付费更换)。
3. T_n: 第 n 次更换的时刻。

(1) 设 $(0, Y_1]$ 内顾客的支出为 c_1,则 $c_1 = \begin{cases} 0 & Y_1 \in (0, w] \\ c \cdot \dfrac{Y_1 - w}{T} & Y_1 \in (w, w+T] \\ c & Y_1 \in (w+T, \infty) \end{cases}$,因此

1. 顾客的期望支出

$$\begin{aligned}
\mathbb{E}[c_1] &= \int_w^{w+T} \frac{c(t-w)}{T} \mathrm{d}F_{Y_1}(t) + \int_{w+T}^{\infty} c dF_{Y_1}(t) \\
&= \frac{c}{T} \int_w^{w+T} (t-w) dG(t) + c\overline{G}(w+T) \\
&= \frac{c}{T} \left(G(t)(t-w) \Big|_w^{w+T} - \int_w^{w+T} G(t) \mathrm{d}t \right) + c\overline{G}(w+T) \\
&= \frac{c}{T} \cdot TG(w+T) - \frac{c}{T} \int_0^T G(w+t) dt + c\overline{G}(w+T) \\
&= cG(w+T) + c\overline{G}(w+T) - \frac{c}{T} \int_0^T G(w+t) dt \\
&= c - \frac{c}{T} \int_0^T G(w+t) dt = \frac{c}{T} \int_0^T \overline{G}(w+T) dt
\end{aligned}$$

2. 顾客付费更换的时间间隔的期望:因为 Y_1 是顾客第一次付费更换的时刻,它等价于 w 时刻后的首次更换时刻:

$$\mathbb{E}[Y_1] = \mathbb{E}[T_{N(w)+1}] = \mathbb{E}[X_1]\mathbb{E}[N(w)+1] = \mu \mathbb{E}[N(w)+1] \qquad \text{(Wald 等式)}$$

根据更新回报定理,顾客购买产品后的单位时间期望支出为 $\dfrac{\mathbb{E}[c_1]}{\mathbb{E}[Y_1]} = \dfrac{\frac{c}{T}\int_0^T \overline{G}(w+T) dt}{\mu \mathbb{E}[N(w)+1]}$

(2) 公司在一个周期 $(0, Y_1]$ 内的成本来自于免费更换新产品,因此

1. 公司的期望成本 $\mathbb{E}[R_1] = c_0(N(w)+1)$
2. 一个付费更换周期的期望时长:$\mathbb{E}[Y_1] = \mu(\mathbb{E}[N(w)+1])$

根据更新回报定理，公司的单位时间期望成本为 $\dfrac{\mathbb{E}[R_1]}{\mathbb{E}[Y_1]} = \dfrac{c_0}{\mu}$

(3) 公司的单位时间净利润为单位时间期望收入 *(*即顾客的单位时间期望支出*)* 减去公司的单位时间期望成本，即 $\dfrac{c\int_0^T \overline{G}(w+T)dt}{T\mu\mathbb{E}[N(w)+1]} - \dfrac{c_0}{\mu}$

交替更新过程

考虑只有两个状态的系统：开（用 1 表示）或关（用 0 表示），系统在 $t=0$ 时是开的且持续开的时间为 Z_1，接着关闭且持续时间为 Y_1，之后又开着持续时间为 Z_2，又关闭时间为 Y_2，如此开关交替重复下去，设 $\{(Z_n, Y_n), n \geq 1\}$ 为独立同分布随机向量序列（允许 Z_i, Y_i 之间不独立），记 $X_n = Z_n + Y_n$，$T_0 = 0, T_n = \sum\limits_{i=1}^{n} X_i$。再记

$$\zeta_t = \begin{cases} 1 & T_n \leq t \leq T_n + Z_{n+1} \\ 0 & T_n + Z_{n+1} \leq t < T_{n+1} \end{cases}$$

则称 $\{\zeta_t, t \geq 0\}$ 为交错更新过程。

- X_n 即第 n 次开-关的持续时间

交替更新定理

设 $H(t) = P(Z_n \leq t)$，$G(t) = P(Y_n \leq t)$，$F(t) = P(X_n \leq t)$，并记 t 时刻系统开着的概率为 $P(t) = P(\zeta_t = 1)$，若 $\mathbb{E}[X_n] < \infty$，则

$$\lim_{t\to\infty} P(t) = \frac{\mathbb{E}[Z_n]}{\mathbb{E}[Z_n] + \mathbb{E}[Y_n]}$$

证明. 记 $X_1 = Z_1 + Y_1$，对第一次更新时刻 X_1 取条件概率得

$$P(t\text{ 时刻系统开着}|X_1 = x) = \begin{cases} P(t-x) & x < t \\ P(Z_1 > t|X_1 = x) & x \geq t \end{cases}$$

从而

$$\begin{aligned} P(t) &= \int_0^\infty P(t\text{ 时刻系统开着}|X_1 = x)\mathrm{d}F(x) = \int_0^t P(t-x)\mathrm{d}F(x) + \int_t^\infty P(Z_1 > t|X_1 = x)\mathrm{d}F(x) \\ &= P(Z_1 > t, X_1 > t) + \int_0^t P(t-x)\mathrm{d}F(x) \\ &= P(Z_1 > t) + \int_0^t P(t-x)\mathrm{d}F(x) \\ &\triangleq \overline{H(t)} + \int_0^t P(t-x)\mathrm{d}F(x) \end{aligned}$$

这个更新方程的解 $P(t)$ 满足

$$P(t) = \overline{H(t)} + \int_0^t \overline{H(t-x)}\mathrm{d}M(x)$$

又因为

1. $\overline{H(t)} = P(Z_1 > t)$ 非负不增
2. $\displaystyle\int_0^\infty \overline{H}(t)\mathrm{d}t = \int_0^\infty P(Z_1 > t)\mathrm{d}t = \mathbb{E}[Z_1] < \infty$ (见第一章，数学期望为尾部概率的积分)

根据关键更新定理，$\displaystyle\lim_{t\to\infty} P(t) = \frac{1}{\mu}\mathbb{E}[Z_1] = \frac{\mathbb{E}[Z_1]}{\mathbb{E}[X_1]} = \frac{\mathbb{E}[Z_1]}{\mathbb{E}[Z_1] + \mathbb{E}[Y_1]}$ □

5 Markov 链

5.1 离散 Markov 链的定义与性质

Markov 链

随机序列 $\{X_n, n \geq 0\}$ 称为 Markov 链，若其取值集合 S 有限或可列，且对 $\forall n \geq 0$ 以及任意状态 $i_0, \cdots, i_{n-1}, i_n, i_{n+1} \in S$，有

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

上式刻画了 Markov 性（即无后效性）。

- **转移概率：** $\forall i, j \in S$，称条件概率 $P(X_{n+1} = j | X_n = i) = p_{ij}(n)$ 为 n 时刻的一步转移概率。
- **n 步转移概率：** $\forall i, j \in S$，称条件概率 $P(X_{m+n} = j | X_m = i) = p_{ij}^{(n)}, m \geq 0, n \geq 1$ 为 Markov 链的 n 步转移概率。相应的 $\mathbf{P}^{(n)} = (p_{ij}^{(n)})$ 为 n 步转移矩阵。
  - 特别地，约定 $p_{ij}^{(0)} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$
- **时齐 Markov 链：** 若 Markov 链的转移概率 $p_{ij}(n)$ 与 n 无关，称 $\{X_n, n \geq 0\}$ 为时齐 Markov 链。

转移矩阵

记 $\mathbf{P} = (p_{ij})$，则称 $\mathbf{P}$ 为 $\{X_n, n \geq 0\}$ 的转移矩阵，其满足两个性质：

1. 非负性：$\forall i, j \in S, p_{ij} \geq 0$
2. 行和为 0：$\forall i \in S, \sum\limits_{j \in S} p_{ij} = 1$

设 $S = \{1, \cdots, s\}$，如果知道 Markov 链的初始分布 $\mathbf{p}_0 = \{P(X_0 = 1), \cdots, P(X_0 = s)\}$，那么可以通过乘转移矩阵实现更新：注意到

$$\forall j \in S, P(X_n = j) = \sum_{i \in S} P(X_n = j | X_{n-1} = i) P(X_{n-1} = i) = \mathbf{p}_{n-1} \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{sj} \end{bmatrix}$$

因此

$$\mathbf{p}_n = \mathbf{p}_{n-1} \mathbf{P}$$

Chapman-Kolmogorov 方程 (C-K 方程)

$\forall i, j \in S, m \geq 0, n \geq 1$，有

1. $p_{ij}^{(m+n)} = \sum\limits_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$

2. $\mathbf{P}^{(n)} = \mathbf{P} \cdot \mathbf{P}^{(n-1)} = \cdots = \mathbf{P}^n$

证明.

$$\begin{aligned}
p_{ij}^{(m+n)} &= P(X_{m+n}=j|X_0=i) = \frac{P(X_{m+n}=j, X_0=i)}{P(X_0=i)} \\
&= \sum_{k\in S} \frac{P(X_{m+n}=j, X_m=k, X_0=i)}{P(X_0=i)} && \text{(Markov 性)} \\
&= \sum_{k\in S} \frac{P(X_{m+n}=j, X_m=k, X_0=i)}{P(X_0=i)} \frac{P(X_m=k, X_0=i)}{P(X_m=k, X_0=i)} \\
&= \sum_{k\in S} P(X_{m+n}=j|X_m=k, X_0=i)\cdot P(X_m=k|X_0=i) \\
&= \sum_{k\in S} P(X_{m+n}=j|X_m=k)\cdot P(X_m=k|X_0=i) = \sum_{k\in S} p_{kj}^{(n)} p_{ik}^{(m)}
\end{aligned}$$

式 (1) 的矩阵形式即 $\mathbf{P}^{(m+n)} = \mathbf{P}^{(m)}\mathbf{P}^{(n)}$，于是 (2) 显然得证。 □

5.2 离散 Markov 链的状态

5.2.1 状态的关系

可达

若存在 $n \geq 0$ 使得 $p_{ij}^{(n)} > 0$，称状态 i 可达状态 j，记作 $i \to j$

互通

若 $i \to j$ 且 $j \to i$，则称 i, j 互通，记作 $i \leftrightarrow j$

互通是一种等价关系

互通是一种等价关系，即满足

- 自返性：$i \leftrightarrow i$
- 对称性：$i \leftrightarrow j$，则 $j \leftrightarrow i$
- 传递性：$i \leftrightarrow j$，$j \leftrightarrow k$，则 $i \leftrightarrow k$

任何两个互通关系归为一类，同一类的状态应该都是互通的，并且一个状态不能同时属于两个不同的类。

证明. 传递性的证明不是显然的，事实上，由 $i \to j$ 与 $j \to k$ 知 $\exists m, n$，s.t. $p_{ij}^{(m)} > 0, p_{jk}^{(n)} > 0$，那么

$$p_{ik}^{(m+n)} = \sum_{j\in S} p_{ij}^{(m)} p_{jk}^{(n)} \geq p_{ij}^{(m)} p_{jk}^{(n)} > 0$$

这说明 $i \to k$。 □

可约

若 Markov 链中只存在一个类，则称其是不可约的，否则是可约的。

比如在赌徒输光问题中，其转移状态矩阵为 **P**。注意到赌徒输光问题中任意两状态 $i, j \in [1, n-1]$ 都是互通的，因此这个 Markov 链可以划分为 3 类：$\{0\}, \{1, \cdots, n-1\}, \{n\}$

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

图 2: gambler's ruin

周期

若集合 $\{n : n \geq 1, p_{ii}^{(n)} > 0\}$ 非空，则称这个集合中元素的最大公约数 $d = d(i)$ 为状态 i 的周期。

- 若 $d > 1$，称 i 是周期的
- 若 $d = 1$，则称 i 是非周期的

特别地，当 $\{n : n \geq 1, p_{ii}^{(n)} > 0\} = \varnothing$ 时，称 i 的周期为无穷大。

- 集合 $\{n : n \geq 1, p_{ii}^{(n)} > 0\}$ 称为**自返步数集**，表示所有从 i 出发，以正概率能返回 i 的步数的集合。

值得注意的是，以正概率从 i 出发再返回 i 的步数一定是周期的倍数，但周期的倍数不一定能以正概率从 i 出发再返回 i。如以下的例子：$\{n : n \geq 1, p_{11}^{(n)} > 0\} = \{4, 6, 8, 10, \cdots\}$，从而周期 $d(1) = 2$，但从状态 1 出发两步不能回到状态 2。

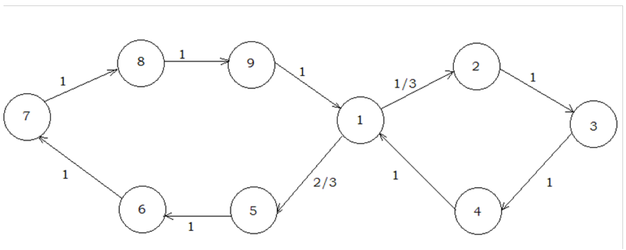


图 3: Markov 链的周期性

周期的互通性

若 $i \leftrightarrow j$，则 $d(i) = d(j)$.

证明. 令 $A = \{n : n \geq 1, p_{ii}^{(n)} > 0\}$，则 $\exists m, n > 0$，使得 $p_{ij}^{(m)} > 0$，$p_{ji}^{(n)} > 0$ (因为 $i \leftrightarrow j$，存在路径 $i \xrightarrow{m} j \xrightarrow{n} i$)。因此 $m + n \in A$，进而 $d(i) | (m + n)$。

再令 $B = \{n : n \geq 1, p_{jj}^{(n)} > 0\}$，$\forall s \in B$，成立 $m + n + s \in A$（因为存在路径 $i \xrightarrow{m} j \xrightarrow{s} j \xrightarrow{n} i$），进而 $d(i) | (m + n + s)$，从而 $d(i) | s$。由 s 的任意性，$d(i)$ 是 B 的公约数，且 $d(j)$ 是 B 的最大公约数，因此 $d(i) | d(j)$。

将上述过程颠倒（即先令 $A=\{n:n\geq 1,p_{jj}^{(n)}>0\},B=\{n:n\geq 1,p_{ii}^{(n)}>0\}$），同理易得 $d(j)|d(i)$，因此 $d(i)=d(j)$. □

5.2.2 状态的分类

常返态与非常返态

$\forall i,j\in S$，以 $f_{ij}^{(n)}$ 表示从 i 出发经 n 步以后首次到达 j 的概率，则

$$f_{ij}^{(0)}=\delta_{ij},f_{ij}^{(n)}=P(X_n=j,X_k\neq j,k=1,\cdots,n-1|X_0=i),n>0$$

令 $f_{ij}=\sum\limits_{n=1}^{\infty}f_{ij}^{(n)}$，

- 若 $f_{jj}=1$，称状态 j 为常返态 (Recurrent state)。
- 若 $f_{jj}<1$，称状态 j 为非常返态或瞬过状态 (Transient state)。

称 $A_n=\{X_n=j,X_k\neq j,k=1,\cdots,n-1|X_0=i\}$ 为首达事件。事件 $\bigcup\limits_{n=1}^{\infty}A_n$ 表示总有一个 n 使得状态 i 经 n 步后可到达 j，注意到**首达事件是不相交的**，即 $A_i\cap A_j=\varnothing,i\neq j$，于是

$$f_{ij}=P(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}P(A_n)$$

因此，f_{ij} 即表示从 i 出发，在有限步内（没有指定具体多少步）能达到 j 的概率。

正常返态与零常返态

记从 i 出发再返回 i 的平均步数为

$$\mu_i=\sum_{n=1}^{\infty}nf_{ii}^{(n)}$$

，则对于常返状态 i，

- 若 $\mu_i<+\infty$，则称 i 为正常返态 (Positive recurrent state)
- 若 $\mu_i=+\infty$，则称 i 为零常返态 (Null recurrent state)
- 若 $\mu_i<+\infty$ 且 $d(i)=1$，则称 i 为遍历状态 (Transitive state)
- 若 $f_{ii}^{(1)}=1\Rightarrow\mu_i=1$（直观上就是总是回到状态 i），则称 i 为吸收状态 (Absorbing state)

可以用一张图来概括 Markov 链中各状态的判别方法：

$$\text{Markov链的状态分类}\begin{cases} \text{常返状态}\ (f_{ii}=1) \begin{cases} \text{正常返状态}\ (f_{ii}=1,\mu_i<+\infty) \quad \supset \text{遍历状态}\ (f_{ii}=1,\mu_i<+\infty,d(i)=1) \quad \supset \text{吸收状态}\ (f_{ii}=1,\mu_i=1,d(i)=1) \\ \text{零常返状态}\ (f_{ii}=1,\mu_i=+\infty) \end{cases} \\ \text{非常返状态}\ (f_{ii}<1) \end{cases}$$

图 4: Markov 链中状态的分类

首达分解定理

$\forall i,j\in S$，$1\le n<\infty$，有

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

证明. 注意到"首达事件不相交"这一性质，由全概率公式

$$p_{ij}^{(n)}=\sum_{l=1}^{n} P(l\text{步首达}j)\cdot P(n-l\text{步返回}j)=\sum_{l=1}^{n} f_{ij}^{(l)}p_{jj}^{(n-l)}$$

□

常返态判别准则

1. 状态 i 为常返状态 $\iff \sum\limits_{n=0}^{\infty} p_{ii}^{(n)}=\infty$
2. 状态 i 为非常返状态 $\iff \sum\limits_{n=0}^{\infty} p_{ii}^{(n)}=\dfrac{1}{1-f_{ii}}<\infty$

证明.

$$\begin{aligned}\sum_{n=0}^{\infty} p_{ii}^{(n)} &= p_{ii}^{(0)}+\sum_{n=1}^{\infty}\sum_{l=1}^{n} f_{ii}^{(l)}p_{ii}^{(n-l)} && (\text{首达分解 } p_{ii}^{(n)})\\ &= 1+\sum_{l=1}^{\infty}\sum_{n=l}^{\infty} f_{ii}^{(l)}p_{ii}^{(n-l)}\\ &= 1+\sum_{l=1}^{\infty} f_{ii}^{(l)}\sum_{m=0}^{\infty} p_{ii}^{(m)}\\ &= 1+f_{ii}\sum_{m=0}^{\infty} p_{ii}^{(m)}\end{aligned}$$

容易看出当 $\sum\limits_{m=0}^{\infty} p_{ii}^{(m)}<\infty$ 时，可以移项操作得到 $\sum\limits_{m=0}^{\infty} p_{ii}^{(m)}=\dfrac{1}{1-f_{ii}}\in(0,+\infty)\iff f_{ii}<1\iff i$ 为非常返状态。

换言之，$\sum\limits_{m=0}^{\infty} p_{ii}^{(m)}<\infty\iff f_{ii}<1$，其逆否命题为 $\sum\limits_{m=0}^{\infty} p_{ii}^{(m)}=\infty\iff f_{ii}=1$。得证。 □

常返态可达状态必互通

常返态可以到达的状态与其互通，且从该到达状态有限步返回的概率为。即：若 $i \to j$，且 i 为常返态，则 $f_{ji} = 1$，自然有 $j \to i$

证明.（反证）假设 $f_{ji} < 1$，则从 j 出发有限步不能回到 i 的概率为 $1 - f_{ji} > 0$，但 $i \to j$，则 $\exists n > 0$ 使得 $p_{ij}^{(n)} > 0$，那么从 i 出发有限步不能回到 i 的概率

$$1 - f_{ii} > p_{ij}^{(n)} \cdot (1 - f_{ji}) > 0$$

这里 $p_{ij}^{(n)} \cdot (1 - f_{ji})$ 是从 i 出发有限步不能回到 i 的一种情况，即 $i \xrightarrow{n} j$，但 $j \nrightarrow i$。由上述不等式知 $f_{ii} < 1$，但这与 i 为常返态矛盾！假设不成立，原命题得证。 □

- 该定理说明：常返态 i 可以到达的状态 j 必然与其互通，**从而同属一个类**。直观上也很好理解，如果 i 选择走到 j，但 j 不能到 i，则天然与 i 是常返的矛盾。

常返态是类性质

常返性是一个类性质。即

1. 若 $i \leftrightarrow j$，则 i, j 同为常返或非常返状态。
2. 若 i, j 同为常返状态，则他们同为正常返或零常返状态。

证明. 对于 1，由于 $i \leftrightarrow j$，$\exists n, m \geq 0$，使得 $p_{ij}^{(n)} > 0, p_{ji}^{(m)} > 0$，根据 Chapman-Kolmogorov 方程，

$$p_{ii}^{(n+l+m)} = \sum_{j,k \in S} p_{ij}^{(n)} p_{jk}^{(l)} p_{ki}^{(m)} \geq p_{ij}^{(n)} p_{jj}^{(l)} p_{ji}^{(m)}$$

两边关于 l 求和

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

类似地

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

上面两公式表示 $\sum_{l=0}^{\infty} p_{ii}^{(l)}$ 和 $\sum_{l=0}^{\infty} p_{jj}^{(l)}$ 是相互控制的，如果一个为 ∞ 或有限值，则另一个也为 ∞ 或有限值。

对于 2，注意到 $p_{ii}^{(n+l+m)} \geq p_{ij}^{(n)} p_{jj}^{(l)} p_{ji}^{(m)} \geq 0$，根据 Markov 链基本极限定理，当 i 为零常返态时，两边关于 $l \to \infty$ 得

$$0 \geq p_{ij}^{(n)} p_{ji}^{(m)} \lim_{l \to \infty} p_{jj}^{(l)} \geq 0$$

可得 $\lim_{l \to \infty} p_{jj}^{(l)} = 0$，即 j 也为零常返态。反之从 j 为零常返态也能推出 i 为零常返态。取逆否命题亦可知 i, j 同为正常返态。得证。 □

以上定理说明了

1. 常返态是一个封闭类。（若 i, j 互通，只要 i 为常返态则 j 必然为常返态）

2. 在常返态内部，零常返态是一个封闭类（若 i,j 互通，只要 i 为零常返态，则 j 必然为零常返态，而不是正常返态）、正常返态也是一个封闭类。

3. 常返态无法到达非常返态，但非常返态可以到达常返态。

5.3 离散 Markov 链的极限性质

5.3.1 $p_{ii}^{(n)}$ 的极限性质

Markov 链基本极限定理

若状态 i 是周期为 d 的常返态，则 $\lim\limits_{n\to\infty} p_{ii}^{(nd)} = \dfrac{d}{\mu_i}$

由此定理出发很容易得到三条推论：

1. 在 i 为常返态的情况下，i 为零常返态 $\iff \mu_i = \infty \iff \lim\limits_{n\to\infty} p_{ii}^{(n)} = 0$
2. 在 i 为常返态的情况下，i 为正常返态 $\iff \mu_i < \infty \iff \lim\limits_{n\to\infty} p_{ii}^{(nd)} = \dfrac{d}{\mu_i}$
3. i 为非常返态 $\Rightarrow \lim\limits_{n\to\infty} p_{ii}^{(n)} = 0$

证明. **证明推论 1：** 若 i 为零常返态，则 $p_{ii}^{(nd)} \to 0$，当 m 不是 d 的倍数时，也有 $p_{ii}^{(m)} = 0$，综合起来有 $p_{ii}^{(n)} \to 0$；反之，若 $p_{ii}^{(n)} \to 0$，如果 i 为正常返态，则 $\mu_i < \infty$，用 Markov 基本极限定理知 $p_{ii}^{(nd)} \to \dfrac{d}{\mu_1} > 0$，矛盾！

证明推论 3： 若 i 为非常返态，则由常返态判别准则知 $\sum\limits_{n=0}^{\infty} p_{ii}^{(n)} < \infty$，这个级数有限说明了求和项 $p_{ii}^{(n)} \to 0$。 □

Markov 链基本极限定理提供了一个判断常返态/非常返态、零常返态/正常返态的方法

- 首先需要看 $\sum\limits_{n=0}^{\infty} p_{ii}^{(n)}$ 是 ∞ 还是 $< \infty$ 作出常返态还是非常返态的判断。
- 然后看 $p_{ii}^{(n)}$ 的极限作出零常返态还是正常返态的判断。

5.3.2 $p_{ij}^{(n)}$ 的极限性质

有关 $p_{ij}^{(n)}$ 的极限性质，要考虑两件事

1. $\lim\limits_{n\to\infty} p_{ij}^{(n)}$ 是否存在
2. $\lim\limits_{n\to\infty} p_{ij}^{(n)}$ 是否与出发状态 i 和到达状态 j 有关

到达状态为非常返态/零常返态

若状态 j 为非常返态或零常返态，则对任意的出发状态 $i \in S$，有 $\lim\limits_{n\to\infty} p_{ij}^{(n)} = 0.$

证明. 由首达分解定理知 $p_{ij}^{(n)} = \sum\limits_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$，从而对 $N < n$，有

$$p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} \leq \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}$$

由于 j 为非常返态或零常返态，$\lim_{n\to\infty} p_{jj}^{(n)}=0$，故而两边令 $n\to\infty$，

$$\lim_{n\to\infty} p_{ij}^{(n)} \leq \sum_{l=N+1}^{\infty} f_{ij}^{(l)} \to 0$$

这是因为 $f_{ij}=\sum_{l=1}^{\infty} f_{ij}^{(l)} \leq 1$，故 $\sum_{l=N+1}^{\infty} f_{ij}^{(l)} \to 0, N\to\infty$。由此 $\lim_{n\to\infty} p_{ij}^{(n)}=0$ □

到达状态为正常返态

若状态 j 为正常返态，且 $d(j)=1$，则如果 $i\leftrightarrow j$，则 $\lim_{n\to\infty} p_{ij}^{(n)}=\frac{1}{\mu_j}$.

证明. 由首达分解定理知 $p_{ij}^{(n)}=\sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$，从而对 $N<n$，有

$$\sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} \leq p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} \leq \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}$$

由于状态 j 为正常返态，故 $\lim_{n\to\infty} p_{jj}^{(n)}=\frac{d(j)}{\mu_j}$。又 $i\leftrightarrow j$，故 $f_{ij}=1$。两边令 $n\to\infty$，

$$\frac{1}{\mu_j}\sum_{l=1}^{N} f_{ij}^{(l)} \leq p_{ij}^{(n)} \leq \frac{1}{\mu_j}\sum_{l=1}^{N} f_{ij}^{(l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}$$

再令 $N\to\infty$ 立刻知 $\lim_{n\to\infty} p_{ij}^{(n)}=\frac{1}{\mu_j}$. □

出发状态为非常返态，到达状态为正常返态

若状态 i 为非常返态，状态 j 为正常返态，$i\to j$ 且 $d(j)=1$，则 $\lim_{n\to\infty} p_{ij}^{(n)}=\frac{f_{ij}}{\mu_j}$.

证明. 由首达分解定理知 $p_{ij}^{(n)}=\sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$，从而对 $N<n$，有

$$\sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} \leq p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} \leq \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}$$

由于状态 j 为正常返态，故 $\lim_{n\to\infty} p_{jj}^{(n)}=\frac{d(j)}{\mu_j}$。又 $i\to j$，故 $f_{ij}<1$ 有限。两边令 $n\to\infty$，

$$\frac{1}{\mu_j}\sum_{l=1}^{N} f_{ij}^{(l)} \leq p_{ij}^{(n)} \leq \frac{1}{\mu_j}\sum_{l=1}^{N} f_{ij}^{(l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}$$

再令 $N\to\infty$ 立刻知 $\lim_{n\to\infty} p_{ij}^{(n)}=\frac{f_{ij}}{\mu_j}$. □

由此，$p_{ij}^{(n)}$ 的极限性质可以用下面这个流程图概括：

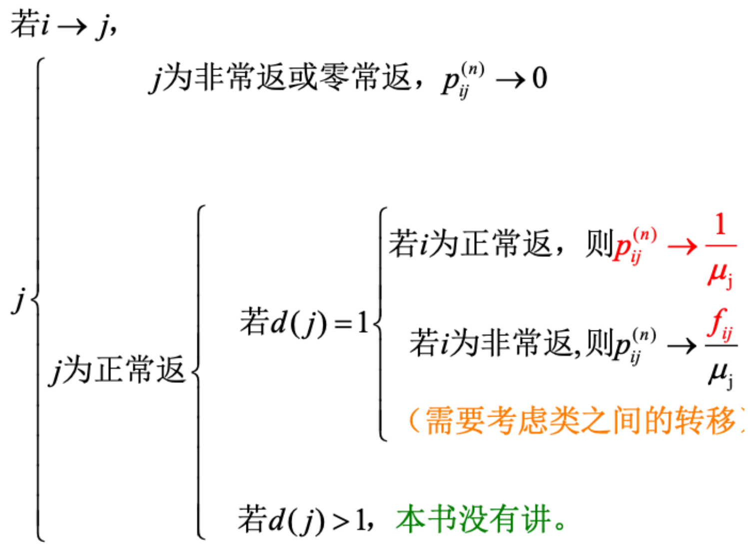


图 5: $p_{ij}^{(n)}$ 的极限性质

有限状态 Markov 链的状态

有限状态 Markov 链中

1. 不会全为非常返态
2. 不会有零常返态

从而不可约有限 Markov 链的一切状态都是正常返的。

证明. 对于 1, 假设 $S=\{1,\cdots,N\}$ 中每个状态均为非常返态, 则对于 $i \to j$, 总有 $p_{ij}^{(n)} \to 0$, 从而 $\lim\limits_{n\to\infty}\sum\limits_{j=1}^{N} p_{ij}^{(n)} = 0$。然而恒成立 $\sum\limits_{j=1}^{N} p_{ij}^{(n)} = 1$，矛盾！假设不成立。
对于 2，假设 S 有零常返态 i，设 C 为零常返态 i 可达的所有状态的集合 $C=\{j : i \to j\}$。因为 i 为 (零) 常返态，由“常返态可达状态必互通”知 $\forall j \in C$，$j \leftrightarrow i$。因此 j 也为零常返态。进而 $p_{ij}^{(n)} \to 0$，从而 $\lim\limits_{n\to\infty}\sum\limits_{j=1}^{N} p_{ij}^{(n)} = 0$。然而恒成立 $\sum\limits_{j=1}^{N} p_{ij}^{(n)} = 1$，矛盾！假设不成立，原命题得证。 □

[推论]：若 Markov 链有一个零常返状态，则必有无限个零常返状态。

5.4 离散 Markov 链的平稳分布

平稳分布 (Stationary Distribution)

对于 Markov 链，称概率分布 $\boldsymbol{\pi}=\{\pi_1,\cdots,\pi_i,\cdots\}$ 为平稳分布，若

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

即

$$\forall j \in S, \pi_j = \sum_{i \in S} \pi_i p_{ij}$$

对于平稳分布，显然有

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}\mathbf{P}^2 = \cdots = \boldsymbol{\pi}\mathbf{P}^n$$

也就是说，$X_0, X_1, \cdots, X_n, \cdots$ 的概率分布都相同。

遍历 Markov 链

称 Markov 链是遍历的，如果所有状态相通且均是周期为 1 的正常返状态。

极限分布 (Asymptotic Distribution)

对于遍历 Markov 链，称概率分布 $\boldsymbol{\pi}^* = \{\pi_1^*, \cdots, \pi_i^*, \cdots\}$ 为极限分布，若

$$\pi_j^* = \lim_{n \to \infty} p_{ij}^{(n)}$$

根据之前关于 $p_{ij}^{(n)}$ 极限性质的论述，容易知道遍历 Markov 链的极限分布为 $\boldsymbol{\pi}^* = \{\frac{1}{\mu_1}, \cdots, \frac{1}{\mu_i}, \cdots\}$。

平稳分布的存在性

对于不可约、周期为 1 的 Markov 链

1. 若它是遍历的，则 $\pi_j = \lim_{n \to \infty} p_{ij}^{(n)} > 0, j \in S$ 是平稳分布且是唯一的平稳分布 $\{\pi_j = \frac{1}{\mu_j}\}$。
2. 若状态都是非正常返的，或都是零常返的，则平稳分布不存在。

证明. 对于 1，能立刻推知 i, j 为周期为 1 的正常返状态，从而 $\pi_j = \lim_{n \to \infty} p_{ij}^{(n)} = \frac{1}{\mu_j}$。由于 $\sum_{i \in S} p_{ij}^{(n)} = 1$，那么

$$1 = \lim_{n \to \infty} \sum_{i \in S} p_{ij}^{(n)} = \sum_{i \in S} \lim_{n \to \infty} p_{ij}^{(n)} = \sum_{i \in S} \pi_j$$

由 C-K 方程知

$$\begin{aligned}
&p_{ij}^{(n+1)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj} \\
\Rightarrow &\lim_{n \to \infty} p_{ij}^{(n+1)} = \lim_{n \to \infty} \sum_{k \in S} p_{ik}^{(n)} p_{kj} = \sum_{k \in S} (\lim_{n \to \infty} p_{ik}^{(n)}) p_{kj} \\
\Rightarrow &\pi_j = \sum_{k \in S} \pi_k p_{kj}
\end{aligned}$$

从而 $\{\pi_j, j \in S\}$ 是平稳分布。假设另外还有一个平稳分布 $\{\tilde{\pi}_j, j \in S\}$，则 $\tilde{\pi}_j = \sum_{k \in S} \tilde{\pi}_k p_{kj}$，从而能在此基础

上归纳得到 $\tilde{\pi}_j = \sum\limits_{k\in S} \tilde{\pi}_k p_{kj}^{(n)}$，两边关于 n 取极限得到

$$\begin{aligned} \tilde{\pi}_j &= \lim_{n\to\infty} \sum_{k\in S} \tilde{\pi}_k p_{kj}^{(n)} \\ \tilde{\pi}_j &= \sum_{k\in S} \tilde{\pi}_k (\lim_{n\to\infty} p_{kj}^{(n)}) \\ \Rightarrow \tilde{\pi}_j &= \sum_{k\in S} \tilde{\pi}_k \pi_j = \pi_j \qquad (\sum_{k\in S} \tilde{\pi}_k = 1) \end{aligned}$$

这说明了 $\{\pi_j, j \in S\}$ 的唯一性。
对于 2，假设存在一个平稳分布 $\{\pi_j, j\in S\}$，由 1 知

$$\pi_j = \sum_{k\in S} \pi_k p_{kj}^{(n)}$$

由于 j 为零常返状态，因此 $p_{kj}^{(n)} \to 0$，从而 $\pi_j = 0$。这就矛盾了！平稳分布首先必须是一个**概率分布**，如果所有的 $j \in S$ 都有 $\pi_j = 0$，显然构不成一个概率分布。假设不成立，原命题成立。 □

这个定理告诉了我们

- 当 Markov 链是不可约遍历链时，极限分布就是平稳分布，而且是唯一的平稳分布。
- 求解 Markov 链**极限分布**的方法：解线性方程组 $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ 得到平稳分布，如果 Markov 链是遍历的，那么极限分布就是平稳分布。
- 求解平均自返步数的公式：**Markov 链是不可约遍历链时**，$\mu_j = \dfrac{1}{\pi_j}$，其中 π_j 是平稳分布。

极限分布不存在但平稳分布存在

极限分布与平稳分布不一定都存在！以下为一个反例说明极限分布不存在，但平稳分布存在。

设转移概率矩阵为 $\mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$，注意到 $\mathbf{P}^n = \begin{cases} \mathbf{P} & n\text{为奇数} \\ \mathbf{I} & n\text{为偶数} \end{cases}$，其极限分布不存在！但是求解

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

立刻得知 $\boldsymbol{\pi} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$，因此该 Markov 链平稳分布存在。
(这是因为这个 Markov 链是不可约正常返但周期为 2 的，平稳分布 $\neq$ 极限分布)

5.5 离散 Markov 链模型的应用

5.5.1 群体消失模型 (分支过程)

考虑一个能产生同类后代的个体组成的群体。每一个体生命结束时以概率 $p_j (j = 0, 1, 2, 3, \dots)$ 产生了 j 个新的后代，与别的个体产生的后代个数相互独立。初始的个体数以 X_0 表示，称为第零代的总数；第零代的后代构成第一代，其总数记为 X_1，第一代的每个个体以同样的分布产生第二代，$\cdots$，一般地，以 X_n 记第 n 代的总数。此 Markov 链 $\{X_n, n = 0, 1, 2, \dots\}$ 称为分支过程。现在假设群体是从单个祖先开始的，即 $X_0 = 1$，则有

$$X_n = \sum_{i=1}^{X_{n-1}} Z_{n-1,i}$$

其中 $Z_{n,i}$ 表示第 n 代的第 i 个成员的后代的个数。我们关心这个 Markov 链的一些性质，包括**期望** $\mathbb{E}(X_n)$，**方差** $\mathsf{Var}(X_n)$ 和**消亡概率** π_0 等。

1. **期望：**首先来考虑第 n 代的平均个体数 $\mathbb{E}[X_n]$，对 X_n 取条件期望，有

$$\begin{aligned} E[X_n] &= E[E[X_n|X_{n-1}]] = \mu E[X_{n-1}] \\ &= \mu^2 E[X_{n-2}] = \cdots = \mu^n \end{aligned}$$

其中 $\mu = \sum_{i=0}^{\infty} i p_i$ 是每个个体的后代个数的均值。从而可以看出，若 $\mu < 1$，则平均个体数单调下降趋于 0。若 $\mu = 1$，各代平均个体数相同。当 $\mu > 1$ 时，平均个体数按指数阶上升至无穷。

2. **方差：**根据条件方差公式

$$\mathsf{Var}(X_n) = \mathbb{E}(\mathsf{Var}(X_n|X_{n-1})) + \mathsf{Var}(\mathbb{E}(X_n|X_{n-1}))$$

给定 X_{n-1} 后，X_n 正是 X_{n-1} 个独立随机变量的和，每个具有分布 $\{p_j, j \geq 0\}$。因此 $\mathbb{E}(X_n|X_{n-1}) = \mu X_{n-1}$，　$\mathsf{Var}(X_n|X_{n-1}) = \sigma^2 X_{n-1}$，于是

$$\mathsf{Var}(X_n) = \mathbb{E}[\sigma^2 X_{n-1}] + \mathsf{Var}[\mu X_{n-1}] = \sigma^2 \mu^{n-1} + \mu^2 \mathsf{Var}(X_{n-1})$$

根据此递推公式可以解出

$$\mathsf{Var}(X_n) = \begin{cases} \sigma^2 \mu^{n-1} \cdot \dfrac{1-\mu^n}{1-\mu} & \mu \neq 1 \\ n\sigma^2 & \mu = 1 \end{cases}$$

3. **种群消亡概率：**下面就来考虑群体最终会消亡的概率 π_0。对第一代个体数取条件，则

$$\begin{aligned} \pi_0 &= P\{\text{群体消亡}\} \\ &= \sum_{j=0}^{\infty} P\{\text{群体消亡}|X_1 = j\} \cdot p_j \\ &= \sum_{j=0}^{\infty} \pi_0^j p_j \end{aligned}$$

上面的第二个等式是因为群体最终灭绝是以第 1 代为祖先的 j 个家族全部消亡，而各家族已经假定为独立的，每一家族灭绝的概率均为 π_0。**直观上，家族消亡与否与 μ 有关。**在不考虑 $p_0 = 1$ 和 $p_0 = 0$ 的平凡情况下（即家族在第零代后就消失或永不消失），我们有以下定理

群体消亡概率 (不考)

设 $0 < p_0 < 1$，则群体消亡概率 $\pi_0 = 1 \iff \mu \leq 1$

考试不考. 令 $F(x) = \sum_{j=0}^{\infty} x^j p_j$，则 $F(x) - x = 0$ 在 $(0,1]$ 上至少有两个解 $x = 1$ 和 $x = \pi_0$。首先，当 $x \in (0,1]$ 时

$$[F(x) - x]'' = \sum_{j=2}^{\infty} j(j-1)x^{j-2} p_j > 0$$

因此 $F(x)-x$ 在 $(0,1]$ 上是严格凸函数。另一方面，当 $x\in(0,1]$ 时

$$[F(x)-x]'=\left(\sum_{j=1}^{\infty} jx^{j-1}p_j\right)-1\leq \sum_{j=1}^{\infty} jp_j-1=\mu-1$$

且 $x=1$ 时，$[F(x)-x]'=\mu-1$，故 $\mu\leq 1 \iff F(x)-x$ 在 $(0,1]$ 单调递减。而 $F(x)-x$ 在 $(0,1]$ 单调递减 $\iff$ 在 $(0,1]$ 上有且仅有一个解，因此解 $x=\pi_0$ 和 $x=1$ 重合，即 $\pi_0=1$。 □

5.5.2 人口结构变化的 Markov 链模型 (考试不考)

考虑社会的教育水平与文化程度的发展变化，可以建立如下模型：将全国所有 16 岁以上的人口分为文盲、初中、高中/中专、大学/大专、中级技术人才、高级技术人才、特级专家等 7 类，结构的变化为升级、退化（如初中文化者会重新变为文盲）、进入（年龄达到 16 岁或移民进入）迁出（死亡或移民国外）。用

- $(n_1(t),n_2(t),\ldots,n_7(t))$ 表示在 t 年各等级的人数
- $N(t)=\sum_{i=1}^{7} n_i(t)$ 表示全社会 16 岁以上人口总数（简称为总人数）
- $R(t)$ 表示第 t 年的总迁入人数
- $W(t)$ 表示第 t 年的总迁出人数
- q_{ij} 表示每年从 i 级转为 j 级的人数在 i 级人数中的百分比
- w_i 表示每年从 i 级迁出的人数在 i 级人数中的百分比
- r_i 表示每年进入 i 级的人数在总人数中的百分比

准转移概率矩阵

$$\mathbf{Q}=(q_{ij})_{7\times 7}$$

不是转移概率矩阵，只是一个准转移概率矩阵 **(每行元素和 ≤ 1)**

这是因为考虑当前该等级与这些人下一步的分流人数关系，我们可以发现

$$\begin{aligned}
n_i(t)&=n_i(t)w_i+q_{ii}n_i(t)+\sum_{j\neq i} q_{ij}n_i(t)\\
\Rightarrow 1&=w_i+\sum_{j=1}^{7} q_{ij}\\
\Rightarrow \sum_{j=1}^{7} q_{ij}&=1-w_i\leq 1
\end{aligned}$$

回顾 Markov 链中概率转移的方式：$P(X_n=j)=\sum_i P(X_n=j|X_{n-1}=i)P(X_{n-1}=i)$。注意到在 $t+1,t$ 时刻，某人属于等级 j 的概率可以用比例 $\dfrac{n_j(t+1)}{N(t+1)},\dfrac{n_j(t)}{N(t)}$ 近似，我们只要能找到类似于以下的关系

$$\frac{n_j(t+1)}{N(t+1)}=\sum_{i=1}^{7} p_{ij}\frac{n_i(t)}{N(t)}$$

就可以反过来推出转移概率矩阵 $\mathbf{P}=(p_{ij})$。配凑过程如下：注意到

- 总人数关系方程 $N(t+1) = N(t) + R(t) - W(t)$
- 各层人数关系方程 $n_j(t+1) = \sum_{i=1}^{7} n_i(t)q_{ij} + r_j R(t)$(注：不需要减去 $w_j n_j(t)$，因为通过 $q_{jj}n_j(t)$，我们已经自动去除了离开等级 j 的人数)

于是可以得到比例关系方程

$$\frac{n_j(t+1)}{N(t+1)} = \frac{N(t)}{N(t+1)}\left(\sum_{i=1}^{7}\frac{n_i(t)q_{ij}}{N(t)} + \frac{R(t)}{N(t)}r_j\right)$$

假设总人口数不变，即 $N(t) = N(t+1)$，那么

$$R(t) = W(t) \Rightarrow \frac{R(t)}{N(t)} = \frac{W(t)}{N(t)} = \frac{\sum_{i=1}^{7} n_i(t)w_i}{N(t)}$$

$$\Rightarrow \frac{n_j(t+1)}{N_j(t+1)} = \sum_{i=1}^{7} q_{ij}\frac{n_i(t)}{N(t)} + \sum_{i=1}^{7} r_j w_i \frac{n_i(t)}{N(t)} = \sum_{i=1}^{7}(q_{ij} + r_j w_i)\frac{n_i(t)}{N(t)}$$

因此转移概率矩阵 $\mathbf{P} = \mathbf{Q} + \mathbf{w}^\top \mathbf{r}$，这里 $\mathbf{r}, \mathbf{w} \in \mathbb{R}^{1\times 7}$。由于 $\{n_i(t), 1 \le t \le 7\}$ 是一个不可约有限状态 Markov 链，其存在唯一的极限分布，因此可以通过控制迁入比例 $\mathbf{r}$/迁出比例 $\mathbf{w}$ 使得长时间后系统能达到某个预设的人口分布。比如：假定给定一个想要的概率分布 π，

$$\boldsymbol{\pi} = \boldsymbol{\pi}(\mathbf{Q} + \mathbf{w}^\top \mathbf{r}) \iff \mathbf{r} = (\boldsymbol{\pi}\mathbf{w}^\top)^{-1}\boldsymbol{\pi}(\mathbf{I} - \mathbf{Q})^{-1}$$

则可使系统的极限分布为 $\boldsymbol{\pi}$。

5.6 连续时间 Markov 链

5.6.1 连续时间 Markov 链的定义与性质

连续时间 Markov 链

对于随机过程 $\{X(t), t \ge 0\}$, S 是离散状态空间, 若 $\forall 0 \le t_0 < t_1 < \cdots < t_{n+1}$ 以及 $i_k \in S, 0 \le k \le n+1$, 有

$$P\{X(t_{n+1}) = i_{n+1}|X(t_n) = i_n, \cdots, X(t_0) = i_0\} = P\{X(t_{n+1}) = i_{n+1}|X(t_n) = i_n\}$$

则称 $\{X(t), t \ge 0\}$ 为连续时间 Markov 链。

- **转移概率：** $\forall s, t \in S$，称条件概率 $P(X(s+t) = j|X(s) = i) = p_{ij}(s,t)$ 为在时刻 s 处于状态 i，经过时间 t 转移到状态 j 的转移概率。
- **时齐连续时间 Markov 链：** 若 $p_{ij}(s,t)$ 与 s 无关，称 $\{X(t), t \ge 0\}$ 为时齐连续时间 Markov 链。简记转移概率 $p_{ij}(s,t) = p_{ij}(t-s)$。

正则 Markov 链

我们称连续时间 Markov 链是正则的，如果它以概率 1 在任意有限时间内只能转移有限次。

接下来的部分全部讨论时齐、正则的连续时间 Markov 链。 转移矩阵 $\mathbf{P}(t) = (p_{ij}(t))_{n\times n}$ 的性质与离散时间 Markov 链类似，有

1. $p_{ij}(t) \geq 0,\ i,j \in S$
2. $\sum\limits_{j\in S} p_{ij}(t) = 1,\ i \in S$
3. Chapman-Kolmogorov 方程: $p_{ij}(t+s) = \sum\limits_{k\in S} p_{ik}(t)p_{kj}(s)$，即 $\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s)$
4. $p_{ij}(0) = \delta_{ij}$，其中 $\delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$，即 $\mathbf{P}(0) = \mathbf{I}$(在转移时间 $t=0$ 内不可能从状态 i 转移到状态 j)

连续性条件

对于时齐连续时间 Markov 链，若

$$\lim_{t\to 0} p_{ij}(t) = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

则称 $\{X(t), t \geq 0\}$ 具有连续性条件。

- 连续性条件即为 $\lim\limits_{t\to 0} \mathbf{P}(t) = \mathbf{I}$

下面指出几种常见的时齐连续时间 Markov 链:

Poisson 过程

根据 Poisson 过程 $\{N(t), t \geq 0\}$ 的独立增量性，$\forall s, t \geq 0$，有

$$\begin{aligned} &P(N(t+s) = j | N(s) = i, N(u) = n_u, 0 \leq u \leq s) \\ =&P(N(t+s) = j | N(s) = i) = \frac{(\lambda t)^{(j-i)}}{(j-i)!} e^{-\lambda t} \end{aligned}$$

因此 Poisson 过程是一个时齐连续时间 Markov 链。其转移概率矩阵满足

$$p_{ij}(t) = \begin{cases} P(N(t+s) = j | N(s) = i) & j \geq i \\ 0 & j < i \end{cases} = \begin{cases} \dfrac{(\lambda t)^{(j-i)}}{(j-i)!} e^{-\lambda t} & j \geq i \\ 0 & j < i \end{cases}$$

Yule 过程 (Pure-birth 过程)

设群体中各个生物体的繁殖是相互独立、强度为 λ 的 Poisson 过程，并且群体中没有死亡。记 $\{X(t), t \geq 0\}$ 为群体中在 t 时刻的生物体总数，则 $\{X(t), t \geq 0\}$ 称为 Yule 过程。

设 $X(0) = 1$，状态空间为 $S = \{1, 2, \cdots,\}$。当群体数目为 i 时，注意到所有生物体的繁殖过程是独立的 Poisson 过程，**因此根据 Poisson 过程的可叠加性**，所有生物体的繁殖过程是强度为 λi 的 Poisson 过程，那么

$$\begin{aligned} &P(X(t+s) = j | X(s) = i, X(u) = n_u, 0 \leq u \leq s) \\ =&P(X(t+s) = j | X(s) = i)) = \frac{(\lambda i t)^{j-i}}{(j-i)!} e^{-\lambda i t} \end{aligned}$$

所以 Yule 过程是一个时齐连续时间 Markov 链。接下来我们求解其转移概率：记 $T_i(i \geq 1)$ 表示群体从 i 增加到 $i+1$ 所需的时间，那么 $T_i \sim \mathsf{Exp}(\lambda i)$，则

$$
\begin{aligned}
P(T_1 \leq t) &= 1 - e^{-\lambda t} \\
P(T_1 + T_2 \leq t) &= \int_0^t P(T_1 + T_2 \leq t | T_1 = x) P(T_1 = x) \mathrm{d}x \\
&= \int_0^t P(T_2 \leq t - x | T_1 = x) P(T_1 = x) \mathrm{d}x \\
&= \int_0^t (1 - e^{-2\lambda(t-x)}) \lambda e^{-\lambda x} \mathrm{d}x = (1 - e^{-\lambda t})^2 \\
&\cdots \\
P(T_1 + \cdots + T_j \leq t) &= (1 - e^{-\lambda t})^j
\end{aligned}
$$

注意到 $\{T_1 + \cdots + T_j \leq t\} \iff \{X(t) \geq j+1 | X(0) = 1\}$，因此

$$
\begin{aligned}
P(X(t) = j | X(0) = 1) &= P(X(t) \geq j | X(0) = 1) - P(X(t) \geq j+1 | X(0) = 1) \\
&= (1 - e^{-\lambda t})^{j-1} - (1 - e^{-\lambda t})^j \\
&= (1 - e^{-\lambda t})^{j-1} e^{-\lambda t}
\end{aligned}
$$

也就是说在从 t 时间内由 1 个生物体转移至 j 个生物体的转移概率 $p_{1j}(t) \sim \mathsf{Geom}(e^{-\lambda t})$。要求 t 时间内由 i 个生物体转移至 j 个生物体的转移概率 $p_{ij}(t)$，等价于求 i 个 i.i.d 的随机变量 p_{1j} 的和，回顾

$$
\mathsf{NB}(k; r, p) = \sum_{i=1}^r \mathsf{Geom}(k; p) = \binom{k-1}{r-1} p^r q^{k-r}
$$

于是

$$
p_{ij} = \binom{j-1}{i-1} (e^{-\lambda t})^i (1 - e^{-\lambda t})^{j-i}
$$

5.6.2 转移速率矩阵

在离散 Markov 链中，成立 $\mathbf{P}^{(n)} = \mathbf{P}^n = e^{n \ln \mathbf{P}}$，那么对连续时间 Markov 链，是否存在类似的关系 $\mathbf{P}(t) = e^{t\mathbf{Q}}$，其中 $\mathbf{Q}$ 为与 t 无关的实数矩阵呢？如果上式存在，那么

$$
\mathbf{P}'(0) = \lim_{t \to 0} \frac{\mathbf{P}(t) - \mathbf{P}(0)}{t} = \lim_{t \to 0} \frac{e^{t\mathbf{Q}} - \mathbf{I}}{t} = \mathbf{Q}
$$

也就是说这样的参数矩阵 $\mathbf{Q}$ 实则是 $\mathbf{P}(t)$ 在 $t = 0$ 处的 (右) 导数

- q_{ij} 反映目标在下一个很小的单位时间内从状态 i 转移到状态 j 的速率。
- q_{ii} 反映目标在下一个很小的单位时间内停留在状态 i 的速率。

基于这样的动机，我们定义转移速率矩阵

转移速率矩阵

对于时齐连续时间 Markov 链，记 **Q** 为转移速率矩阵为满足如下条件的矩阵

$$\lim_{t\to 0}\frac{p_{ij}(t)-p_{ij}(0)}{t}=\begin{cases}\lim\limits_{t\to 0}\dfrac{p_{ii}(t)-1}{t}\triangleq -q_{ii} & i=j\\ \lim\limits_{t\to 0}\dfrac{p_{ij}(t)}{t}\triangleq q_{ij} & i\neq j\end{cases}$$

[注]：**Q** 矩阵中的元素还可以等价表达为

$$p_{ii}(t)=1-q_{ii}t+o(t)$$
$$p_{ij}(t)=q_{ij}t+o(t)$$

关于转移概率矩阵 **Q**，以下不加证明的指出 **Q** 中元素 q_{ij} 存在与否的定理:

转移速率矩阵的存在性

对于时齐连续时间 Markov 链，

- $q_{ii}=\lim\limits_{t\to 0}\dfrac{1-p_{ii}(t)}{t}\leq\infty$
- $q_{ij}=\lim\limits_{t\to 0}\dfrac{p_{ij}(t)}{t}<\infty$

下面介绍转移速率矩阵的性质：

1. q_{ii} **与** $\sum\limits_{j\neq i}q_{ij}$ **的关系：**

   转移速率矩阵

   对无限状态时齐的连续时间 Markov 链

   $$0\leq\sum_{j\neq i}q_{ij}\leq q_{ii}$$

   对有限状态时齐的连续 Markov 链，则

   $$0\leq\sum_{j\neq i}q_{ij}=q_{ii}<\infty$$

2. **保守性：**对于有限状态时齐的连续时间 Markov 链，**Q** 是保守的：(从保守的定义和有限状态时齐连续时间 Markov 链的性质可以直接看出)

   保守

   若 $q_{ii}=\sum\limits_{j\neq i}q_{ij}<\infty$，则称 **Q** 是保守的。

3. **连续时间 Markov 链的自反概率**

自反概率

对于正则的时齐连续时间 Markov 链，其转出去 $(p_{ij}(t_1))$ 再转回来 $(p_{ji}(t_2))$ 的概率是 $t = t_1 + t_2$ 的高阶无穷小

证明. $p_{ij}(t_1) \cdot p_{ji}(t_2) = (q_{ij}t_1 + o(t_1))(q_{ji}t_2 + o(t_2)) \leq q_{ij}q_{ji}t^2 + o(t) = o(t)$ □

4. **连续时间 Markov 链自反概率的拆分**

自反概率的拆分

规定在 t 时间内始终停留在状态 i 的概率为

$$p_{\bar{i}\bar{i}}(t) = P(X(u) = i, 0 \leq u \leq t | X(0) = i)$$

则 t 时间内从状态 i 回到状态 i 的概率为

$$p_{ii}(t) = p_{\bar{i}\bar{i}}(t) + \sum_{j \neq i} p_{ij}(t_1)p_{ji}(t_2) = p_{\bar{i}\bar{i}}(t) + o(t)$$

其中 $t = t_1 + t_2$

证明. t 时间内从状态 i 回到状态 i 的概率 $p_{ii}(t)$ 等于一直停留在状态 i 的概率 $p_{\bar{i}\bar{i}}(t)$ 加上在 t 时间内从状态 i 转移到其他状态 j 再转回来的概率 $\sum\limits_{j \neq i} p_{ij}(t_1)p_{ji}(t_2)$。而后者已证明是 t 的高阶无穷小。 □

5. **连续时间 Markov 链的更新间隔**

连续时间 Markov 链的更新间隔

假定时齐连续时间 Markov 链 $\{X(t), t \geq 0\}$ 在 0 时刻刚刚到达状态 i，记 τ_i 为在 i 停留的时间，则 $\tau_i \sim \text{Exp}(q_{ii})$

证明. 首先先证明 τ_i 服从指数分布。因为

$$\begin{aligned}
&P(\tau_i > t + s | \tau_i > s) \\
=&P(X(u) = i, 0 \leq u \leq s, X(v) = i, s \leq v \leq s + t | X(u) = i, 0 \leq u \leq s) \\
=&P(X(v) = i, s \leq v \leq s + t | X(s) = i) \\
=&P(X(u) = i, 0 \leq u \leq t | X(0) = i) = P(\tau_i > t)
\end{aligned}$$

因此 τ_i 具有无记忆性，这等价于 $\tau_i \sim \text{Exp}(\mu_i)$。其中 μ_i 为指数分布的参数。接下来证明 $\mu_i = q_{ii}$。因为

$$P(\tau_i > t) = e^{-\mu_i t} = 1 - \mu_i t + o(t)$$

又因为

$$P(\tau_i > t) = P(X(u) = i, u \in [0, t] | X(0) = i) = p_{\bar{i}\bar{i}}(t) = 1 - q_{ii}t + o(t)$$

对比系数知 $\mu_i = q_{ii}$ □

由于 $\tau_i \sim \text{Exp}(q_{ii})$，那么在状态 i 停留的平均时长 $\mathbb{E}[\tau_i] = \dfrac{1}{q_{ii}}$，则

- $q_{ii} = \infty \Rightarrow$ 状态 i 为非常返态 (瞬过态)（过程一进入此状态就立即离开）
- $q_{ii} = 0$ 状态 i 为吸收态（过程一旦进入此状态就永远停留在此状态）

5.6.3 Kolmogorov 微分方程

连续时间 Markov 链的转移速率矩阵 $\mathbf{Q}$ 由转移概率矩阵 $\mathbf{P}(0)$ 唯一确定：$\mathbf{Q} = \mathbf{P}'(0)$；反之，**Kolmogorov 微分方程提供了由 Q 求 P(t) 的方法。**

Kolmogorov 微分方程 (有限状态 Markov 链)

设 $\mathbf{Q}$ 为有限状态、时齐、连续时间 Markov 链的转移速率矩阵，$\mathbf{P}(t)$ 为其转移概率矩阵，

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q} \qquad \text{(Kolmogorov 向前方程)}$$

$$\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t) \qquad \text{(Kolmogorov 向后方程)}$$

改写为分量形式

$$p'_{ij}(t) = \sum_{k \neq j} p_{ik}(t) q_{kj} - p_{ij}(t) q_{jj} \qquad \text{(Kolmogorov 向前方程)}$$

$$p'_{ij}(t) = \sum_{k \neq i} q_{ik} p_{kj}(t) - q_{ii} p_{ij}(t) \qquad \text{(Kolmogorov 向后方程)}$$

证明. 由 Chapman-Kolmogorov 方程，

$$p_{ij}(t+s) = \sum_{k \in S} p_{ik}(t) p_{kj}(s)$$

两边关于 t 在 $t = 0$ 处求导，得 **Kolmogorov 向后方程**

$$p'_{ij}(s) = \sum_{k \in S} p'_{ik}(0) p_{kj}(s) = \sum_{k \neq i} q_{ik} p_{kj}(s) - q_{ii} p_{ij}(s) \Rightarrow \mathbf{P}'(s) = \mathbf{Q}\mathbf{P}(s)$$

两边关于 s 在 $s = 0$ 处求导，得 **Kolmogorov 向前方程**

$$p'_{ij}(t) = \sum_{k \in S} p_{ik}(t) p'_{kj}(0) = \sum_{k \neq j} p_{ik}(t) q_{kj} - p_{ij}(t) q_{jj} \Rightarrow \mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$$

□

[注]：在 K-C 方程中，质点经过路径为 $i \xrightarrow{t} k \xrightarrow{s} j$。关于在“前面”的时间 t 求导得到“向前方程”，$\mathbf{P}(t)$ 在 $\mathbf{Q}$ 前面；关于在“后面”的时间 s 求导得到“向后方程”，$\mathbf{P}(s)$ 在 $\mathbf{Q}$ 后面。

Kolmogorov 微分方程 (无限状态 Markov 链)

设 $\mathbf{Q}$ 为无限状态 (状态空间 S 可列)、时齐、连续时间 Markov 链的转移速率矩阵，$\mathbf{P}(t)$ 为其转移概率矩阵，**若 Q 是保守的**，则

$$\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t) \qquad \text{(Kolmogorov 向后方程)}$$

在适当正则的条件下，成立

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q} \qquad \text{(Kolmogorov 向前方程)}$$

Kolmogorov 方程求解 Poisson 过程的转移概率

根据 Poisson 分布的第二个定义

$$\begin{aligned} p_{k,k+1}(h) &= P(N(t+h)-N(t)=1|N(t)=k) = \lambda h + o(h) \\ p_{k,k}(h) &= P(N(t+h)-N(t)=0|N(t)=k) = 1-\lambda h + o(h) \end{aligned}$$

从而转移速率矩阵中 **Q** 的元素为

$$\begin{aligned} q_{kk} &= q_{k,k+1} = \lambda \\ q_{k,k+j} &= 0 & j \geq 2 \\ q_{k,k-j} &= 0 & j \geq 1 \end{aligned}$$

由 Kolmogorov 向后方程

$$p'_{ij}(t) = \sum_{k\neq i} q_{ik}p_{kj}(t) - q_{ii}p_{ij}(t) = q_{i,i+1}p_{i+1,j}(t) - q_{ii}p_{ij}(t)$$

在初始条件 $p_{ii}(0)=1, p_{ij}(0)=0(j\neq i)$ 下，解得

(1) 当 $j=i$ 时，$p'_{ii}(t) = -\lambda p_{ii}(t)$，解得 $p_{ii}(t)=e^{-\lambda t}$

(2) 当 $j=i+1$ 时，$p'_{i,i+1}(t) = \lambda p_{i+1,i+1}(t) - \lambda p_{i,i+1}(t) = \lambda e^{-\lambda t} - \lambda p_{i,i+1}(t)$，解得 $p_{i,i+1}(t) = \lambda t e^{-\lambda t}$

(3) 当 $j=i+2$ 时，$p'_{i,i+2}(t) = \lambda p_{i+1,i+2}(t) - \lambda p_{i,i+2}(t) = \lambda^2 t e^{-\lambda t} - \lambda p_{i,i+2}(t)$，解得 $p_{i,i+2}(t) = \dfrac{(\lambda t)^2}{2}e^{-\lambda t}$

(4) 以此类推，当 $i=j+n$ 时，$p_{i,i+n}(t) = \dfrac{(\lambda t)^n}{n!}e^{-\lambda t}$

Kolmogorov 方程求解 Yule 过程的转移概率

由于 Yule 过程中祖先单一，故 $X(0)=1$。当 t 时刻有 k 个生物体时，所有生物体的繁殖过程可以看作强度为 λk 的 Poisson 过程，因此

$$\begin{aligned} p_{k,k+1}(h) &= P(X(t+h)-X(t)=1|X(t)=k) = \lambda kh + o(h) \\ p_{k,k}(h) &= P(X(t+h)-X(t)=0|X(t)=k) = 1-\lambda kh + o(h) \\ p_{k,k+j}(h) &= P(X(t+h)-X(t)\geq 2|X(t)=k) = o(h) & j\geq 2 \\ p_{k,k-j}(h) &= P(X(t+h)-X(t)<0|X(t)=k) = 0 & j \geq 1 \end{aligned}$$

所以转移速率矩阵 **Q** 中的元素为

$$\begin{aligned} q_{k,k+1} &= q_{k,k} = \lambda k \\ q_{k,k+j} &= 0 & j \geq 2 \\ q_{k,k-j} &= 0 & j \geq 1 \end{aligned}$$

根据 Kolmogorov 向前方程

$$
\begin{aligned}
p_{ij}'(t) &= \sum_{k\neq j} p_{ik}(t)q_{kj} - p_{ij}(t)q_{jj} = p_{i,j-1}(t)q_{j-1,j} - p_{ij}(t)q_{jj} \\
\Rightarrow p_{ii}'(t) &= -i\lambda p_{ii}(t) \\
\Rightarrow p_{ij}'(t) &= (j-1)\lambda p_{i,j-1}(t) - j\lambda p_{ij}(t)
\end{aligned}
$$

在初始条件 $p_{ii}(0)=1, p_{ij}(0)=0(j\neq i)$ 下，我们希望证明 $p_{ij}(t)=\binom{j-1}{i-1}(e^{-\lambda t})^i(1-e^{-\lambda t})^{j-i}$。
(1) $j=i$ 时，从 $p_{ii}'(t)=-i\lambda p_{ii}(t)$ 容易解得 $p_{ii}(t)=e^{-\lambda it}=\binom{i-1}{i-1}(e^{-\lambda t})^i(1-e^{-\lambda t})^{i-i}$。
(2) $j>i$ 时证明采取数学归纳法：设上述转移概率结论在 $j-1$ 时成立，即

$$
p_{i,j-1}(t) = \binom{j-2}{i-1}(e^{-\lambda t})^i(1-e^{-\lambda t})^{j-1-i}
$$

则 j 时，

$$
\begin{aligned}
p_{ij}'(t) &= (j-1)\lambda p_{i,j-1}(t) - j\lambda p_{ij}(t) \\
&= (j-1)\lambda\binom{j-2}{i-1}(e^{-\lambda t})^i(1-e^{-\lambda t})^{j-1-i} - j\lambda p_{ij}(t) \\
\Rightarrow p_{ij}'(t) + j\lambda p_{ij}(t) &= (j-1)\lambda\binom{j-2}{i-1}(e^{-\lambda t})^i(1-e^{-\lambda t})^{j-1-i} \\
\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{j\lambda t}p_{ij}(t)\right) &= e^{\lambda jt}(j-1)\lambda\binom{j-2}{i-1}e^{-\lambda it}(1-e^{-\lambda t})^{j-1-i} \\
&= \lambda(j-1)\binom{j-2}{i-1}e^{\lambda(j-i)t}(1-e^{-\lambda t})^{j-1-i} \\
&= \lambda(j-i)\binom{j-1}{i-1}e^{\lambda(j-i)t}(1-e^{-\lambda t})^{j-1-i}
\end{aligned}
$$

对上式两边关于 t 积分即得 $p_{ij}(t)=\binom{j-1}{i-1}(e^{-\lambda t})^i(1-e^{-\lambda t})^{j-i}$

Kolmogorov 方程求解生灭 (Birth-death) 过程的转移概率

设群体中所有生物体的繁殖是强度为 λ_i 的 Poisson 过程,同时所有生物体的死亡也是强度为 μ_i 的 Poisson 过程。记 $\{X(t),t\geq 0\}$ 为群体中在 t 时刻的生物体总数，则 $\{X(t),t\geq 0\}$ 称为生灭过程。等价的定义也就是

$$
\begin{aligned}
p_{i,i+1}(h) &= P(X(t+h)-X(t)=1|X(t)=i) = \lambda_i h + o(h) \\
p_{i,i-1}(h) &= P(X(t+h)-X(t)=-1|X(t)=i) = \mu_i h + o(h) \\
p_{i,i}(h) &= P(X(t+h)-X(t)=0|X(t)=i) = 1-(\lambda_i+\mu_i)h + o(h) \\
p_{ij}(h) &= o(h) \qquad\qquad |i-j|\geq 2
\end{aligned}
$$

所以转移速率矩阵 **Q** 中的元素为

$$
\begin{aligned}
q_{ii} &= \mu_i + \lambda_i \\
q_{i,i+1} &= \lambda_i, \quad q_{i,i-1} = \mu_i \\
q_{i,j} &= 0 \qquad\qquad |i-j| \geq 2
\end{aligned}
$$

从而可以写出生灭过程的 Kolmogorov 向前向后方程

$$p'_{ij}(t) = p_{i,j-1}(t)\lambda_{j-1} - p_{ij}(t)(\mu_i + \lambda_i) + p_{i,j+1}(t)\mu_{j+1} \qquad \text{(向前方程)}$$

$$p'_{ij}(t) = \mu_i p_{i-1,j}(t) - (\mu_i + \lambda_i)p_{ij}(t) + \lambda_i p_{i+1,j}(t) \qquad \text{(向后方程)}$$

5.6.4 稳定状态下的状态概率分析

状态概率平衡方程

若 $\lim\limits_{t\to\infty} p_{ij}(t)$ 存在，记 $P_j = \lim\limits_{t\to\infty} p_{ij}(t) = \lim\limits_{t\to\infty} P(X(t) = j|X(0) = i)$(平稳分布)，则成立

$$q_{jj}P_j = \sum_{k\neq j} P_k q_{kj}$$

证明. 由 Kolmogorov 向前方程

$$p'_{ij}(t) = \sum_{k\neq j} p_{ik}(t)q_{kj} - p_{ij}(t)q_{jj}$$

由于 $t\to\infty$ 时系统稳定，$p'_{ij}(t)\to 0$，于是在上式两边同时取 $t\to\infty$ 得

$$0 = \sum_{k\neq j} P_k q_{kj} - P_j q_{jj}$$

□

这个定理其实表现的是一种“流量守恒”原理：注意到

- LHS $q_{jj}P_j$: 由于 $q_{jj} = \sum\limits_{j\neq k} q_{jk}$ 表示从状态 j 流出的速率，因此 $q_{jj}P_j$ 可理解为状态 j 的流出率。
- RHS $\sum\limits_{k\neq j} P_k q_{kj}$: q_{kj} 表示从其他状态 $k\neq j$ 流入状态 j 的速率，因此 $\sum\limits_{k\neq j} P_k q_{kj}$ 可理解为状态 j 的流入率。

如果系统处于平稳状态，那么系统中的流入率等于流出率，达到“统计平衡”。

M/M/S 排队系统

顾客的来到是参数为 λ 的 Poisson 过程，服务员数量为 s 个，每个顾客接受服务的时间服从参数为 μ 的指数分布。遵从先来先服务，若服务员没有空闲就排队的原则，以 $X(t)$ 记 t 时刻系统中顾客数目，则 $\{X(t), t\geq 0\}$ 是一个生灭过程（来到看作出生，离去看作死亡）

- 来到过程是恒定参数为 λ 的 Poisson 过程。

- 离去过程是变参数为 μ_n 的 Poisson 过程。其中 μ_n 记系统中有 n 个顾客时的离去率

$$\mu_n = \begin{cases} n\mu & 1 \leq n \leq s \\ s\mu & n > s \end{cases}$$

以下我们考虑 M/M/1 排队系统，即只有一个服务员的情况。我们关心这个系统的极限分布。记系统中恰有 n 个顾客的**长程 (或极限) 概率**为 $P_n = \lim\limits_{t\to\infty} P(X(t)=n)$

- 比如说 $P_0 = 0.3$ 表示长时间下，系统中有 30% 的时间没有顾客，$P_1 = 0.2$ 表示系统中有 20% 的时间只有一位顾客

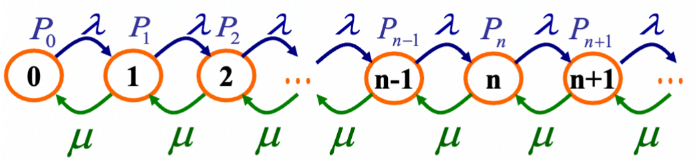


图 6: M/M/1 排队模型

由于 $\mu_n = n\mu$ 为定值，根据状态概率平衡方程，可以写出如下等式

$$\begin{cases} P_0\lambda = P_1\mu & n=0 \\ P_n(\lambda+\mu) = P_{n-1}\lambda + P_{n+1}\mu & 1 \leq n \leq s \end{cases}$$

第一个等式是边界条件，表示流入状态 0 和流出状态 0 的速率相等；第二个等式是递推条件，表示状态 n 的流入速率等于流出速率。容易解得 $\dfrac{P_n}{P_{n-1}} = \dfrac{\lambda}{\mu}$，那么 $P_n = P_0\left(\dfrac{\lambda}{\mu}\right)^n$。由于 $\sum\limits_{n=0}^{\infty} P_n = 1$，不妨记**服务强度** $\rho = \dfrac{\lambda}{\mu}$，那么当 $\rho \in [0,1)$，时，M/M/1 过程存在**唯一的平稳分布**

$$\begin{aligned} 1 &= \sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} P_0\rho^n = P_0\sum_{n=0}^{\infty}\rho^n = \frac{P_0}{1-\rho} \\ \Rightarrow & P_0 = 1-\rho \qquad \text{(系统空闲的概率)} \\ \Rightarrow & \{P_j, j\in S\} = \{\rho^j(1-\rho)\}_{j\in S} \end{aligned}$$

6 鞅

6.1 鞅的定义与性质

鞅 (Martinagle) 的背景来源于赌博：赌博者对他在一系列赌博后获得最大期望收益的策略感兴趣。假设一个赌博者正在进行一系列赌博，每次输赢的概率都是 $\frac{1}{2}$。记 $\{Y_n, n \geq 1\}$ 为每次赌博的结果，显然 $\{Y_n\}$ i.i.d. 且 $P(Y_n = 1) = P(Y_n = -1) = \frac{1}{2}$。如果**赌博者采用的赌博策略依赖于前面的赌博结果**，那么他在第 n 次时的赌博策略 (即下注金额) 可以用下面的随机变量序列表示

$$b_n = b_n(Y_1, Y_2, \cdots, Y_{n-1})$$

若赌赢则获利 b_n，否则输掉 b_n。设 X_0 为该赌博者的初始赌资，则他在第 n 次赌博后的赌资为 $X_n = X_0 + \sum_{i=1}^{n} Y_i b_i$。赌博者关心在 n 次赌博后，第 $n+1$ 次赌博的期望赌资，即

$$\begin{aligned}\mathbb{E}[X_{n+1}|X_n] &= \mathbb{E}[X_{n+1}|Y_1, Y_2, \cdots, Y_n] \\ &= \mathbb{E}[X_n|Y_1, Y_2, \cdots, Y_n] + \mathbb{E}[Y_{n+1}b_{n+1}|Y_1, Y_2, \cdots, Y_n]\end{aligned}$$

注意到 X_n, b_{n+1} 本身都由 $Y_1, \cdots, Y_n$ 决定，且 $\{Y_n\}$ i.i.d，因此上式可化为

$$\begin{aligned}\mathbb{E}[X_{n+1}|X_n] &= X_n + b_{n+1}\mathbb{E}[Y_{n+1}|Y_1, Y_2, \cdots, Y_n] \\ &= X_n + b_{n+1}\mathbb{E}[Y_{n+1}] \\ &= X_n\end{aligned}$$

也就是说，**对于公平赌博而言，赌博者在每次赌博后的期望赌资是不变的。任何赌博者都无法通过改变赌博策略将公平赌博变为有利的赌博 (即 $\mathbb{E}[X_{n+1}|X_n] > X_n$)**。由此，我们引出鞅的概念。

鞅 (Martinagle)

设有两个过程 $\{X_n, n \geq 0\}$ 和 $\{Y_n, n \geq 0\}$，若对于任意 $n \geq 0$，都有

1. $\mathbb{E}[|X_n|] < \infty$
2. $\mathbb{E}[X_{n+1}|Y_0, Y_1, \cdots, Y_n] = X_n$

则称 $\{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的鞅。

下面给出几种常见的鞅：

零均值独立随机变量和 (赌博)

设 $\{X_n, n \geq 1\}$ 为 i.i.d 序列，$X_0 = 0$，且 $\forall n \geq 0, \mathbb{E}[|X_n|] < \infty, \mathbb{E}[X_n] = 0$，则 $\{S_n = \sum_{i=1}^{n} X_i, n \geq 0\}$ 是关于 $\{X_n\}$ 的鞅。

证明. 1. 验证条件 1：$\mathbb{E}[|S_n|] \leq \sum_{i=1}^{n} \mathbb{E}[|X_i|] < \infty$

2. 验证条件 2

$$\begin{aligned}\mathbb{E}[S_{n+1}|X_0,X_1,\cdots,X_n] &= \mathbb{E}[S_n+X_n|X_0,X_1,\cdots,X_n]\\ &= \mathbb{E}[S_n|X_0,X_1,\cdots,X_n]+\mathbb{E}[X_n|X_0,X_1,\cdots,X_n]\\ &= S_n+0=S_n\end{aligned}$$

□

加倍赌博

考虑如下赌注翻倍的公平赌博：记 Y_n 为第 n 次赌博的结果，$P(Y_n=1)=P(Y_n=-1)=\dfrac{1}{2}$，$W_n$ 为第 n 次赌博后所输/所赢的总金额，$W_0=0$。每次抛硬币之前的赌注都比上一次翻一倍，直到赢了赌博即停。则 $\{W_n\}$ 是关于 $\{Y_n\}$ 的鞅。

证明. 值得注意的是，上述赌博具有突然停止的特点，$|W_n|$ 从赢了之后起就不再变化 (最多只能赢 1 块钱)，即 $P(W_{n+1}=1||W_n|=1)=1$。下验证该过程满足鞅的条件：

1. 验证条件 1

$$\begin{aligned}\mathbb{E}[|W_n|] &= \mathbb{E}[|W_n||Y_1=1]P(Y_1=1)+\mathbb{E}[|W_n||Y_1=-1]P(Y_1=-1)\\ &= \frac{1}{2}+\frac{1}{2}\mathbb{E}[|W_n||Y_1=-1]\\ &= \frac{1}{2}+\frac{1}{2}\left(\mathbb{E}[|W_n||Y_1=-1,Y_2=1]P(Y_2=1)+\mathbb{E}[|W_n||Y_1=-1,Y_2=-1]P(Y_2=-1)\right)\\ &= \frac{1}{2}+\frac{1}{2}\left(\frac{1}{2}+\frac{1}{2}\mathbb{E}[|W_n||Y_1=-1,Y_2=-1]\right)\\ &= \frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^2}\mathbb{E}[|W_n||Y_1=-1,Y_2=-1]\\ &= \cdots=\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^n}+\frac{1}{2^n}\mathbb{E}[|W_n||Y_1=-1,Y_2=-1,\cdots,Y_n=-1]\\ &= \left(1-\frac{1}{2^n}\right)+\frac{1}{2^n}(|-1-2-\cdots-2^{n-1}|)\\ &= \left(1-\frac{1}{2^n}\right)+\frac{1}{2^n}(2^n-1)=2-\frac{1}{2^{n-1}}<\infty\end{aligned}$$

2. 验证条件 2: 如果第 n 次赌博赢了，那么 $\mathbb{E}[W_{n+1}|W_n=1]=1=W_n$，如果第 n 次赌博没赢

$$\begin{aligned}\mathbb{E}[W_{n+1}|W_n<0] &= \mathbb{E}[W_{n+1}|W_n<0,Y_{n+1}=1]P(Y_{n+1}=1)+\mathbb{E}[W_{n+1}|W_n<0,Y_{n+1}=-1]P(Y_{n+1}=-1)\\ &= \frac{1}{2}\cdot 1+\frac{1}{2}\cdot(-2^{n+1}+1)=-2^n+1=W_n\end{aligned}$$

无论如何都有 $\mathbb{E}[W_{n+1}|W_n]=\mathbb{E}[W_{n+1}|Y_1,\cdots,Y_n]=W_n$

□

Polya 坛模型

坛子中有 b 只黑球，r 只红球，随机取出一只，把原球放回，并加进与抽出球同色的球 c 只。再摸第二次，这样下去共摸了 n 次。记 X_n 为第 n 次抽取后坛子中的红球数，$X_0=r$，M_n 为第 n 次抽取后红球

所占的比例，则 $\{M_n\}$ 是关于 $\{X_n\}$ 的鞅。

证明. 容易发现 $\{X_n\}$ 为一个非时齐的 Markov 链，转移概率为

$$P(X_{n+1}=k+c|X_n=k)=\frac{k}{nc+b+r}$$
$$P(X_{n+1}=k|X_n=k)=\frac{nc+b+r-k}{nc+b+r}$$

从而第 n 次抽取后红球所占比例 $M_n=\dfrac{X_n}{nc+b+r}$。注意到

$$\begin{aligned}\mathbb{E}[X_{n+1}|X_n=k]&=(k+c)\frac{k}{nc+b+r}+k\frac{nc+b+r-k}{nc+b+r}\\&=k\cdot\frac{c+nc+b+r}{nc+b+r}\\\Rightarrow\mathbb{E}[X_{n+1}|X_n]&=X_n\cdot\frac{c+nc+b+r}{nc+b+r}\end{aligned}$$

1. 验证条件 1：

$$\mathbb{E}[|M_n|]=\frac{1}{nc+b+r}\mathbb{E}[|X_n|]<1<\infty$$

2. 验证条件 2

$$\begin{aligned}\mathbb{E}[M_{n+1}|X_1,\cdots,X_n]=\mathbb{E}[M_{n+1}|X_n]&=\mathbb{E}\left[\frac{X_{n+1}}{(n+1)c+b+r}|X_n\right]\\&=\frac{1}{(n+1)c+b+r}\mathbb{E}[X_{n+1}|X_n]\\&=\frac{1}{(n+1)c+b+r}X_n\cdot\frac{c+nc+b+r}{nc+b+r}\\&=M_n\end{aligned}$$

故 $\{M_n\}$ 是关于 $\{X_n\}$ 的鞅。 □

6.2 上鞅与下鞅

鞅描述的是公平的赌博,而上鞅和下鞅分别描述了不利赌博 (即 $\mathbb{E}[X_{n+1}|X_n]\le X_n$) 和有利赌博 (即 $\mathbb{E}[X_{n+1}|X_n]\ge X_n$)。

上鞅

设有两个过程 $\{X_n,n\ge 0\}$ 和 $\{Y_n,n\ge 0\}$，若对于任意 $n\ge 0$，都有

1. $\mathbb{E}[X_n^-]>-\infty$，其中 $X_n^-=\min\{0,X_n\}$
2. $\mathbb{E}[X_{n+1}|Y_0,Y_1,\cdots,Y_n]\le X_n$
3. X_n 是 $Y_0,Y_1,\cdots,Y_n$ 的函数

则称 $\{X_n,n\ge 0\}$ 是关于 $\{Y_n,n\ge 0\}$ 的上鞅。

下鞅

设有两个过程 $\{X_n, n \geq 0\}$ 和 $\{Y_n, n \geq 0\}$，若对于任意 $n \geq 0$，都有

1. $\mathbb{E}[X_n^+] < +\infty$，其中 $X_n^+ = \max\{0, X_n\}$
2. $\mathbb{E}[X_{n+1}|Y_0, Y_1, \cdots, Y_n] \geq X_n$
3. X_n 是 $Y_0, Y_1, \cdots, Y_n$ 的函数

则称 $\{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的下鞅。

鞅/上鞅/下鞅的性质

1. $\{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的上鞅 $\iff \{-X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的下鞅。
2. $\{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的鞅 $\iff \{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的上鞅和下鞅。
3. 设 $\{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的 (下) 鞅，则
   (a) $\forall k \geq 0$，$\mathbb{E}[X_{n+k}|Y_0, Y_1, \cdots, Y_n](\geq) = X_n$
   (b) (期望单调递增) $\forall 0 \leq k \leq n$，$\mathbb{E}[X_n](\geq) = \mathbb{E}[X_k](\geq) = \mathbb{E}[X_0]$

证明. 下证明性质 3 的 (a)(b)。对于 (a)：当 $k = 0, 1$ 时由定义易知。设定理对于 $k \geq 1$ 时成立，则对 $k+1$ 时

$$\begin{aligned}
\mathbb{E}[X_{n+k+1}|Y_0, \cdots, Y_n] &= \mathbb{E}[\mathbb{E}[X_{n+k+1}|Y_0, \cdots, Y_{n+k}]|Y_0, \cdots, Y_n] \\
&(\geq) = \mathbb{E}[X_{n+k}|Y_0, \cdots, Y_n] && \text{(下鞅性质)} \\
&(\geq) = X_n && \text{(归纳假设)}
\end{aligned}$$

由数学归纳法知 (a) 对一切 $k \geq 0$ 成立。

对于 (b)：$\forall 0 \leq k \leq n$，由 (a) 知

$$\mathbb{E}[X_n|Y_0, \cdots, Y_k](\geq) = X_k$$

对其取重期望公式得到

$$\mathbb{E}[X_n] = \mathbb{E}[\mathbb{E}[X_n|Y_0, \cdots, Y_k]](\geq) = \mathbb{E}[X_k]$$

类似地，$\mathbb{E}[X_k] \geq \mathbb{E}[X_0]$，故 (b) 对一切 $0 \leq k \leq n$ 成立。 □

由上述性质容易看出看出**下鞅列的期望是单调递增**的，**上鞅列的期望是单调递减**的。在动态多期随机决策应用中，这种基于期望的单调性，通常能帮助我们找到决策在平均意义上的最优值或上下界。

6.2.1 鞅的凸变换

条件 Jensen 不等式

设 $\phi(x)$ 为 $\mathbb{R}$ 上的凸函数，随机变量 X 满足:

1. $\mathbb{E}[|X|] < \infty$
2. $\mathbb{E}[|\phi(X)|] < \infty$

$Y_0, \cdots, Y_n$ 为独立随机变量序列，则

$$\mathbb{E}[\phi(X)|Y_0, \cdots, Y_n] \geq \phi(\mathbb{E}[X|Y_0, \cdots, Y_n])$$

[注]：本质上就是 Jensen 不等式 $\mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[X])$ 的条件形式。

鞅的凸变换

若 $\{X_n, n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的鞅，$\phi(x)$ 为凸函数，且 $\forall n, \mathbb{E}[\phi(X_n)^+] < \infty$，则 $\{\phi(X_n), n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的下鞅。

[注]：鞅经凸变换变为下鞅，下鞅经凸变换仍为下鞅。
[注]：$(\cdot)^2, e^{\cdot}$ 这些常见的变换都是凸变换。

证明. 由条件 Jensen 不等式，对于任意 $n \geq 0$，有

$$\mathbb{E}[\phi(X_{n+1})|Y_0, \cdots, Y_n] \geq \phi(\mathbb{E}[X_{n+1}|Y_0, \cdots, Y_n]) = \phi(X_n)$$

故 $\{\phi(X_n), n \geq 0\}$ 是关于 $\{Y_n, n \geq 0\}$ 的下鞅。 □

6.3 鞅的停时定理

停时 (stopping time)

称**随机函数** T 是关于随机过程 $\{X_n, n \geq 0\}$ 的停时，若 T 在 $\{0, 1, \cdots, \infty\}$ 中取值，且

$$\forall n \geq 0, \{T = n\} \in \sigma(X_0, X_1, \cdots, X_n)$$

即事件 $\{T = n\}$ 是 $\{X_0, X_1, \cdots, X_n\}$ 的函数。

- 我们称 T 为随机函数而非随机变量，是因为 T 并不需要满足概率和为 1 的性质。
- 停时也被称为 Markov 时间，即事件 $\{T = n\}$ 只依赖于 n 时刻之前的信息 $\{X_0, X_1, \cdots, X_n\}$，而不依赖于 n 之后的过程。

以下为停时的一些例子

1. 固定时间 $T \equiv k$ 是停时。
2. 首达时间为停时。即过程 $\{X_n, n \geq 0\}$ 首次达到某一状态空间子集 A 的时间 $T(A) = \min\{n \geq 0 : X_n \in A\}$ 是停时。因为首达时为 n 可以用如下方式表示：$\{T(A) = n\} = \{X_1 \notin A, \cdots, X_{n-1} \notin A, X_n \in A\}$

3. 对任何固定的 k，过程第 k 次到达某一状态空间子集 A 的时间 $T_k(A)$ 是停时。

停时的初等性质

若 T_1, T_2 是关于 $\{X_n, n \geq 0\}$ 的停时，则 $T_1 + T_2, \min\{T_1, T_2\}, \max\{T_1, T_2\}, \min\{T_1, k\}$($k$ 为固定自然数) 也是停时。

证明. 先证明下面的命题：以下三者等价

$$\begin{aligned}
\{T = n\} \in \sigma(X_0, X_1, \cdots, X_n) &\iff \{T > n\} \in \sigma(X_0, X_1, \cdots, X_n) \\
&\iff \{T \leq n\} \in \sigma(X_0, X_1, \cdots, X_n)
\end{aligned}$$

这是因为 $\{T \leq n\} = \bigcup\limits_{k=0}^{n} \{T = k\}$，$\{T > n\} = \Omega - \{T \leq n\}$，$\{T = n\} = \{T \leq n\} - \{T \leq n-1\}$，这三个集合是互相控制的，所以一个属于 $\sigma(X_0, X_1, \cdots, X_n)$，另外两个也属于 $\sigma(X_0, X_1, \cdots, X_n)$。

基于此，我们证明停时的初等性质。因为

$$\begin{aligned}
\{T_1 + T_2 = n\} &= \bigcup_{k=0}^{n} \{T_1 = k, T_2 = n - k\} \\
&= \bigcup_{k=0}^{n} \{T_1 = k\} \cap \{T_2 = n - k\} \in \sigma(X_0, X_1, \cdots, X_n) \\
\{\max\{T_1, T_2\} = n\} &= \{T_1 = n, T_2 \leq n\} \cup \{T_2 = n, T_1 \leq n\} \in \sigma(X_0, X_1, \cdots, X_n) \\
\{\min\{T_1, T_2\} = n\} &= \{T_1 = n, T_2 \geq n\} \cup \{T_2 = n, T_1 \geq n\} \in \sigma(X_0, X_1, \cdots, X_n) \\
\{\min\{T_1, k\} = n\} &= \{T_1 = n\} \cup \{T_1 \geq n\} \in \sigma(X_0, X_1, \cdots, X_n)
\end{aligned}$$

故 $T_1 + T_2, \min\{T_1, T_2\}, \max\{T_1, T_2\}$ 也是停时。 □

在引入鞅的停时定理前，我们先引入鞅的**有界停时定理**。设 $\{M_n \geq 0\}$ 为关于 $\{X_n, n \geq 0\}$ 的鞅，由先前停时的性质知，若 T 为 $\{X_n, n \geq 0\}$ 的停时，**则** $T \wedge n = \min\{T, n\}$ 也为 $\{X_n, n \geq 0\}$ 的停时，且此时成立 $T \leq n$，**我们就称这样的 T 为有上界 n 的停时。**

鞅的有界停时定理 (证明可以不会，但须知定理结论 & 应用)

设 $\{M_n \geq 0\}$ 为关于 $\{X_n, n \geq 0\}$ 的 (下) 鞅。T 为关于 $\{X_n, n \geq 0\}$ 的停时，则

$$\begin{aligned}
\mathbb{E}[M_{T \wedge n} | X_0] (\leq) &= M_0 \\
\mathbb{E}[M_{T \wedge n}] (\leq) &= \mathbb{E}[M_0]
\end{aligned}$$

[注]：换句话说，鞅在 (有上界的) **停时下的期望等于初始时刻的期望**。

证明. (这里只证明第二个结论) 先引入一个引理

设 $\{M_n, n \geq 0\}$ 为关于 $\{X_n, n \geq 0\}$ 的 (下) 鞅，T 为关于 $\{X_n, n \geq 0\}$ 的停时，则 $\forall n \geq k$，有

$$\mathbb{E}[M_n I_{\{T=k\}}] (\geq) = \mathbb{E}[M_k I_{\{T=k\}}]$$

证明. 由全期望公式以及鞅/上鞅/下鞅的性质 3(a)

$$\begin{aligned}\mathbb{E}[M_n I_{\{T=k\}}] &= \mathbb{E}[\mathbb{E}[M_n I_{\{T=k\}}|X_0, X_1, \cdots, X_k]] \\ &= \mathbb{E}[I_{\{T=k\}}\mathbb{E}[M_k|X_0, X_1, \cdots, X_k]] \\ (\geq) &= \mathbb{E}[M_k I_{\{T=k\}}]\end{aligned}$$

□

接下来证明鞅的有界停时定理 (的第二个结论)。因为停时 T 有上界 n，当 $T \leq n$ 时 $T \wedge n = T$，当 $T > n$ 时 $T \wedge n = n$，故

$$\begin{aligned}\mathbb{E}[M_{T\wedge n}] &= \mathbb{E}\left[\sum_{k=0}^{n-1} M_T I_{\{T=k\}}\right] + \mathbb{E}\left[\sum_{k=n}^{\infty} M_n I_{\{T=k\}}\right] = \mathbb{E}\left[\sum_{k=0}^{n-1} M_T I_{\{T=k\}}\right] + \mathbb{E}\left[M_n I_{\{T\geq n\}}\right] \\ &= \mathbb{E}\left[\sum_{k=0}^{n-1} M_k I_{\{T=k\}}\right] + \mathbb{E}\left[M_n I_{\{T\geq n\}}\right] \\ (\text{引理}) &= \mathbb{E}\left[\sum_{k=0}^{n-1} M_n I_{\{T=k\}}\right] + \mathbb{E}\left[M_n I_{\{T\geq n\}}\right] \\ &= \mathbb{E}[M_n(I_{\{T\geq n\}} + \sum_{k=0}^{n-1} I_{\{T=k\}})] = \mathbb{E}[M_n] \\ (\text{鞅的性质 3(b)}) &= \mathbb{E}[M_0]\end{aligned}$$

□

然而，许多问题并不满足 $T \leq n$ 这一严格条件，例如在赌博问题中，赌博者不能确定在某一确定的时刻 n 之前停止赌博，但可以保证这场赌博不会无限期的延续下去。因此我们将限制放松到 $P(T < \infty) = 1$(以概率 1 可以保证停止)，引入鞅的停时定理。

鞅的停时定理 (证明可以不会，但须知定理结论 & 应用)

设 $\{M_n \geq 0\}$ 为关于 $\{X_n, n \geq 0\}$ 的鞅。T 为关于 $\{X_n, n \geq 0\}$ 的停时，且满足

1. $P(T < \infty) = 1$
2. $\mathbb{E}[|M_T|] < \infty$
3. $\lim\limits_{n\to\infty} \mathbb{E}[|M_n| I_{\{T>n\}}] = 0$

则有 $\mathbb{E}[M_T] = \mathbb{E}[M_0]$.

[注]：这是一个更强的结论，鞅在**停时下的期望等于初始时刻的期望**。

回顾一下之前介绍的“加倍下注赌博”的例子，我们已经证明了第 n 次赌博后的财富 $\{W_n, n \geq 0\}$ 是关于赌博结果 $\{Y_n, n \geq 0\}$ 的鞅。赢得赌博的时刻 T 为 $\{X_n, n \geq 0\}$ 的停时 (因为 $T \sim \text{Geom}(1/2)$)，那么是否成立 $\mathbb{E}[M_T] = \mathbb{E}[M_0]$?

1. 条件 1 和条件 2 显然成立。

2. 条件 3 不满足，因为

$$\begin{aligned}\mathbb{E}[|W_n|I_{\{T>n\}}] &= \mathbb{E}[|W_n|I_{\{T>n\}}|T>n]P(T>n)+\mathbb{E}[|W_n|I_{\{T>n\}}|T\le n]P(T\le n)\\ &= \mathbb{E}[|W_n||T>n]P(T>n)\\ &= (2^n-1)\frac{1}{2^n}\to 1\neq 0\end{aligned}$$

事实上，因为赌博最终以赢得 1 结束，$\mathbb{E}[W_T]=1\neq 0=\mathbb{E}[W_0]$。

鞅的停时定理的条件 3: $\mathbb{E}[|M_nI_{\{T>n\}}|]\to 0$ 有时难以验证，于是可以用一些**包含条件 3 的、易验证的条件来代替**。首先我们要引入**一致可积鞅**的概念。

一致可积序列 (不考)

称一族随机变量 $\{X_n,n\ge 0\}$ 是一致可积的，若

$$\lim_{M\to\infty}\sup_{n\ge 0}\mathbb{E}[|X_n|I_{\{|X_n|>M\}}]=0$$

- $\{M_n,n\ge 0\}$ 是关于 $\{X_n,n\ge 0\}$ 的鞅 + $\{M_n,n\ge 0\}$ 一致可积 $\Rightarrow$ $\{M_n,n\ge 0\}$ 是关于 $\{X_n,n\ge 0\}$ 的一致可积鞅。

鞅的停时定理 (一致可积条件)(不考)

设 $\{M_n\ge 0\}$ 为关于 $\{X_n,n\ge 0\}$ 的**一致可积鞅**。T 为关于 $\{X_n,n\ge 0\}$ 的停时，且满足

1. $P(T<\infty)=1$
2. $\mathbb{E}[|M_T|]<\infty$

则有 $\mathbb{E}[M_T]=\mathbb{E}[M_0]$.

然而，一致可积条件也并不容易加以验证，故此处指出两条常用的、验证 $\{X_n,n\ge 0\}$ 是一致可积序列的充分条件

1. 若存在 $C<\infty$ 使得 $\forall n\ge 0,\mathbb{E}[X_n^2]<C$，则 $\{X_n,n\ge 0\}$ 是一致可积序列。
2. 设 $\{M_n,n\ge 0\}$ 是关于 $\{X_n,n\ge 0\}$ 的鞅，若存在非负随机变量 Y 满足 $\mathbb{E}[Y]<\infty$ 且 $\forall n\ge 0,|M_n|<Y$，则 $\{M_n,n\ge 0\}$ 是一致可积鞅。

6.3.1 下鞅最大值不等式 (考试不考不等式 & 不等式相关证明)

下鞅最大值不等式

设 $\{X_n,n\ge 0\}$ 是下鞅，且 $\forall n\ge 0,X_n\ge 0$，则 $\forall\lambda>0$，有

$$\lambda P(\max_{0\le k\le n}X_k\ge\lambda)\le\mathbb{E}[X_n]$$

回顾 Markov 不等式

$$P(X\ge\lambda)\le\frac{\mathbb{E}[X]}{\lambda}\Rightarrow P(\max_{0\le k\le n}X_k\ge\lambda)\le\frac{\mathbb{E}\left[\max_{0\le k\le n}X_k\right]}{\lambda}$$

由非负下鞅条件，有 $0 \le \mathbb{E}[X_k] \le \mathbb{E}[X_n] \Rightarrow \max\limits_{0\le k\le n} \mathbb{E}[X_k] = \mathbb{E}[X_n]$，从而下鞅最大值不等式可转写为

$$P(\max_{0\le k\le n} X_k \ge \lambda) \le \frac{\max\limits_{0\le k\le n} \mathbb{E}[X_k]}{\lambda}$$

然而，由 Jensen 不等式

$$\mathbb{E}\left[\max_{0\le k\le n} X_k\right] \ge \max_{0\le k\le n} \mathbb{E}[X_k]$$

下鞅最大值不等式提供了一个更紧的上界，其不能由 Markov 不等式直接得到。

6.4 鞅收敛定理 (考试不考)

鞅收敛定理

设 $\{M_n, n \ge 0\}$ 是关于 $\{X_n, n \ge 0\}$ 的鞅，且存在常数 $C < \infty$ 使得 $\forall n \ge 0, \mathbb{E}[|M_n|] < C$，则存在随机变量 M_∞ 使得

$$\lim_{n\to\infty} M_n = M_\infty$$

鞅收敛定理

设 $\{M_n, n \ge 0\}$ 是关于 $\{X_n, n \ge 0\}$ 的一致可积鞅，则存在随机变量 M_∞ 使得

$$\lim_{n\to\infty} M_n = M_\infty$$

且 $\mathbb{E}[M_\infty] = \mathbb{E}[M_0]$。

6.5 连续鞅 (考试不考)

连续鞅

设 $\{X(t), t \ge 0\}$ 为一连续随机过程，记 $\mathcal{F}_t = \sigma(X(s), 0 \le s \le t)$，若

1. $\forall t \ge 0$，有 $\mathbb{E}[|X(t)|] < \infty$
2. $\forall s, t \ge 0$，有 $\mathbb{E}[X(t+s)|\mathcal{F}_t] = X(t)$，a.s.
3. $\forall t \ge 0$，$X(t)$ 是 $\mathcal{F}_t$ 可测的

则称 $\{X(t), t \ge 0\}$ 为鞅。

Markov 性

设 $\{X(t), t \ge 0\}$ 为一连续随机过程，记 $\mathcal{F}_t = \sigma(X(s), 0 \le s \le t)$，若 $\forall s, t \ge 0$

$$p(X(t+s)|\mathcal{F}_t) = p(X(t+s)|X(t)), \text{ a.s.}$$

则称 $\{X(t), t \ge 0\}$ 具有 Markov 性。

[注]: 这与之前章节关于连续 Markov 链的定义是一致的，换句话说，$\forall t_1 < \cdots < t_n$，在给定 $X(t_1), \cdots, X(t_{n-1})$ 的条件下，$X(t_n)$ 或条件概率密度 (或条件分布) 只依赖于 $X(t_{n-1})$，只是这里使用 σ 代数来描述 Markov 性。

7 Brown 运动

7.1 Brown 运动的定义与基本性质

在介绍 Brown 运动之前，有必要指出其和随机游走的关系。考虑一直线上的简单的、对称的随机游动，设质点每过 Δt 时间就随机地以概率 $p=\dfrac{1}{2}$ 向左或向右移动 Δx。记 $X_i=\begin{cases}1 & \text{向右}\\ -1 & \text{向左}\end{cases}$，那么 $\mathbb{E}[X_i]=0, \mathsf{Var}(X_i)=1$。$t$ 时刻质点的位置 $X(t)$ 可以表示为

$$X(t)=\Delta x\sum_{i=1}^{[t/\Delta t]}X_i$$

于是

$$\mathbb{E}[X(t)]=0, \mathsf{Var}(X(t))=\Delta x^2\sum_{i=1}^{[t/\Delta t]}\mathsf{Var}(X_i)=\left[\frac{t}{\Delta t}\right]\Delta x^2$$

我们认为质点是随机的连续运动的，即 $\Delta t\to 0$，注意到

- 如果 $\Delta x=k\Delta t$，则当 $\Delta t\to 0$ 时，$\mathsf{Var}(X(t))\sim\Delta t\to 0$，从而 $X(t)=0$ a.s.，即质点的位置几乎处处时刻为原点。
- 如果 $\Delta x=k(\Delta t)^{\frac{1}{3}}$，则当 $\Delta t\to 0$ 时，$\mathsf{Var}(X(t))\sim\dfrac{1}{(\Delta t)^{\frac{1}{3}}}\to\infty$。这意味着质点**在有限时间内可以远离出发点无穷远，这种情况对于连续运动的粒子显然不合理。**
- 如果 $\Delta x=k\sqrt{\Delta t}$，则当 $\Delta t\to 0$ 时，$\mathsf{Var}(X(t))\to k^2t$。这样质点的位置的方差是 t 的线性函数，这种情况是合理的。

我们沿用 $\Delta x=k\sqrt{\Delta t}$ 的假设，另一方面，注意到 $X(t)=\Delta x\sum_{i=1}^{[t/\Delta t]}X_i$ 是 i.i.d. 的随机变量和，于是当 $\Delta t\to 0$ 时，依**中心极限定理**，当 $\Delta t\to 0$ 时，

$$\frac{\sum_{i=1}^{[t/\Delta t]}\Delta xX_i-0}{k\sqrt{t}}\xrightarrow{d}\mathcal{N}(0,1)\iff X(t)\xrightarrow{d}\mathcal{N}(0,k^2t)$$

Brown 运动就基于这样的背景，它可以被看作是步长为 $c\sqrt{\Delta t}$ 的随机游走的极限。

Brown 运动 (Brownian Motion)

若随机过程 $\{B(t),t\geq 0\}$ 满足

1. $B(0)=0$
2. $\{B(t),t\geq 0\}$ 有平稳独立增量
3. $\forall t, B(t)\sim\mathcal{N}(0,\sigma^2t)$

则称 $\{B(t),t\geq 0\}$ 为 **Brown 运动**，其中 σ 为常数。

不失一般性，接下来全部讨论 $\sigma=1$ 的标准 Brown 分布。Brown 运动有以下更加明确的等价定义：

Brown 运动

若始于 x 的随机过程 $\{B^x(t), t \geq 0\}$ 满足

1. **路径的连续性**：$B(t)$ 是关于 t 的连续函数
2. **独立增量**：$\forall 0 \leq t_1 < t_2 < \cdots < t_n, B(t_1), B(t_2) - B(t_1), \cdots, B(t_n) - B(t_{n-1})$ 相互独立
3. **平稳增量**：$\forall 0 \leq s < t, B(t) - B(s) \sim \mathcal{N}(0, t-s)$

则称 $\{B^x(t), t \geq 0\}$ 为 **Brown 运动**。

[注]：容易知道始于 x 的 Brown 运动都可以转化为始于 0 的 Brown 运动 $B(t)$，即 $B(t) = B^x(t) - x$，这称为 Brown 运动的**空间平移不变性**。

Brown 运动增量的偶数阶矩

Brown 运动增量的二阶、四阶矩为

$$\mathbb{E}[(B(s+t) - B(s))^2] = t, \mathbb{E}[(B(s+t) - B(s))^4] = 3t^2$$

证明. 回顾：服从 $\mathcal{N}(0, t)$ 的随机变量 X 的 k 阶矩为

$$\mathbb{E}[X^k] = \begin{cases} 0, & k\text{为奇数} \\ (k-1)!!t^{\frac{k}{2}}, & k\text{为偶数} \end{cases}$$

代入公式易知 $\mathbb{E}[(B(s+t) - B(s))^2] = t, \mathbb{E}[(B(s+t) - B(s))^4] = 3t^2$。 □

Brown 运动的二阶变差

对任意 $[0, t]$ 上的分割 $0 = t_0 < t_1 < \cdots < t_n = t$，$\delta_n = \max\limits_{1 \leq i \leq n} (t_i - t_{i-1})$，定义 Brown 运动的二阶变差为

$$S_n(t) = \lim_{\delta_n \to 0} \sum_{i=1}^{n} (B(t_i) - B(t_{i-1}))^2$$

Brown 运动二阶变差的收敛性

Brown 运动二阶变差 $S_n(t)$ 均方收敛于 t，即

$$\lim_{\delta_n \to 0} \mathbb{E}[(S_n(t) - t)^2] = 0$$

证明. 令 $\Delta t_i = t_i - t_{i-1}$，$\Delta B(t_i) = B(t_i) - B(t_{i-1})$，则 $\Delta B(t_i) \sim \mathcal{N}(0, \Delta t_i)$。根据此前 Brown 运动增量的偶数阶矩的结论

$$\mathbb{E}[(\Delta B(t_i))^2] = \Delta t_i, \mathbb{E}[(\Delta B(t_i))^4] = 3(\Delta t_i)^2$$

于是

$$
\begin{aligned}
\mathbb{E}[S_n(t)] &= \sum_{i=1}^{n}\mathbb{E}[(\Delta B(t_i))^2] = \sum_{i=1}^{n}\Delta t_i = t \\
\mathrm{Var}[S_n(t)] &= \sum_{i=1}^{n}\mathrm{Var}[(\Delta B(t_i))^2] = \sum_{i=1}^{n}\mathbb{E}[(\Delta B(t_i))^4] - (\mathbb{E}[(\Delta B(t_i))^2])^2 \\
&= \sum_{i=1}^{n}3(\Delta t_i)^2 - (\Delta t_i)^2 = \sum_{i=1}^{n}2(\Delta t_i)^2 \le 2\delta_n\sum_{i=1}^{n}\Delta t_i \\
&= 2t\delta_n \to 0
\end{aligned}
$$

因此

$$
\mathbb{E}[(S_n(t)-t)^2] = \mathrm{Var}[S_n(t)-t] = \mathrm{Var}[S_n(t)] \to 0
$$

□

7.2 Gauss 过程

Gauss 过程

如果对于任何 $n \ge 1, t_1, \cdots, t_n$，随机向量 $(X(t_1), \cdots, X(t_n))$ 服从多元正态分布，则称 $\{X(t), t \ge 0\}$ 为 **Gauss 过程**。

Brown 运动是 Gauss 过程

Brown 运动 $\{B(t), t \ge 0\}$ 是 Gauss 过程。

证明. $\forall n \ge 1, t_1, \cdots, t_n$，$B(t_1), \cdots, B(t_n)$ 都可以表示为独立的正态随机变量

$$
B(t_1), B(t_2) - B(t_1), \cdots, B(t_n) - B(t_{n-1})
$$

的线性组合，从而 $B(t_1), \cdots, B(t_n)$ 服从多元正态分布，从而 $\{B(t), t \ge 0\}$ 是 Gauss 过程。 □

Brown 运动的数字特征

1. $\{B(t), t \ge 0\}$ 是 Brown 运动 $\Rightarrow \mathbb{E}[B(t)] = 0, \mathrm{Cov}(B(s), B(t)) = \min\{s, t\}$
2. $\{B(t), t \ge 0\}$ 是 Brown 运动 $\Leftarrow \{B(t), t \ge 0\}$ 是满足 $\mathbb{E}[B(t)] = 0, \mathrm{Cov}(B(s), B(t)) = \min\{s, t\}$，且**样本轨道连续**的 Gauss 过程

证明. 只证明 (1) 的协方差公式成立。对于 $s < t$

$$
\begin{aligned}
\mathrm{Cov}(B(s), B(t)) &= \mathrm{Cov}(B(s), B(s) + B(t) - B(s)) \\
&= \mathrm{Var}(B(s)) + \mathrm{Cov}(B(s), B(t) - B(s)) = s
\end{aligned}
$$

对于 $t < s$，同理可得 $\mathrm{Cov}(B(s), B(t)) = t$。 □

Brown 运动关于时间步 t 的积分

求 $\int_0^1 B(t)\mathrm{d}t$ 的概率分布。这里 $\int_0^1 B(t)\mathrm{d}t$ 表示 Riemann 积分。

解. 首先要明确为什么 $\int_0^1 B(t)\mathrm{d}t$ 是一个随机变量。实际上，其完整定义为 $\int_0^1 B(t,\omega)\mathrm{d}t$，因此 $B(t,\omega)$ 关于 t 积分后是一个关于样本点 ω 的随机变量。又因为 $B(t)$ 是处处连续的，因此其 *Riemann* 积分必存在。

对 $[0,1]$ 作任意分割 $0=t_0<t_1<\cdots<t_n=1$，记 $\Delta t_i=t_i-t_{i-1}, \delta_n=\max\limits_{1\le i\le n}\Delta t_i$，则

$$\int_0^1 B(t)\mathrm{d}t=\lim_{\delta_n\to 0}\sum_{i=1}^n B(t_i)\Delta t_i$$

我们已知 $B(t)$ 关于 t 的 *Riemann* 积分必存在，故取 $t_i=\frac{i}{n}$，则

$$\int_0^1 B(t)\mathrm{d}t=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n B\left(\frac{i}{n}\right)$$

注意到 $B(\frac{i}{n})\sim\mathcal{N}(0,\frac{i}{n})$，因此 $\int_0^1 B(t)\mathrm{d}t$ 必为零均值正态随机变量。

$$\begin{aligned}
\mathit{Var}(\int_0^1 B(t)\mathrm{d}t) &= \mathit{Cov}\left(\int_0^1 B(t)\mathrm{d}t,\int_0^1 B(s)\mathrm{d}s\right)=\mathbb{E}\left[\int_0^1 B(t)\mathrm{d}t\int_0^1 B(s)\mathrm{d}s\right]\\
&=\mathbb{E}\left[\int_0^1\int_0^1 B(t)B(s)\mathrm{d}t\mathrm{d}s\right]=\int_0^1\int_0^1\mathbb{E}[B(t)B(s)]\mathrm{d}t\mathrm{d}s\\
&=\int_0^1\int_0^1\min(s,t)\mathrm{d}s\mathrm{d}t=\frac{1}{3}
\end{aligned}$$

所以 $\int_0^1 B(t)\mathrm{d}t\sim\mathcal{N}(0,\frac{1}{3})$.

7.3 Brown 运动的 Markov 性和鞅性 (考试不考)

Brown 运动的鞅性质

设 $\{B(t),t\ge 0\}$ 为 Brown 运动，则

1. $\{B(t),t\ge 0\}$ 为鞅。
2. $\{B^2(t)-t,t\ge 0\}$ 为鞅。
3. $\forall u,\exp\left(uB(t)-\frac{1}{2}u^2t\right)$ 为鞅。

证明. 由 Brown 运动的独立增量性知，$B(t+s)-B(t)$ 与 $\mathcal{F}_t=\sigma(B(s),0\le s\le t)$ 独立，从而对**任意函数** g，有

$$\mathbb{E}[g(B(t+s)-B(t))|\mathcal{F}_t]=\mathbb{E}[g(B(t+s)-B(t))]$$

对于 (1)

$$\begin{aligned}\mathbb{E}[B(t+s)|\mathcal{F}_t] &= \mathbb{E}[B(t+s)-B(t)+B(t)|\mathcal{F}_t] \\ &= \mathbb{E}[B(t+s)-B(t)|\mathcal{F}_t]+\mathbb{E}[B(t)|\mathcal{F}_t] \\ &= B(t)+\mathbb{E}[B(t+s)-B(t)] = B(t) \qquad \text{得证}\end{aligned}$$

对于 (2)，首先 $\mathbb{E}[B^2(t)]=t \Rightarrow \mathbb{E}[|B^2(t)-t|]=0<\infty$，再注意到

$$\begin{aligned}\mathbb{E}[B^2(t+s)|\mathcal{F}_t] &= \mathbb{E}[(B(t+s)-B(t)+B(t))^2|\mathcal{F}_t] \\ &= \mathbb{E}[(B(t+s)-B(t))^2|\mathcal{F}_t]^2+\mathbb{E}[B^2(t)|\mathcal{F}_t]-2\mathbb{E}[B(t)(B(t+s)-B(t))|\mathcal{F}_t] \\ &= \mathbb{E}[(B(t+s)-B(t))^2]+B^2(t)-2\mathbb{E}[B(t)]\mathbb{E}[B(t+s)-B(t)] \\ &= s+B^2(t)\end{aligned}$$

$$\Rightarrow \mathbb{E}[B^2(t+s)-(t+s)|\mathcal{F}_t] = s+B^2(t)-(s+t) = B^2(t)-t \qquad \text{得证}$$

对于 (3)，首先根据 $B(t)\sim\mathcal{N}(0,t)$ 的矩母函数知 $\mathbb{E}[\exp(uB(t))]=\exp(\frac{tu^2}{2})$，于是 $\mathbb{E}[|\exp(uB(t))-\frac{tu^2}{2}|]=0<\infty$，再注意到

$$\begin{aligned}\mathbb{E}[\exp(uB(t+s))|\mathcal{F}_t] &= \mathbb{E}[\exp(uB(t+s)-uB(t)+uB(t))|\mathcal{F}_t] \\ &= \exp(uB(t))\mathbb{E}[\exp(uB(t+s)-uB(t))] = \exp(uB(t))\exp(\frac{su^2}{2})\end{aligned}$$

$$\Rightarrow \mathbb{E}[\exp(uB(t+s)-\frac{1}{2}u^2(t+s))|\mathcal{F}_t] = \exp(uB(t))\exp(\frac{su^2}{2})\exp(-\frac{1}{2}u^2(s+t)) = \exp(uB(t)-\frac{1}{2}u^2t) \qquad \text{得证}$$

□

Brown 运动的 Markov 性

设 $\{B(t),t\geq 0\}$ 是 Brown 运动，则 $\{B(t),t\geq 0\}$ 具有 Markov 性。

证明. **不妨利用分布函数 $F_X(x)$ 与矩母函数 $\phi(t)$ 一一对应的性质来证明这一点。**对于任意 $s<t$，有

$$\begin{aligned}\mathbb{E}[\exp(uB(t+s))|\mathcal{F}_t] &= \mathbb{E}[\exp(uB(t+s)-uB(t)+uB(t))|\mathcal{F}_t] \\ &= \mathbb{E}[\exp(uB(t))|B(t)]\cdot\mathbb{E}[\exp(uB(t+s)-uB(t))|B(t)] \\ &= \mathbb{E}[\exp(uB(t+s))|B(t)]\end{aligned}$$

$\mathcal{F}_t$ 条件下 $\{B(t),t\geq 0\}$ 的矩母函数和 $B(t)$ 条件下相同，所以 $\{B(t),t\geq 0\}$ 具有 Markov 性。 □

7.4 与 Brown 运动有关的概率分布

7.4.1 首次击中 x 的时刻 T_x

首次击中 x 的时刻 T_x 的概率分布

记 $T_x=\inf\{t:B(t)=x,t\geq 0\}$，则当 $B(0)=0,x>0$ 时，

$$P(T_x\leq t) = 2P(B(t)\geq x) = 2\left(1-\Phi(\frac{x}{\sqrt{t}})\right)$$

证明. 考虑在 t 时刻之前是否击中过 x

$$P(B(t)\geq x)=P(B(t)\geq x|T_x\leq t)P(T_x\leq t)+P(B(t)\geq x|T_x>t)P(T_x>t)$$

(1) 若 t 之后首次击中 x，则 t 时刻 $B(t)$ 的位置一定小于 x，即 $P(B(t)\geq x|T_x>t)=0$。
(2) 若 t 之前首次击中 x，则 t 时刻 $B(t)$ 的位置一定大于 x 的概率为 $\frac{1}{2}$(因为 $B(t)\sim\mathcal{N}(0,t)$)，即 $P(B(t)\geq x|T_x\leq t)=\frac{1}{2}$。
所以 $P(B(t)\geq x)=\frac{1}{2}P(T_x\leq t)$，即 $P(T_x\leq t)=2P(B(t)\geq x)=2\left(1-\Phi(\frac{x}{\sqrt{t}})\right)$。 □

Brown 运动的常返性

$$P(T_x<\infty)=1,\mathbb{E}[T_x]=\infty$$

[注]: Brown 运动以概率 1 击中 x，但它的平均时间是无穷，这类似于 Markov 链中的零常返态：$f_{jj}=1$，但 $\mu_j=\infty$。故上述性质也称为 Brown 运动的常返性。

证明. 当 $x>0$ 时，

$$P(T_x\leq t)=2P(B(t)\geq x)=\frac{2}{\sqrt{2\pi t}}\int_x^\infty e^{-\frac{u^2}{2t}}\mathrm{d}u=\frac{2}{\sqrt{2\pi}}\int_{\frac{x}{\sqrt{t}}}^\infty e^{-\frac{u^2}{2}}\mathrm{d}u$$

又注意到 $\{T_x<\infty\}=\bigcup\limits_{t>0}\{T_x\leq t\}$，且事件集 $\{T_x\leq t\}$ 单调递增，根据概率的下连续性知，$P(\bigcup\limits_{t>0}\{T_x\leq t\})=\lim\limits_{t\to\infty}P(T_x\leq t)$。从而

$$P(T_x<\infty)=\lim_{t\to\infty}P(T_x\leq t)=\frac{2}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{u^2}{2}}\mathrm{d}u=1$$

对 $P(T_x\leq t)=\frac{2}{\sqrt{2\pi t}}\int_x^\infty e^{-\frac{u^2}{2t}}\mathrm{d}u=2\left(1-\Phi\left(\frac{x}{\sqrt{t}}\right)\right)$ 关于 t 求导可得

$$f_{T_x}(t)=-2\Phi'\left(\frac{x}{\sqrt{t}}\right)\frac{x}{-2t^{\frac{3}{2}}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2t}}\frac{x}{t^{\frac{3}{2}}}=\frac{x}{\sqrt{2\pi}}t^{-\frac{3}{2}}e^{-\frac{x^2}{2t}}$$

而

$$\begin{aligned}\mathbb{E}[T_x]&=\int_0^\infty P(T_x>t)\mathrm{d}t=\int_0^\infty\left(1-\frac{2}{\sqrt{2\pi}}\int_{x/\sqrt{t}}^\infty e^{-\frac{u^2}{2}}\mathrm{d}u\right)\mathrm{d}t\\&=\frac{2}{\sqrt{2\pi}}\int_0^\infty\left(\sqrt{\frac{\pi}{2}}-\int_{x/\sqrt{t}}^\infty e^{-\frac{u^2}{2}}\mathrm{d}u\right)\mathrm{d}t=\frac{2}{\sqrt{2\pi}}\int_0^\infty\int_0^{x/\sqrt{t}}e^{-\frac{u^2}{2}}\mathrm{d}u\mathrm{d}t\\&=\frac{2}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{u^2}{2}}\mathrm{d}u\int_0^{x^2/u^2}\mathrm{d}t=\frac{2x^2}{\sqrt{2\pi}}\int_0^\infty\frac{1}{u^2}e^{-\frac{u^2}{2}}\mathrm{d}u\\&\geq\frac{2x^2}{\sqrt{2\pi}}\int_0^\infty\frac{1}{u^2}\mathrm{d}u=\infty\end{aligned}$$

当 $x<0$ 时，由对称性易知 T_x **与** T_{-x} **同分布**，此时

$$P(T_x\leq t)=\frac{2}{\sqrt{2\pi}}\int_{\frac{-x}{\sqrt{t}}}^\infty e^{-\frac{u^2}{2t}}\mathrm{d}u$$

$$f_{T_x}(t)=\frac{-x}{\sqrt{2\pi}}t^{-\frac{3}{2}}e^{-\frac{x^2}{2t}},\ u>0$$

类似仍能得到 $P(T_x<\infty)=1,\mathbb{E}[T_x]=\infty$ 的结论。 □

Brown 运动最大值最小值的分布

若 $x>0$，则

$$\{\max_{0\le s\le t} B(s)\ge x\} \iff \{T_x\le t\}$$

若 $x<0$，则

$$\{\min_{0\le s\le t} B(s)\le x\} \iff \{T_x\le t\}$$

[注]：当 $x>0$ 时，$P(\sup\limits_{0\le s\le t} B(s)\ge x)=P(T_x\le t)=2P(B(t)\ge x)$，这也被称为 **Levy 等式**。

证明. 根据 Brown 运动的连续性，$x>0$ 时，t 之前的最大值大于等于 $x \iff t$ 之前击中 x；$x<0$ 时，t 之前的最小值小于等于 $x \iff t$ 之前击中 x。 □

Brown 运动处处不可导 (考试不考)

对于 Brown 运动 $\{B(t), t\ge 0\}$，$\forall t_0\ge 0$

$$P\left(\limsup_{t\to t_0}\frac{B(t)-B(t_0)}{t-t_0}=+\infty\right)=1$$

证明. 只需证 $\forall c>0$，有

$$P\left(\sup_{t_0\le t\le t_0+\delta}\frac{B(t)-B(t_0)}{t-t_0}\ge c\right)\to 1,\ \delta\to 0$$

事实上，由 Levy 等式，$P(\sup\limits_{0\le s\le t} B(s)\ge x)=2P(B(t)\ge x)$，那么

$$\begin{aligned}
P\left(\sup_{t_0\le t\le t_0+\delta}\frac{B(t)-B(t_0)}{t-t_0}\ge c\right) &\ge P\left(\sup_{t_0\le t\le t_0+\delta} B(t)-B(t_0)\ge c\delta\right)\\
&=2P\left(B(t)-B(t_0)\ge c\delta\right)\\
&=2\frac{1}{\sqrt{2\pi}}\int_{c\sqrt{\delta}}^{\infty} e^{-\frac{u^2}{2}}\mathrm{d}u\to 1,\ \delta\to 0
\end{aligned}$$

□

7.4.2 arcsin 律与 arccos 律

Brown 运动至少一次越过横轴的概率

设 $\{B^x(t), t\ge 0\}$ 为始于 x 的 Brown 运动，则 $B^x(t)$ 在 $(0,t)$ 中至少有一个零点的概率为：

$$\frac{|x|}{\sqrt{2\pi}}\int_0^t u^{-\frac{3}{2}}e^{-\frac{x^2}{2u}}\mathrm{d}u$$

证明. 若 $x<0$，令 $A=\{\exists u,\ \text{s.t. } B^x(u)=0, u\in(0,t)\}$，则
$A\iff$ 从 x 出发，t 之前击中 0

$\iff$ 从 0 出发，t 之前击中 $-x$

$\iff \{T_{-x} \le t\} \iff \{T_x \le t\}$，所以 $P(A) = P(T_x \le t) = \dfrac{-x}{\sqrt{2\pi}} \displaystyle\int_0^t u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} \mathrm{d}u.$

若 $x > 0$，类似地可得 $P(A) = P(T_x \le t) = \dfrac{x}{\sqrt{2\pi}} \displaystyle\int_0^t u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} \mathrm{d}u$ □

Brown 运动的 $\arccos$ 律

设 $\{B^y(t), t \ge 0\}$ 为 Brown 运动，则 $B^y(t)$ 在 (a,b) 中至少有一个零点的概率为 $\dfrac{2}{\pi}\arccos\sqrt{\dfrac{a}{b}}$。

证明. 仍然令 $A(a,b) = \{\exists u,\ \text{s.t. } B(u) = 0, u \in (a,b)\}$，则可以先确定 a 时刻的位置，再考虑穿越。且注意到 $B(a) \sim \mathcal{N}(0,a)$，于是

$$
\begin{aligned}
P(A(a,b)) &= \int_{-\infty}^{\infty} P(A(a,b)|B(a) = x) f_{B(a)}(x) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} P(A(0,b-a)|B(0) = x) f_{B(a)}(x) \mathrm{d}x && \text{(Markov 性)} \\
&= \int_{-\infty}^{\infty} \left(\int_0^{b-a} \frac{|x|}{\sqrt{2\pi}} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} \mathrm{d}u \right) \cdot \frac{1}{\sqrt{2\pi a}} e^{-\frac{x^2}{2a}} \mathrm{d}x \\
&= \frac{2}{\sqrt{2\pi a}} \int_0^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2a}} \int_0^{b-a} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} \mathrm{d}u \mathrm{d}x \\
&= \frac{1}{\sqrt{a}\pi} \int_0^{b-a} u^{-\frac{3}{2}} \int_0^{\infty} x e^{-x^2\left(\frac{1}{2a} + \frac{1}{2u}\right)} \mathrm{d}x \mathrm{d}u \\
&= \frac{1}{\sqrt{a}\pi} \int_0^{b-a} u^{-\frac{3}{2}} \frac{au}{a+u} \mathrm{d}u \\
&= \frac{\sqrt{a}}{\pi} \int_0^{b-a} \frac{u^{-\frac{1}{2}}}{a+u} \mathrm{d}u = \frac{2}{\pi} \arccos\sqrt{\frac{a}{b}}
\end{aligned}
$$

□

Brown 运动的 $\arcsin$ 律

设 $\{B^y(t), t \ge 0\}$ 为 Brown 运动，则 $B^y(t)$ 在 (a,b) 没有零点的概率为 $\dfrac{2}{\pi}\arcsin\sqrt{\dfrac{a}{b}}$。

证明.

$$
\begin{aligned}
P(\text{没有零点}) &= 1 - P(\text{至少有一个零点}) \\
&= 1 - \frac{2}{\pi}\arccos\sqrt{\frac{a}{b}} = \frac{2}{\pi}\arcsin\sqrt{\frac{a}{b}}
\end{aligned}
$$

□

7.5 Brown 运动的变体

7.5.1 Brown 桥

在许多实际问题中，往往是在指定 $X(0)=x$ 和 $X(t_0)=y$ 的条件下，研究中间过程的情形，即 $\{X(t), 0 \le t \le t_0 | X(0)=x, X(t_0)=y\}$ 的性质。考虑标准 Brown 运动 $\{B(t), t \ge 0\}, B(0)=0$，记

$$X(t) = B(t) + x + \frac{t}{t_0}(y - x - B(t_0))$$

容易发现过程 $\{X(t), t \ge 0\}$ 的任何路径必经过 $(0,x)$ 和 (t_0, y)，就好像“一座桥的两端”一样。通过坐标变换，可以将 t 规范到 $[0,t_0]=[0,1]$ 上，且 $X(0)=0, X(1)=1$

$$X(t) = B(t) - tB(1)$$

Brown 桥

设 $\{B(t), t \ge 0\}$ 为 Brown 桥，令

$$B^*(t) = B(t) - tB(1),\ 0 \le t \le 1$$

则称 $\{B^*(t), 0 \le t \le 1\}$ 为 **Brown 桥**。

[注]：Brown 桥也是 Gauss 过程，因为

$$\begin{pmatrix} B^*(t_1) \\ \vdots \\ B^*(t_n) \end{pmatrix} = \begin{pmatrix} B(t_1) \\ \vdots \\ B(t_n) \end{pmatrix} - B(1) \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$$

Brown 桥的数字特征有

1. $\mathbb{E}[B^*(t)] = 0$
2. $\forall 0 \le s,t \le 1$，$\mathrm{Cov}(B^*(s), B^*(t)) = \min\{s,t\} - st$

证明.

$$\begin{aligned}
\mathbb{E}[B^*(t)] &= \mathbb{E}[B(t) - tB(1)] = 0 \\
\mathrm{Cov}(B^*(s), B^*(t)) &= \mathbb{E}[B^*(s)B^*(t)] \\
&= \mathbb{E}[(B(s) - sB(1))(B(t) - tB(1))] \\
&= \mathbb{E}[B(s)B(t)] - s\mathbb{E}[B(1)B(t)] - t\mathbb{E}[B(1)B(s)] + st\mathbb{E}[B(1)B(1)] \\
&= \min(s,t) - 2st + st = \min(s,t) - st
\end{aligned}$$

□

Brown 桥的等价定义

设 $\{B^*(t), t \ge 0\}$ 为 Gauss 过程，且 $\forall 0 \le s,t \le 1$，$\mathbb{E}[B^*(t)] = 0, \mathrm{Cov}(B^*(s), B^*(t)) = \min\{s,t\} - st$，则 $\{B^*(t), t \ge 0\}$ 为 Brown 桥。

7.5.2 有吸收值的 Brown 运动

有吸收值的 Brown 运动

若 $x>0$，令

$$Z(t)=\begin{cases}B(t), & t<T_x\\ x, & t\geq T_x\end{cases}$$

则称 $\{Z(t),t\geq 0\}$ 为**有吸收值 x 的 Brown 运动**。

显然 $Z(t)$ 是混合型随机变量，我们考虑其分布函数 $F_Z(y)=P(Z(t)\leq y)$

1. $y>x$ 时，$F_Z(y)=P(Z(t)\leq y)=1$
2. $y=x$ 时，$F_Z(y)=P(Z(t)=x)=P(T_x\leq t)=\dfrac{2}{\sqrt{2\pi t}}\displaystyle\int_x^{\infty}e^{-\frac{u^2}{2t}}\mathrm{d}u$
3. $0<y<x$ 时，

$$\begin{aligned}
P(Z(t)\leq y)&=P(B(t)\leq y,T_x>t)\\
&=P(B(t)\leq y)-P(B(t)\leq y,T_x\leq t) && \text{(到过 } x \text{ 且 } t \text{ 时刻小于 } y\text{)}\\
&=P(B(t)\leq y)-P(B(t)\geq 2x-y,T_x\leq t) && \text{(到达 } x \text{ 后下降 } x-y \iff \text{ 到达 } x \text{ 后上升 } x-y\text{)}\\
&=P(B(t)\leq y)-P(B(t)\geq 2x-y) && \text{(因为 } B(t)>2x-y \text{ 能推出 } B(t)>x\text{)}\\
& && \text{(所以 } \{B(t)>2x-y\}\subset\{B(t)>x\}\subset\{T(x)\leq t\}\text{)}\\
&=P(B(t)\leq y)-P(B(t)\leq y-2x) && \text{(Brown 运动对称性)}\\
&=\frac{1}{\sqrt{2\pi t}}\int_{y-2x}^{y}e^{-\frac{u^2}{2t}}\mathrm{d}u
\end{aligned}$$

综上

$$F_Z(y)=\begin{cases}1, & y>x\\ \dfrac{2}{\sqrt{2\pi t}}\displaystyle\int_x^{\infty}e^{-\frac{u^2}{2t}}\mathrm{d}u, & y=x\\ \dfrac{1}{\sqrt{2\pi t}}\displaystyle\int_{y-2x}^{y}e^{-\frac{u^2}{2t}}\mathrm{d}u, & 0<y<x\end{cases}$$

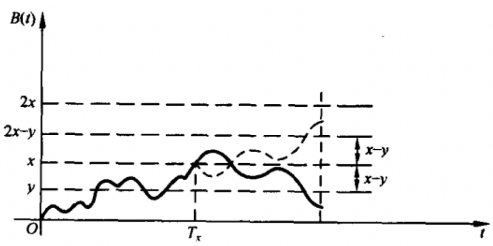


图 7: 带吸收值 x 的 Brown 运动

7.5.3 在原点反射的 Brown 运动

在原点反射的 Brown 运动

若 $\{B(t), t \geq 0\}$ 为 Brown 运动，定义

$$Y(t) = |B(t)|, t \geq 0$$

为**在原点反射的 Brown 运动**。

考虑其分布函数 $F_Y(y) = P(Y(t) \leq y)$

$$\begin{aligned} F_Y(y) &= P(Y(t) \leq y) = P(-y \leq B(t) \leq y) \\ &= P(B(t) \leq y) - P(B(t) \leq -y) = 2P(B(t) \leq y) - 1 \\ &= \frac{2}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{-\frac{u^2}{2t}} \mathrm{d}u - 1 \end{aligned}$$

7.5.4 几何 Brown 运动

几何 Brown 运动

若 $\{B(t), t \geq 0\}$ 为 Brown 运动，定义

$$X(t) = e^{B(t)}, t \geq 0$$

为**几何 Brown 运动**。

根据正态随机变量 $B(t) \sim \mathcal{N}(0, t)$ 的矩母函数 $\mathbb{E}[e^{uB(t)}] = e^{\frac{1}{2}u^2 t}$，我们可以求几何 Brown 运动的数字特征：

1. $\mathbb{E}[X(t)] = \mathbb{E}[e^{B(t)}] = e^{\frac{1}{2}t}$
2. $\mathrm{Var}(X(t)) = \mathbb{E}[e^{2B(t)}] - \mathbb{E}[e^{B(t)}]^2 = e^{2t} - e^t$

几何 Brown 运动常用来建模**相对变化为 i.i.d. 随机变量的过程**，例如股票价格。设 $X(n)$ 为 n 时刻股票的价格，我们认为收益率 $\dfrac{X(n)}{X(n-1)} = R(n)$，$n \geq 1$ 是 i.i.d 的。不妨设 $X(0) = 1$，则

$$\begin{aligned} X(n) &= \frac{X(n)}{X(n-1)} \cdots \frac{X(1)}{X(0)} X(0) = R(n) \cdots R(1) \\ \Rightarrow \ln X(n) &= \sum_{i=1}^{n} \ln R(i) \end{aligned}$$

当 $n \to \infty$，由中心极限定理知 $\ln X(n) \to \mathcal{N}(0, n)$，其中假设 $\mu = \mathbb{E}[\ln R(i)] = 0$，$\sigma^2 = \mathrm{Var}[\ln R(i)] = 1$。于是 $\{\ln X(n), n \geq 0\}$ 渐进为 Brown 运动，股价 $\{X(n), n \geq 0\}$ 渐进为几何 Brown 运动。