应用随机过程学习笔记

Julique, Xuancrane

gene_introne@163.com

期待您的批评与指正!

2023 年 9 月 23 日

目录

学习路线

本笔记主要框架参考的是“张波、商豪、邓军（2020）《应用随机过程》第 5 版，中国人民大学出版社”，非此教材中的内容均来自课堂讲解和课下讨论.

首先，随机过程可以视为“概率论和微积分的综合”，简单来看，随机过程

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

是定义在 $T \times \Omega$ 上的二元函数.

- 固定 t，$X(t_0,\omega)$ 是随机变量 X，对应概率论；
- 固定 ω，$X(t,\omega_0)$ 是一元函数 $f(t)$，对应微积分.

一般而言，学习有两条线路：

- 概率论 $\to$ 随机过程 $\to$ 实变函数、测度论 $\to$ 随机分析
- 数理统计 $\to$ 计量经济学、时间序列分析

其次，本笔记按照个人学习习惯，将学习内容分点陈述并且分为定义、定理（引理、推论、命题等统称为定理）、注记和例子，其中**定义**的专有名词用红色加粗字体显示，注记中**重要的内容**用蓝色加粗字体显示.

与前一版对比，本笔记更新了一系列有趣或有助于理解的练习题，并附上详细解答.

本笔记无偿提供，仅供学习参考使用，请不要以商业目的传播，谢谢！

第一章 预备知识

本章主要介绍的是（高等）概率论的基本知识.

- 其中比较重要的是概率空间、几个常见的分布、数字特征、矩母函数、多元正态分布、条件期望的性质（尤其是第 1-6 条）和卷积的性质. 此外，也会经常运用初等概率论中的全概率公式和 Bayes 公式.
- 多数定理的详细证明需要使用更加丰富和深刻的数学知识（实变函数、测度论等），以应用为目的的学习者不必深究.

1.1 概率空间

定义 1.1.1. 设 Ω 是一个样本空间（或任意一个集合），$\mathcal{F}$ 是 Ω 的某些子集组成的集合族，如果满足：

1. $\Omega \in \mathcal{F}$;
2. 若 $A \in \mathcal{F}$，则 $A^c = \Omega \backslash \mathrm{A} \in \mathcal{F}$;
3. 若 $A_n \in \mathcal{F}, n = 1, 2, \cdots$，则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$;

则称 $\mathcal{F}$ 为 σ **代数**. 称 $(\Omega, \mathcal{F})$ 为**可测空间**，称 $\mathcal{F}$ 中的元素为**事件**.

定理 1.1.2. （定理）σ 代数的其他性质.

1. 对可列个无限交封闭：若 $A_n \in \mathcal{F}, n = 1, 2, \cdots$，则
$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$$
证明.
$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c \in \mathcal{F}$$
□

2. 对有限交、有限并封闭：若 $A_n \in \mathcal{F}, n = 1, 2, \cdots$，则
$$\bigcap_{n=1}^{k} A_n \in \mathcal{F}, \bigcup_{n=1}^{k} A_n \in \mathcal{F}$$

证明.

$$\bigcap_{n=1}^{k} A_n = A_1 \cap A_2 \cap \cdots \cap A_k \cap \Omega \cap \Omega \cdots \in \mathcal{F}$$

$$\bigcup_{n=1}^{k} A_n = A_1 \cup A_2 \cup A_k \cup \cdots \cup \varnothing \cup \varnothing \cdots \in \mathcal{F}$$

□

3. 对集合的上下极限封闭：若 $A_n \in \mathcal{F}, n = 1, 2, \cdots$，则

$$\overline{\lim_{n\to\infty}} A_n \in \mathcal{F}, \ \underline{\lim_{n\to\infty}} A_n \in \mathcal{F}$$

证明.

$$\overline{\lim_{n\to\infty}} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \in \mathcal{F}$$

$$\underline{\lim_{n\to\infty}} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m \in \mathcal{F}$$

□

定义 1.1.3. 对于 Ω 上的任一非空集类 $\mathcal{C}$(由 Ω 中一些子集构成的集合)，存在包含 $\mathcal{C}$ 的最小 σ 代数，称为 **$\mathcal{C}$ 生成的 σ 代数**，记作 $\sigma(\mathcal{C})$，即，

$$\bigcap \{\mathcal{H} | \ \mathcal{H}\text{为包含}\mathcal{C}\text{的}\sigma\text{代数}\}$$

定义 1.1.4. （定义）设 $(\Omega, \mathcal{F})$ 为可测空间，$\mathbb{P}$ 是定义在 $\mathcal{F}$ 上的实值函数，如果满足

1. 非负性：$\forall A \in \mathcal{F}, 0 \leqslant \mathbb{P}(A) \leqslant 1$；
2. 规范性：$\mathbb{P}(\Omega) = 1$；
3. 可列可加性：对 $\mathcal{F}$ 中两两不相容的事件 $A_1, A_2, \cdots$，有

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

则称 $\mathbb{P}$ 是 $(\Omega, \mathcal{F})$ 上的概率，称 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间，$\mathbb{P}(A)$ 为事件 A 的概率.

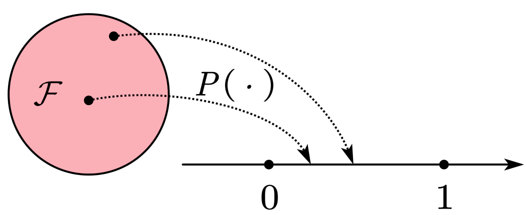

图 1.1. 概率函数

注记 1.1.5. 为什么定义 $\mathcal{F}$? ——为了定义概率.

1. 若 $A \in \mathcal{F}$，$\mathbb{P}(A)$ 才有意义.
2. 若 $A \notin \mathcal{F}$，而，但 $\mathbb{P}(A)$ 没有意义.

例子 1.1.6. 设 $\Omega = \{1,2,3,4\}$，$\mathcal{F}_1 = \{\Omega, \varnothing, \{1\}, \{2,3,4\}\}$，对概率空间 $(\Omega, \mathcal{F}_1, \mathbb{P})$，我们定义 $\mathbb{P}\{\{1\}\} = \frac{1}{2}, \mathbb{P}(\varnothing) = 0$, $\mathbb{P}\{\Omega\} = 1, \mathbb{P}\{\{2,3,4\}\} = \frac{1}{2}$，虽然有 $\{1,2\} \subset \Omega$，但 $\mathbb{P}\{\{1,2\}\}$ 没有意义.

例子 1.1.7. 试写出 $\Omega = \{1,2,3\}$ 的所有 σ 代数.

首先，最小的 σ 代数是

$$\mathcal{F}_1 = \{\Omega, \varnothing\}$$

除此之外，为了满足 σ 代数的性质，可以写出：

$$\mathcal{F}_2 = \{\Omega, \varnothing, \{1\}, \{2,3\}\}$$

$$\mathcal{F}_3 = \{\Omega, \varnothing, \{2\}, \{1,3\}\}$$

$$\mathcal{F}_4 = \{\Omega, \varnothing, \{3\}, \{1,2\}\}$$

最后，最大的 σ 代数包含了 Ω 的所有子集，它是

$$\mathcal{F}_5 = \{\Omega, \varnothing, \{1\}, \{2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2\}\}$$

1.2 随机变量和分布函数

定义 1.2.1. 设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是完备的概率空间，X 是定义在 Ω 上取值于实数集 $\mathbb{R}$ 的函数，称 X 是 **$\mathcal{F}$ 上（或 $\mathcal{F}$-可测的）的随机变量**，简称**随机变量**，如果 $\forall x \in \mathbb{R}$，有

$$\{\omega : X(\omega) \leqslant x\} \in \mathcal{F}$$

定义 1.2.2. 如果 X 是 $\mathcal{F}$ 上的随机变量，函数

$$F(x) = \mathbb{P}\{\omega : X(\omega) \leqslant x\}, -\infty < x < +\infty$$

称为**随机变量 X 的分布函数**.

定义 1.2.3. 两个随机变量 X 和 Y，称它们是**等价的**，如果

$$\mathbb{P}\{\omega \in \Omega : X(\omega) \neq Y(\omega)\} = 0$$

定理 1.2.4. （定理）几个等价的命题：

1. X 是随机变量;
2. $\forall x\in\mathbb{R}$, $\{\omega: X(\omega)\geqslant x\}\in\mathcal{F}$.
3. $\forall x\in\mathbb{R}$, $\{\omega: X(\omega)> x\}\in\mathcal{F}$.
4. $\forall x\in\mathbb{R}$, $\{\omega: X(\omega)< x\}\in\mathcal{F}$.

定理 1.2.5. 联合分布函数的性质:

1. $F(x_1,x_2,\cdots,x_d)$ 对每个变量单调不减;
2. $F(x_1,x_2,\cdots,x_d)$ 对每个变量都是右连续的;
3. 对 $i=1,2,\cdots d$,

$$\lim_{x_i\to-\infty}F(x_1,x_2,\cdots,x_d)=0$$

$$\lim_{x_i\to+\infty}F(x_1,x_2,\cdots,x_d)=1$$

定义 1.2.6. 如果

$$f(x_1,x_2,\cdots,x_d)=\frac{\partial^d F}{\partial x_1\partial x_2\cdots\partial x_d}$$

对所有的 $(x_1,x_2,\cdots,x_d)\in\mathbb{R}^d$ 存在, 则称 $f(x_1,x_2,\cdots,x_d)$ 为 $\boldsymbol{X}$ 的**联合概率密度函数**.

定义 1.2.7. $1\leqslant k_1<k_2<\cdots<k_n\leqslant d$, 则 $X_{k_1},X_{k_2},\cdots,X_{k_n}$ 的**边际分布函数**定义为

$$F_{X_{k_1},X_{k_2},\cdots,X_{k_n}}(x_{k_1},x_{k_2},\cdots,x_{k_d})=F(\infty,\cdots,\infty,x_{k_1},\infty,\cdots,\infty,x_{k_n},\infty,\cdots,\infty)$$

1.3 数字特征、矩母函数和特征函数

注记 1.3.1. 期望的统一表达式:

$$\mathbb{E}(X)=\int_{-\infty}^{+\infty}x\mathrm{d}F(x)$$

定义 1.3.2. 若随机变量 X 的分布函数为 $F_X(x)$, 则称 e^{tX} 的期望 (如果存在) 为 X 的**矩母函数**, 即

$$\phi_X(t)=\mathbb{E}\left[e^{tX}\right]=\int_\Omega e^{tX(\omega)}\mathbb{P}(\mathrm{d}\omega)=\int_{-\infty}^{+\infty}e^{tx}\mathrm{d}F_X(x)$$

例子 1.3.3. 求 $X\sim N(\mu,\sigma^2)$ 的矩母函数.

$$\phi_X(t)=\mathbb{E}\left[e^{tX}\right]=\int_{-\infty}^{+\infty}\frac{e^{tx}}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$$

令 $u=\frac{x-\mu}{\sigma}$, 得

$$\text{上式}=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}+t\sigma u+\mu t}\mathrm{d}u=e^{\mu t+\frac{\sigma^2}{2}t^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{(u-\sigma t)^2}{2}}\mathrm{d}u=e^{\mu t+\frac{\sigma^2}{2}t^2}$$

最后一步是换元后，根据 Γ 函数的性质得到，即

$$2\int_0^{+\infty} e^{-\frac{x^2}{2}}\mathrm{d}x=\sqrt{2}\left(2\int_0^{+\infty} y^{2\cdot\frac{1}{2}-1}e^{-y^2}\mathrm{d}y\right)=\sqrt{2}\Gamma\left(\frac{1}{2}\right)=\sqrt{2\pi}$$

综上，

$$\phi_X(t)=\mathbb{E}\left[e^{tX}\right]=e^{\mu t+\frac{\sigma^2}{2}t^2}$$

例子 1.3.4. 矩母函数的应用 1：求 n 阶矩.（如果对 $\phi(t)$ 的求导运算和求期望运算可以交换次序）

$$\phi'(t)=\mathbb{E}\left[Xe^{tX}\right]$$

$$\phi''(t)=\mathbb{E}\left[X^2e^{tX}\right]$$

$$\vdots$$

$$\phi^{(n)}(t)=\mathbb{E}\left[X^ne^{tX}\right]$$

令 $t=0$，则 $\mathbb{E}[X^n]=\phi^{(n)}(0)$.

例子 1.3.5. 矩母函数的应用 2：与矩母函数相关的一个不等式

$$\mathbb{P}\{\xi\geqslant t\}\leqslant\inf_{\lambda\geqslant 0}e^{-\lambda t}\mathbb{E}\left(e^{\lambda\xi}\right)$$

简单推导：显然地，

$$\mathbb{P}\{\xi\geqslant t\}=\mathbb{P}\{\xi\lambda\geqslant t\lambda\}=\mathbb{P}\left\{e^{\xi\lambda}\geqslant e^{t\lambda}\right\}\leqslant e^{-\lambda t}\mathbb{E}\left(e^{\lambda\xi}\right)$$

其中 $\lambda>0$，最后一步是根据 Markov 不等式[1]

$$\mathbb{P}\{\xi\geqslant t\}\leqslant\frac{\mathbb{E}(\xi)}{t},\xi\geqslant 0$$

容易改写为

$$\mathbb{P}\{\xi\geqslant t\}\leqslant\inf_{\lambda\geqslant 0}e^{-\lambda t}\mathbb{E}\left(e^{\lambda\xi}\right)$$

这个不等式右端含有矩母函数，这是一个很好的不等式，也是一个在找“尾概率”时常用的技巧.

特别地，当 $\xi\sim N(\mu,\sigma^2)$，不等式的形式就变成

$$\mathbb{P}\{\xi\geqslant t\}\leqslant\inf_{\lambda\geqslant 0}e^{-\lambda t}e^{\mu t+\frac{\sigma^2}{2}t^2}=\inf_{\lambda\geqslant 0}e^{(\mu-\lambda)t+\frac{\sigma^2}{2}t^2}$$

[1]Markov 不等式的证明：对非负随机变量 ξ，

$$t\mathbb{P}\{\xi\geqslant t\}=\int_t^{\infty}t\mathrm{d}F(x)\leqslant\int_t^{\infty}x\mathrm{d}F(x)\leqslant\int_0^{\infty}x\mathrm{d}F(x)=\mathbb{E}(x)$$

定理 1.3.6. 当矩母函数存在时，矩母函数**唯一地**决定分布.

注：但是矩母函数不一定存在

定义 1.3.7. 若随机变量 X 的分布函数为 $F_X(x)$，则称 e^{itX} 的期望为 X 的**特征函数**，即

$$\psi_X(t)=\mathbb{E}\left(e^{itX}\right)=\int_\Omega e^{itX(\omega)}\mathbb{P}(\mathrm{d}\omega)=\int_{-\infty}^{+\infty}e^{itx}\mathrm{d}F_X(x)$$

注记 1.3.8. 如果 $F_X(x)$ 有密度 $f(x)$，则特征函数 $\psi_X(t)$ 就是 $f(x)$ 的 Fourier 变换：

$$\psi_X(t)=\int_{-\infty}^{+\infty}e^{itx}f(x)\,\mathrm{d}x$$

定理 1.3.9. 特征函数的性质：

1. 有界性：$|\psi(t)|\leqslant 1=\psi(0)$；
2. 共轭对称性：$\psi(-t)=\overline{\psi(t)}$；
3. 一致连续性：$|\psi(t+h)-\psi(t)|\leqslant\int_{-\infty}^{+\infty}\left|e^{ihx}-1\right|\mathrm{d}F(x)$
4. 设 $Y=a+bX$，则 Y 的特征函数是 $\psi_Y(t)=e^{ibt}\psi_X(at)$；
5. 两个相互独立的随机变量之和的特征函数等于它们特征函数之积；
6. 非负定性：对于 $\forall n\in\mathbb{N}^+,\forall t_1,t_2,\cdots,t_n\in\mathbb{R},\forall\lambda_1,\lambda_2,\cdots,\lambda_n\in\mathbb{C}$，有
$$\sum_{k=1}^{n}\sum_{j=1}^{n}\psi(t_k-t_j)\lambda_k\bar{\lambda}_j\geqslant 0$$
7. 设随机变量 X 的 n 阶矩存在，则它的特征函数可微分 n 次，且当 $k\leqslant n$ 时，有
$$\psi^{(k)}(0)=i^k\mathbb{E}\left(X^k\right)$$

定理 1.3.10. （唯一性定理）特征函数与分布函数**相互唯一**确定.

定理 1.3.11. 若 $\boldsymbol{X}\sim N(\mathbf{0},\boldsymbol{I}_d)$，则 $\boldsymbol{X}$ 的任一线性组合 $\boldsymbol{Y}=\boldsymbol{A}_{n\times d}\boldsymbol{X}+\boldsymbol{\mu}$ 服从 n 维正态分布 $N\left(\boldsymbol{\mu},\boldsymbol{A}\boldsymbol{A}^T\right)$.

定理 1.3.12. 若 $\boldsymbol{Y}\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$，则

$$\boldsymbol{A}\boldsymbol{Y}+\boldsymbol{b}\sim N\left(\boldsymbol{A}\boldsymbol{\mu}+\boldsymbol{b},\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T\right)$$

1.4 收敛性

定义 1.4.1. 设 $\{X_n,n\geqslant 1\}$ 为随机变量序列，称随机变量序列 $\{X_n,n\geqslant 1\}$**几乎必然收敛（或以概率 1 收敛）**于 X，记为 $X_n\to X$, a.s. 或 $X_n\xrightarrow{\text{a.s.}}X$，若存在随机变量 X 使得

$$\mathbb{P}\left\{\omega\in\Omega:X(\omega)=\lim_{n\to\infty}X_n(\omega)\right\}=1$$

定义 1.4.2. 设 $\{X_n, n \geqslant 1\}$ 为随机变量序列，称随机变量序列 $\{X_n, n \geqslant 1\}$**依概率收敛**于 X，记为 $X_n \xrightarrow{\mathbb{P}} X$，若存在随机变量 X 使得 $\forall \varepsilon > 0$，有

$$\lim_{n\to\infty} \mathbb{P}\{|X_n - X| \geqslant \varepsilon\} = 0$$

定义 1.4.3. 设随机变量序列 $\{X_n, n \geqslant 1\} \subset \mathcal{L}^p$, $X \in \mathcal{L}^p$, $p \geqslant 1$, 称随机变量序列 $\{X_n, n \geqslant 1\}$ **p 次平均收敛**于 X，或称 $\{X_n\}$ 在 $\mathcal{L}^p$ 中强收敛于 X，若有

$$\lim_{n\to\infty} \mathbb{E}\left(|X_n - X|^p\right) = 0$$

当 $p = 2$ 时，称为**均方收敛**.

定义 1.4.4. 设 $\{F_n(x)\}$ 是分布函数列，如果存在一个单调不减函数 $F(x)$ 使得在 $F(x)$ 的所有**连续点** x 上，均有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

则称 $\{F_n(x)\}$**弱收敛**于 $F(x)$，记为 $F_n(x) \xrightarrow{\mathrm{W}} F(x)$.

记随机变量 X_n, X 的分布函数分别为 $F_n(x)$,$F(x)$,如果 $F_n(x) \xrightarrow{\mathrm{W}} F(x)$,那么称 $\{X_n, n \geqslant 1\}$ **依分布收敛**于 X，记为 $X_n \xrightarrow{\mathrm{L}} X$.

定理 1.4.5. 收敛性的两条定理：

1. 随机变量序列 $X_n \xrightarrow{\text{a.s.}} X$ 的充分必要条件是 $\forall \varepsilon > 0$，有

$$\mathbb{P}\left\{\limsup_{n\to\infty} \left(|X_n - X| \geqslant \varepsilon\right)\right\} = 0$$

2. 随机变量序列 $X_n \xrightarrow{\mathbb{P}} X$ 的充分必要条件是 $\{X_n\}$ 的任意子序列都包含几乎必然收敛于 X 的子序列.

注记 1.4.6. 收敛性的关系：

1. 几乎必然收敛 $\Longrightarrow$ 依概率收敛 $\Longrightarrow$ 依分布收敛；
2. p 次平均收敛 $\Longrightarrow$ 依概率收敛 $\Longrightarrow$ 依分布收敛；
3. 几乎必然收敛与 p 次平均收敛之间没有蕴含关系；
4. 若 X 为质点（Point mass）分布,[2]则依分布收敛 $\Longrightarrow$ 依概率收敛.

1.5 条件概率，条件期望和独立性

定义 1.5.1. 独立性的几种定义：

[2]即 $\mathbb{P}(X = a) = 1$，记作 $X \sim \delta_a$.

1. 设对于 n 个事件 $A_1, A_2, \cdots, A_n$，如果对任意的 $k\,(1 \leqslant k \leqslant n)$，以及任意的 $1 \leqslant i_1 < i_2 < \cdots < i_n \leqslant k$，有
$$\mathbb{P}\left(\bigcap_{j=1}^{m} A_{k_j}\right) = \prod_{j=1}^{m} \mathbb{P}\left(A_{k_j}\right)$$
则称 $A_1, A_2, \cdots, A_n$ 是**相互独立**的.
2. 设 $\{A_i, i \in I\}$ 是一族事件，若对 I 的任意有限子集 $\{i_1, i_2, \cdots, i_k\} \neq \varnothing$，有
$$\mathbb{P}\left(\bigcap_{j=1}^{m} A_{i_j}\right) = \prod_{j=1}^{m} \mathbb{P}\left(A_{i_j}\right)$$
则称 $\{A_i, i \in I\}$ 是**相互独立**的.
3. $\{\mathcal{A}_i, i \in I\}$ 是一族事件类，如果对 I 的任意有限子集 $\{i_1, i_2, \cdots, i_k\} \neq \varnothing$，任意 $A_{i_j} \in \mathcal{A}_{i_j}$ 有
$$\mathbb{P}\left(\bigcap_{j=1}^{m} A_{i_j}\right) = \prod_{j=1}^{m} \mathbb{P}\left(A_{i_j}\right)$$
则称 $\{\mathcal{A}_i, i \in I\}$ 是**独立事件类**.
4. 设 $\{X_i, i \in I\}$ 是 Ω 上的一族随机变量，如果 σ 代数族
$$\{\sigma\left(X_i\right), i \in I\}$$
是独立事件类，则称 $\{X_i, i \in I\}$**相互独立**.

定理 1.5.2. 独立性的性质：

1. 随机变量 $X_1, X_2, \cdots, X_n$ 独立的充分必要条件是
$$F\left(x_1, x_2, \cdots, x_n\right) = F_{X_1}\left(x_1\right) F_{X_2}\left(x_2\right) \cdots F_{X_n}\left(x_n\right)$$
2. 若随机变量 $X_1, X_2, \cdots, X_n \in \mathcal{L}^1$ 是独立的，则
$$\mathbb{E}\left[\prod_{k=1}^{n} X_k\right] = \prod_{k=1}^{n} \mathbb{E}\left[X_k\right]$$
3. 若随机变量 $X_1, X_2, \cdots, X_n \in \mathcal{L}^2$ 是独立的，则
$$\operatorname{Var}\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n} \operatorname{Var}\left[X_k\right]$$

定义 1.5.3. 对有界函数 $G(x)$ 和单调函数 $F(x)$，可以定义**卷积**
$$F * G(x) = \int_{-\infty}^{+\infty} G(x-t)\,\mathrm{d}F(t)$$

定理 1.5.4. 卷积的运算性质：

1. 交换律：

$$F_1 * F_2(x) = F_2 * F_1(x)$$

2. 结合律：

$$(F * G) * H(x) = F * (G * H)(x)$$

3. 分配律：

$$F * (G + H)(x) = F * G(x) + F * H(x)$$

定义 1.5.5. 设 $X_k\,(k = 1, 2, \cdots, n)$ 是独立同分布于 F 的随机变量，令 $S_0 = 0, S_n = \sum_{k=1}^{n} X_k, n = 1, 2\cdots$，则称 S_n 的分布 F_n 为 n **重卷积**，它满足

$$F_n(x) = F * F_{n-1}(x), n = 1, 2, \cdots$$

其中 $F_0(x) = \begin{cases} 0 & , x < 0 \\ 1 & , x \geqslant 0 \end{cases}$.

定义 1.5.6. 设事件 B 发生的概率 $\mathbb{P}(B) > 0$，则事件 B 发生的条件下事件 A 发生的**条件概率**为

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

定义 1.5.7. （全概率公式）设 $\{B_n\}$ 是 Ω 的一个分割，且 $\forall n, \mathbb{P}(B_n) > 0$，如果 $A \in \mathcal{F}$，则

$$\mathbb{P}(A) = \mathbb{P}(A(B_1 \cup B_2 \cup \cdots \cup B_n)) = \sum_{i=1}^{n} \mathbb{P}(A|\,B_i)\,\mathbb{P}(B_i)$$

定理 1.5.8. （Bayes 公式）设 $\{B_n\}$ 是 Ω 的一个分割，且 $\forall n, \mathbb{P}(B_n) > 0$，如果 $\mathbb{P}(A) > 0$，则

$$\mathbb{P}(B_k|\,A) = \frac{\mathbb{P}(B_k)\,\mathbb{P}(A|\,B_k)}{\sum_n \mathbb{P}(B_n)\,\mathbb{P}(A|\,B_n)}, n \geqslant 1$$

定理 1.5.9. 条件期望的重要性质：

1. 迭代期望律：

$$\mathbb{E}_{\mathcal{G}}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$$

2. 若 X 是 $\mathcal{G}$-可测的，则

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X], \text{ a.s.}$$

3. 若 a, b 为实数，$X, Y, aX + bY$ 的期望均存在，则

$$\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}], \text{ a.s.}$$

4. 设 X, XY 的期望均存在，且 Y为$\mathcal{G}$-可测的，则

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}], \text{ a.s.}$$

5. 若 X 与 $\mathcal{G}$ 相互独立（即 $\sigma(X)$ 与 $\mathcal{G}$ 相互独立），则

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X], \text{ a.s.}$$

6. 若 $\mathcal{G}_1, \mathcal{G}_2$ 是两个子 σ 代数，使得 $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$，则

$$\mathbb{E}[\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1], \text{ a.s.}$$

也叫 σ 代数的"小吃大原则".

7. 设 $\mathcal{G} = \{\varnothing, \Omega\}$，则

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X], \text{ a.s.}$$

8. 若 $X \leqslant Y$, a.s.，则

$$\mathbb{E}[X|\mathcal{G}] \leqslant \mathbb{E}[Y|\mathcal{G}], \text{ a.s.}$$

9. 设 $X^+ = \max(X, 0)$，$X^- = \max\{0, -X\}$，则

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}\left[X^+\middle|\mathcal{G}\right] - \mathbb{E}\left[X^-\middle|\mathcal{G}\right], \text{ a.s.}$$

10. 绝对值不等式：

$$|\mathbb{E}[X|\mathcal{G}]| \leqslant \mathbb{E}[|X||\mathcal{G}], \text{ a.s.}$$

11. 设 $0 \leqslant X_n \uparrow X$, a.s.，则

$$\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}], \text{ a.s.}$$

12. 若 X, Y 相互独立，函数 $g(x, y)$ 使得 $\mathbb{E}[|g(X, Y)|] < +\infty$，则有

$$\mathbb{E}[g(X, Y)|Y] = \mathbb{E}[g(X, y)]|_{y=Y}, \text{ a.s.}$$

注：这里 $\mathbb{E}[g(X, y)]|_{y=Y}$ 的含义是，先将 y 视为常数，求得期望 $\mathbb{E}[g(X, y)]$ 后，再将随机变量 Y 代入到 y 的位置.

定义 1.5.10. （多变量的条件密度）设 $f(x_1, x_2, \cdots, x_n)$ 是随机变量 $X_1, X_2, \cdots, X_d$ 的联合密度函数，则 $X_1, X_2, \cdots, X_k$ 在给定 $X_{k+1}, X_{k+2}, \cdots, X_d$ 时的**条件密度**定义为

$$\begin{aligned}
&f_{1,2,\cdots,k}(u_1, u_2, \cdots, u_k| x_{k+1}, x_{k+2}, \cdots, x_d) \\
=& \frac{f(u_1, u_2, \cdots, u_k, x_{k+1}, x_{k+2}, \cdots, x_d)}{\int_{\mathbb{R}}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} f(y_1, y_2, \cdots, y_k, x_{k+1}, x_{k+2}, \cdots, x_d)\,\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n}
\end{aligned}$$

1.6 练习

练习 1.1. 一矿工被困在有三个门的矿井中，第一个门通一隧道，沿此隧道走 2 小时可到达安全区；第二个门通一隧道，沿此隧道走 3 小时可回到原矿井中；第三个门通一隧道，沿此隧道走 5 小时可回到原矿井中. 假定此矿工总是等可能地在三个门中选择一个，用 X 表示矿工到达安全区所用的时间，求 X 的均值.

Solution. 记 Y_i 为事件"矿工第一次选择了第 i 个门"，则

$$\mathbb{E}\left[X\middle|\,\sigma\left(Y_1,Y_2,Y_3\right)\right]=\sum_{i=1}^{3}\mathbb{E}\left[X\middle|\,Y_i\right]\mathbb{I}_{\{Y=i\}}$$

两边同时取期望，得

$$\mathbb{E}X=\mathbb{E}\left[\mathbb{E}\left[X\middle|\,\sigma\left(Y_1,Y_2,Y_3\right)\right]\right]=\sum_{i=1}^{3}\mathbb{E}\left[X\middle|\,Y_i\right]\mathbb{P}\left(Y_i\right)$$

即

$$\begin{aligned}\mathbb{E}X&=\sum_{i=1}^{3}\mathbb{E}\left[X\middle|\,Y_i\right]\mathbb{P}\left(Y_i\right)\\&=\frac{1}{3}\left(\mathbb{E}\left[X\middle|\,Y_1\right]+\mathbb{E}\left[X\middle|\,Y_2\right]+\mathbb{E}\left[X\middle|\,Y_3\right]\right)\\&=\frac{1}{3}\left(2+3+\mathbb{E}X+5+\mathbb{E}X\right)\end{aligned}$$

所以

$$\mathbb{E}X=10$$

□

练习 1.2. 不断投掷一个质地均匀的硬币，得到正面或反面的概率均为 $\frac{1}{2}$. 当连续两次投掷都得到正面时，停止投掷. 记投掷的次数为 N，求 N 的均值.

Solution. 当完成两次投掷时，有四种情况：正正、正反、反正、反反.

这四个事件是互不相交的，且构成对样本空间的一个划分. 但注意到，当我们得到"反正"时，并不能让我们回到最初的境况（最初我们没有获得正面，可以类比矿工回到最初的位置）.

于是考虑划分：正正、正反、反 (?)

进而，类似地，我们有

$$\mathbb{E}N=\frac{1}{4}\left[2+\left(2+\mathbb{E}N\right)\right]+\frac{1}{2}\left[1+\mathbb{E}N\right]$$

所以

$$\mathbb{E}N=6$$

□

练习 1.3. Let $X_1, X_2, \cdots$ be a sequence of random variables such that

$$\mathbb{P}\left(X_n = \frac{1}{n}\right) = 1 - \frac{1}{n^2} \quad \text{and} \quad \mathbb{P}\left(X_n = n\right) = \frac{1}{n^2}$$

Does X_n converge in probability? Does X_n converge in quadratic mean?

Solution. First, check $\forall \varepsilon > 0, \exists n > \mathbb{N}$, such that $n > \frac{1}{\varepsilon}$

$$\mathbb{P}\left(|X_n - 0| > \varepsilon\right) < \mathbb{P}\left(|X_n - 0| > \frac{1}{n}\right) = \frac{1}{n^2} \to 0$$

Thus, X_n converges to 0 in probability.

Second,

$$\mathbb{E}\left[|X_n - 0|^2\right] = \mathbb{E}X_n^2 = \frac{1}{n^2}\left(1 - \frac{1}{n^2}\right) + n^2 \cdot \frac{1}{n^2} = 1 + \frac{1}{n^2}\left(1 - \frac{1}{n^2}\right) \to 1$$

Thus, X_n does not converge to 0 in quadratic mean. □

练习 1.4. Let $\lambda_n = \frac{1}{n}$ for $n = 1, 2, \cdots$. Let $X_n \sim \text{Poisson}(\lambda_n)$.

1. Show that $X_n \xrightarrow{\mathbb{P}} 0$.
2. Show that $Y_n := nX_n \xrightarrow{\mathbb{P}} 0$.

Proof. For X_n, check

$$\mathbb{P}\left(|X_n - 0| > \varepsilon\right) = \mathbb{P}\left(X_n > \varepsilon\right) \leqslant \mathbb{P}\left(X_n > 0\right) = \mathbb{P}\left(X_n \neq 0\right) = 1 - e^{-\frac{1}{n}} \to 0$$

$$\mathbb{P}\left(|nX_n - 0| > \varepsilon\right) = \mathbb{P}\left(X_n > \frac{\varepsilon}{n}\right) \leqslant \mathbb{P}\left(X_n > 0\right) = \mathbb{P}\left(X_n \neq 0\right) = 1 - e^{-\frac{1}{n}} \to 0$$

Then we complete the proof.

Note that we can also compute

$$\mathbb{E}\left[|X_n - 0|^2\right] = \mathbb{E}X_n^2 = \text{Var}\left(X_n\right) - \left(\mathbb{E}X_n\right)^2 = \frac{1}{n} - \frac{1}{n^2} \to 0$$

then X_n converges to 0 in quadratic mean and thus converges to 0 in probability.

For $Y_n = nX_n$, we may check

$$\mathbb{E}(Y_n^2) = \text{Var}(Y_n) + \mathbb{E}(Y_n)^2 = n^2\lambda_n + n^2\lambda_n^2 \nrightarrow 0$$

but we do not have a straightforward quadratic mean convergence on Y_n.

We do not have a straightforward $\mathcal{L}_1$ convergence either:

$$\mathbb{E}(|Y_n|) = \mathbb{E}(Y_n) = n\lambda_n = 1$$

□

第二章　随机过程的基本概念和基本类型

本章主要介绍随机过程的基本定义和概念，建立一个一般性、可供分析的框架。实际中应用较多的内容集中在后续章节的具有特殊性质的随机过程上，因此本章内容只需了解即可.

2.1　基本概念

注记 2.1.1. 什么是随机过程?

1. 随机过程是概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的一族随机变量，$\{X(t), t \in T\}$，其中 t 是属于参数集 T 的参数.
2. 随机过程是概率论和微积分的整合，随机过程 $\{X(t, \omega), t \in T, \omega \in \Omega\}$ 是定义在 $T \times \Omega$ 上的二元函数.
   - $X(t_0, \omega)$ 是随机变量 X，对应概率论；
   - $X(t, \omega_0)$ 是一元函数 $f(t)$，对应微积分.

2.2　有限维分布与 Kolmogorov 定理

定义 2.2.1. 对任意有限个 $t_1, t_2, \cdots, t_n \in T$，**随机过程的 n 维分布**定义为

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

定义 2.2.2. 随机过程的**所有**的一维、二维、$\cdots$、n 维分布等等的全体

$$\{F_{t_1,t_2,\cdots,t_n}(x_1, x_2, \cdots, x_n), t_1, t_2, \cdots, t_n \in T, n \geqslant 1\}$$

称为**随机过程 $\{X(t), t \in T\}$ 的有限分布族**.

定理 2.2.3. 分布族的性质:

1. 对称性：对 $(1,2,\cdots,n)$ 的任一排列 $(j_1,j_2,\cdots,j_n)$ 有

$$\begin{aligned}&F_{t_{j_1},t_{j_2},\cdots,t_{j_n}}\left(x_{j_1},x_{j_2},\cdots,x_{j_n}\right)\\=&\mathbb{P}\left\{X\left(t_{j_1}\right)\leqslant x_{j_1},X\left(t_{j_2}\right)\leqslant x_{j_2},\cdots,X\left(t_{j_n}\right)\leqslant x_{j_n}\right\}\\=&\mathbb{P}\left\{X\left(t_1\right)\leqslant x_1,X\left(t_2\right)\leqslant x_2,\cdots,X\left(t_n\right)\leqslant x_n\right\}\\=&F_{t_1,t_2,\cdots,t_n}\left(x_1,x_2,\cdots,x_n\right)\end{aligned}$$

2. 相容性：对 $m<n$，有

$$F_{t_1,t_2,\cdots,t_m,t_{m+1},\cdots,t_n}\left(x_1,x_2,\cdots,x_m,\infty,\infty,\cdots,\infty\right)=F_{t_1,t_2,\cdots,t_m}\left(x_1,x_2,\cdots,x_m\right)$$

定理 2.2.4. (Kolmogorov 定理) 分布函数族 $\left\{F_{t_1,t_2,\cdots,t_n}\left(x_1,x_2,\cdots,x_n\right),t_1,t_2,\cdots,t_n\in T,n\geqslant 1\right\}$ 如果满足对称性和相容性，则必存在一个随机过程 $\{X\left(t\right),t\in T\}$，使得这个分布函数族恰好是该随机过程的有限维分布族.

注记 2.2.5. Kolmogorov 定理说明，随机过程的有限维分布族是随机过程概率特征的完整描述，它是证明随机过程存在的有力工具，但是在刻画随机过程方面并不好用.

定义 2.2.6. 对于随机过程 $\{X\left(t\right),t\in T\}$，如果

$$\mu_X\left(t\right)=\mathbb{E}\left[X\left(t\right)\right]$$

存在，则称其为过程的**均值函数**.

定义 2.2.7. 对于随机过程 $\{X\left(t\right),t\in T\}$，如果 $\forall t\in T,\mathbb{E}\left[X^2\left(t\right)\right]$ 存在，则称随机过程 $\{X\left(t\right),t\in T\}$ 为**二阶矩过程**.

定义 2.2.8. 对二阶矩过程的几个定义.

1. 协方差函数：对于二阶矩过程 $\{X\left(t\right),t\in T\}$，称

$$\gamma\left(t_1,t_2\right)=\mathbb{E}\left[\left(X\left(t_1\right)-\mu_X\left(t_1\right)\right)\left(X\left(t_2\right)-\mu_X\left(t_2\right)\right)\right]$$

为过程的**协方差函数**.

2. 方差函数：对于二阶矩过程 $\{X\left(t\right),t\in T\}$，称

$$\operatorname{Var}\left[X\left(t\right)\right]=\gamma\left(t,t\right)$$

为过程的**方差函数**.

3. 自相关函数：对于二阶矩过程 $\{X\left(t\right),t\in T\}$，称

$$R_X\left(s,t\right)=\mathbb{E}\left[X\left(s\right)X\left(t\right)\right],s,t\in T$$

为过程的**自相关函数**.

注记 2.2.9. 由 Cauchy-Schwartz 不等式

$$[\mathbb{E}(XY)]^2 \leqslant \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)$$

知，二阶矩过程的协方差函数和自相关函数一定存在.

此外，协方差函数、均值函数、自相关函数满足：

$$\gamma_X(s,t) = R_X(s,t) - \mu_X(s)\,\mu_X(t)$$

2.3 随机过程的基本类型

定义 2.3.1. 如果随机过程 $\{X(t), t \in T\}$, 对**任意**的 $t_1, t_2, \cdots, t_n \in T$ 和**任意**的 h（使得 $t_i + h \in T$）有

$$(X(t_1+h), X(t_2+h), \cdots X(t_n+h))\, d\, (X(t_1), X(t_2), \cdots, X(t_n))$$

则称 $\{X(t), t \in T\}$ 是**严平稳的**.

注记 2.3.2. 严平稳意味着有限维分布关于时间平移不变.

定义 2.3.3. 宽平稳过程或二阶平稳过程：如果随机过程 $\{X(t), t \in T\}$,

1. 所有二阶矩都存在；
2. 有均值函数 $\mathbb{E}[X(t)] = \mu$；
3. 协方差函数 $\gamma(t,s)$ 只与时间差 $t - s$ 有关；

则称 $\{X(t), t \in T\}$ 为**宽平稳过程**或**二阶平稳过程**.

注记 2.3.4. 关于宽平稳过程的注意点：

1. $\gamma(s,t) = \gamma(0, t-s), t, s \in T$ 可以记为 $\gamma(t-s)$；
2. $\gamma(t)$ 是偶函数；
3. $\gamma(0) = \mathrm{Var}[X(t)]$；
4. $|\gamma(\tau)| \leqslant \gamma(0)$.

证明. 由 Cauchy-Schwartz 不等式，

$$|\mathbb{E}[X(t)X(s)]| \leqslant \sqrt{\mathbb{E}[X^2(t)]\,\mathbb{E}[X^2(s)]}$$

因此

$$|\gamma(t-s)| + \mu_X^2 \leqslant \sqrt{(\mathrm{Var}[X^2(t)] + \mu_X^2)(\mathrm{Var}[X^2(s) + \mu_X^2])} = \gamma(0) + \mu_X^2$$

令 $\tau = t - s$，即得证. □

注记 2.3.5. $\gamma(\tau)$ 具有非负定性，即对任意时刻 t_k 和实数 a_k，$k=1,2,\cdots,N$ 有

$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \gamma(t_i - t_j) \geqslant 0$$

写成矩阵形式，就是

$$\boldsymbol{a}^{\mathrm{T}}\,\boldsymbol{a} \geqslant 0$$

或者

$$(a_1, a_2, \cdots, a_n)\begin{pmatrix} \gamma(t_1-t_1) & \gamma(t_2-t_1) & \cdots & \gamma(t_1-t_n) \\ \gamma(t_2-t_1) & \gamma(t_2-t_2) & \cdots & \gamma(t_2-t_n) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(t_n-t_1) & \gamma(t_n-t_2) & \cdots & \gamma(t_n-t_n) \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \geqslant 0$$

注记 2.3.6. 严平稳过程与宽平稳过程的关系：

1. 严平稳过程不一定是宽平稳过程，因为不能保证二阶矩存在；
2. 宽平稳过程也不一定是严平稳过程，因为宽平稳过程只能保证一阶矩和二阶矩不随时间推移而改变，却不能保证有限维分布随时间推移不变.

例子 2.3.7. 比较两个特殊的随机过程：

1. $\{X_n, n \geqslant 0\}$ 为独立同分布的随机变量序列，$\mathbb{E}[X_n^2] < \infty$，$\mathbb{E}[X_n] = \mu$，$n = 0,1,2,\cdots$，由大数定律知
$$\frac{1}{n}(X_0 + X_1 + \cdots + X_{n-1}) \xrightarrow{\mathbb{P}} \mu$$
2. 设随机过程 $\{Y_n = Y, n \geqslant 0\}$,其中 Y 是随机变量,$\mathbb{E}[Y^2] < \infty$,那么 $\frac{1}{n}(Y_0 + Y_1 + \cdots + Y_{n-1}) = Y$ 这里对时间平均后，随机性并没有发生变化. [1]

注记 2.3.8. 一次观察获得一条样本路径，可以估计均值

$$\hat{\mu} = \frac{1}{n}(X_1(t) + X_2(t) + \cdots + X_n(t))$$

也可以估计协方差

$$\hat{\gamma}(\tau) = \frac{1}{n}\sum_{k=1}^{n}[[X_k(t+\tau) - \hat{\mu}][X_k(t) - \hat{\mu}]]$$

如果需要对随机过程通过**一次观察**就能观察到 μ 和 $\gamma(\tau)$，需要平稳过程加上遍历性定理.

定义 2.3.9. 设 $\{X(t), -\infty < t < +\infty\}$ 是平稳过程，如果

$$\bar{X} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} X(t)\,\mathrm{d}t = \mu$$

[1] 主要原因在于 Y_n 之间不再独立，很明显 $\mathrm{Cov}(Y_i, Y_j) = Var(Y)$，$Y_i$和$Y_j$ 一般是相关的.

或当参数空间为 $T=\mathbb{Z}$ 时，若

$$\bar{X}=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{k=-N}^{N}X(t)=\mu$$

则称过程 $\{X(t),t\in T\}$ 的**均值具有遍历性**.

注记 2.3.10. 均方意义下的极限：

$$\lim_{T\to\infty}\mathbb{E}\left[\left|\frac{1}{2T}\int_{-T}^{T}X(t)\,\mathrm{d}t-\mu\right|^2\right]=0$$

注意辨析：

1. 数列的极限 $\lim_{n\to\infty}a_n=a$，用绝对距离 $d=|a_n-a|$ 来刻画收敛；
2. 随机变量之间的收敛 $\mathbb{P}\{\omega|\,\xi(\omega)-\eta(\omega)\neq 0\}=0$，不用 $d=|\xi-\eta|$ 刻画，因为随机变量之差还是随机变量，讨论“两个随机变量随机地靠近”是没有意义的，而应该用均方距离 $d(\xi,\eta)=\mathbb{E}\left(|\xi-\eta|^2\right)$ 来刻画.

定义 2.3.11. 设 $\{X(t),-\infty<t<+\infty\}$ 是平稳过程，如果

$$\bar{\gamma}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}[X(t)-\mu]\,[X(t+\tau)-\mu]\,\mathrm{d}t=\gamma(\tau)$$

或当参数空间为 $T=\mathbb{Z}$ 时，若

$$\bar{\gamma}(\tau)=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{k=-N}^{N}[X(t)-\mu]\,[X(t+\tau)-\mu]=\gamma(\tau)$$

则称过程 $\{X(t),t\in T\}$ 的**协方差具有遍历性**.

定义 2.3.12. 随机过程的均值和协方差函数都具有遍历性，则称此**随机过程具有遍历性**.

定义 2.3.13. 三个概念.

1. 平稳增量过程的定义：设随机过程 $\{X(t),t\in T\}$，如果对任意的 t_1,t_2，有

$$X(t_1+h)-X(t_1)\overset{d}{=}X(t_2+h)-X(t_2)$$

   则称 $\{X(t),t\in T\}$ 为**平稳增量过程**.
2. 独立增量过程的定义：如果对任意 $t_1,t_2,\cdots,t_n\in T$，$t_1<t_2<\cdots<t_n$，随机变量 $X(t_2)-X(t_1),X(t_3)-X(t_2),\cdots,X(t_n)-X(t_{n-1})$ 是相互独立的，则称 $\{X(t),t\in T\}$ 为**独立增量过程**.
3. 平稳独立增量过程的定义：兼有独立增量和平稳增量的过程称为**平稳独立增量过程**.

注记 2.3.14. 平稳独立增量过程的均值函数 $\mu(t)$、方差函数 $\sigma^2(t)$、自相关函数 $R(s,t)$ 均是线性函数.

证明. 设随机过程 $\{X(t), t \geqslant 0\}$ 是平稳独立增量过程，不失一般性地，设 $X(0)=0$，则

$$
\begin{aligned}
\mu(t+s) &= \mathbb{E}[X(t+s)] \\
&= \mathbb{E}[X(t+s)-X(s)+X(s)-X(0)] \\
&= \mathbb{E}[X(s+t)-X(s)]+\mathbb{E}[X(s)-X(0)] \\
&= \mu(t)+\mu(s)
\end{aligned}
$$

$$
\begin{aligned}
\sigma^2(t+s) &= \operatorname{Var}[X(t+s)] \\
&= \operatorname{Var}[X(t+s)-X(s)+X(s)-X(0)] \\
&= \operatorname{Var}[X(t+s)-X(s)]+\operatorname{Var}[X(s)-X(0)] \\
&= \operatorname{Var}[X(t)]+\operatorname{Var}[X(s)] \\
&= \sigma^2(t)+\sigma^2(s)
\end{aligned}
$$

由于 $R(s,t)=\mathbb{E}[X(s)X(t)]$，由于 s 和 t 的地位对称，可先固定 s，即证明 $R(s_0,t)$ 是关于 t 的线性函数.

$$
\begin{aligned}
R(s_0,t+h) &= \mathbb{E}[X(s_0)X(t+h)] \\
&= X(s_0)\mathbb{E}[X(t+h)] \\
&= X(s_0)\mu(t+h) \\
&= X(s_0)\mu(t)+X(s_0)\mu(h) \\
&= R(s_0,t)+R(s_0,h)
\end{aligned}
$$

□

注记 2.3.15. 以下简单提出遍历性定理的内容，这些内容不需要掌握，仅仅知道有这个概念即可，因为今后能处理的过程都满足遍历性定理，如果不满足遍历性定理，也没有可靠的工具来处理.

时间不能倒流，在现实中，只能根据一条样本轨道估计过程的性质. 如果某过程不满足遍历性定理，那么根据一个样本估计总体，估计的结果一定是不理想的.

定理 2.3.16. (均值遍历性定理)

1. 设 $\{X_t, -\infty<t<+\infty\}$ 是平稳过程，其均值函数为 μ，协方差函数是 $\gamma(\tau)$，则该过程的均值有遍历性的**充分必要条件**是

$$
\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)\gamma(\tau)\mathrm{d}\tau=0
$$

2. 设 $\{X_n, n=0,\pm1,\pm2,\cdots\}$ 是平稳过程，其均值函数为 μ，协方差函数是 $\gamma(\tau)$，则该过程的均值有遍历性的**充分必要条件**是

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\tau=0}^{N-1}\gamma(\tau)=0$$

定理 2.3.17. 两个推论.

1. 若
$$\int_{-\infty}^{+\infty}|\gamma(\tau)|\,\mathrm{d}\tau<\infty$$
则均值遍历性定理成立.
2. 对于平稳序列而言，如果
$$\lim_{\tau\to\infty}\gamma(\tau)=0$$
则均值遍历性定理成立.

定理 2.3.18. (协方差函数遍历性定理) 设 $\{X_t,-\infty<t<+\infty\}$ 是平稳过程，其均值函数为零，则其协方差函数具有遍历性的**充分必要条件**是

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau_1}{2T}\right)\left(B(\tau_1)-\gamma^2(\tau)\right)\mathrm{d}\tau_1=0$$

其中 $B(\tau_1)=\mathbb{E}\left[X(t+\tau_1+\tau)X(t+\tau_1)X(t+\tau)X(t)\right]$.

注记 2.3.19. 对协方差函数遍历性的考察可以转换为对均值函数遍历性的考察：对平稳过程 $\{X(t),t\in T\}$ 的协方差函数的遍历性定理，可以考虑 $\{Y_\tau(t),t\in T\}$ 的均值遍历性，其中

$$Y_\tau(t)=[X(t+\tau)-m][X(t)-m]$$

注记 2.3.20. 两个等价关系：

1. 均值函数具有遍历性 $\Longleftrightarrow \mathrm{Var}\left(\bar{X}\right)=0$;
2. 协方差函数具有遍历性 $\Longleftrightarrow \mathrm{Var}\left[\bar{\gamma}(\tau)\right]=0$.

2.4 练习

练习 2.1. 证明. □

第三章 Poisson 过程

本章需要掌握 Poisson 过程的定义、Poisson 过程的判定条件、与 Poisson 过程相联系的分布，并了解非齐次 Poisson 过程、复合 Poisson 过程和条件 Poisson 过程的定义、性质及应用.

3.1 Poisson 过程的概念

注记 3.1.1. 复习.

1. Poisson 分布：若 $\xi \sim \mathrm{Poi}\,(\lambda)$，则
$$\mathbb{P}\left\{\xi = n\right\} = \frac{\lambda^n e^{-\lambda}}{n!}, n = 0, 1, 2, \cdots$$
2. Poisson 分布的数字特征：$\mathbb{E}\,(\xi) = \lambda$，$\mathrm{Var}\,(\xi) = \lambda$.
3. Poisson 分布可以由二项分布近似：如果 $\eta \sim B\,(n, p)$，则
$$\lim_{n\to\infty} \mathbb{P}\left\{\eta = k\right\} = \frac{\lambda^k}{k!} e^{-\lambda}$$

证明.
$$\begin{aligned}
\mathbb{P}\left\{\eta = k\right\} &= \begin{pmatrix} n \\ k \end{pmatrix} p^k \left(1 - p\right)^{n-k} = \frac{n\,(n-1)\cdots(n+1-k)}{k!} p^k \left(1 - p\right)^{n-k} \\
&= \frac{n\,(n-1)\cdots(n+1-k)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad \left(\text{令} p = \frac{\lambda}{n}\right) \\
&= \frac{\lambda^k}{k!} \left[1 \cdot \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)\right] \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}
\end{aligned}$$

进而
$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left\{\eta = k\right\} &= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left[1 \cdot \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)\right] \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \\
&= \frac{\lambda^k}{k!} e^{-\lambda}
\end{aligned}$$

□

定义 3.1.2. 随机过程 $\{N(t), t \geqslant 0\}$ 称为**计数过程**，如果 $N(t)$ 表示从时刻 0 到 t 某一特定事件 A 发生的次数，并满足

1. $N(t) \geqslant 0$ 且 $N(t) \in \mathbb{Z}$;
2. $s < t$ 时，$N(t) \geqslant N(s)$，且 $N(t) - N(s)$ 表示在 $(s, t]$ 时间内事件 A 发生的次数.

定义 3.1.3. 计数过程 $\{N(t), t \geqslant 0\}$ 称为参数为 $\lambda\,(\lambda > 0)$ 的**Poisson 过程**，如果

1. $N(0) = 0$;
2. 过程有独立增量;
3. 对任意的 $s, t \geqslant 0$，有 $N(t+s) - N(s) \sim \mathrm{Poi}(\lambda t)$，即

$$\mathbb{P}\{N(t+s) - N(s)\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, n = 0, 1, 2, \cdots \tag{3.1}$$

定义 3.1.4. 由于 $\mathbb{E}[N(t)] = \lambda t$，可以认为 λ 是单位时间内发生的事件的平均次数，称 λ 为 Poisson 过程的**强度**或**速率**.（实际应用中注意时间单位）

定义 3.1.5. (Poisson 过程的等价定义)计数过程 $\{N(t), t \geqslant 0\}$ 称为参数为 $\lambda\,(\lambda > 0)$ 的**Poisson 过程**，如果

1. $N(0) = 0$;
2. 过程有平稳独立增量;
3. 存在 $\lambda > 0$，当 $h \downarrow 0$ 时，在 h 内事件发生 1 次的概率

$$\mathbb{P}\{N(t+h) - N(t) = 1\} = \lambda h + o(h)$$

4. 当 $h \downarrow 0$ 时，在 h 内事件发生 2 次以上的概率

$$\mathbb{P}\{N(t+h) - N(t) \geqslant 2\} = o(h)$$

注记 3.1.6. 等价定义的含义：把 $[0, t]$ 划分为 n 个相等的时间区间，则当 $n \to \infty$ 时，事件在每个小区间内发生 2 次以上的概率趋于 0，事件发生一次的概率 $p \approx \lambda\frac{t}{n}$，显然 p 很小，而事件不发生的概率为 $1 - p \approx 1 - \lambda\frac{t}{n}$，这恰好是一次 Bernoulli 试验.

定理 3.1.7. 定义 3.1.3 与定义 3.1.5 等价，即满足等价定义 3.1.5 的计数过程一定是 Poisson 过程（满足原始定义 3.1.3 ），反之，Poisson 过程一定满足等价定义 3.1.5.

证明. 由(3.1)易得原始定义 3.1.3 $\Longrightarrow$ 等价定义 3.1.5 .

下面证明等价定义 3.1.5 $\Longrightarrow$ 原始定义 3.1.3，记

$$P_n(t) = \mathbb{P}\{N(t) = n\}, n = 0, 1, 2, \cdots$$

$$P(h) = \mathbb{P}\{N(h) \geqslant 1\} = \sum_{n=1}^{\infty} P_n(h) = 1 - P_1(h)$$

由于事件 $\{N(t+h)=0\} \Longleftrightarrow \{N(t+h)-N(t)=0, N(t)=0\}$ 和独立增量性，有

$$\begin{aligned} P_0(t+h) &= \mathbb{P}\{N(t+h)=0\} = \mathbb{P}\{N(t+h)-N(t)=0, N(t)=0\} \\ &= \mathbb{P}\{N(t+h)-N(t)=0\}\mathbb{P}\{N(t)=0\} = P_0(h)P_0(t) \\ &= (1-\lambda h + o(h))P_0(t) \end{aligned}$$

所以

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

令 $h \to 0$ 可以得到微分方程

$$P'_0(t) = \lambda P_0(t), P_0(0) = 1$$

解之，得

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

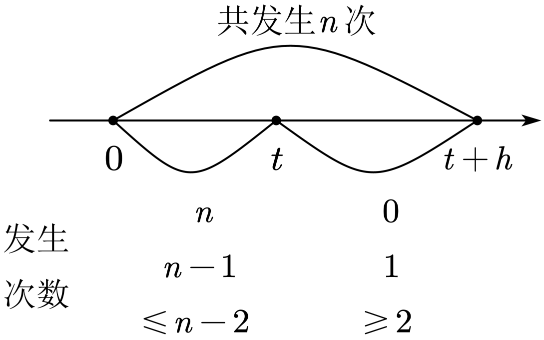

图 3.1. 分类讨论

当 $n \geqslant 1$ 时，应用全概率公式和独立增量性，可得

$$\begin{aligned} P_n(t+h) =& \mathbb{P}\{N(t+h)=n\} \\ =& \mathbb{P}\{N(t)=n, N(t+h)-N(t)=0\} \\ & + \mathbb{P}\{N(t)=n-1, N(t+h)-N(t)=1\} \\ & + \mathbb{P}\{N(t+h)=n, N(t+h)-N(t) \geqslant 2\} \\ =& P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + o(h) \\ =& P_n(t)(1-\lambda h + o(h)) + P_{n-1}(t)(\lambda h + o(h)) + o(h) \\ =& (1-\lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h) \end{aligned}$$

那么

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

令 $h \to 0$ 可以得到递推关系

$$P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t), P_0(t) = e^{-\lambda t}$$

用数学归纳法容易得到

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

$\square$

注记 3.1.8. Poisson 过程的一种变式：事件 A 的发生强度为 λ，但每次发生只能以概率 p 记录下来（发生和记录独立），以 $M(t)$ 表示到 t 时刻被记录下来的事件总数，则 $\{M(t), t \geqslant 0\}$ 是一个强度为 λp 的 Poisson 过程.

证明. 根据定义 3.1.3，由

$$\begin{aligned}
\mathbb{P}(M(t) = m) &= \sum_{n=0}^{\infty} \mathbb{P}(M(t) = m | N(t) = m+n) \cdot \mathbb{P}(M(t) = m | N(t) = m+n) \\
&= \sum_{n=0}^{\infty} \begin{pmatrix} m+n \\ m \end{pmatrix} p^m (1-p)^n \frac{(\lambda t)^{m+n}}{(m+n)!} e^{-\lambda t} \\
&= e^{-\lambda t} p^m (\lambda t)^m \sum_{n=0}^{\infty} \frac{(m+n)!}{m!n!} (1-p)^n \frac{(\lambda t)^n}{(m+n)!} \\
&= \frac{e^{-\lambda t} (\lambda p t)^m}{m!} \sum_{n=0}^{\infty} \frac{(\lambda (1-p) t)^n}{n!} \\
&= \frac{e^{-\lambda t} (\lambda p t)^m}{m!} e^{\lambda (1-p) t} \\
&= \frac{e^{-\lambda p t} (\lambda p t)^m}{m!}
\end{aligned}$$

容易验证 $\{M(t), t \geqslant 0\}$ 是一个强度为 λp 的 Poisson 过程. $\square$

3.2 与 Poisson 过程相联系的若干分布

定义 3.2.1. **经验分布函数**为

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{x_i \leqslant t\}$$

它的跃度是 $\frac{1}{n}$，其中示性函数 $\mathbb{I}_A(\omega) = \begin{cases} 0, & \omega \notin A \\ 1, & \omega \in A \end{cases}$.

注记 3.2.2. Poisson 过程的样本路径是跃度为 1 的阶梯函数：

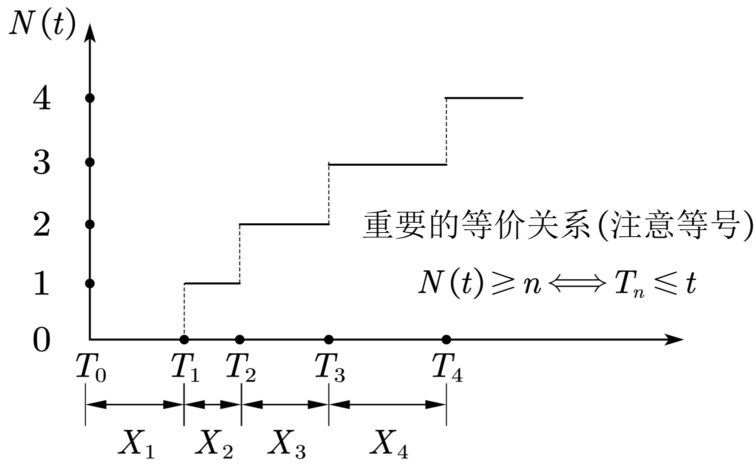

图 3.2. Poisson 过程的样本路径

定理 3.2.3. $X_n, n=1,2,\cdots$ 相互独立，且服从参数为 λ 的指数分布，概率密度是

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

均值 $\mathbb{E}[X_n]=\frac{1}{\lambda}$，方差 $\mathrm{Var}[X_n]=\frac{1}{\lambda^2}$.

证明. 考虑 X_1 的分布，

$$F_{X_1}(t)=\mathbb{P}\{X_1\leqslant t\}=1-\mathbb{P}\{X_1>t\}=1-\mathbb{P}\{N(t)=0\}=1-e^{-\lambda t}$$

再考虑 X_2 的分布，

$$\begin{aligned}\mathbb{P}(X_2>t|\,X_1=s)&=\mathbb{P}(N(s+t)-N(s)=0|\,X_1=s)\\&=\mathbb{P}(N(s+t)-N(s)=0)\\&=e^{-\lambda t}\end{aligned}$$

与 X_1 的取值无关，因此 X_1,X_2 相互独立. 类似地，可以归纳证明该定理. □

定理 3.2.4. $T_n=\sum_{i=1}^n X_i, n=1,2,\cdots$ 服从参数为 n 和 λ 的 Γ 分布，概率密度是

$$f(t)=\begin{cases}\frac{\lambda^n}{\Gamma(n)}x^{n-1}e^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}, n>0,\lambda>0$$

均值 $\mathbb{E}[T_n]=\frac{n}{\lambda}$，方差 $\mathrm{Var}[T_n]=\frac{n}{\lambda^2}$.

证明 1. 由 Γ 分布的可加性即可得证，这里再证明一次 Γ 分布的可加性. 可求 X_i 的矩母函数 $M(t):=\mathbb{E}\left(e^{tX_i}\right)=\frac{\lambda}{\lambda-t}$, 和参数为 n 和 λ 的 Γ 分布的矩母函数为 $M_n(t):=\left(\frac{\lambda}{\lambda-t}\right)^n$. 而

$$M_n(t)=\left(\frac{\lambda}{\lambda-t}\right)^n=\prod_{i=1}^n\mathbb{E}\left[e^{tX_i}\right]=\mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right]=\mathbb{E}\left[e^{tT_n}\right]$$

则 $T_n=\sum_{i=1}^{n}X_i, n=1,2,\cdots$ 服从参数为 n 和 λ 的 Γ 分布. □

证明 2. 因为 $N(t)\geqslant n \Longleftrightarrow T_n\leqslant t$，则

$$F_{T_n}(t)=\mathbb{P}\{T_n\leqslant t\}=\mathbb{P}\{N(t)\geqslant n\}=\sum_{j=n}^{\infty}e^{-\lambda t}\frac{(\lambda t)^j}{j!}$$

求导得

$$f_{T_n}(t)=-\sum_{j=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^j}{j!}+\sum_{j=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!}=\lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}=\frac{\lambda^n}{\Gamma(n)}t^{n-1}e^{-\lambda t}$$

□

定义 3.2.5. (Poisson 过程的又一等价定义) 计数过程 $\{N(t), t\geqslant 0\}$ 称为参数为 $\lambda\,(\lambda>0)$ 的 **Poisson 过程**，如果每次事件发生的时间间隔 $X_1, X_2, \cdots$ 相互独立，且服从参数为 λ 的指数分布.

定理 3.2.6. 已知 $[0,t]$ 内事件 A **只发生 1 次的条件下**，A 发生的时刻在 $[0,t]$ 上是均匀分布，即

$$\mathbb{P}\{T_1\leqslant s|\,N(t)=1\}=\frac{s}{t}$$

证明.

$$\begin{aligned}\mathbb{P}\{T_1\leqslant s|\,N(t)=1\}&=\frac{\mathbb{P}\{T_1\leqslant s, N(t)=1\}}{\mathbb{P}\{N(t)=1\}}=\frac{\mathbb{P}\{N(s)=1, N(t)-N(s)=0\}}{\mathbb{P}\{N(t)=1\}}\\&=\frac{\mathbb{P}\{N(s)=1\}\mathbb{P}\{N(t)-N(s)=0\}}{\mathbb{P}\{N(t)=1\}}\\&=\frac{st\cdot e^{-\lambda s}\times e^{-\lambda(t-s)}}{\lambda t\cdot e^{-\lambda t}}=\frac{s}{t}\end{aligned}$$

□

定理 3.2.7. 在 **已知 $N(t)=n$ 的条件下**，事件发生的 n 个时刻 $T_1, T_2, \cdots, T_n$ 的联合密度是 (均匀分布)

$$f(t_1,t_2,\cdots,t_n)=\frac{n!}{t^n}, 0<t_1<t_2<\cdots<t_n<t \tag{3.2}$$

证明. 设 $0<t_1<t_2<\cdots<t_n<t_{n+1}=t$，取 h_i 充分小，使得 $t_i+h_i<t_{i+1}$，其中 $i=1,2,\cdots,n$. 于是，有

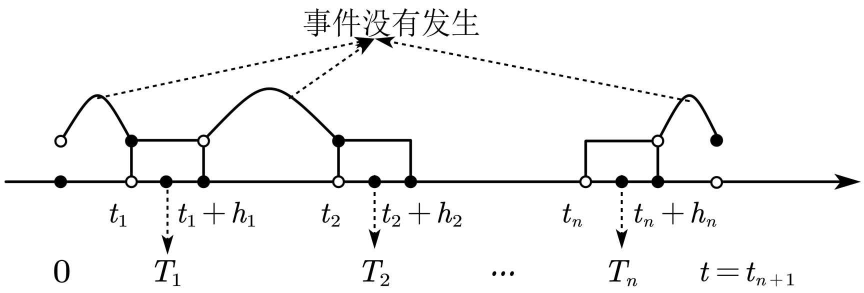

图 3.3. n 次事件发生时刻的条件分布

$$\begin{aligned}
&\mathbb{P}\left\{t_i < T_i \leqslant t_{i+1}, i=1,2,\cdots,n \middle| N(t)=n\right\}\\
=&\frac{\mathbb{P}\left\{N(t_i+h_i)-N(t_i)=1, N(t_{i+1})-N(t_i+h_i)=0, i=1,2,\cdots,n, N(t_1)=0\right\}}{\mathbb{P}\left\{N(t)=n\right\}}\\
=&\frac{\mathbb{P}\left\{N(t_1)=0\right\}\prod_{i=1}^n\left[\mathbb{P}\left\{N(t_i+h_i)-N(t_i)=1\right\}\mathbb{P}\left\{N(t_{i+1})-N(t_i+h_i)=0\right\}\right]}{\mathbb{P}\left\{N(t)=n\right\}}\\
=&\frac{\exp\left\{-\lambda\left(t-\sum_{i=1}^n h_i\right)\right\}\prod_{i=1}^n \lambda h_i e^{-\lambda h_i}}{\frac{e^{-\lambda t}(\lambda t)^n}{n!}}\\
=&\frac{n!}{t^n}\prod_{i=1}^n h_i
\end{aligned}$$

那么根据多元随机变量的密度函数定义，[1] 可得

$$\begin{aligned}
f(t_1,t_2,\cdots,t_n) &= \lim_{\substack{h_i\to 0\\ i=1,2,\cdots,n}} \frac{\mathbb{P}\left\{t_i \leqslant T_i \leqslant t_i+h_i, i=1,2,\cdots,n \middle| N(t)=n\right\}}{\prod_{i=1}^n h_i}\\
&= \frac{n!}{t^n}, 0<t_1<t_2<\cdots<t_n<t
\end{aligned}$$

□

注记 3.2.8. 由于分布(3.2)恰好是在 $[0,t]$ 区间上服从均匀分布的 n 个相互独立的随机变量 $Y_1,Y_2,\cdots,Y_n$ 的顺序统计量 $Y_{(1)},Y_{(2)},\cdots,Y_{(n)}$ 的联合分布. 所以，在已知 $[0,t]$ 内发生了 n 次事件的前提下，各次事件发生的时刻 $T_1,T_2,\cdots,T_n$（不排序）可以看作是相互独立的、且服从 $[0,t]$ 上均匀分布的随机变量.

定理 3.2.9. (时变的 Poisson 过程) 如果事件 A 在 s 时刻发生被记录到的概率为 $P(s)$（发生和记录独立），以 $M(t)$ 表示在 t 时刻被记录到的事件数. 这样，$M(t)$ 不再具有平稳增量，

[1]简单地，从二元随机变量的情形来看，由累积分布函数推得概率密度函数的过程是求导，推导这个分布的时候，只不过将求导写成了求极限的过程，就如

$$f(x,y)=\lim_{\Delta x\to 0,\Delta y\to 0}\frac{F(x+\Delta x,y+\Delta y)}{\Delta x\cdot\Delta y}=\lim_{\Delta x\to 0,\Delta y\to 0}\frac{P\left\{x\leqslant X\leqslant x+\Delta x, y\leqslant Y\leqslant y+\Delta y\right\}}{\Delta x\cdot\Delta y}$$

故它不再是 Poisson 过程，但是在任意给定 t 的情况下，$M(t)$ 仍服从 Poisson 分布，参数是 λtp，其中

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

证明. 将某次事件发生"记录下来"看作一次"Bernoulli 实验成功"，那么

$$\mathbb{P}(M(t)=m) = \sum_{k=0}^{\infty} \mathbb{P}\{M(t)=m|\,N(t)=m+k\}\,\mathbb{P}\{N(t)=m+k\}$$

其中

$$\mathbb{P}\{M(t)=m|\,N(t)=m+k\} = \begin{pmatrix} m+k \\ m \end{pmatrix} p^m(1-p)^k$$

并且

$$\begin{aligned} p &= \mathbb{P}\{A\text{在}\,[0,t]\,\text{发生且被记录}|\,N(t)=m+k\}\ (\text{事件互不相容}) \\ &= \int_0^t \mathbb{P}\{A\text{在}s\text{时刻发生且被记录}|\,N(t)=m+k\}\,\mathrm{d}s\ (\text{发生与记录独立}) \\ &= \int_0^t \mathbb{P}\{A\text{在}s\text{时刻发生}|\,N(t)=m+k\}\cdot\mathbb{P}\{A\text{在}s\text{时刻被记录}\}\,\mathrm{d}s \\ &= \int_0^t \frac{1}{t}P(s)\,\mathrm{d}s \end{aligned}$$

所以

$$\begin{aligned} \mathbb{P}(M(t)=m) &= \sum_{k=0}^{\infty} \begin{pmatrix} m+k \\ m \end{pmatrix} p^m(1-p)^k \frac{(\lambda t)^{m+k}e^{-\lambda t}}{(m+k)!} \\ &= \sum_{k=0}^{\infty} \frac{(m+k)!}{m!\cdot k!} p^m(1-p)^k \frac{(\lambda t)^{m+k}e^{-\lambda t}}{(m+k)!} \\ &= \frac{(\lambda tp)^m}{m!}e^{-\lambda t}\sum_{k=0}^{\infty}\frac{1}{k!}(1-p)^k(\lambda t)^k \\ &= \frac{(\lambda tp)^m}{m!}e^{-\lambda t}\cdot e^{\lambda t(1-p)} = \frac{(\lambda tp)^m}{m!}e^{-\lambda tp} \end{aligned}$$

其中 $p = \int_0^t \frac{1}{t}P(s)\,\mathrm{d}s$.

□

3.3 Poisson 过程的推广

定义 3.3.1. 计数过程 $\{N(t), t \geqslant 0\}$ 称为强度函数为 $\lambda(t)\,(\lambda(t)>0, t\geqslant 0)$ 的**非齐次 Poisson 过程**，如果

1. $N(0)=0$;

2. 过程有独立增量；
3. $\mathbb{P}\{N(t+h)-N(t)=1\}=\lambda(t)h+o(h)$；
4. $\mathbb{P}\{N(t+h)-N(t)\geqslant 2\}=o(h)$.

定义 3.3.2. （非齐次 Poisson 过程的等价定义）计数过程 $\{N(t),t\geqslant 0\}$ 称为强度函数为 $\lambda(t)$（$\lambda(t)>0,t\geqslant 0$）的**非齐次 Poisson 过程**，如果

1. $N(0)=0$；
2. 过程有独立增量；
3. 对任意实数 $t\geqslant 0,s\geqslant 0,N(t+s)-N(t)$ 是参数为

$$m(t+s)-m(t)=\int_t^{t+s}\lambda(u)\,\mathrm{d}u$$

的 Poisson 分布.

定理 3.3.3. （齐次和非齐次的 Poisson 过程的转换定理）设 $\{N(t),t\geqslant 0\}$ 是一个强度函数为 $\lambda(t)>0\,(t\geqslant 0)$ 的非齐次 Poisson 过程，那么对任意 $t\geqslant 0$，令

$$N^*(t)=N\left(m^{-1}(t)\right)$$

则 $\{N^*(t)\}$ 是一个强度为 1 的 Poisson 过程.

证明. 由 $\lambda(t)>0$, 则 $m(t)=\int_0^t\lambda(s)\,\mathrm{d}s>0$ 且单调, 所以 $m^{-1}(t)$ 存在. 容易知道 $N^*(0)=0$ 而且过程有平稳独立增量. [2] 令 $v(t)=m^{-1}(t)$ 则

$$N^*(t)=N\left(m^{-1}(t)\right)=N(v(t))$$

设 $v=m^{-1}(t),v+h'=m^{-1}(t+h)$，则由[3]

$$h=t+h-t=m(v+h')-m(v)=\int_v^{v+h'}\lambda(s)\,\mathrm{d}s=\lambda(v)h'+o(h')\ \ (h\to 0)$$

---

[2]独立增量很容易知道，$N^*(t)$ 也具有平稳增量，是因为

$$N^*(t+s)-N^*(t)=N\left(m^{-1}(t+s)\right)-N\left(m^{-1}(t)\right)$$

由于非齐次 Poisson 过程的定义，可以知道

$$N^*(t+s)-N^*(t)=N\left(m^{-1}(t+s)\right)-N\left(m^{-1}(t)\right)\sim\mathrm{Poi}\left(\tilde{\lambda}\right)$$

其中，$\tilde{\lambda}=m\left[m^{-1}(t+s)\right]-m\left[m^{-1}(t)\right]=s$，所以分布仅与增量有关.

[3]$\int_v^{v+h'}\lambda(s)\,\mathrm{d}s=\lambda(v)h'+o(h\prime)$ 是由于积分区间 h' 极小.

得到

$$\begin{aligned}\lim_{h\to0^+}\frac{\mathbb{P}\left\{N^*\left(t+h\right)-N^*\left(t\right)=1\right\}}{h}&=\lim_{h\prime\to0^+}\frac{\mathbb{P}\left\{N\left(v+h\prime\right)-N\left(v\right)=1\right\}}{\lambda\left(v\right)h\prime+o\left(h\prime\right)}\\&=\lim_{h\prime\to0^+}\frac{\lambda\left(v\right)h\prime+o\left(h\prime\right)}{\lambda\left(v\right)h\prime+o\left(h\prime\right)}\\&=1\end{aligned}$$

即

$$\mathbb{P}\left\{N^*\left(t+h\right)-N^*\left(t\right)=1\right\}=h+o\left(h\right)$$

同理可证

$$\lim_{h\to0^+}\frac{\mathbb{P}\left\{N^*\left(t+h\right)-N^*\left(t\right)\geqslant2\right\}}{h}=0$$

即

$$\mathbb{P}\left\{N^*\left(t+h\right)-N^*\left(t\right)\geqslant2\right\}=o\left(h\right)$$

□

注记 3.3.4. 这种变换可看作是“刻度的变换”，将原始线性刻度的时间重新线性化或非线性化了，就好像是换了一个时钟来计时. 以一些简单的例子来看，如果 $\lambda(t)=\lambda$，则 $m(t)=\lambda t$，而且 $m^{-1}(t)=\frac{t}{\lambda}$，因此 $N^*(t)=N\left(\frac{t}{\lambda}\right)$，显然它是一个以 1 为强度的 Poisson 过程，这个变换就是令 $v_1=\frac{t}{\lambda}$ 作为 $N^*(t)$ 的时间刻度，量纲不发生变化：如果 $\lambda(t)=t$，那么 $m(t)=\frac{t^2}{2}$，$m^{-1}(t)=\sqrt{2t}$，且 $N^*(t)=N\left(\sqrt{2t}\right)$，这相当于令 $v_2=\sqrt{2t}$ 作为 $N^*(t)$ 的时间刻度，量纲发生了变化.

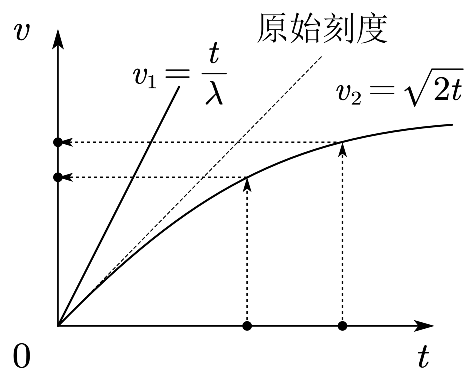

图 3.4. 齐次和非齐次的 Poisson 过程的转换

定义 3.3.5. 称随机过程 $\{X(t),t\geqslant0\}$ 为**复合随机过程**，如果对 $\forall t\geqslant0, X(t)$ 可以表示为

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

其中 $\{N(t),t\geqslant0\}$ 是一个 Poisson 过程，$Y_1,Y_2,\cdots$ 是一族独立同分布的随机变量，并且也独立于过程 $\{N(t),t\geqslant0\}$.

注记 3.3.6. 复合 Poisson 过程不一定是计数过程，但是当 $Y_i \equiv c, i=1,2,\cdots$ 其中 c 为常数时，可以化为 Poisson 过程.

定理 3.3.7. (复合 Poisson 过程的性质) 如果 $\left\{X(t)=\sum_{i=1}^{N(t)} Y_i, t \geqslant 0\right\}$ 是复合 Poisson 过程，Poisson 过程 $\{N(t), t \geqslant 0\}$ 的强度为 λ，则

1. $\{X(t), t \geqslant 0\}$ 有独立增量;
2. 若 $\mathbb{E}\left[Y_i^2\right]<+\infty$，则

$$\mathbb{E}[X(t)]=\lambda t \cdot \mathbb{E} Y_1$$

$$\operatorname{Var}[X(t)]=\lambda t \cdot \mathbb{E}\left[Y_1^2\right]$$

证明. 证明如下.

1. 令 $0 \leqslant t_1<t_2<\cdots<t_n$，则

$$X(t_k)-X(t_{k-1})=\sum_{i=N(t_{k-1})+1}^{N(t_k)} Y_i, k=1,2 \cdots, n$$

由于 $\{N(t)\}$ 的独立增量性，以及各 Y_i 之间的独立性容易得出 $X(t)$ 的独立增量性.

2. 利用矩母函数便于求 n 阶矩的特性:

$$\begin{aligned}
\phi(u) &= \mathbb{E}\left[e^{u X_t}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{u X_t} \mid N(t)=n\right] \cdot \mathbb{P}\{N(t)=n\} \text{ (迭代期望律)} \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{u(Y_1+Y_2+\cdots+Y_n)} \mid N(t)=n\right] \cdot e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\
&= \sum_{n=0}^{\infty} \prod_{i=1}^{n} \mathbb{E}\left(e^{u Y_i}\right) \cdot e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\
&= \sum_{n=0}^{\infty}\left(\mathbb{E}\left[e^{u Y_1}\right]\right)^n \cdot e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\
&= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{\left(\lambda t E\left[e^{u Y_1}\right]\right)^n}{n!} \\
&= \exp\left\{\lambda t\left[\mathbb{E}\left(e^{u Y_1}\right)-1\right]\right\}
\end{aligned}$$

因此

$$\phi'(u)=\lambda t\left(\mathbb{E}\left[Y_1 e^{u Y_1}\right]\right) \phi(u)$$

$$\phi''(u)=\lambda t\left(\mathbb{E}\left[Y_1^2 e^{u Y_1}\right]\right) \phi(u)+\lambda^2 t^2\left(\mathbb{E}\left[Y_1 e^{u Y_1}\right]\right)^2 \phi(u)$$

容易知道

$$\mathbb{E}[X(t)]=\phi'(0)=\lambda t E[Y_1]$$

$$\operatorname{Var}[X(t)]=\phi''(0)-[\phi'(0)]^2=\lambda t E\left[Y_1^2\right]$$

□

定义 3.3.8. 设个体风险因素是随机变量 $\Lambda > 0$,在 $\Lambda = \lambda$ 的条件下,如果计数过程 $\{N(t), t \geqslant 0\}$ 是参数为 λ 的 Poisson 过程,则称 $\{N(t), t \geqslant 0\}$ 为**条件 Poisson 过程**.

注记 3.3.9. 条件 Poisson 过程下,如果 Λ 的分布是 G,那么随机选择一个个体,由全概率公式可以知道,在长度 t 的时间区间内发生 n 次事件的概率为

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

定理 3.3.10. (条件 Poisson 过程的性质) 如果 $\{N(t), t \geqslant 0\}$ 为条件 Poisson 过程,且有 $\mathbb{E}[\Lambda^2] < \infty$,则

1. $\mathbb{E}[N(t)] = t\mathbb{E}[\Lambda]$;
2. $\mathrm{Var}[N(t)] = t^2 \mathrm{Var}[\Lambda] + t\mathbb{E}[\Lambda]$.

证明. 证明如下.

1. 根据迭代期望律,容易有

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

2. 由方差的公式和迭代期望律,有

$$\begin{aligned}
\mathrm{Var}[N(t)] &= \mathbb{E}\left[N^2(t)\right] - [\mathbb{E}(N(t))]^2 \\
&= \mathbb{E}_{\Lambda}\left[\mathbb{E}\left(N^2(t)\middle|\Lambda\right)\right] - t^2[\mathbb{E}(\Lambda)]^2 \\
&= \mathbb{E}_{\Lambda}\left[\Lambda^2 t^2 + \Lambda t\right] - t^2[\mathbb{E}(\Lambda)]^2 \\
&= t^2\left(\mathbb{E}\left[\Lambda^2\right] - [\mathbb{E}(\Lambda)]^2\right) + t\mathbb{E}[\Lambda] \\
&= t^2 \mathrm{Var}(\Lambda) + t\mathbb{E}[\Lambda]
\end{aligned}$$

□

3.4 练习

练习 3.1. 设 $\{N(t), t \geqslant 0\}$ 是参数为 λ 的泊松过程, 对于 $s < t$ 时, 求

1. $\mathbb{P}\{N(s) = 1, N(t) = 2\}$
2. $\mathbb{E}[N(s)|N(t) = 2]$

解答. 主要根据独立增量性质求解.

1.

$$
\begin{aligned}
\mathbb{P}\left\{N\left(s\right)=1,N\left(t\right)=2\right\} &= \mathbb{P}\left\{N\left(s\right)=1,N\left(t\right)-N\left(s\right)=1\right\} \\
&= \mathbb{P}\left\{N\left(s\right)=1\right\}\mathbb{P}\left\{N\left(t\right)-N\left(s\right)=1\right\} \\
&= \lambda^2\left(t-s\right)s\cdot e^{-\lambda t}
\end{aligned}
$$

2. 先求 $N\left(s\right)|\,N\left(t\right)=2$ 的分布：

$$
\begin{aligned}
\mathbb{P}\{\,N\left(s\right) &= kN\left(t\right)=2 = \frac{\mathbb{P}\left\{N\left(s\right)=k,N\left(t\right)=2\right\}}{\mathbb{P}\left\{N\left(t\right)=2\right\}} \\
&= \frac{\mathbb{P}\left\{N\left(s\right)=k\right\}\cdot\mathbb{P}\left\{N\left(t\right)-N\left(s\right)=2-k\right\}}{\mathbb{P}\left\{N\left(t\right)=2\right\}} \\
&= \frac{\frac{(\lambda s)^k}{k!}e^{-\lambda s}\times\frac{(\lambda(t-s))^{2-k}}{(2-k)!}e^{-\lambda(t-s)}}{\frac{(\lambda t)^2}{2!}e^{-\lambda t}} \\
&= \begin{pmatrix} 2 \\ k \end{pmatrix}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{2-k},0\leqslant k\leqslant 2
\end{aligned}
$$

所以

$$
N\left(s\right)|\,N\left(t\right)=2\sim\mathrm{Bernoulli}\left(2,\frac{s}{t}\right)
$$

进而

$$
\mathbb{E}\left[N\left(s\right)|\,N\left(t\right)=2\right]=\frac{2s}{t}
$$

$\square$

第四章　更新过程

更新过程是 Poisson 过程的推广, Poisson 过程是更新过程的一类特殊情况. 对于事件发生的时间间隔 $X_i, i=1,2,\cdots$

- Poisson 过程中，$X_i \sim \exp(\lambda)$；
- 更新过程中，X_i的分布函数是$F(x)$.

本章要求理解更新过程定义、掌握更新密度、更新过程等概念，Wald 等式的证明及应用.

4.1　更新过程定义及其若干分布

定义 4.1.1. 设 $\{X_n, n=1,2,\cdots\}$ 是一列**独立同分布**的**非负**随机变量，它们的分布函数是 $F(x)$，令 $T_n=\sum_{i=1}^{n} X_i\ (n\geqslant 1), T_0=0$，把由

$$N(t)=\sup\{n: T_n\leqslant t\}$$

定义的计数过程称为**更新过程**.

为了避免 $F(x)$ 为常函数，$F(x)$ 需要满足 $F(0)=\mathbb{P}\{X_n=0\}\neq 1$ 以及均值:

$$\mu=\mathbb{E}[X_n]=\int_0^{\infty} x\mathrm{d}F(x), 0<\mu\leqslant+\infty$$

此外，称 $M(t)=\mathbb{E}[N(t)]$ 为**更新函数**.

注记 4.1.2. 对于更新过程，有限时间内最多只能发生有限次更新，无穷多次更新只能在无限长的时间内发生，或者给定 t，有

$$\mathbb{P}\{N(t)<\infty\}=1$$

事实上，由大数定律知道

$$\frac{\sum_{n=1}^{n} X_n}{n}=\frac{T_n}{n}\xrightarrow{\mathbb{P}}\mu$$

这说明 $n\to+\infty$ 当且仅当 $T_n\to+\infty$，只能在无限长的时间里成立.

定理 4.1.3.

$$M(t)=\sum_{n=1}^{\infty}F_n(t)$$

其中 F_n 是 F 的 n 重卷积.

证明. 根据重要的等价关系

$$N(t)\geqslant n\Longleftrightarrow T_n\leqslant t$$

有

$$\begin{aligned}\mathbb{P}\{N(t)=n\}&=\mathbb{P}\{N(t)\geqslant n\}-\mathbb{P}\{N(t)\geqslant n+1\}\\&=\mathbb{P}\{T_n\leqslant t\}-\mathbb{P}\{T_{n+1}\leqslant t\}\\&=\mathbb{P}\left\{\sum_{i=1}^{n}X_i\leqslant t\right\}-\mathbb{P}\left\{\sum_{i=1}^{n+1}X_i\leqslant t\right\}\\&=F_n(t)-F_{n+1}(t)\end{aligned}$$

因此，由期望的定义可得

$$\begin{aligned}M(t)=\mathbb{E}[N(t)]&=\sum_{n=1}^{\infty}nP\{N(t)=n\}\\&=\sum_{n=1}^{\infty}n[F_n(t)-F_{n+1}(t)]\\&=\sum_{n=1}^{\infty}nF_n(t)-\sum_{n=1}^{\infty}nF_{n+1}(t)\\&=F_1(t)+\sum_{n=2}^{\infty}nF_n(t)-\sum_{m=2}^{\infty}(m-1)F_m(t)\\&=F_1(t)+\sum_{n=2}^{\infty}nF_n(t)-\sum_{m=2}^{\infty}mF_m(t)+\sum_{n=2}^{\infty}F_n(t)\\&=\sum_{n=1}^{\infty}F_n(t)\end{aligned}$$

□

定理 4.1.4. $M(t)$ 是 t 的不减函数，且对 $t\in[0,+\infty),M(t)<+\infty$.

证明. $M(t)$ 的单调性容易通过 $N(t)$ 也是不减的得到. 下面证明有限性.

首先，证明 $\exists b>0,F(b)<1$：由更新过程关于 F 的非平凡性规定，有 $F(0)<1$，即 $\mathbb{P}\{X_n=0\}<1$，与之等价的是 $\mathbb{P}\{X_n>0\}>0$，则

$$\exists a>0,\mathbb{P}\{X_n\geqslant a\}>0\text{ 或}\mathbb{P}\{X_n<a\}<1$$

而 $F(a) = \mathbb{P}\{X_n \leqslant a\} = \mathbb{P}\{X_n < a\} + \mathbb{P}\{X_n = a\}$，为了使得右边能够小于 1，就要防止 $\mathbb{P}\{X_n = a\} = \mathbb{P}\{X_n \geqslant a\}$ 导致右边等于 1. 因此可以取 $b \in (0,a)$ 使得 $F(b) < 1$. 显然，对任意固定的 t，总能选取正整数 k，使得 $kb \geqslant t$. 所以[1]

$$\{T_k \leqslant t\} \subset \{T_k \leqslant kb\} \subset \{X_1 > b, X_2 > b, \cdots, X_k > b\}^c$$

于是

$$\mathbb{P}\{T_k \leqslant t\} \leqslant 1 - \mathbb{P}\{X_1 > b, X_2 > b, \cdots, X_k > b\} = 1 - [1 - F(b)]^k = 1 - \beta \tag{4.1}$$

其中，令 $\beta = [1 - F(b)]^k$. 将 T_{mk} 作分解，即 $T_{mk} = \sum_{i=1}^{m} \left[T_{ik} - T_{(i-1)k}\right]$，那么

$$\{T_{mk} \leqslant t\} \subset \left\{T_k - T_0 \leqslant t, T_{2k} - T_k \leqslant t, \cdots, T_{mk} - T_{(m-1)k} \leqslant t\right\} \tag{4.2}$$

由更新区间的独立同分布，得到

$$\mathbb{P}\left\{T_k - T_0 \leqslant t, T_{2k} - T_k \leqslant t, \cdots, T_{mk} - T_{(m-1)k} \leqslant t\right\} = [\mathbb{P}\{T_k \leqslant t\}]^m \tag{4.3}$$

综合(4.1)(4.2)(4.3)得到

$$\mathbb{P}\{T_{mk} \leqslant t\} \leqslant [\mathbb{P}\{T_k \leqslant t\}]^m \leqslant (1-\beta)^m$$

对任意 $j \geqslant 0$，有

$$\{T_{mk+j} \leqslant t\} \subset \{T_{mk} \leqslant t\}$$

因此，对左边的 k 个片段和，有 $\sum_{n=mk}^{(m+1)k-1} \mathbb{P}\{T_n \leqslant t\} \leqslant kP\{T_{mk} \leqslant t\}$ 最后综上，对 k 个片段和进行放缩

$$\begin{aligned}
M(t) = \sum_{n=1}^{\infty} F_n(t) &= \sum_{n=1}^{\infty} \mathbb{P}\{T_n \leqslant t\} \\
&= \sum_{n=1}^{k-1} \mathbb{P}\{T_n \leqslant t\} + \sum_{n=k}^{\infty} \mathbb{P}\{T_n \leqslant t\} \\
&\leqslant \sum_{n=1}^{k-1} \mathbb{P}\{T_n \leqslant t\} + \sum_{n=1}^{\infty} kP\{T_{mk} \leqslant t\} \\
&\leqslant \sum_{n=1}^{k-1} \mathbb{P}\{T_n \leqslant t\} + \sum_{n=1}^{\infty} k(1-\beta)^m \\
&= \sum_{n=1}^{k-1} \mathbb{P}\{T_n \leqslant t\} + \frac{k}{\beta} < \infty
\end{aligned}$$

□

[1]集合 $\{X_1 > b, X_2 > b, \cdots, X_k > b\}^c$ 代表 $X_1, X_2, \cdots X_k$ 中，至少有一个小于 b.

4.2 更新方程及其应用

定义 4.2.1. 在 $M(t)$ 的导数存在的条件下，其导数 $m(t)$ 称为**更新密度**. 如果求导和求和能交换顺序，那么

$$m(t)=\sum_{n=1}^{\infty} f_n(t)$$

其中 $f_n(t)$ 是 $F_n(t)$ 的密度函数.

定理 4.2.2. $M(t)$ 和 $m(t)$ 分别满足积分方程

$$M(t)=F(t)+\int_0^t M(t-s)\,\mathrm{d}F(s) \tag{4.4}$$

$$m(t)=f(t)+\int_0^t m(t-s)\,f(s)\,\mathrm{d}s \tag{4.5}$$

其中 $f(t)=F'(t)$.

证明. 由 $M(t)$ 的性质，有

$$\begin{aligned}
M(t)&=\sum_{n=1}^{\infty} F_n(t)\\
&=F(t)+\sum_{n=2}^{\infty} F_n(t)\\
&=F(t)+\left[\sum_{n=2}^{\infty}(F_{n-1}*F)\right](t)\\
&=F(t)+\left[\sum_{n=1}^{\infty}(F_n*F)\right](t)\\
&=F(t)+[M*F](t)
\end{aligned}$$

其中 $[M*F](t)=\int_0^t M(t-s)\,\mathrm{d}F(s)$，这可以根据卷积的定义得到，即

$$\begin{aligned}
[M*F](t)&=\int_{-\infty}^{+\infty} M(t-s)\,\mathrm{d}F(s)\\
&=\int_{-\infty}^{0} M(t-s)\,\mathrm{d}F(s)+\int_0^t M(t-s)\,\mathrm{d}F(s)+\int_t^{+\infty} M(t-s)\,\mathrm{d}F(s)\\
&=0+\int_0^t M(t-s)\,\mathrm{d}F(s)+0\\
&=\int_0^t M(t-s)\,\mathrm{d}F(s)
\end{aligned}$$

- 因为当 $s<0$ 时, $F(s)=0$，所以在 $(-\infty,0)$ 上的积分 $\int_{-\infty}^0 M(t-s)\,\mathrm{d}F(s)$ 为 0;
- 同样地，当 $s>t$ 时,$M(t-s)=0$，所以在 $(t,+\infty)$ 上的积分 $\int_t^{+\infty} M(t-s)\,\mathrm{d}F(s)$ 为 0.

这就证得了积分方程(4.4)，积分方程(4.5)在此基础上求导即得. □

定义 4.2.3. 称一积分方程为**更新方程**，如果它具有以下形式

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

其中 $H(t), F(t)$ 均为已知，并且当 $t<0$ 时，$H(t)=F(t)=0$.

特别地，当 $H(t)$ 在任何区间上有界时，称上述方程为**适定更新方程**（Proper Renewal Equation），简称为**更新方程**.

定理 4.2.4. （更新方程解的存在及唯一性定理）设更新方程中 $H(t)$ 有界，则方程在有限区间内存在**唯一**的**有界**解

$$K(t) = H(t) + \int_0^t H(t-s)\,\mathrm{d}M(s) \tag{4.6}$$

其中 $M(t)=\sum_{n=1}^{\infty} F_n(t)$ 是分布函数 F 的更新函数.

证明. [2] 证明分为三步，分别证明(4.6)是解、有界性和唯一性.

Step 1. 证明(4.6)是解：从(4.4)的证明中可以知道 $M(t)$ 和 $F(t)$ 满足的一个关系：

$$M(t) = F(t) + F * M(t)$$

(4.14) 将(4.6)作变形，得

$$\begin{aligned} K(t) &= H(t) + M * H(t) \\ &= H(t) + (F + F*M) * H(t) \\ &= H(t) + F * (H + M*H)(t) \\ &= H(t) + F * K(t) \end{aligned}$$

因此，(4.6)是更新方程的解.

Step 2. 证明(4.6)有界：根据 $M(t)$ 有界且单调不减，以及 $H(t)$ 的有界性，可以知道，对

[2]该证明需要用到卷积的相关性质，这里复习一下：假设 $B(t)$ 是 **单调增加** 的右连续函数，且 $B(0)=0$，$C(t), C_1(t), C_2(t)$ 均为光滑有界函数（保证卷积的定义），则有

- $\max\limits_{0\leqslant t\leqslant T}|B*C(t)| \leqslant \max\limits_{0\leqslant t\leqslant T}|C(t)|\cdot B(T)$；
- 分配律：$B*C_1(t) + B*C_2(t) = B*(C_1+C_2)(t)$；
- 结合律：$C_1*(C_2*B)(t) = (C_1*C_2)*B(t)$.

$\forall T > 0$，有

$$\begin{aligned}\sup_{0\leqslant t\leqslant T}|K(t)| &= \sup_{0\leqslant t\leqslant T}\left[|H(t)| + \int_0^t |H(t-s)|\,\mathrm{d}M(s)\right] \\ &\leqslant |H(t)| + \int_0^T |H(T-s)|\,\mathrm{d}M(s) \\ &\leqslant \sup_{0\leqslant t\leqslant T}\left[|H(t)| + |H(t)|\cdot M(T)\right] \\ &= \sup_{0\leqslant t\leqslant T}|H(t)|\left[1+M(T)\right] < \infty\end{aligned}$$

其中因为 $M(t)$ 非负，就不必加绝对值了.

Step 3. 证明(4.6)唯一：设除了 $K(t)$ 外，更新方程还有一个有界的解 $\tilde{K}(t)$，于是

$$\tilde{K}(t) = H(t) + F * \tilde{K}(t)$$

该关系进行迭代，可以得到

$$\begin{aligned}\tilde{K}(t) &= H(t) + F * \tilde{K}(t) \\ &= H(t) + F * \left(H + F * \tilde{K}\right)(t) = H(t) + F * H(t) + F_2 * \tilde{K}(t) \\ &= H(t) + F * H(t) + F_2 * \left(H + F * \tilde{K}\right)(t) \\ &= H(t) + \left(\sum_{k=1}^{2} F_k\right) * H(t) + F_2 * \tilde{K}(t)\end{aligned}$$

反复迭代 n 次，可以归纳得到

$$\tilde{K}(t) = H(t) + \left(\sum_{k=1}^{n-1} F_k\right) * H(t) + F_n * \tilde{K}(t) \tag{4.7}$$

如果一直如此迭代下去，则需要研究(4.7)右边的极限，根据更新函数 $M(t)$ 的定义，有

$$\lim_{n\to\infty}\sum_{k=1}^{n-1} F_k(t) = M(t)$$

现在要证明唯一性，只需要证明(4.7)右边第三项趋于零，即 $\lim_{n\to\infty} F_n * \tilde{K}(t) = 0$. 注意到，对 $\forall t$ 有

$$\left|F_n * \tilde{K}(t)\right| = \left|\int_0^t \tilde{K}(t-s)\,\mathrm{d}F_n(x)\right| \leqslant \sup_{0\leqslant x\leqslant t}\left|\tilde{K}(t-x)\right|\cdot F_n(t)$$

由于 $\tilde{K}(t)$ 是有界解，所以 $\sup\limits_{0\leqslant x\leqslant t}\left|\tilde{K}(t-x)\right| < \infty$，又因为正项级数 $M(t) = \sum_{n=1}^{\infty} F_n(t) < \infty$，所以 $\lim_{n\to\infty} F_n(t) = 0$，从而

$$\lim_{n\to\infty}\left|F_n * \tilde{K}(t)\right| = 0$$

对(4.7)令 $n \to \infty$，即为所证. □

定理 4.2.5. (Wald 等式) 对更新过程，设 $\mathbb{E}[X_i] < \infty, i = 1, 2, \cdots$，有

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

证明 *1.* 仅根据更新过程的性质证明.

由更新过程增量的独立同分布和迭代期望律，有

$$\begin{aligned}
\mathbb{E}\left[T_{N(t)+1}\right] &= \sum_{k=0}^{\infty} \mathbb{E}\left[\sum_{i=1}^{N(t)+1} X_i \middle| N(t) = k\right] \cdot \mathbb{P}\{N(t) = k\} \\
&= \sum_{k=0}^{\infty} \mathbb{E}\left[\sum_{i=1}^{k+1} X_i\right] \cdot \mathbb{P}\{N(t) = k\} \\
&= \sum_{k=0}^{\infty} (k+1)\,\mathbb{E}[X_1]\,\mathbb{P}\{N(t) = k\} \\
&= \mathbb{E}[X_1]\left[\sum_{k=0}^{\infty} kP\{N(t) = k\} + \sum_{k=0}^{\infty} \mathbb{P}\{N(t) = k\}\right] \\
&= \mathbb{E}[X_1]\,[\mathbb{E}[N(t)] + 1] \\
&= \mathbb{E}[X_1]\,\mathbb{E}[N(t) + 1]
\end{aligned}$$

□

证明 *2.* 使用更新方程证明. 对 $T_{N(t)+1}$ 关于 X_1 取条件期望，得

$$\mathbb{E}\left[T_{N(t)+1}\middle| X_1 = x\right] = \begin{cases} x & , x > t \\ x + \mathbb{E}\left[T_{N(t-x)+1}\right] & , x \leqslant t \end{cases}$$

下面正式构造更新方程. 令 $K(t) = \mathbb{E}\left[T_{N(t)+1}\right]$，则

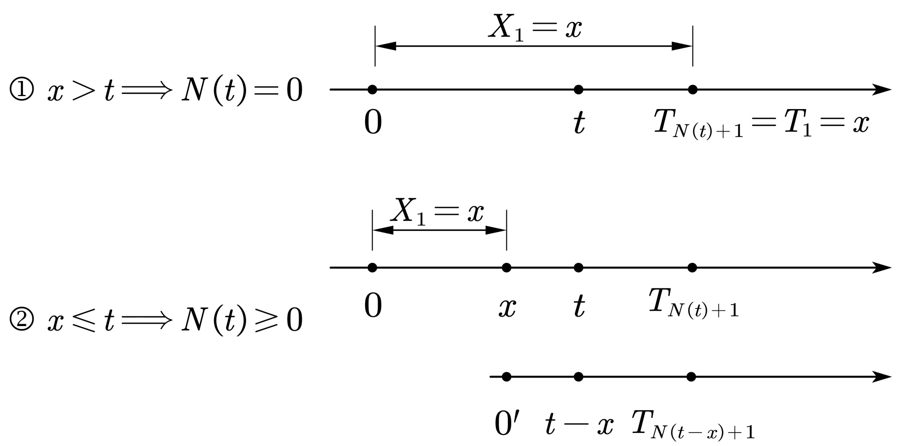

图 4.1. 证明 Wald 等式

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

这是一个特殊的更新方程，因为 $H(t)=\mathbb{E}[X_1]$ 是个常函数. 由更新方程解的存在性定理，容易得到解

$$
\begin{aligned}
K(t) &= \mathbb{E}[X_1]+\int_0^t \mathbb{E}[X_1]\,\mathrm{d}M(s)\\
&= \mathbb{E}(X_1)[1+M(t)]\\
&= \mathbb{E}(X_1)[1+\mathbb{E}[N(t)]]\\
&= \mathbb{E}(X_1)\cdot\mathbb{E}[N(t)+1]
\end{aligned}
$$

□

4.3 更新定理

注记 4.3.1. 提出问题：我们知道 Poisson 过程是特殊的更新过程. 对强度为 λ 的 Poisson 过程，更新函数为 $M(t)=\mathbb{E}[N(t)]=\lambda t$，有关系式

$$\frac{M(t)}{t}=\frac{1}{\mathbb{E}[X_n]}=\lambda$$

本节的目的就是探究一般的更新过程有无类似关系.

定理 4.3.2. (Feller 初等更新定理) 记 $\mu=\mathbb{E}[X_n]$，则

$$\frac{M(t)}{t}\longrightarrow\frac{1}{\mu}, t\to\infty$$

特别地，若 $\mu=\infty$，$\frac{M(t)}{t}\to 0$.

证明. 当 $\mu<\infty$ 时，由于

$$T_{N(t)+1}>t$$

两边取期望，结合 Wald 等式，有

$$\mathbb{E}\left[T_{N(t)+1}\right]=\mathbb{E}(X_1)\,\mathbb{E}[N(t)+1]=\mu[M(t)+1]>t$$

所以

$$\frac{M(t)}{t} > \frac{1}{\mu} - \frac{1}{t}$$

两边取下极限，得

$$\lim_{t\to\infty}\inf\frac{M(t)}{t} > \frac{1}{\mu} \tag{4.8}$$

下面使用一些技巧，构造一个有界的更新过程．固定常数 M，对 X_n 作一个“截断”，即令

$$X_n^c = \begin{cases} X_n, & X_n \leqslant M \\ M, & X_n > M \end{cases}$$

则 $X_n^c, n = 1, 2, \cdots$ 就确定了一个新的更新过程，令

$$T_n^c = \sum_{i=1}^{n} X_i^c, N(t)^c = \sup\{n | T_n^c \leqslant t\}$$

因为 $X_n^c \leqslant M$，则

$$T_{N^c(t)+1}^c = t + X_n^c \leqslant t + M$$

两边求期望，结合 Wald 等式，得

$$\mathbb{E}\left[T_{N^c(t)+1}^c\right] = \mathbb{E}(X_1^c)\,\mathbb{E}[N^c(t)+1] = \mu_M[M(t)^c + 1] \leqslant t + M$$

其中 $\mu_M = \mathbb{E}[X_1^c], M(t)^c = \mathbb{E}[N(t)^c]$，因此

$$\frac{M(t)^c}{t+M} \leqslant \frac{1}{\mu_M} - \frac{1}{t+M}$$

两边取上极限，并且因为

$$X_n^c \leqslant X_n \Longrightarrow T_n^c \leqslant T_n \Longrightarrow N(t)^c \geqslant N(t) \Longrightarrow M(t)^c \geqslant M(t)$$

于是

$$\lim_{t\to\infty}\sup\frac{M(t)}{t} \leqslant \lim_{t\to\infty}\sup\frac{M(t)^c}{t} \leqslant \frac{1}{\mu_M} \tag{4.9}$$

令 $M \to \infty$，可以知道 $X_n^c \to X_n \Longrightarrow \mathbb{E}[X_n^c] \to \mathbb{E}[X_n] \Longrightarrow \mu_M \to \mu$，结合(4.8)与(4.9)两边夹，即得所证.

当 $\mu = \infty$ 时，可以同样地证明. □

注记 4.3.3. 提出问题：Feller 初等更新定理说明 $t \to \infty$ 时，$M(t) \approx \frac{t}{\mu}$，可以猜想当 $t \to \infty$ 时，是否对于 $\forall h$，有

$$M(t+h) - M(t) \to \frac{t+h}{\mu} - \frac{t}{\mu} = \frac{h}{\mu}$$

定义 4.3.4. 称随机变量 X 服从**格点分布**，或称随机变量 X 的分布 F 是**格点的**，如果存在 $d \geqslant 0$，使得

$$\sum_{n=0}^{\infty} \mathbb{P}\left\{X = nd\right\} = 1$$

并称满足上述条件的最大的 d 为格点分布的周期.

定理 4.3.5. (Blackwell 更新定理) 对于更新过程，记 $\mu = \mathbb{E}\left[X_n\right]$，

1. 若 X_i 不服从格点分布，则对一切 $a \geqslant 0$，当 $t \to \infty$ 时，有

$$M\left(t+a\right) - M\left(t\right) \to \frac{a}{\mu}$$

2. 若 X_i 服从格点分布，且周期是 d，当 $t \to \infty$ 时，有

$$\mathbb{P}\left\{\text{在}nd\text{处发生更新}\right\} \to \frac{d}{\mu}$$

第五章　Markov 链

独立增量过程 (Poisson 过程、第七章中的维纳过程) 都是 Markov 过程. 本章要求理解 Markov 链、转移概率、随机矩阵等基本概念；掌握 n 步转移概率、n 步转移矩阵等概念及 C-K 方程；重点掌握状态的分类及相关性质；理解极限定理、不变分布与极限分布；掌握连续时间 Markov 链、转移概率和 Kolmogorov 微分方程.

5.1　基本概念

定义 5.1.1. 随机过程 $\{X_n, n \geqslant 0\}$ 称为**Markov 链**，如果

1. 它只取有限个或可列个值；
2. 对 $\forall n \geqslant 0$，以及任意状态 $i, j, i_0, i_1, \cdots, i_{n-1}$ 有

$$\mathbb{P}\left\{ \underbrace{X_{n+1} = j}_{\text{未来}} \middle| \underbrace{X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}}_{\text{历史}}, \underbrace{X_n = i}_{\text{现在}} \right\} = \mathbb{P}\left\{ X_{n+1} = j \middle| X_n = j \right\}$$

其中 $X_n = i$ 表示过程在时刻 n 处于状态 i.

定义 5.1.2. 几个概念.

- **状态空间**：$\{X_n, n \geqslant 0\}$ 所有可能取值组成的集合，记为 $\mathcal{S}$，它含有限个或可列个元素.
- **一步转移概率** p_{ij}：

$$p_{ij} = \mathbb{P}\left\{ X_{n+1} = j \middle| X_n = j \right\}$$

- **时齐 Markov 链**：转移概率 p_{ij} 只与状态 i和j 有关的 Markov 链. (一般情况下，转移概率还可能与时刻 n 或者上一步状态相关).
- 转移概率矩阵 $\boldsymbol{P} = (p_{ij})_{n \times n}$ 的性质：

  1. 元素非负：$p_{ij} \geqslant 0, i, j \in \mathcal{S}$；
  2. 行和为 1：$\sum_{j \in \mathcal{S}} p_{ij} = 1, \forall i \in \mathcal{S}$.

  具有这两条性质的矩阵称为**随机矩阵**.

- n **步转移概率**：称条件概率

$$p_{ij}^{(n)}=\mathbb{P}\left\{X_{m+n}=j|\,X_m=j\right\},i,j\in S,m\geqslant 0,n\geqslant 1$$

为 Markov 链的 n 步转移概率. 相应地 $\boldsymbol{P}^{(n)}=\left(p_{ij}^{(n)}\right)$ 称为 n 步转移矩阵，它们只关心从状态 i 经过 n 步后转移到状态 j 的前后状态 i和j（**而且包含了中间所有可能的情况**），对中间的 $n-1$ 步转移经过的状态没有要求.

定理 5.1.3. C-K 方程（Chapman-Kolmogorov Equation）：对一切 $n,m\geqslant 0$，$i,j\in\mathcal{S}$ 有

1.
$$p_{ij}^{(m+n)}=\sum_{k\in\mathcal{S}}p_{ik}^{(m)}p_{kj}^{(n)} \tag{5.1}$$

2.
$$\boldsymbol{P}^{(n)}=\boldsymbol{P}\cdot\boldsymbol{P}^{(n-1)}=\boldsymbol{P}\cdot\boldsymbol{P}\cdot\boldsymbol{P}^{(n-2)}=\cdots=\boldsymbol{P}^n \tag{5.2}$$

证明. 首先，

$$p_{ij}^{(m+n)}=\mathbb{P}\left\{X_{m+n}=j|\,X_0=i\right\}=\frac{\mathbb{P}\left\{X_{m+n}=j,X_0=i\right\}}{\mathbb{P}\left\{X_0=i\right\}}$$

再根据全概率公式,[1]

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

根据矩阵的乘法法则，有 $\boldsymbol{P}^{(m+n)}=\boldsymbol{P}^{(m)}\cdot\boldsymbol{P}^{(n)}$：

$$\begin{aligned}
\boldsymbol{P}^{(m+n)}&=\left(p_{ij}^{(m+n)}\right)=\left(\sum_{k\in\mathcal{S}}p_{ik}^{(m)}p_{kj}^{(n)}\right)\\
&=\begin{pmatrix}
\sum_{k\in\mathcal{S}}p_{1k}^{(m)}p_{k1}^{(n)} & \sum_{k\in\mathcal{S}}p_{1k}^{(m)}p_{k2}^{(n)} & \cdots & \sum_{k\in\mathcal{S}}p_{1k}^{(m)}p_{kn}^{(n)}\\
\sum_{k\in\mathcal{S}}p_{2k}^{(m)}p_{k1}^{(n)} & \sum_{k\in\mathcal{S}}p_{2k}^{(m)}p_{k2}^{(n)} & \cdots & \sum_{k\in\mathcal{S}}p_{2k}^{(m)}p_{kn}^{(n)}\\
\vdots & \vdots & \ddots & \vdots\\
\sum_{k\in\mathcal{S}}p_{nk}^{(m)}p_{k1}^{(n)} & \sum_{k\in\mathcal{S}}p_{nk}^{(m)}p_{k2}^{(n)} & \cdots & \sum_{k\in\mathcal{S}}p_{nk}^{(m)}p_{kn}^{(n)}
\end{pmatrix}=\boldsymbol{P}^{(m)}\cdot\boldsymbol{P}^{(n)}
\end{aligned}$$

而(5.2)可以通过递推证得.

□

[1] $\mathbb{P}\left\{B\right\}=\sum_{i=1}^{n}\mathbb{P}\left\{B\cap A_i\right\}$，其中 $\bigcup_{i=1}^{n}A_i=\Omega,A_i\cap A_j=\oslash\ (i\neq j)$.

注记 5.1.4. 转移矩阵包含了中间的所有情况：从状态 i 经过 n 步到达状态 j 的概率 $p_{ij}^{(n)}$ 是从状态 i 不超过 n 步到达状态 j 的概率，或者说 $\boldsymbol{P}^{(n)}$ 包含了 $\boldsymbol{P}, \boldsymbol{P}^{(2)}, \cdots, \boldsymbol{P}^{(n-1)}$ 的所有信息. 所以计算至多 n 步的概率时，$p_{ij}^{(1)} + p_{ij}^{(2)} + \cdots + p_{ij}^{(n)}$ 是错误的.

例子 5.1.5. 甲乙两人进行某种比赛，设每局甲胜的概率是 p，乙胜的概率是 q，和局的概率是 r，$p+q+r=1$. 设每局比赛后，胜者记 “+1” 分，负者记 “-1” 分，和局不记分，且当两人中有一人获得 2 分时结束比赛. 以 X_n 表示比赛至第 n 局时甲获得的分数，则 $\{X_n, n=1,2,\cdots\}$ 为时齐 Markov 链，求在甲当前获得 1 分的情况下，不超过两局可结束比赛的概率.

解答. 一步转移矩阵为

$$\boldsymbol{P} = \begin{array}{c} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ q & r & p & 0 & 0 \\ 0 & q & r & p & 0 \\ 0 & 0 & q & r & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

那么

$$\boldsymbol{P}^{(2)} = \boldsymbol{P} \cdot \boldsymbol{P} = \begin{array}{c} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ q+rq & r^2+pq & 2pr & 0 & 0 \\ q^2 & 2rq & r^2+2pq & 2pr & p^2 \\ 0 & q^2 & 2qr & r^2+pq & p+pr \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

则甲当前获得 1 分的情况下，不超过两局可结束比赛的概率为

$$p_{1,2}^{(2)} + p_{1,-2}^{(2)} = p + pr$$

注意并不是

$$p_{1,2}^{(1)} + p_{1,-2}^{(1)} + p_{1,2}^{(2)} + p_{1,-2}^{(2)}$$

□

5.2 状态的分类及性质

定义 5.2.1. 一些概念.

- 可达：称状态 i **可达**状态 j $(i, j \in \mathcal{S})$，如果 $\exists n \geqslant 0, p_{ij}^{(n)} > 0$，记为 $i \to j$.
- 互通：称状态 i 与 j**互通**，如果 $i \to j$，同时 $j \to i$，记为 $i \leftrightarrow j$.

注记 5.2.2. 互通是一种等价关系：

1. 自返性：$i \leftrightarrow i$；
2. 对称性：若 $i \leftrightarrow j$，则 $j \leftrightarrow i$；
3. 传递性：若 $i \leftrightarrow k, k \leftrightarrow j$，则 $i \leftrightarrow j$.

- 类：任何两个互通的状态归为一**类**. 特别地，每个吸收态自成一类.

  注记 5.2.3. 同在一类的状态是互通的，并且任何一个状态不能同时处于两个不同的类.

- **可约**的定义：
  - 只存在一个类的 Markov 链，称它是**不可约**的.
  - 反之，如果存在两个及以上的类，则称它是**可约**的.
- 状态的周期：若集合 $\left\{n \middle| n \geqslant 1, p_{ii}^{(n)} > 0\right\}$ 非空，则称它的最大公约数 $d = d(i)$ 为状态 i 的**周期**.
- 状态的周期性：
  - 如果 $d > 1$，则称状态 i 是**周期**的；
  - 若 $d = 1$，则称状态 i 是**非周期**的；
  - 特别地，如果集合 $\left\{n \middle| n \geqslant 1, p_{ii}^{(n)} > 0\right\} = \varnothing$，则状态 i 的**周期无穷大**.

$$\left\{n \middle| n \geqslant 1, p_{ii}^{(n)} > 0\right\} \begin{cases} \text{空集, 周期无穷大} \\ \text{非空} \begin{cases} d = 1, \text{非周期的} \\ d > 1, \text{周期的} \end{cases} \end{cases}$$

注记 5.2.4. 状态 i 有周期 d，但并不是对所有的 n，都有 $p_{ii}^{(nd)} > 0$，即不能说从状态 i 经过 nd 步一定能返回状态 i. 但是可以证明当 n 充分大之后一定有 $p_{ii}^{(nd)} > 0$.

定理 5.2.5. 若状态 i, j 属于同一类，那么 $d(i) = d(j)$.

证明. 因为状态 i, j 属于同一类，所以 $i \leftrightarrow j$，即存在 $m, n > 0$，使得 $p_{ij}^{(m)} > 0, p_{ij}^{(n)} > 0$，那么 $p_{ii}^{(m+n)} = \sum_{k \in \mathcal{S}} p_{ik}^{(m)} p_{ki}^{(n)} \geqslant p_{ij}^{(m)} p_{ji}^{(n)} > 0$.

对所有使得 $p_{jj}^{(s)} > 0$ 的 s，有 $p_{ii}^{(m+n+s)} \geqslant p_{ij}^{(m)} p_{jj}^{(s)} p_{ji}^{(n)} > 0$，所以 i 的周期 $d(i)$ 应该满足 $\frac{n+m}{d(i)} \in \mathbb{Z}, \frac{m+n+s}{d(i)} \in \mathbb{Z}$，所以 $\frac{s}{d(i)} \in \mathbb{Z}$.

而 $p_{jj}^{(s)} > 0$ 说明 $s \in A = \left\{n \middle| n \geqslant 1, p_{jj}^{(n)} > 0\right\}$，那么 $\frac{s}{d(j)} \in \mathbb{Z}$. 而且，$d(j)$ 是集合 A 元素的最大公约数，而 $d(i)$ 只是它们的一个约数，所以 $\frac{d(j)}{d(i)} \in \mathbb{Z}$.

同样地，可以证明 $\frac{d(i)}{d(j)} \in \mathbb{Z}$，所以 $d(i) = d(j)$.

□

定义 5.2.6. **首达概率**：从 i 出发经过 n 步首次达到 j 的概率 $f_{ij}^{(n)}$.

$$f_{ij}^{(n)} = \mathbb{P}\{X_n = j, X_k \neq j, k = 1, 2, \cdots, n-1 | X_0 = i\}, n \geqslant 1$$

特别地，$f_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$.

定义 5.2.7. 令 $f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$，

- 如果 $f_{ii} = 1$，称状态 i 为**常返状态**；
- 如果 $f_{ii} < 1$，称状态 i 为**非常返**或**瞬过状态**.

注记 5.2.8. f_{ij} 的含义：从 i 出发，有限步内可以到达 j 的**概率**（$0 < f_{ij} \leqslant 1$）.

- 如果 i 为常返状态，说明在有限步内，过程将以概率 1 重新返回 i；
- 如果 i 为非常返状态，说明过程以概率 $1 - f_{ii} > 0$ 不再返回 i，或说从 i 滑过.

定义 5.2.9. 几个概念.

- 对于**常返状态** i，定义由 i 出发再返回 i 的**平均步数**（时间）：

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

- 对于**常返状态** i，
  - 若 $\mu_i < +\infty$，则称 i 为**正常返状态**；
  - 若 $\mu_i = +\infty$，则称 i 为**零常返状态**.
- 遍历状态：若 i 是正常返状态，且是非周期的，则称之为**遍历状态**.
- 吸收状态：若 i 是遍历状态，且 $f_{ii}^{(1)} = 1$，则称之为**吸收状态**. 此时 $\mu_i = 1$.

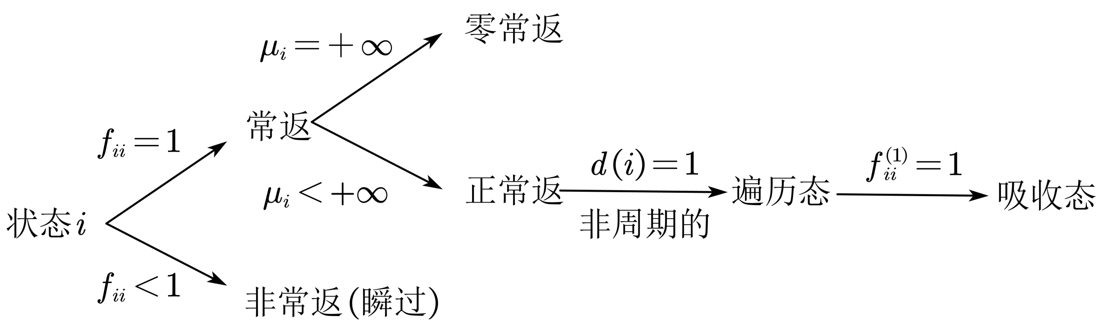

图 5.1. 状态的分类

引理 5.2.10. (C-K 方程的变式)

$$f_{ij}^{(l+1)} = \sum_{k \neq j} f_{ik}^{(1)} f_{kj}^{(l)}$$

证明. (以下的各式中，$t=1,2,\cdots,l.$)

$$
\begin{aligned}
f_{ij}^{(l+1)} &= \mathbb{P}\left\{X_{l+1}=j, X_t\neq j \mid X_0=i\right\}\\
&= \sum_{k\neq j}\frac{\mathbb{P}\left\{X_{l+1}=j, X_t\neq j, X_0=i, X_1=k\right\}}{\mathbb{P}\left\{X_0=i\right\}}\\
&= \sum_{k\neq j}\frac{\mathbb{P}\left\{X_{l+1}=j, X_t\neq j, X_0=i, X_1=k\right\}}{\mathbb{P}\left\{X_1=k, X_0=i\right\}}\cdot\frac{\mathbb{P}\left\{X_1=k, X_0=i\right\}}{\mathbb{P}\left\{X_0=i\right\}}\\
&= \sum_{k\neq j}\mathbb{P}\left\{X_{l+1}=j, X_t\neq j \mid X_0=i, X_1=k\right\}\cdot\mathbb{P}\left\{X_1=k \mid X_0=i\right\}\\
&= \sum_{k\neq j}\mathbb{P}\left\{X_{l+1}=j, X_t\neq j \mid X_1=k\right\}\cdot\mathbb{P}\left\{X_1=k \mid X_0=i\right\}\\
&= \sum_{k\neq j}f_{kj}^{(l)}p_{ik}^{(1)} = \sum_{k\neq j}f_{ik}^{(1)}f_{kj}^{(l)}
\end{aligned}
$$

□

引理 5.2.11. 对任意状态 i,j 和 $1\leqslant n<+\infty$，有

$$
p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)}p_{jj}^{(n-l)} \tag{5.3}
$$

证明. 用数学归纳法和 C-K 方程及其变式容易得

当 $n=1$ 时，右边 $=f_{ij}^{(1)}p_{jj}^{(0)}=p_{ij}^{(1)}=$ 左边；

当 $n=2$ 时，右边 $=f_{ij}^{(1)}p_{jj}^{(1)}+f_{ij}^{(2)}p_{jj}^{(0)}=f_{ij}^{(1)}+f_{ij}^{(2)}=p_{ij}^{(2)}=$ 左边；

假设对 $n-1$，有 $p_{ij}^{(n-1)}=\sum_{l=1}^{n-1}f_{ij}^{(l)}p_{jj}^{(n-1-l)}$，由 C-K 方程，有

$$
\begin{aligned}
\text{左边} = p_{ij}^{(n)} &= \sum_{k\in\mathcal{S}} p_{ik}p_{kj}^{(n-1)}\\
&= p_{ij}^{(1)}p_{jj}^{(n-1)} + \sum_{k\in\mathcal{S},k\neq j} p_{ik}p_{kj}^{(n-1)}\\
&= f_{ij}^{(1)}p_{jj}^{(n-1)} + \sum_{k\in\mathcal{S},k\neq j} f_{ik}^{(1)}\left(\sum_{l=1}^{n-1} f_{kj}^{(l)}p_{jj}^{(n-1-l)}\right) \quad \left(\text{因为} f_{ik}^{(1)}=p_{ik}\right)\\
&= f_{ij}^{(1)}p_{jj}^{(n-1)} + \sum_{l=1}^{n-1}\left(\sum_{k\in\mathcal{S},k\neq j} p_{ik}f_{kj}^{(l)}\right)p_{jj}^{(n-1-l)} \quad \left(\text{因为} f_{ij}^{(l+1)}=\sum_{k\neq j}f_{ik}^{(1)}f_{kj}^{(l)}\right)\\
&= f_{ij}^{(1)}p_{jj}^{(n-1)} + \sum_{l=1}^{n-1} f_{ij}^{(l+1)}p_{jj}^{(n-1-l)}\\
&= f_{ij}^{(1)}p_{jj}^{(n-1)} + \sum_{l=2}^{n} f_{ij}^{(l)}p_{jj}^{(n-l)}\\
&= \sum_{l=1}^{n} f_{ij}^{(l)}p_{jj}^{(n-l)} = \text{右边}
\end{aligned}
$$

这个式子表明了 n 步转移概率 $p_{ij}^{(n)}$ 与 n 步首达概率 $f_{ij}^{(n)}$ 的关系，与 C-K 方程一起反映了 Markov 链中各种转移概率的关系，也和 C-K 方程一样，体现了“解构”的思想. □

定理 5.2.12. 状态 i 为常返状态，**当且仅当** $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}} = \infty$；

状态 i 为非常返状态时，**当且仅当** $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}} < +\infty$.

证明. 由(5.3)可知

$$\begin{aligned}\sum_{n=0}^{\infty} p_{ii}^{(n)} &= p_{ii}^{(0)} + \sum_{n=1}^{\infty}\left(\sum_{l=1}^{n} f_{ii}^{(l)} p_{ii}^{(n-l)}\right) \quad (\text{注意} n \geqslant l)\\ &= 1 + \sum_{l=1}^{\infty}\left(\sum_{n=l}^{\infty} f_{ii}^{(l)} p_{ii}^{(n-l)}\right) \quad (\text{令} m = n - l \geqslant 0)\\ &= 1 + \left(\sum_{l=1}^{\infty} f_{ii}^{(l)}\right)\left(\sum_{m=0}^{\infty} p_{ii}^{(m)}\right)\end{aligned}$$

因此

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

所以

$$i\text{是常返的} \Longleftrightarrow f_{ii} = 1 \Longleftrightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$$

$$i\text{是非常返的} \Longleftrightarrow f_{ii} < 1 \Longleftrightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}} < +\infty$$

□

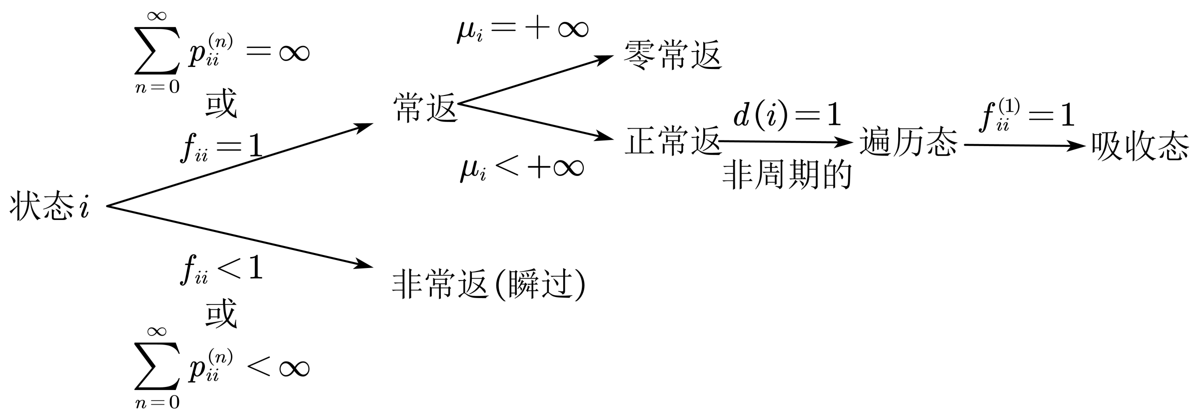

图 5.2. 状态的分类

引理 5.2.13. 若 $i \leftrightarrow j$，且 i 为常返状态 $(f_{ii} = 1)$，则 $f_{ji} = 1$.

定理 5.2.14. 常返性是一个类性质.

证明. 本节只能证明若 $i \leftrightarrow j$，则 i, j 同为常返或非常返.

由 $i \leftrightarrow j \Longrightarrow \exists m, n \geqslant 0,$ 使得 $p_{ij}^{(n)} > 0, p_{ji}^{(m)} > 0.$

由 C-K 方程，有

$$p_{ii}^{(n+m+l)} \geqslant p_{ij}^{(n)} p_{jj}^{(l)} p_{ji}^{(m)} p_{jj}^{(n+m+l)} \geqslant p_{ji}^{(m)} p_{ii}^{(l)} p_{ij}^{(n)}$$

分别关于 l 求和，得

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

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

可见，$\sum_{l=0}^{\infty} p_{jj}^{(l)}$ 和 $\sum_{l=0}^{\infty} p_{ii}^{(l)}$ 相互控制，同为无穷或有限. 即，

- $(*)$ 为无穷 $\Longrightarrow (\Delta\Delta)$ 为无穷 $\Longrightarrow (\Delta)$ 为无穷；
- (Δ) 为无穷 $\Longrightarrow (**)$ 为无穷 $\Longrightarrow (*)$ 为无穷；

反之，

- $(*)$ 为有限 $\Longrightarrow (**)$ 为有限 $\Longrightarrow (\Delta)$ 为有限；
- (Δ) 为有限 $\Longrightarrow (\Delta\Delta)$ 为有限 $\Longrightarrow (*)$ 为有限.

□

注记 5.2.15. 说明。

- 同属一类的状态 i, j，它们同为常返状态或非常返状态.
- 当同属一类的状态 i, j 同为常返状态时，它们又同为正常返和零常返状态.

注记 5.2.16. Stirling 公式：$n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \quad (n \to +\infty)$

5.3 极限定理及平稳分布

例子 5.3.1. 设 Markov 链有两个状态 1 和 2，它的转移矩阵为

$$\boldsymbol{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, 0 < p, q, p+q < 1$$

求 $\lim_{n\to\infty} \boldsymbol{P}^{(n)}$ 与回转的平均步数 μ_1, μ_2，并判断状态 1 和 2 的常返性.

解答. 由 $\boldsymbol{P}^{(n)} = \boldsymbol{P}^n$，考虑将矩阵 $\boldsymbol{P}$ 作分解为 $\boldsymbol{P} = \boldsymbol{QDQ}^{-1}$.

$$|\boldsymbol{P} - \lambda \mathbb{E}| = \begin{vmatrix} 1-p-\lambda & p \\ q & 1-q-\lambda \end{vmatrix} = (1-p-\lambda)(1-q-\lambda) - pq$$

$$= \lambda^2 + (p+q-2)\lambda + 1 - p - q = (\lambda - 1)[\lambda - (1-p-q)]$$

解得特征值为 $\lambda_1=1,\lambda_2=1-p-q$.

对 $\lambda_1=1$，$\boldsymbol{P}-\lambda_1\mathbb{E}=\begin{pmatrix}-p & p\\ q & -q\end{pmatrix}\Longrightarrow\boldsymbol{\xi}_1=\begin{pmatrix}1\\1\end{pmatrix}$；

对 $\lambda_2=1-p-q$，$\boldsymbol{P}-\lambda_2\mathbb{E}=\begin{pmatrix}q & p\\ q & p\end{pmatrix}\Longrightarrow\boldsymbol{\xi}_2=\begin{pmatrix}p\\-q\end{pmatrix}$；

所以 $\boldsymbol{Q}=\begin{pmatrix}1 & p\\ 1 & -q\end{pmatrix}$，而$\boldsymbol{D}=\boldsymbol{Q}^{-1}\boldsymbol{P}\boldsymbol{Q}=\begin{pmatrix}1 & 0\\ 0 & 1-p-q\end{pmatrix}$，因此

$$\boldsymbol{Q}^{-1}=\frac{\boldsymbol{Q}^*}{|\boldsymbol{Q}|}=\frac{1}{-(p+q)}\begin{pmatrix}-q & -1\\ -p & 1\end{pmatrix}^T=\frac{1}{p+q}\begin{pmatrix}q & p\\ 1 & -1\end{pmatrix}$$

进而有

$$\begin{aligned}\boldsymbol{P}^{(n)}=\boldsymbol{P}^n=\boldsymbol{Q}\boldsymbol{D}^n\boldsymbol{Q}^{-1}&=\frac{1}{p+q}\begin{pmatrix}1 & p\\ 1 & -q\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & (1-p-q)^n\end{pmatrix}\begin{pmatrix}q & p\\ 1 & -1\end{pmatrix}\\ &=\frac{1}{p+q}\begin{pmatrix}q+p(1-p-q)^n & p-p(1-p-q)^n\\ q-q(1-p-q)^n & p+q(1-p-q)^n\end{pmatrix}\end{aligned}$$

取极限得 $\lim_{n\to\infty}\boldsymbol{P}^{(n)}=\lim_{n\to\infty}\boldsymbol{P}^n=\frac{1}{p+q}\begin{pmatrix}q & p\\ q & p\end{pmatrix}$.

从 1 出发返回 1 的首达概率 $f_{11}^{(1)}=1-p$，$f_{11}^{(2)}=pq$，容易得到，当 $n\geqslant 2$ 时，有 $f_{11}^{(n)}=p(1-q)^{n-2}q$. 所以

$$f_{11}=\sum_{n=1}^{\infty}f_{11}^{(n)}=1-p+\sum_{n=2}^{\infty}p(1-q)^{n-2}q=1-p+pq\cdot\frac{1}{q}=1$$

因此状态 1 是常返的，而 $1\leftrightarrow 2$，所以 2 也是常返的.

再计算平均步数，有

$$\mu_1=\sum_{n=1}^{\infty}nf_{11}^{(n)}=1-p+\sum_{n=2}^{\infty}np(1-q)^{n-2}q=1-p+pq\sum_{n=0}^{\infty}\left[n(1-q)^n+2(1-q)^n\right]$$

这是一个幂级数求和，由 $\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n,x\in(-1,1)$, 易知 $\frac{x}{(1-x)^2}=\sum_{n=0}^{\infty}nx^n,x\in(-1,1)$, 所以

$$\mu_1=1-p+pq\left(\frac{1-q}{q^2}\right)+2pq\frac{1}{q}=1-p+\frac{p-pq}{q}+2p=1+\frac{p}{q}<+\infty$$

因此状态 1 和 2 都是正常返的，而且 $d(1)=1$，所以他们都是遍历态（另外，$f_{11}^{(1)}\neq 1$，可以说明 1 是非吸收态).

同样地，容易得到从 2 出发返回 2 的首达概率 $f_{22}^{(1)}=1-q$，当 $n\geqslant 2$ 时，有 $f_{22}^{(n)}=q(1-p)^{n-2}p$.

$$f_{22}=\sum_{n=1}^{\infty}f_{22}^{(n)}=1-q+\sum_{n=2}^{\infty}q(1-p)^{n-2}p=1-q+qp\cdot\frac{1}{p}=1$$

所以对平均步数，有

$$\begin{aligned}\mu_2 &= \sum_{n=1}^{\infty} n f_{22}^{(n)} = 1-q+\sum_{n=2}^{\infty} nq\,(1-p)^{n-2}\,p \\ &= 1-q+qp\sum_{n=0}^{\infty}\left[n\,(1-p)^n+2\,(1-p)^n\right] \\ &= 1-q+qp\frac{1-p}{p^2}+2qp\frac{1}{p}=1+\frac{q}{p}<+\infty\end{aligned}$$

仅从状态 2 的计算也可以得出，状态 2 是常返中的正常返，并且是遍历的非吸收态. □

定理 5.3.2. 若状态 i 是周期 d 的常返状态，则

$$\lim_{n\to\infty} p_{ii}^{(nd)}=\frac{d}{\mu_i}$$

特别地，当 $\mu_i=+\infty$ 时，$\lim_{n\to\infty} p_{ii}^{(nd)}=\frac{d}{\mu_i}=0$.

证明. 不加证明地给出. □

推论 5.3.3. 在状态 i 是常返状态的条件下，状态 i 是零常返状态，**当且仅当**

$$\lim_{n\to\infty} p_{ii}^{(n)}=0$$

证明. i 是零常返的 $\Longrightarrow \mu_i=+\infty$，当 $\frac{m}{d(i)}\in\mathbb{Z}$ 时，$\lim_{n\to\infty} p_{ii}^{(nd)}=\frac{d}{\mu_i}=0$，当 $\frac{m}{d(i)}\notin\mathbb{Z}$ 时，$p_{ii}^{(m)}=0$，综合来看 $\lim_{m\to\infty} p_{ii}^{(m)}=0$；

反之，如果 $\lim_{m\to\infty} p_{ii}^{(m)}=0$，用反证法，假设 i 是正常返状态，那么 $\mu_i<\infty$，当 $\frac{m}{d(i)}\in\mathbb{Z}$ 时，$\lim_{n\to\infty} p_{ii}^{(nd)}=\frac{d}{\mu_i}<\infty$，说明子列不收敛于 0. □

推论 5.3.4. 若状态 i 是非常返状态，则 $\sum_{n=1}^{\infty} p_{ii}^{(n)}<\infty$，那么

$$\lim_{n\to\infty} p_{ii}^{(n)}=0$$

定理 5.3.5. 常返性是一个类性质.

证明. 定理 5.2.14 证明了若 $i\leftrightarrow j$，则 i,j 同为常返或非常返. 现在证明若 i,j 同为常返状态，则 i,j 同为正常返或零常返.

若 i 为零常返状态，那么 $\lim_{n\to\infty} p_{ii}^{(n)}=0$，而由 C-K 方程，

$$p_{ii}^{(n+m+l)}\geqslant p_{ij}^{(n)}p_{jj}^{(m)}p_{ji}^{(l)}\geqslant 0$$

令 $m\to+\infty$，则

$$\lim_{m\to\infty} p_{ii}^{(n+m+l)}\geqslant p_{ij}^{(n)}p_{ji}^{(l)}\left(\lim_{m\to\infty} p_{jj}^{(m)}\right)\geqslant 0$$

那么 $\lim_{n\to\infty} p_{jj}^{(n)}=0$，所以 j 也是零常返状态.

同理，如果 j 是零常返状态，可以知道 i 也是零常返状态.

注意到常返状态只分为正常返和零常返，再根据逆否命题与原命题的等价性，命题得证.

□

定理 5.3.6. 若 j 为非常返状态或零常返状态，则 $\forall i \in \mathcal{S}$，有

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0$$

其次，若 j 是周期为 d 的正常返状态，则 $\forall i \leftrightarrow j, i \in \mathcal{S}$，有

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{d}{\mu_j}$$

证明. 分两部分证明.

1. 首先，若 j 为非常返状态或零常返状态，由(5.3)知道 $p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$ 对 $N < n$ 有

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

   先固定 N，令 $n \to \infty$，由于 $p_{jj}^{(n-l)} \to 0$，所以 $\sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} \to 0$，那么

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

   再令 $N \to \infty$，由于 $f_{ij} = \sum_{l=1}^{\infty} f_{ij}^{(l)} \leqslant 1$，由 Cauchy 收敛准则，这个级数的尾部一定收敛于 0，即 $\lim_{n\to\infty} \sum_{l=N+1}^{\infty} f_{ij}^{(l)} = 0$，再由两边夹法则，有 $\lim_{n\to\infty} p_{ij}^{(n)} = 0$.

2. 其次，若 j 是周期为 d 的正常返状态，对 $1 \leqslant N < n$ 有

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

   先固定 N，令 $n \to \infty$，由 $\lim_{n\to\infty} \sum_{l=N+1}^{\infty} f_{ij}^{(l)} = 0$，可知

$$\lim_{n\to\infty} \sup \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} \leqslant \lim_{n\to\infty} \sup p_{ij}^{(n)} \leqslant \lim_{n\to\infty} \sup \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)}$$

   结合 $\lim_{n\to\infty} p_{jj}^{(nd)} = \frac{d}{\mu_j}$ ，可知

$$\frac{d}{\mu_j} \sum_{l=1}^{N} f_{ij}^{(l)} \leqslant \lim_{n\to\infty} p_{ij}^{(n)} \leqslant \frac{d}{\mu_j} \sum_{l=1}^{N} f_{ij}^{(l)}$$

   再令 $N \to \infty$，因为 i 与 j 互通，而且 j 是正常返的，再根据引理 5.2.13，有

$$f_{ij} = \sum_{l=1}^{\infty} f_{ij}^{(l)} = 1$$

   所以 $\lim_{n\to\infty} p_{ij}^{(n)} = \frac{d}{\mu_j}$.

□

推论 5.3.7. [2] $\forall i,j\in\mathcal{S}$，有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}=\begin{cases}0, & j\text{为非常返或零常返状态}\\ \frac{d}{\mu_j}, & j\text{为正常返状态}, i\leftrightarrow j\end{cases}$$

推论 5.3.8. 有限状态的 Markov 链，不可能全为非常返状态，也不可能有零常返状态.

证明.（反证法）设状态空间为 $\mathcal{S}=\{1,2,\cdots,N\}$.

1. 假设 N 个状态都是非常返的，对状态 i：
   (a) 如果 $i\to j$，可知 $\lim_{n\to\infty}p_{ij}^{(n)}=0$；
   (b) 如果 $i\nrightarrow j$，可知 $\forall n, p_{ij}^{(n)}=0$. 因此 $\lim_{n\to\infty}\sum_{j=1}^{N}p_{ij}^{(n)}=0$，这与 $\sum_{j=1}^{N}p_{ij}^{(n)}=1$ 矛盾.
2. 假设 $\mathcal{S}$ 中有零常返状态, 设为 i, 令 $\mathcal{C}=\{j:i\to j\}$ 为 i 的可达状态集, $\Longrightarrow\sum_{j\in\mathcal{C}}p_{ij}^{(n)}=1$, 而且 $\forall j\in\mathcal{C}$, $j\to i$, 否则从 i 出发不能有限步回到 i, 这与 i 是常返状态矛盾 $\Longrightarrow i\leftrightarrow j$, 由于常返性是类性质 $\Longrightarrow j$ 也是零常返. 那么 $\lim_{n\to\infty}p_{ij}^{(n)}=0\Longrightarrow\lim_{n\to\infty}\sum_{j\in\mathcal{C}}p_{ij}^{(n)}=0$，这与 $\sum_{j\in\mathcal{C}}p_{ij}^{(n)}=1$ 矛盾.

□

推论 5.3.9. 不可约 (只有一个类) 的有限 Markov 链是正常返的.

推论 5.3.10. 若 Markov 链有一个零常返状态，则必有无限个零常返状态.

定义 5.3.11. 对于 Markov 链，概率分布 $\{p_j,j\in\mathcal{S}\}$ 称为**平稳分布**，如果

$$p_j=\sum_{i\in\mathcal{S}}p_ip_{ij}$$

定理 5.3.12. 若 Markov 链的初始分布 $\mathbb{P}\{X_0=j\}=p_j$ 是平稳分布，那么从此以后，分布都是平稳的.

证明. 因为初始分布是平稳分布，所以 $\mathbb{P}\{X_0=j\}=p_j=\sum_{i\in\mathcal{S}}p_ip_{ij}$，于是

$$\mathbb{P}\{X_1=j\}=\sum_{i\in\mathcal{S}}\mathbb{P}\{X_1=j|\,X_0=i\}\cdot\mathbb{P}\{X_0=i\}=\sum_{i\in\mathcal{S}}p_ip_{ij}=p_j$$

以此类推，可以知道 $X_n,n=0,1,2,\cdots$ 的分布都是相同的. □

定义 5.3.13. 如果所有状态都互通而且周期均为 1，那么称 Markov 链是**遍历**的.

[2]该推论容易由数列的极限性质推得，数列 $\{a_n\}$ 的极限存在且为 a，那么数列 $\left\{\frac{a_1+a_2+\cdots+a_n}{n}\right\}$ 的极限也存在且为 a.

定义 5.3.14. 对于**遍历**的 Markov 链，极限

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

称为 Markov 链的**极限分布**.

定理 5.3.15. 对于不可约（1 个类）且非周期 ($d=1$) 的 Markov 链，

1. 若它是遍历的，则极限分布 $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j}\,(j \in \mathcal{S})$ 就是平稳分布，并且是唯一的平稳分布；
2. 若状态**都是**非常返的或**都是**零常返的，则平稳分布不存在.

证明. 分情况讨论.

1. 若它是遍历的. 首先可知极限分布存在.

   (a) 先证明它是平稳分布：

   $$\sum_{j\in\mathcal{S}} p_{ij}^{(n)} = 1 \Longrightarrow \lim_{n\to\infty} \sum_{j\in\mathcal{S}} p_{ij}^{(n)} = 1 \Longrightarrow \sum_{j\in\mathcal{S}} \lim_{n\to\infty} p_{ij}^{(n)} = 1 \Longrightarrow \sum_{j\in\mathcal{S}} \pi_i = 1$$

   再由 C-K 方程，有 $p_{ij}^{(n+1)} = \sum_{k\in\mathcal{S}} p_{ik}^{(n)} p_{kj}$，两边取极限

   $$\lim_{n\to\infty} p_{ij}^{(n+1)} = \lim_{n\to\infty} \sum_{k\in\mathcal{S}} p_{ik}^{(n)} p_{kj} = \sum_{k\in\mathcal{S}} \left(\lim_{n\to\infty} p_{ik}^{(n)}\right) p_{kj}$$

   这与 $\pi_j = \sum_{k\in\mathcal{S}} \pi_k p_{kj}$ 等价，说明了分布 π_j 的平稳性.

   (b) 再证明平稳分布的唯一性：

   如果除了 π_j 之外，还有一个平稳分布 $\tilde{\pi}_j = \sum_{k\in\mathcal{S}} \tilde{\pi}_k p_{kj}$，归纳可知

   $$\tilde{\pi}_j = \sum_{k\in\mathcal{S}} \tilde{\pi}_k p_{kj}^{(n)}, n = 1, 2, \cdots$$

   两端取极限，有 $\tilde{\pi}_j = \sum_{i\in\mathcal{S}} \tilde{\pi}_i \left(\lim_{n\to\infty} p_{ij}^{(n)}\right) = \sum_{i\in\mathcal{S}} \tilde{\pi}_i \pi_j = \pi_j \sum_{i\in\mathcal{S}} \tilde{\pi}_i = \pi_j$.

2. 若它的状态都是非常返的或都是零常返的.

   假设存在一个平稳分布 $\{\pi_j, j \in \mathcal{S}\}$ 则

   $$\pi_j = \sum_{i\in\mathcal{S}} \pi_i p_{ij}^{(n)}, n = 1, 2, \cdots \tag{5.4}$$

   由于 $\lim_{n\to\infty} p_{ij}^{(n)} = 0$，对(5.4)取极限后 $\pi_j = \sum_{i\in\mathcal{S}} \pi_i \left(\lim_{n\to\infty} p_{ij}^{(n)}\right) = 0$，这显然不满足概率的归一性，于是假设不成立.

□

注记 5.3.16. 极限分布定理的应用：对于有限状态且遍历的 Markov 链，可以容易根据方程 $\pi_j = \sum_{i\in\mathcal{S}} \pi_i p_{ij}$，解得 $\pi_j = \frac{1}{\mu_j}$，从而求出平均回转时间 μ_i.

例子 5.3.17. 转移矩阵为 $\boldsymbol{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$，容易知道，它是有限状态且遍历的，所以要求 μ_1 和 μ_2，需解线性方程组

$$\boldsymbol{P}^T\boldsymbol{\pi} = \boldsymbol{\pi} \Longleftrightarrow \begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix}^T \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

可以解得 $p\pi_1 = q\pi_2$，再根据规范性条件 $\pi_1 + \pi_2 = 1$，所以 $\pi_1 = \frac{q}{p+q}, \pi_2 = \frac{p}{p+q} \Longrightarrow \mu_1 = 1 + \frac{p}{q}, \mu_2 = 1 + \frac{q}{p}$.

5.4 Markov 链的应用

定义 5.4.1. **分支过程**指的是一个 Markov 链 $\{X_n, n = 0, 1, 2, \cdots\}$. 它描述的是一个能产生同类后代的个体组成的群体，每一个个体生命结束后以概率 $p_j\,(j = 0, 1, 2, \cdots)$ 产生 j 个新的后代，且个体间相互独立. X_0 表示初始的个体总数，$X_n\,(n = 1, 2, \cdots)$ 表示第 n 代个体的总数.

定义 5.4.2. **单个祖先的分支过程**：即 $X_0 = 1$，那么

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

其中 $Z_{n-1,i}$ 表示第 n 代的第 i 个成员的后代个数. 以下考虑的都是单个祖先的分支过程.

注记 5.4.3. 第 n 代的平均个体数：

$$\mathbb{E}\left(X^n\right) = \mathbb{E}\left[\mathbb{E}\left(X_n | X_{n-1}\right)\right] = \mu\mathbb{E}\left(X_{n-1}\right) = \mu^2\mathbb{E}\left(X_{n-2}\right) = \cdots = \mu^n$$

其中 $\mu = \sum_{i=0}^{\infty} ip_i$ 是每个个体后代数目的均值.

1. 若 $\mu < 1$，则平均个体数单调递减趋于零；
2. 若 $\mu = 1$，各代平均个体数相同；
3. 若 $\mu > 1$，平均个体数单调递增至无穷.

定理 5.4.4. (关于群体消亡概率 π_0 的定理) $0 < p_0 < 1$，则 $\pi_0 = 1 \Longleftrightarrow \mu \leqslant 1$.（不考虑平凡情况，因为当 $p_0 = 1$ 时，群体在第 0 代后就消失；而 $p_0 = 0$ 时，群体永不消失）

证明. 对第 1 代个体数取条件概率，因为群体的最终灭绝是以第 1 代为祖先的 j 个家族全部消亡为前提的，而 j 个家族互相独立，每一个家族灭绝的概率均为 π_0，所以

$$\pi_0 = \mathbb{P}\{\text{群体消亡}\} = \sum_{j=1}^{\infty} \mathbb{P}\{\text{群体消亡} | X_1 = j\} \times p_j = \sum_{j=0}^{\infty} \pi_0^j p_j$$

令 $F(\pi_0) := \sum_{j=0}^{\infty} \pi_0^j p_j$，显然，$F(\pi_0) = \pi_0$ 是 $y = x$ 与 $y = F(x)$ 的交点横坐标. 实际上，有那么很显然，$F(1) = \sum_{j=0}^{\infty} p_j = 1$，即 $(1,1)$ 是交点.

1. 当 $p_0 + p_1 = 1$ 时，$F(\pi_0) = p_0 + \pi_0 p_1$ 是一条直线.
2. 当 $p_0 + p_1 < 1$ 时，有

$$F(0) = p_0$$
$$F'(x) = \sum_{j=1}^{\infty} j x^{j-1} p_j > 0, 0 < x < 1$$
$$F''(x) = \sum_{j=1}^{\infty} j(j-1) x^{j-2} p_j > 0, 0 < x < 1$$

因此，$F(x)$ 是过点 $(0, p_0)$、单调增加的凸函数.

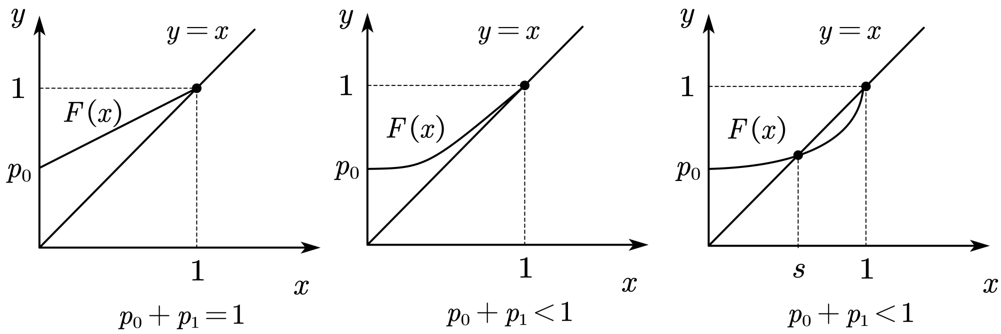

图 5.3. 三种情况

1. 当 $p_0 + p_1 = 1$ 时，$F(x) = x$ 只有一个解 $\pi_0 = 1$，此时家族一定灭绝.
2. 当 $p_0 + p_1 < 1$ 时，有两种情况：
   (a) $\forall s \in (0,1), F(s) < s$，此时 $F(x) = x$ 仅有唯一解；
   (b) $\exists s \in (0,1), F(s) = s$，此时 $F(x) = x$ 有两个解 s 或 1.

   下面证明在第 2 种情况下，π_0 必定取值为 s. 要证 $\pi_0 = s$，只需证明 π_0 的取值是方程 $F(x) = x$ 的最小解. 利用归纳法，设方程的解为 π，当 $n = 1$ 时，有

$$\pi = \sum_{j=0}^{\infty} \pi^j p_j \geqslant \pi^0 p_0 = p_0 = \mathbb{P}\{X_1 = 0\}$$

   假设 $\pi \geqslant \mathbb{P}\{X_n = 0\}$，则

$$\mathbb{P}\{X_{n+1} = 0\} = \sum_{j=0}^{\infty} \mathbb{P}\{X_{n+1} = 0 | X_1 = j\} \times p_j = \sum_{j=0}^{\infty} [\mathbb{P}\{X_n = j\}]^j \times p_j \leqslant \sum_{j=0}^{\infty} \pi_0^j \times p_j = \pi$$

   所以，对一切 n，有 $\pi \geqslant \mathbb{P}\{X_n = 0\}$，两边对 n 取极限得

$$\pi \geqslant \lim_{n\to\infty} \mathbb{P}\{X_n = 0\} = \mathbb{P}\{群体最终灭绝\} = \pi_0$$

   因此 π_0 的取值应该为 s.

下面讨论 μ 的情况：

1. 当 $p_0+p_1=1$ 时，$\mu=\sum_{j=0}^{\infty}jp_j=p_1<1$；
2. 当 $p_0+p_1<1$ 时，有两种情况：
   (a) $\pi_0=1$ 时，由图形可知 $F'(1)\leqslant 1$，而 $F'(1)=\sum_{j=0}^{\infty}jp_j=\mu$，所以此时 $\mu\leqslant 1$；
   (b) $\pi_0=s$ 时，易得 $F'(1)=\mu>1$.

综上所述，有

$$\pi_0=1\Longleftrightarrow\mu\leqslant 1$$

这就证明了该定理.

□

5.5 连续时间 Markov 链

注记 5.5.1. 前面几节讨论的是离散时间和离散状态空间的 Markov 链，本节讨论的是连续时间和离散状态空间的 Markov 链.

待完善.

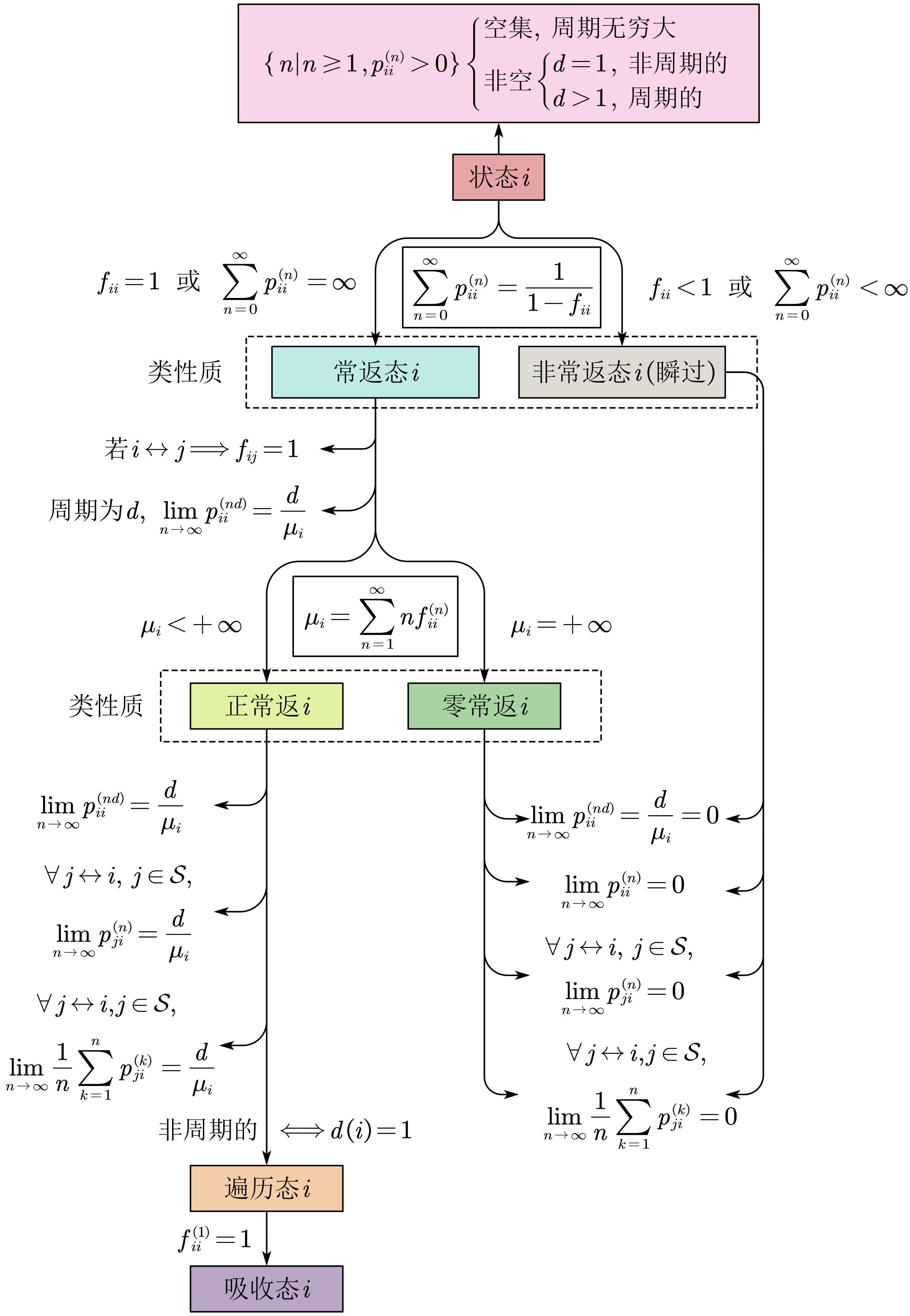

图 5.4. Markov 链的状态分类及其性质

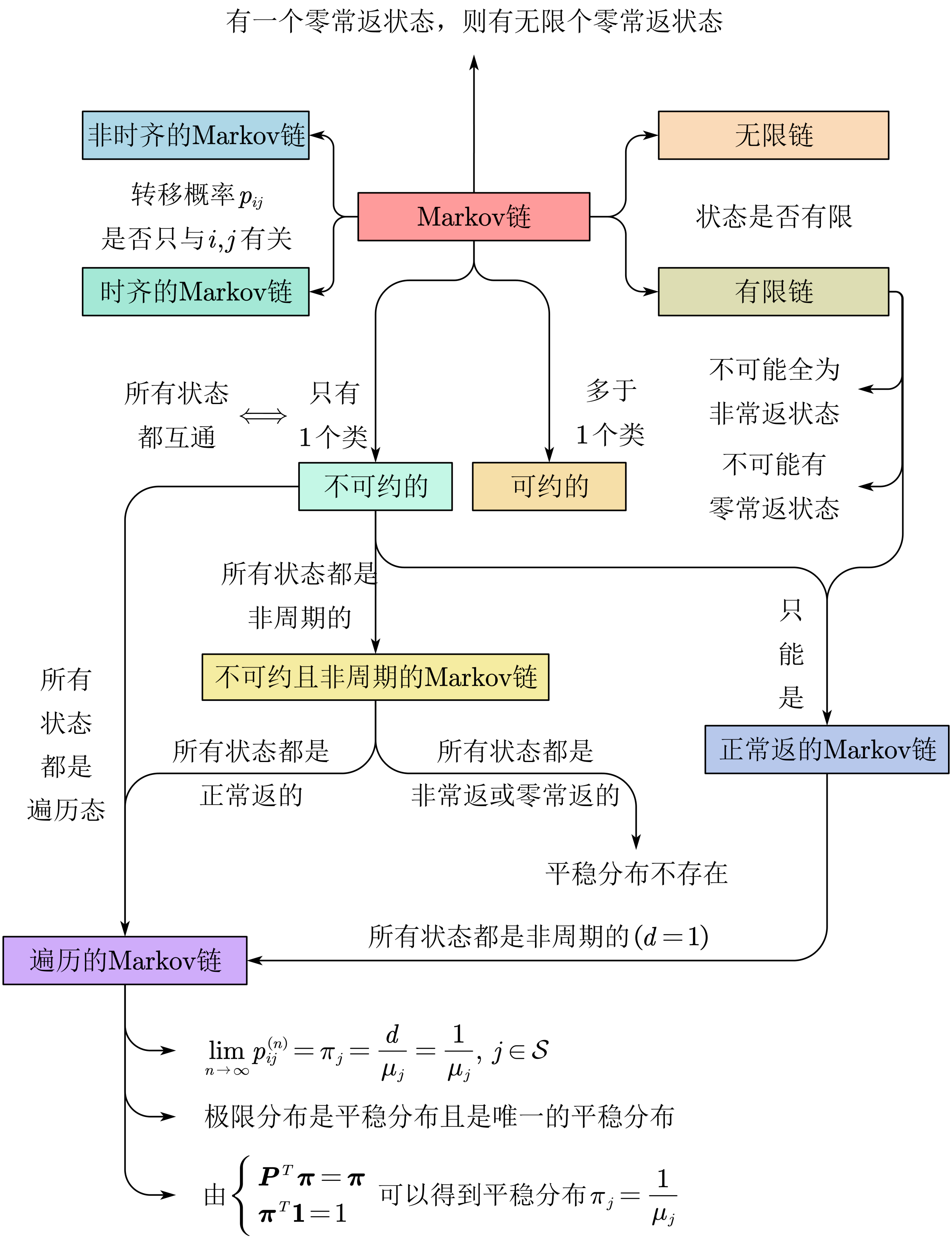

图 5.5. Markov 链的分类及其性质

第六章　鞅

本章介绍另一类在金融中广泛应用的随机过程——鞅. 要求理解鞅的基本概念（包括上鞅和下鞅）；掌握停时的概念、鞅的停时定理；了解一致可积性；理解鞅收敛定理；掌握连续鞅的定义积相关性质.

6.1　基本概念

注记 6.1.1. “公平”的赌博：意味着每次赌博的输赢机会是均等的，并且赌博策略是依赖于前面的赌博结果的.

任何赌博者都不能改变赌博策略使得“公平”的赌博变成有利于自己的赌博.

$$\mathbb{E}\left[X_{n+1}|Y_1, Y_2, \cdots, Y_n\right] = X_n$$

定义 6.1.2. 下鞅、上鞅和鞅.

1. 随机过程 $\{X_n, n \geqslant 0\}$ 称为**关于 $\{Y_n, n \geqslant 0\}$ 的下鞅**（sub-martingale），如果对于 $n \geqslant 0$ 满足：
   (a) X_n 由 $(Y_0, Y_1, \cdots, Y_n)$ 决定；
   (b) $\mathbb{E}\left[X_j^+\right] < \infty$，其中 $X_n^+ = \max\{0, X_n\}$；
   (c) $\mathbb{E}\left[X_{n+1}|Y_1, Y_2, \cdots, Y_n\right] \geqslant X_n$.
2. 上鞅的定义：随机过程 $\{X_n, n \geqslant 0\}$ 称为**关于 $\{Y_n, n \geqslant 0\}$ 的上鞅**（sup-martingale），如果对于 $n \geqslant 0$ 满足：
   (a) X_n 由 $(Y_0, Y_1, \cdots, Y_n)$ 决定；
   (b) $\mathbb{E}\left[X_j^-\right] < \infty$，其中 $X_n^- = \max\{0, -X_n\}$；
   (c) $\mathbb{E}\left[X_{n+1}|Y_1, Y_2, \cdots, Y_n\right] \leqslant X_n$.
3. 随机过程 $\{X_n, n \geqslant 0\}$ 称为**关于 $\{Y_n, n \geqslant 0\}$ 的鞅**（martingale），如果 $\{X_n, n > 0\}$ 既是 $\{Y_n, n > 0\}$ 的上鞅也是它的下鞅.

注记 6.1.3. 赌博的公平与鞅.

1. 下鞅：有利于自己的赌博策略；
2. 上鞅：不利于自己的赌博策略；
3. 鞅：公平的赌博，平均而言，赌博者在下一次赌博结束时的赌资（随机变量）等于现在的赌资. .

定义 6.1.4. 在完备的概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上,我们称 $\{\mathcal{F}_n, n \geqslant 0\}$ 是σ **子代数流**,如果 $\{\mathcal{F}_n, n \geqslant 0\}$ 是 $\mathcal{F}$ 上的一列 σ 子代数，并且使得 $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n, n \geqslant 0$.

注记 6.1.5. σ 代数流是元素为集合族（$\mathcal{F}$）的集合族. [1]

$$
\begin{array}{ccccccccc}
\mathcal{F}_n & \Longrightarrow & \sigma\{Y_0\} & \subset & \sigma\{Y_0, Y_1\} & \subset \cdots \subset & \sigma\{Y_1, Y_2, \cdots, Y_n\} \\
 & & \downarrow & & \downarrow & & \downarrow \\
\{\mathcal{F}_n,\ n \geqslant 0\} & \Longrightarrow & \{\mathcal{F}_1 & \subset & \mathcal{F}_2 & \subset \cdots \subset & \mathcal{F}_n\}
\end{array}
$$

图 6.1. σ 代数流

定义 6.1.6. 对随机过程 $\{X_n, n \geqslant 0\}$，如果 $\forall x \in \mathbb{R}, \{X_n \leqslant x\} \in \mathcal{F}_n$，那么称 $\{X_n, n \geqslant 0\}$ 是 $\mathcal{F}_n$**可测的**.

定义 6.1.7. 随机过程 $\{X_n, n \geqslant 0\}$ 称为 $\{\mathcal{F}_n, n \geqslant 0\}$**适应的**，如果对于 $\forall n \geqslant 0, X_n$ 是 $\mathcal{F}_n$ 可测的（或称 $\{X_n, \mathcal{F}_n, n \geqslant 0\}$ 是**适应列**).

注记 6.1.8. 将"X_n由 $(Y_0, Y_1, \cdots, Y_n)$ 决定"或"X_n是 $(Y_0, Y_1, \cdots, Y_n)$ 的函数"重新定义:$\{X_n, n \geqslant 0\}$ 是 $\{\mathcal{F}_n, n \geqslant 0\}$ 适应的.（令 $\mathcal{F}_n = \sigma\{Y_1, Y_2, \cdots, Y_n\}, n \geqslant 0$，则 $\{\mathcal{F}_n, n \geqslant 0\}$ 是一个 σ 子代数流）

定义 6.1.9. 设 $\{\mathcal{F}_n, n \geqslant 0\}$ 是 $\mathcal{F}$ 中的单调递增的子 σ 代数列，随机过程 $\{X_n, n \geqslant 0\}$ 称为**关于 $\{Y_n, n \geqslant 0\}$ 的鞅**，如果它满足：

1. $\{X_n, n \geqslant 0\}$ 是 $\{\mathcal{F}_n, n \geqslant 0\}$ 适应的；
2. $\mathbb{E}[|X_n|] < \infty$；
3. $\forall n \geqslant 0, \mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$.

定理 6.1.10. 关于鞅的几个结论（试证）：

- 适应列 $\{X_n, \mathcal{F}_n, n \geqslant 0\}$ 是下鞅 $\Longleftrightarrow \{-X_n, \mathcal{F}_n, n \geqslant 0\}$ 是上鞅.
- 如果 $\{X_n, \mathcal{F}_n\}, \{Y_n, \mathcal{F}_n\}$ 是两个下鞅 $\Longrightarrow \{aX_n + bY_n, \mathcal{F}_n\}$ 也是下鞅.

[1]一个集合，集合的元素也是集合，即元素为集合的集合

- 如果 $\{X_n, \mathcal{F}_n\}, \{Y_n, \mathcal{F}_n\}$ 是两个下鞅(或上鞅)$\Longrightarrow \{\max(X_n, Y_n), \mathcal{F}_n\}$(或 $\{\min(X_n, Y_n), \mathcal{F}_n\}$)也是上鞅（或下鞅）.

定理 6.1.11. (条件 Jensen 不等式) 设 $\varphi(x)$ 为实数集 $\mathbb{R}$ 上的凸函数，[2]随机变量 M 满足

1. $\mathbb{E}[|M|] < \infty$;
2. $\mathbb{E}[|\varphi(M)|] < \infty$.

那么 $\mathbb{E}[\varphi(M)|\mathcal{F}_n] \geqslant \varphi[\mathbb{E}(M|\mathcal{F}_n)]$，其中 $\mathcal{F}_n$ 是任意递增的 σ 代数列.

注记 6.1.12. 条件 Jensen 不等式的特殊情况：如果 $\varphi(x) = x^2$，那么 Jensen 不等式的形式是 $\mathbb{E}[X^2|\mathcal{F}_n] \geqslant [\mathbb{E}(X|\mathcal{F}_n)]^2$ 或 $\mathrm{Var}(X|\mathcal{F}_n) \geqslant 0$.

注记 6.1.13. 利用条件 Jensen 不等式构造新的下鞅：设 $\{M_n, n \geqslant 0\}$ 是关于 $\{\mathcal{F}_n, n \geqslant 0\}$ 的鞅（或下鞅），$\varphi(x)$ 为 $\mathbb{R}$ 上的凸函数，且 $\forall n \geqslant 0, \mathbb{E}\left[|\varphi(M_n)^+|\right] < \infty$，那么 $\{\varphi(M_n), n \geqslant 0\}$ 是关于 $\{\mathcal{F}_n, n \geqslant 0\}$ 的鞅（或下鞅）.

特别地，

- $\{|M_n|, n \geqslant 0\}$ 是下鞅；
- 若 $\forall n \geqslant 0, \mathbb{E}[M_n^2] < \infty$，那么 $\{M_n^2, n \geqslant 0\}$ 是下鞅.

6.2 鞅的停时定理

注记 6.2.1. 我们知道鞅的性质：对于关于 $\{X_n, n \geqslant 0\}$ 的鞅 $\{M_n, n \geqslant 0\}$，$\forall n \geqslant 0, \mathbb{E}[M_n] = \mathbb{E}[M_0]$.

因为 $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$，对两边关于 $\mathcal{F}_n$ 取期望，那么 $\mathbb{E}[M_{n+1}] = \mathbb{E}[M_n]$，递推可得这一性质.

下面，鞅的停时定理目标是寻找使得 $\mathbb{E}[M_T] = \mathbb{E}[M_0]$ 成立的条件，其中 T 是随机变量，即在随机的时刻结束赌局，此时的赌本与赌博者开始时的赌本一样.

定义 6.2.2. 设 $\{X_n, n \geqslant 0\}$ 是一个随机变量序列，称**随机函数**T 是关于过程 $\{X_n, n \geqslant 0\}$ 的**停时**，如果

1. T 在 $\{0, 1, 2, \cdots, \infty\}$ 上取值；
2. 对每一个 $n \geqslant 0$，$\{T = n\} \in \sigma(X_0, X_1, \cdots, X_n)$.

注记 6.2.3. 停时概念的内涵：对于事件 $\{T = n\}$ 或 $\{T \neq n\}$ 都应该由 n 时刻及其之前的信息完全确定.

---

[2]定义在有穷域无穷开区间 I 上的函数 $\varphi(x)$，称它是**凸函数**，如果 $\forall x, y \in I, 0 < \alpha < 1$，有 $\alpha\varphi(x) + (1-\alpha)\varphi(y) \geqslant \varphi[\alpha x + (1-\alpha)x]$.

即赌博者决定何时结束赌博，仅仅以他过去的输赢来判断．而不能说：如果下一局赢了就停止赌博．

例子 6.2.4. 停时的几个例子．

- 确定时刻 $T=n$ 是一个停时．
- 首达时刻 $T(A)=\inf\{n,X_n\in A\}$[3]是一个停时，它表示 $\{X_n,n\geqslant 0\}$ 首次进入集合 A 的时刻．[4]
- 如果 T和S 是两个停时，那么 $T+S,\min(T,S),\max(T,S)$ 都是停时．

证明．T和S 是停时，那么 $T,S\in\mathcal{F}_n=\sigma(X_0,X_1,\cdots,X_n)$，由 $\mathcal{F}_n$ 的性质，[5]有

1. $\{T\leqslant n\}=\bigcup_{k=0}^{n}\{T=k\}\in\mathcal{F}_n$；
2. $\{T>n\}=\overline{\{T\leqslant n\}}\in\mathcal{F}_n$；
3. $\{T=n\}=\{T\leqslant n\}-\{T\leqslant n-1\}=\{T\leqslant n\}\bigcap\overline{\{T\leqslant n-1\}}\in\mathcal{F}_n$.

根据这种思路，可知 $\{T+S=n\}=\bigcup_{t=0}^{n}\{T=t\}\bigcup\{S=n-t\}\in\mathcal{F}_n$，所以 $T+S$ 是停时．

$$\begin{aligned}\{\min(T,S)=n&=\{n|n\leqslant T,n\leqslant S,S=n\text{ 或}T=n\}\\&=\{T\geqslant n\}\bigcap\{S\geqslant n\}\bigcap\left(\{S=n\}\bigcup\{T=n\}\right)\in\mathcal{F}_n\end{aligned}$$

$$\begin{aligned}\{\max(T,S)=n&=\{n|n\geqslant T,n\geqslant S,S=n\text{ 或}T=n\}\\&=\{T\leqslant n\}\bigcap\{S\leqslant n\}\bigcap\left(\{S=n\}\bigcup\{T=n\}\right)\in\mathcal{F}_n\end{aligned}$$

所以 $T+S,\min(T,S),\max(T,S)$ 都是停时． □

定理 6.2.5. (有界停时定理) 设 $\{M_n,n\geqslant 0\}$ 是一个关于 $\{X_n,n\geqslant 0\}$ 的鞅，T 是一个关于 $\{X_n,n\geqslant 0\}$ 的停时, 并且 $T\leqslant K$ (有界停时), 设 $\mathcal{F}_n=\sigma(X_0,X_1,\cdots,X_n)$, 则有 $\mathbb{E}[M_T|\mathcal{F}_0]=M_0$，等式两边取期望后，有 $\mathbb{E}[M_T]=\mathbb{E}[M_0]$.

证明．将 M_T 改写为

$$M_T=\sum_{j=0}^{K}M_j\mathbb{I}_{\{T=j\}}$$

---

[3]约定 $T(\oslash)=\inf\{n,X_n\in\oslash\}=\infty$.

[4]$\{T(A)=n\}=\{X_0\notin A,X_1\notin A,\cdots,X_{n-1}\notin A,X_n\in A\}$，所以 $\{T(A)=n\}$ 完全由 $X_0,X_1,\cdots,X_n$ 所决定，或者 $\{T(A)=n\}\in\sigma(X_0,X_1,\cdots,X_n)$.

[5]若 $A_0,A_1,\cdots,A_n\in\mathcal{F}$，那么它们的可列交和可列并也在 $\mathcal{F}$ 中；$\Omega\in\mathcal{F}$；如果 $A\in\mathcal{F}$，那么$\bar{A}\in\mathcal{F}$.

然后利用停时的结论以及 $\{T=K\} \Longleftrightarrow \{T>K-1\}$，可以得到

$$\begin{aligned}
\mathbb{E}\left[M_T|\mathcal{F}_{K-1}\right] &= \mathbb{E}\left[\sum_{j=0}^{K-1} M_j \mathbb{I}_{\{T=j\}}\middle|\mathcal{F}_{K-1}\right] + \mathbb{E}\left[M_K \mathbb{I}_{\{T=K\}}\middle|\mathcal{F}_{K-1}\right] \\
&= \sum_{j=0}^{K-1} M_j \mathbb{I}_{\{T=j\}} + \mathbb{I}_{\{T>K-1\}} \mathbb{E}\left[M_K|\mathcal{F}_{K-1}\right] \\
&= \sum_{j=0}^{K-1} M_j \mathbb{I}_{\{T=j\}} + \mathbb{I}_{\{T>K-1\}} M_{K-1} \\
&= \sum_{j=0}^{K-2} M_j \mathbb{I}_{\{T=j\}} + \mathbb{I}_{\{T>K-2\}} M_{K-1}
\end{aligned}$$

由于 $\mathcal{F}_{K-2} \subset \mathcal{F}_{K-1} \subset \mathcal{F}_K$，关于 $\mathcal{F}_{K-2}$ 取条件期望，[6]有

$$\begin{aligned}
\mathbb{E}\left[M_T|\mathcal{F}_{K-2}\right] &= \mathbb{E}\left[\mathbb{E}\left[M_T|\mathcal{F}_{K-1}\right]|\mathcal{F}_{K-2}\right] \\
&= \mathbb{I}_{\{T>K-3\}} M_{K-2} + \sum_{j=0}^{K-3} M_j \mathbb{I}_{\{T=j\}}
\end{aligned}$$

不断重复上述运算，有

$$\mathbb{E}\left[M_T|\mathcal{F}_0\right] = \mathbb{I}_{\{T\geqslant 0\}} M_0 = M_0$$

□

定理 6.2.6. (鞅停时定理) $\mathcal{F}_n = \sigma(X_0, X_1, \cdots, X_n)$，设 $\{M_n, n \geqslant 0\}$ 是一个关于 $\{X_n, n \geqslant 0\}$ 的鞅（或者说 $\{M_n, n \geqslant 0\}$ 是一个关于 $\{\mathcal{F}_n = \sigma(X_0, X_1, \cdots, X_n)\}$ 的鞅），T 是一个关于 $\{X_n, n \geqslant 0\}$ 的停时，且满足

1. $\mathbb{P}\{T<\infty\}=1$;
2. $\mathbb{E}[|M_T|]<\infty$;
3. $\lim_{n\to\infty} \mathbb{E}\left[|M_n| \mathbb{I}_{\{T>n\}}\right]=0$;

则有

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

注记 6.2.7. 鞅停时定理的确定思路：

1. 有界停时定理是鞅停时定理的一种特殊情况．有界停时定理中，存在确定的 K，使得 $\mathbb{P}\{T \leqslant K\}=1$(赌博者确定在某一局后一定会停止赌博);而鞅停时定理中,$\mathbb{P}\{T<\infty\}=1$，这只能保证过程以概率 1 停止（赌博者只能保证赌局不会无限期延续）.

[6]条件期望的性质：若 $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$，那么 $E\left[E\left[X|\mathcal{G}_2\right]\ |\mathcal{G}_1\right] = E\left[X|\mathcal{G}_1\right]$.

2. 将 M_T 与有界停时联系起来. 令 $T_n=\min\{T,n\}\leqslant n$，那么[7]

$$M_T = M_{T_n} + M_T\mathbb{I}_{\{T>n\}} - M_n\mathbb{I}_{\{T>n\}}$$

其中，显然 $T_n\leqslant n$ 是有界停时，且 $T_0=\min\{T,0\}=0$，由有界停时定理，有

$$\mathbb{E}\left[M_{T_n}\right] = \mathbb{E}\left[M_0\right]$$

3. 设定条件 $\mathbb{P}\{T<\infty\}=1$ 和 $\mathbb{E}\left[|M_T|\right]<\infty$ 使得第二项趋于 0.
   首先需要保证 $\mathbb{P}\{T<\infty\}=1 \Longrightarrow \lim_{n\to\infty}\mathbb{P}\{T>n\}=0$，并且要限制 $\mathbb{E}\left[M_T\right]$ 的范围，才能使得第二项收敛于 0.
4. 设定条件 $\lim_{n\to\infty}\mathbb{E}\left[|M_n|\,\mathbb{I}_{\{T>n\}}\right]=0$ 使得第三项趋于 0.

注记 6.2.8. 鞅停时定理的现实意义：在公平的博弈中, 无论使用何种赌博策略,[8] 赌博者都不可能赢.

定理 6.2.9. (另一种形式的鞅停时定理) $\{M_n, n\geqslant 0\}$ 是一个关于 $\{\mathcal{F}_n=\sigma\left(X_0,X_1,\cdots,X_n\right)\}$ 的鞅，T 是一个关于 $\{X_n,n\geqslant 0\}$ 的停时，且满足

1. 鞅本身取值有界，即存在 $K\geqslant 0$ 使得 $|M_n|<K$；
2. $\mathbb{P}\{T<\infty\}=1$；那么

$$\mathbb{E}\left[M_T\right] = \mathbb{E}\left[M_0\right]$$

定理 6.2.10. (上鞅的停时定理) 设 $\{M_n, n\geqslant 0\}$ 是一个关于 $\{X_n,n\geqslant 0\}$ 的上鞅，T 是关于 $\{X_n,n\geqslant 0\}$ 的停时，$T_n=\min\left(T,n\right)$，设存在一个非负随机变量 W，使得

1. $\mathbb{E}\left[W\right]<\infty$；
2. $\forall n\geqslant 0, M_{T_n}\geqslant -W$；

则有

$$\mathbb{E}\left[M_0\right] \geqslant \mathbb{E}\left[M_T\mathbb{I}_{\{T<\infty\}}\right]$$

特别地，如果停时再满足

3. $\mathbb{P}\{T<\infty\}=1$

则有

$$\mathbb{E}\left[M_0\right] \geqslant \mathbb{E}\left[M_T\right]$$

注记 6.2.11. 上鞅停时定理说明，在一定条件下，把 $\mathbb{E}\left[M_0\right]\geqslant\mathbb{E}\left[M_n\right]$ 中的 n 换成随机变量 T，不 4 等式 $\mathbb{E}\left[M_0\right]\geqslant\mathbb{E}\left[M_T\right]$ 仍然成立.

---

[7] $T_n=\min\{T,n\}$ 意味着当 $T_n=T\Longrightarrow T\leqslant n\Longrightarrow I_{\{T>n\}}=0$；反之 $T_n=n\Longrightarrow T\geqslant n\Longrightarrow I_{\{T>n\}}=1$.

[8] 指改变策略从而改变停止赌博的时间.

6.3 一致可积性

定理 6.3.1. 如果随机变量 X 满足 $\mathbb{E}[|X|]<\infty$，则 $\forall\varepsilon>0$，$\exists\delta>0$，当 $\mathbb{P}(A)<\delta$ 时，有

$$\mathbb{E}[|X_n|\,\mathbb{I}_A]<\varepsilon$$

证明. 设 X 的分布函数是 F，由于 $\mathbb{E}[|X|]<\infty$，那么

$$\mathbb{E}[|X|]=\int_0^n |x|\,\mathrm{d}F(x)+\int_n^{+\infty}|x|\,\mathrm{d}F(x)=\mathbb{E}\left[|X|\,\mathbb{I}_{\{|X|\leqslant n\}}\right]+\mathbb{E}\left[|X|\,\mathbb{I}_{\{|X|>n\}}\right]<\infty$$

要使得这个反常积分收敛，积分的右尾应该等于零，即

$$\lim_{n\to\infty}\mathbb{E}\left[|X|\,\mathbb{I}_{\{|X|>n\}}\right]=\lim_{n\to\infty}\int_n^{+\infty}|x|\,\mathrm{d}F(x)=0 \tag{6.1}$$

设 $\mathbb{P}\{|X|>n\}=\delta$，$A$ 是另外一个发生概率为 δ 的事件，$\mathbb{P}(A)=\delta$.

而且，事件 $\{|X|>n\}$ 是从 $+\infty$ 处向原点截所得到的，里面包含了所有使得事件概率为 δ 的所有 $|x|$ 最大值.

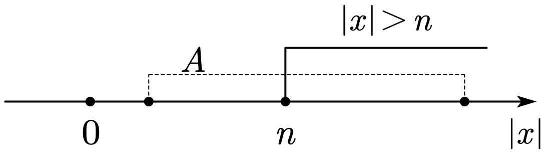

图 6.2. 一致可积性

因此 $\mathbb{E}[|X|\,\mathbb{I}_A]=\int_0^{+\infty}|x|\,\mathbb{I}_A\mathrm{d}F(x)=\int_A|x|\,\mathrm{d}F(x)\leqslant\int_n^{+\infty}|x|\mathrm{d}F(x)$，所以 $\mathbb{E}[|X|\,\mathbb{I}_A]\leqslant\mathbb{E}\left[|X|\,\mathbb{I}_{\{|X|>n\}}\right]$.

如果 $\mathbb{P}(A)<\delta$，那么 $\mathbb{E}[|X|\,\mathbb{I}_A]<\mathbb{E}\left[|X|\,\mathbb{I}_{\{|X|>n\}}\right]$，由(6.1)可知，

$$\lim_{n\to\infty}\mathbb{E}[|X|\,\mathbb{I}_A]\leqslant\lim_{n\to\infty}\mathbb{E}\left[|X|\,\mathbb{I}_{\{|X|>n\}}\right]=0 \tag{6.2}$$

将(6.2)用 $\varepsilon-\delta$ 的语言表示出来，就是需要证明的内容. □

定义 6.3.2. 一列随机变量 $X_0,X_1,\cdots,X_n$，称它们**一致可积**，如果 $\forall\varepsilon>0$，$\exists\delta>0$，使得 $\forall A$，当 $\mathbb{P}(A)<\delta$ 时，

$$\mathbb{E}[|X_n|\,\mathbb{I}_A]<\varepsilon$$

对任意 n 成立.

注记 6.3.3. 一致可积性的关键：

1. δ 不能依赖于 n；
2. $\mathbb{E}[|X_n|\mathbb{I}_A]<\varepsilon$ 对任意 n 成立.

注记 6.3.4. 一致可积性的目的是给出一个比鞅停时定理中定理6.2.6条件 3 更强但易于验证的条件.

如果 $\{M_n, n \geqslant 0\}$ 是一个关于 $\{X_n, n \geqslant 0\}$ 的一致可积鞅，T 是停时且 $\mathbb{P}\{T < \infty\} = 1$，那么 $\lim_{n\to\infty} \mathbb{P}\{T > n\} = 0$，根据一致可积的定义，可以得到定理6.2.6条件 3，即

$$\lim_{n\to\infty} \mathbb{E}\left[|M_n| \mathbb{I}_{\{T>n\}}\right] = 0$$

定理 6.3.5. (停时定理) 设 $\{M_n, n \geqslant 0\}$ 是一个关于 $\{X_n, n \geqslant 0\}$ 的一致可积鞅，T 是关于 $\{X_n, n \geqslant 0\}$ 的停时，满足

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

那么

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

定理 6.3.6. 两个一致可积的充分条件.

- 一致可积的充分条件 1：假设 $X_0, X_1, \cdots, X_n$ 是一列随机变量，并且存在常数 $C < \infty$，使得 $\mathbb{E}[X_n^2] < C$ 对于所有的 n 成立，则该序列一致可积.
- 一致可积的充分条件 2：设 $\{M_n\}$ 是关于 $\{\mathcal{F}_n\}$ 的鞅，如果存在一个非负随机变量 Y，满足 $\mathbb{E}(Y) < \infty$ 且 $\forall n, |M_n| < Y$，则 $\{M_n\}$ 是一致可积鞅.

6.4 鞅收敛定理

定理 6.4.1. (鞅收敛定理) 设 $\{M_n, n \geqslant 0\}$ 是一个关于 $\{X_n, n \geqslant 0\}$ 的鞅，并且存在常数 $C < \infty$，使得 $\forall n, \mathbb{E}[|M_n|] < C$，则当 $n \to \infty$，序列 $\{M_n\}$ 收敛到一个随机变量 M_∞.

推论 6.4.2. 如果 $\{M_n, n \geqslant 0\}$ 是一个关于 $\{X_n, n \geqslant 0\}$ 的一致可积鞅，则 $\{M_n\}$ 的极限存在 ($\lim_{n\to\infty} M_n = M_\infty$)，且

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

6.5 连续鞅

注记 6.5.1. 回顾：随机过程 $\{X(t), t \geqslant 0\}$ 称为 $\{\mathcal{F}_t\}$适应的，如果对每一个 $t \geqslant 0$，X_t 为 $\mathcal{F}_t$ 可测的，即 $\forall B \in \mathcal{B}$，有 $X_t^{-1}(B) \in \mathcal{F}_t$.

定义 6.5.2. 一个适应过程 $\{X_t, t \geqslant 0\}$ 称为**关于 $\{\mathcal{F}_t\}$ 的连续鞅**，如果

每个 X_t 可积，即 $\mathbb{E}(|X_t|) < \infty$；

对一切 $0 \leqslant s < t$，有

$$\mathbb{E}\left(X_t \mid \mathcal{F}_s\right) = X_s, \text{ a.s.}$$

特别地，当 $\mathcal{F}_t = \sigma\left(X_u, 0 \leqslant u \leqslant t\right)$，即包含一切形如 $\{X_s \leqslant s\}$ $(s \leqslant t, x \in \mathbb{R})$ 的事件的最小 σ 代数，那么对关于 $\{\mathcal{F}_t\}$ 的鞅 $\{X_t, t \geqslant 0\}$，有

$$\mathbb{E}\left(X_t \mid X_u, 0 \leqslant u \leqslant s\right) = X_s, \text{a.s.}$$

简称 $\{X_t\}$ 为**鞅**.

定理 6.5.3. 若随机过程 $\{X_t\}$ 是鞅，则对 $t > 0$，有

$$\mathbb{E}\left(X_t\right) = \mathbb{E}\left[\mathbb{E}\left(X_t \mid X_0\right)\right] = \mathbb{E}\left(X_0\right)$$

定义 6.5.4. 非负广义**随机函数**τ（$\tau : \Omega \to [0, \infty]$）称为$\mathcal{F}_t$ **停时**，如果 $\mathbb{P}\{\tau < \infty\} = 1$，并且对一切 $t \geqslant 0, \{\tau \leqslant t\}$ 是 $\mathcal{F}_t$ 可测的，即

$$\{\tau \leqslant t\} \in \mathcal{F}_t$$

特别地，当 $\mathcal{F}_t = \sigma\left(X_u, 0 \leqslant u \leqslant t\right)$，称 τ 为**关于随机过程 $\{X_t, t \geqslant 0\}$ 的停时**.

若存在常数 $k > 0$ 使得 $\mathbb{P}\{\tau \leqslant k\} = 1$，则称 τ 为**有界停时**.

定理 6.5.5. （有界停时定理）若 τ 是有界停时，则有

$$\mathbb{E}\left(X_\tau\right) = \mathbb{E}\left(X_0\right)$$

定理 6.5.6. 设 $\{X_t, t \geqslant 0\}$ 是一个鞅，并且对 $\forall t \geqslant 0, X_t \geqslant 0$（或者称为非负鞅），则存在几乎处处收敛的有限极限，即

$$\lim_{t \to \infty} X_t = X_\infty, \text{ a.s.}$$

第七章　Brown 运动

Brown 运动是鞅、Markov 过程、高斯过程的特例．本章要求理解 Brown 运动的基本概念与性质；掌握 Gauss 过程；掌握 Brown 运动的鞅性质、Brown 运动的 Markov 性及 Brown 运动的最大值分布；了解 Brown 桥、有吸收值的 Brown 运动、有原点反射的 Brown 运动、几何 Brown 运动、有漂移的 Brown 运动．

7.1　基本概念

注记 7.1.1. Brown 运动的来源：一个粒子在直线上做简单的随机游动，每隔 Δt 时间内等概率地向左或向右移动 Δx 的距离，t 时刻粒子的位置 $X(t)$ 为

$$X(t) = \Delta x \left(X_1 + X_2 + \cdots + X_{\frac{t}{\Delta t}} \right)$$

设诸 X_i 相互独立，且

$$\mathbb{P}\{X_i = 1\} = \mathbb{P}\{X_i = -1\} = \frac{1}{2}$$

那么

$$\mathrm{Var}[X(t)] = (\Delta x)^2 \frac{t}{\Delta t}$$

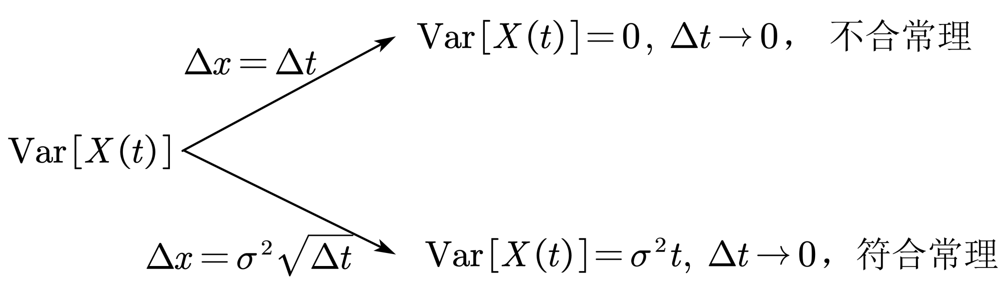

图 7.1. Brown 运动的来源

定义 7.1.2. 随机过程 $\{X(t), t \geqslant 0\}$ 如果满足

1. $X(0) = 0$;
2. $\{X(t), t \geqslant 0\}$ 有平稳独立增量;

3. $\forall t > 0, X(t)$ 服从正态分布 $N(0, \sigma^2 t)$.

则称 $\{X(t), t \geqslant 0\}$ 为**Brown 运动**或**Wiener 过程**，常记为 $\{B(t), t \geqslant 0\}$ 或 $\{W(t), t \geqslant 0\}$. 特别地，

- $\sigma = 1$ 的 Brown 运动称为**标准 Brown 运动**. 如果 $\sigma \neq 1$，可以通过 $\left\{\frac{X(t)}{\sigma}, t \geqslant 0\right\}$ 将它化为标准 Brown 运动.
- 若 $B(0) = x$，称 $B(t)$ 为**始于 x 的 Brown 运动**，记为 $\{B^x(t)\}$.

定义 7.1.3. 称随机过程 $\{X(t), t \geqslant 0\}$ 是**空间齐次**的，如果它的有限维分布[1]是空间平移不变的，即

$$\mathbb{P}\{X(t_1) \leqslant x_1, X(t_2) \leqslant x_2, \cdots, X(t_n) \leqslant x_n | X(0) = 0\} = \mathbb{P}\{X(t_1) \leqslant x_1 + x, X(t_2) \leqslant x_2 + x, \cdots, X(t_n$$

定理 7.1.4. Brown 运动的性质：

1. 正态增量：$B(t) - B(s) \sim N(0, t-s)$，特别地，当 $s = 0$ 时 $B(t) \sim N(0, t)$；
2. 独立增量：$B(t) - B(s)$ 独立于过去的状态 $B(u)$，$0 \leqslant u \leqslant s$；
3. 连续路径：$B(t), t \geqslant 0$ 是 t 的连续函数；
4. 空间齐性：$B^x(t) - x = B^0(t)$.

定义 7.1.5. 设过程从 x 开始，$B(0) = x$，则 $B(t) \sim N(x, t)$，于是

$$P_x\{B(t) \in (a, b)\} = \int_a^b \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} \mathrm{d}y$$

概率 P_x 的下标表示过程从 x 开始. 被积函数

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$$

称为 Brown 运动的**转移概率密度**.

定理 7.1.6. 任意 Brown 运动的有限维分布：

$$\begin{aligned} & P_x\{B(t_1) \leqslant x_1, B(t_2) \leqslant x_2, \cdots, B(t_n) \leqslant x_n\} \\ = & \int_{-\infty}^{x_1} p_{t_1}(x, y_1) \mathrm{d}y_1 \int_{-\infty}^{x_2} p_{t_2 - t_1}(y_1, y_2) \mathrm{d}y_2 \cdots \int_{-\infty}^{x_n} p_{t_n - t_{n-1}}(y_{n-1}, y_n) \mathrm{d}y_n \end{aligned} \tag{7.1}$$

证明. 我们先计算条件概率

$$\begin{aligned} P\{B(t+s) \leqslant b | B(s) = x\} &= P\{B(t+s) - B(s) \leqslant b - x\} \\ &= \int_{-\infty}^{b-x} \frac{1}{\sqrt{2\pi t}} e^{-\frac{u^2}{2t}} \mathrm{d}u = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} \mathrm{d}y \\ &= \int_{-\infty}^{b} p_t(x, y) \mathrm{d}y \end{aligned}$$

[1]随机过程是基于有限维分布给出的.

再将(7.1)中联合概率改写成条件概率的形式：

$$
\begin{aligned}
&P_x\left\{B\left(t_1\right)\leqslant x_1, B\left(t_2\right)\leqslant x_2,\cdots,B\left(t_n\right)\leqslant x_n\right\}\\
=&P_x\left\{B\left(t_n\right)\leqslant x_n\middle| B\left(t_i\right)\leqslant x_i, i=1,2,\cdots,n-1\right\}\cdot P_x\left\{B\left(t_i\right)\leqslant x_i, i=1,2,\cdots,n-1\right\}\\
=&P_x\left\{B\left(t_n\right)\leqslant x_n\middle| B\left(t_{n-1}\right)\leqslant x_{n-1}\right\}\cdot P_x\left\{B\left(t_{n-1}\right)\leqslant x_{n-1}\middle| B\left(t_{n-2}\right)\leqslant x_{n-1}\right\}\\
&\cdots P_x\left\{B\left(t_2\right)\leqslant x_2\middle| B\left(t_1\right)\leqslant x_1\right\}\cdot P_x\left\{B\left(t_1\right)\leqslant x_1\right\}\\
=&\int_{-\infty}^{x_1} p_{t_1}\left(x,y_1\right)\mathrm{d}y_1\int_{-\infty}^{x_2} p_{t_2-t_1}\left(y_1,y_2\right)\mathrm{d}y_2\cdots\int_{-\infty}^{x_n} p_{t_n-t_{n-1}}\left(y_{n-1},y_n\right)\mathrm{d}y_n
\end{aligned}
$$

证毕.

另外，不用条件概率也可以证明(7.1)，仅根据 Brown 运动的独立增量性即可. 首先我们可以得到

$$
\begin{aligned}
&P_x\left\{B\left(t_1\right)\leqslant x_1, B\left(t_2\right)\leqslant x_2,\cdots,B\left(t_n\right)\leqslant x_n\right\}\\
=&P_x\left\{B\left(t_1\right)-B(0)\leqslant x_1-x_0, B\left(t_2\right)-B\left(t_1\right)\leqslant x_2-B\left(t_1\right),\cdots,B\left(t_n\right)-B\left(t_{n-1}\right)\leqslant x_n-B\left(t_{n-1}\right)\right\}\quad(*)
\end{aligned}
$$

令 $u_i=B\left(t_i\right)-B\left(t_{i-1}\right), B\left(t_0\right)=x, i=1,2,\cdots,n$，它们是互相独立的正态随机变量，所以

$$
\begin{aligned}
(*)=&P_x\left\{u_1\leqslant x_1-x_0, u_2\leqslant x_2-x_0-u_1,\cdots,u_n\leqslant x_n-x_0-\sum_{i=1}^{n-1}u_i\right\}\\
=&\int_{-\infty}^{x_1-x_0}\int_{-\infty}^{x_2-x_0-u_1}\cdots\int_{-\infty}^{x_n-x_0-\sum_{i=1}^{n-1}u_i} f\left(u_1,u_2,\cdots,u_n\right)\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_n\\
=&\int_{-\infty}^{x_1-x_0}\frac{1}{\sqrt{2\pi t_1}}e^{-\frac{u_1^2}{2t_1}}\mathrm{d}u_1\int_{-\infty}^{x_2-x_0-u_1}\frac{1}{\sqrt{2\pi\left(t_2-t_1\right)}}e^{-\frac{u_2^2}{2\left(t_2-t_1\right)}}\mathrm{d}u_2\cdots\\
&\int_{-\infty}^{x_n-x_0-\sum_{i=1}^{n-1}u_i}\frac{1}{\sqrt{2\pi\left(t_n-t_{n-1}\right)}}e^{-\frac{u_n^2}{2\left(t_n-t_{n-1}\right)}}\mathrm{d}u_n\quad(**)
\end{aligned}
$$

再令 $y_i=u_i+x_0+\sum_{j=1}^{i-1}u_{j-1}, x_0=x, u_0=0, i=1,2,\cdots,n$，那么 $y_{i+1}-y_i=u_{i+1}$，因此有

$$
\begin{aligned}
(**)=&\int_{-\infty}^{x_1}\frac{1}{\sqrt{2\pi t_1}}e^{-\frac{\left(y_1-x\right)^2}{2t_1}}\mathrm{d}y_1\int_{-\infty}^{x_2}\frac{1}{\sqrt{2\pi\left(t_2-t_1\right)}}e^{-\frac{\left(y_2-y_1\right)^2}{2\left(t_2-t_1\right)}}\mathrm{d}y_2\cdots\int_{-\infty}^{x_n}\frac{1}{\sqrt{2\pi\left(t_n-t_{n-1}\right)}}e^{-\frac{\left(y_n-y_{n-1}\right)^2}{2\left(t_n-t_{n-1}\right)}}\mathrm{d}y_n\\
=&\int_{-\infty}^{x_1} p_{t_1}\left(x,y_1\right)\mathrm{d}y_1\int_{-\infty}^{x_2} p_{t_2-t_1}\left(y_1,y_2\right)\mathrm{d}y_2\cdots\int_{-\infty}^{x_n} p_{t_n-t_{n-1}}\left(y_{n-1},y_n\right)\mathrm{d}y_n
\end{aligned}
$$

□

定义 7.1.7. 当 $\{t_i^n\}_{i=0}^n$ 遍取 $[0,t]$ 的分割，其**模**定义为

$$\delta_n=\max_{0\leqslant i\leqslant n-1}\left(t_{i+1}^n-t_i^n\right)\to 0$$

Brown 运动的**二次变差** $[B,B](t)$ 定义为依概率收敛意义下的极限

$$[B,B](t):=[B,B]([0,t])=\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}\left|B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right|^2$$

而**变差**即为极限 $\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}\left|B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right|$.

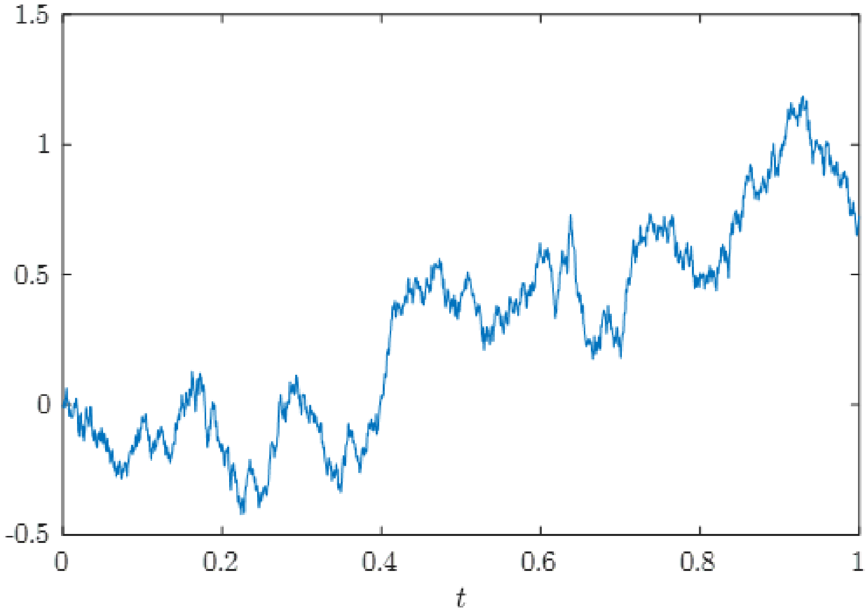

图 7.2. Brown 运动一条样本轨道

定理 7.1.8. Brown 运动样本轨道 $B(t), 0 \leqslant t \leqslant T$（的性质）

1. 是关于 t 的连续函数；

证明. $\forall t \in (0, T), \forall \Delta t > 0$, 当 $\Delta t \to 0$,

$$\mathbb{E}\left[B(t+\Delta t) - B(t)\right]^2 = \Delta t \to 0$$

$$\operatorname{Var}\left[B(t+\Delta t) - B(t)\right]^2 = 3(\Delta t)^2 \to 0$$

则

$$B(t+\Delta t) \xrightarrow{\mathcal{L}^2} B(t) \Longrightarrow B(t+\Delta t) \xrightarrow{\text{a.s.}} B(t)$$

而

$$B(t+\Delta t) - B(t) \stackrel{d}{=} B(t) - B(t-\Delta t)$$

则类似地，

$$B(t-\Delta t) \xrightarrow{\text{a.s.}} B(t)$$

□

2. 在任意区间（无论区间多么小）上都是不单调的；

直觉. 例如 $B(t+\Delta t) - B(t)$ 是随机变量，无法确定其的正负情况. □

3. 任意点都是不可微的；

直觉. $B(t+\Delta t)-B(t)\overset{d}{=}z\sqrt{\Delta t},\ \ z\sim\mathcal{N}(0,1)$. 那么 $\lim_{\Delta t\to 0}\frac{z\sqrt{\Delta t}}{\Delta t}$ 不可能收敛到有限值. □

4. 对任意 t，在 $[0,t]$ 上的二次变差 $[B,B](t)=t$；

证明. 取区间 $[0,t]$ 的分割 $\{t_i^n\}_{i=0}^n$ 使得 $\sum_n\delta_n<\infty$，记

$$S_n=\sum_{i=0}^{n-1}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2$$

则

$$\mathbb{E}\left(S_n\right)=\sum_{i=0}^{n-1}\mathbb{E}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2=\sum_{i=0}^{n-1}\left(t_{i+1}^n-t_i^n\right)=t$$

那么[2]

$$\begin{aligned}\operatorname{Var}\left(S_n\right)&=\operatorname{Var}\left\{\sum_{i=0}^{n-1}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2\right\}\\&=\sum_{i=0}^{n-1}\operatorname{Var}\left\{\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2\right\}\\&=\sum_{i=0}^{n-1}\operatorname{Var}\left\{\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2\right\}\\&=\sum_{i=0}^{n-1}\mathbb{E}\left\{\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^4\right\}\\&=\sum_{i=0}^{n-1}3\left(t_{i+1}^n-t_i^n\right)^2\leqslant 3\max_i\left\{t_{i+1}^n-t_i^n\right\}t=3t\delta_n\end{aligned}$$

所以

$$\sum_{n=1}^{\infty}\operatorname{Var}\left(S_n\right)=3t\sum_n\delta_n<\infty$$

由单调收敛定理，得

$$\sum_{n=1}^{\infty}\operatorname{Var}\left(S_n\right)=\mathbb{E}\sum_{n=1}^{\infty}\left[S_n-\mathbb{E}\left(S_n\right)\right]^2=\mathbb{E}\left[\sum_{n=1}^{\infty}\left(S_n-t\right)^2\right]<\infty$$

因此

$$\sum_{i=1}^{\infty}\left(S_n-t\right)^2<\infty,\ \text{a.s.}$$

于是

$$S_n-t\to 0,\ \text{a.s.}$$

---

[2]正态分布的 n 阶中心距 $\mathbb{E}(\xi-\mu)^n=\begin{cases}0, & n\text{为奇数}\\(n-1)!!\sigma^n, & n\text{为偶数}\end{cases}$, 其中 $\xi\sim\mathcal{N}\left(\mu,\sigma^2\right)$.

故

$$[B,B](t)=t$$

□

5. 在任意区间（无论区间多么小）上都是无限变差的.

证明. (Proof by making a contradiction) Assume there exist $A\in\mathcal{F}$, such that $\mathbb{P}(A)>0$ and for every $\omega\in A$,

$$\limsup_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\left|W\left(t_{j+1}\right)-W\left(t_j\right)\right|(\omega)<\infty$$

Then,

$$\begin{aligned}\sum_{j=0}^{n-1}\left[W\left(t_{j+1}\right)-W\left(t_j\right)\right]^2 &\leqslant \max_{0\leqslant k\leqslant n-1}\left|W\left(t_{k+1}\right)-W\left(t_k\right)\right|\cdot\sum_{j=0}^{n-1}\left|W\left(t_{j+1}\right)-W\left(t_j\right)\right| \\ &\leqslant \max_{0\leqslant k\leqslant n-1}\left|W\left(t_{k+1}\right)-W\left(t_k\right)\right|\cdot\limsup_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\left|W\left(t_{j+1}\right)-W\left(t_j\right)\right| \\ &\to 0\end{aligned}$$

since by uniform continuity of continuous functions over a closed interval,

$$\lim_{\|\Pi\|\to 0}\max_{0\leqslant k\leqslant n-1}\left|W\left(t_{j+1}\right)-W\left(t_j\right)\right|(\omega)=0$$

However, we know

$$\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\left[W\left(t_{j+1}\right)-W\left(t_j\right)\right]^2=T,\ \text{a.s.}$$

Therefore, $\sum_{j=0}^{n-1}\left|W\left(t_{j+1}\right)-W\left(t_j\right)\right|$ is not bounded. □

7.2 Gauss 过程

引理 7.2.1. 如果随机变量 $\boldsymbol{Z}_{p\times 1}\sim\mathcal{N}_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$，那么 $\boldsymbol{A}_{m\times p}\boldsymbol{Z}_{p\times 1}\sim\mathcal{N}_m\left(\boldsymbol{A\mu},\boldsymbol{A\Sigma A}^T\right)$，即有 $\mathbb{E}(\boldsymbol{AZ})=\boldsymbol{A}\mathbb{E}(\boldsymbol{Z}),\mathrm{Cov}(\boldsymbol{AZ},\boldsymbol{AZ})=\boldsymbol{A}\mathrm{Cov}(\boldsymbol{Z},\boldsymbol{Z})\boldsymbol{A}^T$，其中 $\boldsymbol{A}_{m\times p}$ 是常数矩阵.

引理 7.2.2. 如果相互独立的随机变量 X,Y 满足 $X\sim\mathcal{N}\left(\mu_1,\sigma_1^2\right),Y\sim\mathcal{N}\left(\mu_2,\sigma_2^2\right)$，那么 $(X,X+Y)\sim\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$，其中 $\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$，$\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1^2 & \sigma_1^2\\ \sigma_1^2 & \sigma_1^2+\sigma_2^2\end{pmatrix}$. [3]

[3]均值向量容易证明，下面简单验证一下协方差矩阵：由于 $\begin{pmatrix}X\\ Y\end{pmatrix}\sim\mathcal{N}\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0\right)$ 或 $\mathcal{N}\left[\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{pmatrix}\right]$，而 $X+Y\sim\mathcal{N}\left(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2\right)$，所以 $\mathrm{Cov}(X,X+Y)=\mathrm{Var}(X)+\mathrm{Cov}(X,Y)=\sigma_1^2$，$\mathrm{Var}(X+Y)=\sigma_1^2+\sigma_2^2$.

定理 7.2.3. Brown 运动是均值函数为 $m(t)=0$，协方差函数为 $\gamma(s,t)=\min(t,s)$ 的 Gauss 过程.

证明. 均值为零容易得到，下面证明计算协方差函数. 协方差函数

$$\gamma(s,t)=\mathrm{Cov}[B(t),B(s)]=\mathbb{E}[B(t)B(s)]$$

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

若 $t\geqslant s$，同理，$\gamma(s,t)=s$. 即得证. □

例子 7.2.4. （协方差函数的应用）求样本轨道函数和的分布.

欲求 $\frac{1}{4}\left[B\left(\frac{1}{4}\right)+B\left(\frac{1}{2}\right)+B\left(\frac{3}{4}\right)+B(1)\right]$ 的分布，可以令随机变量

$$\boldsymbol{X}=\frac{1}{4}\left(B\left(\frac{1}{4}\right),B\left(\frac{1}{2}\right),B\left(\frac{3}{4}\right),B(1)\right)^{\mathrm{T}}$$

所以 $\boldsymbol{X}\sim N(0,\boldsymbol{\Sigma}=(\sigma_{ij}))$，其中 $\sigma_{ij}=\min(i,j)$，$\boldsymbol{\Sigma}=\frac{1}{64}\begin{pmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{pmatrix}$,

再令 $\boldsymbol{A}=(1,1,1,1)$，所以 $\boldsymbol{AX}\sim N\left(0,\boldsymbol{A\Sigma A}^T\right)$，这个方差是一个二次型，容易计算得到 $\boldsymbol{AX}\sim N\left(0,\frac{15}{32}\right)$.

注记 7.2.5. 所有逼近和的分布都是零均值的正态分布.

由于函数连续是黎曼可积的充分条件，且 Brown 运动的样本轨道 $B(t)$ 是连续的，所以对每个路径来说，Riemann 积分都存在.

由 Riemann 积分的定义，可以从逼近和 $\sum_{i=1}^{n}B(t_i)\Delta t_i$ 的极限分布得到，这里是对区间 $[0,t]$ 作分割，$t_i\in[0,t],\Delta t_i=t_{i+1}-t_i$. 在例子7.2.4中，我们对 $[0,1]$ 作分割，并取 $n=4,\Delta t_i=\frac{1}{n}$, 即$t_i=\frac{i}{n}$. 一般地，可以类似地证明该命题.

例子 7.2.6. 求 $\int_0^t B(s)\,\mathrm{d}s$ 的期望和方差.

$$\begin{aligned}\mathbb{E}\left[\int_0^t B(s)\,\mathrm{d}s\right] &= \int_0^t \mathbb{E}[B(s)]\,\mathrm{d}s = 0\\ \mathrm{Var}\left[\int_0^t B(s)\,\mathrm{d}s\right] &= \mathrm{Cov}\left(\int_0^t B(s_1)\,\mathrm{d}s_1, \int_0^t B(s_2)\,\mathrm{d}s_2\right)\\ &= \mathbb{E}\left[\int_0^t B(s_1)\,\mathrm{d}s_1 \int_0^t B(s_2)\,\mathrm{d}s_2\right]\\ &= \int_0^t \int_0^t \mathbb{E}[B(s_1)B(s_2)]\mathrm{d}s_1\mathrm{d}s_2\\ &= \int_0^t \int_0^t \min\{s_1, s_2\}\mathrm{d}s_1\mathrm{d}s_2\\ &= 2\int_0^t \mathrm{d}s_1 \int_0^{s_1} s_2\mathrm{d}s_2\\ &= \frac{t^3}{3}\end{aligned}$$

进一步地，我们知道 $\int_0^t B(s)\,\mathrm{d}s \sim \mathcal{N}(0, \frac{t^3}{3})$.

7.3 Brown 运动的鞅性质

定理 7.3.1. 设 $\{B(t)\}$ 是 Brown 运动，则

1. $\{B(t)\}$ 是鞅;
2. $\{B^2(t) - t\}$ 是鞅;
3. $\forall u \in \mathbb{R}, \left\{e^{uB(t) - \frac{tu^2}{2}}\right\}$ 是鞅.

证明. 根据 Brown 运动的独立增量性，容易知道对任意的函数 $g(x)$ 有

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

由 $B(t) \sim N(0,t)$ 可知随机变量 $B(t)$ 可积，$\mathbb{E}[B(t)] = 0$，接着对于 $g(x) = x$，有

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

证得 (1).

再由 $\mathbb{E}[B^2(t)] = t < \infty$，所以 $B^2(t)$ 可积，对于 $g(x) = x^2$，有

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

两端同时减去 $t+s$，则证得 (2).

由 $B(t) \sim N(0,t)$ 的矩母函数 $\mathbb{E}\left[e^{uB(t)}\right] = e^{\frac{t}{2}u^2} < \infty$，这说明 $e^{uB(t)}$ 可积，而且 $\mathbb{E}\left[e^{uB(t)-\frac{tu^2}{2}}\right] = 1$，对于 $g(x) = e^{ux}$，有

$$\begin{aligned}\mathbb{E}\left[e^{uB(t+s)}\middle|\mathcal{F}_t\right] &= \mathbb{E}\left[e^{uB(t)+u[B(t+s)-B(t)]}\middle|\mathcal{F}_t\right]\\ &= e^{uB(t)}\mathbb{E}\left[e^{u[B(t+s)-B(t)]}\middle|\mathcal{F}_t\right]\\ &= e^{uB(t)}\mathbb{E}\left[e^{u[B(t+s)-B(t)]}\right]\\ &= e^{uB(t)}e^{\frac{s}{2}u^2}\end{aligned}$$

两端同时乘以 $e^{-\frac{s+t}{2}u^2}$，则证得 (3).

从以上证明可以总结得到，证明鞅的重要思想是利用 Brown 运动独立增量的性质，其他的无非就是变换 $g(x)$. □

7.4 Brown 运动的 Markov 性

定义 7.4.1. 设 $\{X(t), t \geqslant 0\}$ 是一个连续随机过程，如果对 $\forall t, s > 0$ 有

$$\mathbb{P}\{X(t+s) \leqslant y|\mathcal{F}_t\} = \mathbb{P}\{X(t+s) \leqslant y|X(t)\}, a.s.$$

其中 $\mathcal{F}_t = \sigma\{X(u), 0 \leqslant u \leqslant t\}$，则称 $\{X(t)\}$ 为**Markov 过程**，这种性质称为**Markov 性**(指未来的事件仅与当下有关，而与过去无关).

连续状态下 Markov 过程的转移概率定义为在时刻 s 处于状态 x 的条件下，过程在时刻 t 的**分布函数**$F(y,t,x,s) = \mathbb{P}\{X(t) \leqslant y|X(s) = x\}$.

定理 7.4.2. Brown 运动具有 Markov 性.

证明. 由于矩母函数可以唯一确定分布函数，考虑使用矩母函数验证. 由 Brown 运动的鞅性，可知

$$\mathbb{E}\left[e^{uB(t+s)}\middle|\mathcal{F}_t\right] = e^{uB(t)}e^{\frac{s}{2}u^2} \tag{7.2}$$

又因为 $B(t+s) - B(t) \sim N(0,s)$，所以(7.2)左边可以化为关于 $B(t)$ 的条件期望形式，即

$$\mathbb{E}\left[e^{uB(t+s)}\middle|\mathcal{F}_t\right] = e^{uB(t)}\mathbb{E}\left[e^{u[B(t+s)-B(t)]}\middle|B(t)\right] = \mathbb{E}\left[e^{uB(t+s)}\middle|B(t)\right]$$

这说明在给定过去和现在 $\mathcal{F}_t$ 的条件下，未来事件的分布仅与当下 $B(t)$ 有关. □

注记 7.4.3. Brown 运动的转移概率是正态的:

$$\begin{aligned} F(y,t,x,s) &= \mathbb{P}\{B(t) \leqslant y \mid B(s)=x\} \\ &= \mathbb{P}\{B(t)-B(s) \leqslant y-x \mid B(s)=x\} \\ &= \mathbb{P}\{B(t)-B(s) \leqslant y-x\} \\ &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(u-x)^2}{2(t-s)}} \mathrm{d}u \end{aligned}$$

注记 7.4.4. Brown 运动的转移概率满足 $F(y,t,x,s)=F(y,t-s,x,0)$,即

$$\mathbb{P}\{B(t) \leqslant y \mid B(s)=x\} = \mathbb{P}\{B(t-s) \leqslant y \mid B(0)=x\}$$

这称为 Brown 运动的**时齐性**,即分布不随时间平移而变化.

另外,根据 Brown 运动的有限维分布可以知道,

$$\begin{aligned} &\mathbb{P}_x\{B(t_1) \leqslant x_1, B(t_2) \leqslant x_2, \cdots, B(t_n) \leqslant x_n\} \\ =& \int_{-\infty}^{x_1} p_{t_1}(x,y_1)\,\mathrm{d}y_1 \int_{-\infty}^{x_2} p_{t_2-t_1}(y_1,y_2)\,\mathrm{d}y_2 \cdots \int_{-\infty}^{x_n} p_{t_n-t_{n-1}}(y_{n-1},y_n)\,\mathrm{d}y_n \\ =& \mathbb{P}_x\{B(t_1+s) \leqslant x_1, B(t_2+s) \leqslant x_2, \cdots, B(t_n+s) \leqslant x_n\} \end{aligned}$$

所以,Brown 运动的所有有限维分布都是时齐的.

当 $s=0$ 时,$F(y,t,x,0)=p_t(x,y)=\frac{1}{\sqrt{2\pi t}}e^{-\frac{(y-x)^2}{2t}}$.

定义 7.4.5. 如果非负随机变量 T 可以取无穷值,即 $T:\Omega \to [0,\infty]$,并且 $\forall t$ 有 $\{T>t\} \in \mathcal{F}_t = \sigma\{B(u), 0 \leqslant u \leqslant t\}$,则称 T 为**关于 Brown 运动 $\{B(t), t \geqslant 0\}$ 的停时**.

定理 7.4.6. (Brown 运动的强 Markov 性定理)设 T 是关于 Brown 运动 $\{B(t)\}$ 的有限停时,记

$$\mathcal{F}_T = \left\{A \in \mathcal{F} : A \bigcap \{T \leqslant t\} \in \mathcal{F}_t, \forall t \geqslant 0\right\}$$

则

$$\mathbb{P}\{B(T+t) \leqslant y \mid \mathcal{F}_T\} = \mathbb{P}\{B(T+t) \leqslant y \mid B(T)\},\ \text{a.s.}$$

注记 7.4.7. 如果定义

$$\hat{B}(t) := B(T+t) - B(T)$$

则 $\hat{B}(t)$ 是始于 0 且独立于 $\mathcal{F}_T$ 的 Brown 运动.

7.5 Brown 运动的最大值变量及反正弦律

定理 7.5.1. 三个概率.

1. 设 $\{B^x(t)\}$ 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
$$\frac{|x|}{\sqrt{2\pi}}\int_0^t u^{-\frac{3}{2}}e^{-\frac{x^2}{2u}}\mathrm{d}u$$
2. $B^y(t)$ 是 Brown 运动，$B^y(t)$ 在区间 (a,b) 中至少有一个零点的概率为
$$\frac{2}{\pi}\arccos\sqrt{\frac{a}{b}}$$
3. $B^y(t)$ 是 Brown 运动，$B^y(t)$ 在区间 (a,b) 中没有零点的概率为
$$\frac{2}{\pi}\arcsin\sqrt{\frac{a}{b}}$$

7.6 Brown 运动的几种变化

定义 7.6.1. 设 $\{B(t),t\geqslant 0\}$ 是 Brown 运动，令
$$B^*(t)=B(t)-tB(1),0\leqslant t\leqslant 1$$
称随机过程 $\{B^*(t),0\leqslant t\leqslant 1\}$ 是**Brown 桥**.

注记 7.6.2. 当 $0\leqslant s\leqslant t\leqslant 1$，有
$$\mathbb{E}[B^*(t)]=0$$
$$\begin{aligned}\mathbb{E}[B^*(s)B^*(t)]&=\mathbb{E}[B(s)-sB(1)(B(t)-tB(1))]\\&=\mathbb{E}\left[B(s)B(t)-stB^2(1)-tB(s)B(1)-sB(t)B(1)\right]\\&=\min(s,t)-st-ts+st\\&=s(1-t)\end{aligned}$$

注记 7.6.3. Brown 桥的特点：$B^*(0)=B^*(1)=0$，过程的起点和终点像固定的一样，与桥状相似.

定义 7.6.4. 设 T_x 为 Brown 运动 $\{B(t)\}$ 首次击中 x 的时间，$x>0$，即
$$T_x=\inf\{t>0,B(t)=x\}$$
令
$$Z(t)=\begin{cases}B(t), & t<T_x\\ x, & t\geqslant T_x\end{cases}$$
则 $\{Z(t),t\geqslant 0\}$ 是击中 x 后永远停留在那里的 Brown 运动，称为**有吸收值 x 的 Brown 运动**.

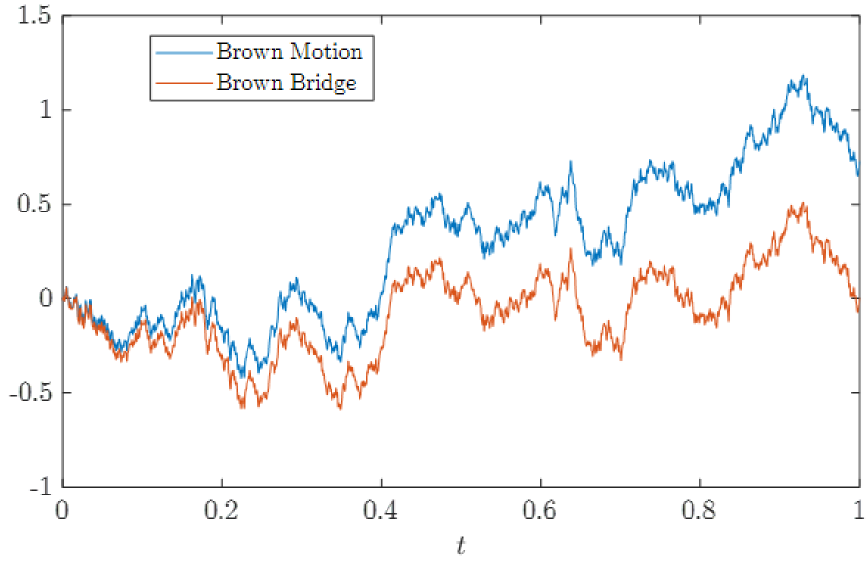

图 7.3. Brown 桥

注记 7.6.5. 随机变量 $Z(t)$ 的分布有离散和连续两个部分:

1. 离散部分:

$$\mathbb{P}\{Z(t)=x\}=\mathbb{P}\{T_x\leqslant t\} = \frac{2}{\sqrt{2\pi t}}\int_x^{\infty} e^{-\frac{y^2}{2t}}\mathrm{d}y$$

2. 连续部分:

$$\mathbb{P}\{Z(t)\leqslant y\}=\frac{1}{\sqrt{2\pi t}}\int_{y-2x}^{y} e^{-\frac{u^2}{2t}}\mathrm{d}u$$

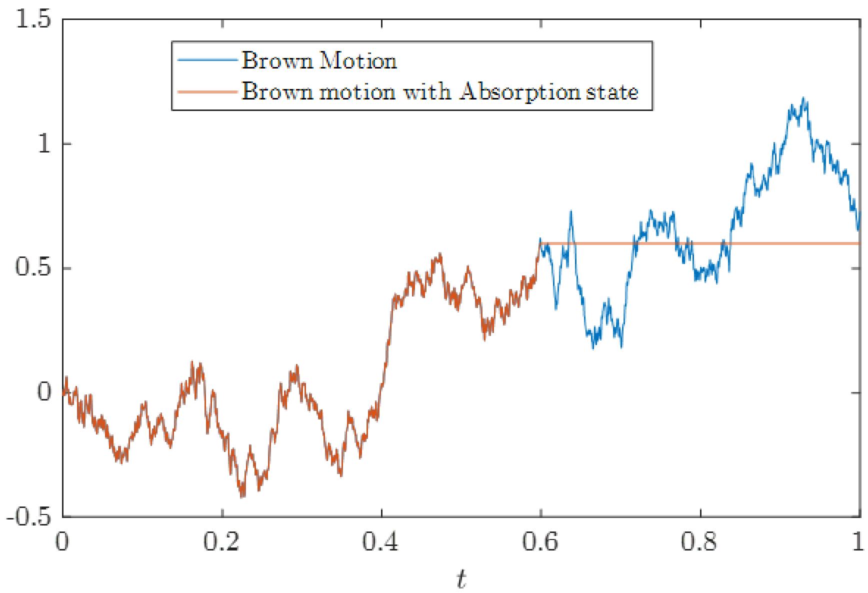

图 7.4. 有吸收值的 Brown 运动

定义 7.6.6. 令

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

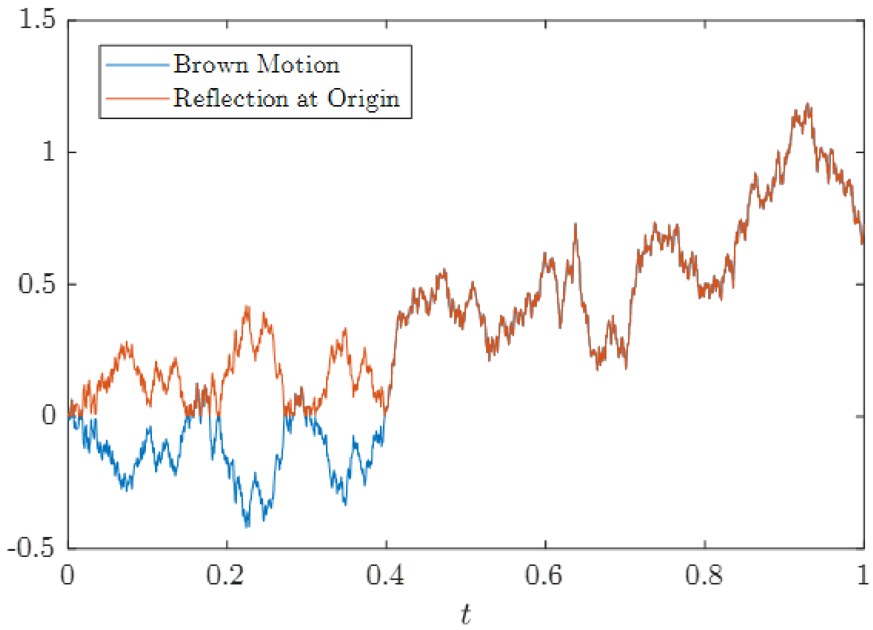

图 7.5. 在原点反射的 Brown 运动

定义的过程 $\{Y(t), t \geqslant 0\}$ 称为**在原点反射的 Brown 运动**.

注记 7.6.7. 随机变量 $Y(t)$ 的分布是

$$\begin{aligned}\mathbb{P}\{Y(t) \leqslant y\} &= \mathbb{P}\{B(t) \leqslant y\} - \mathbb{P}\{B(t) \leqslant -y\} \\ &= 2P\{B(t) \leqslant y\} - 1 \\ &= \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{u^2}{2t}} \mathrm{d}u - 1, y > 0\end{aligned}$$

定义 7.6.8. 令

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

称 $X(t)$ 为**几何 Brown 运动**.

注记 7.6.9. 几何 Brown 运动的均值函数与方差函数：因为 Brown 运动的矩母函数是

$$\mathbb{E}\left[e^{sB(t)}\right] = e^{t\frac{s^2}{2}}$$

那么对几何 Brown 运动，有

$$\mathbb{E}[X(t)] = \mathbb{E}\left[e^{B(t)}\right] = e^{\frac{t}{2}}$$

$$\begin{aligned}\operatorname{Var}[X(t)] &= \mathbb{E}\left[X^2(t)\right] - (\mathbb{E}[X(t)])^2 \\ &= \mathbb{E}\left[e^{2B(t)}\right] - e^t \\ &= e^{2t} - t\end{aligned}$$

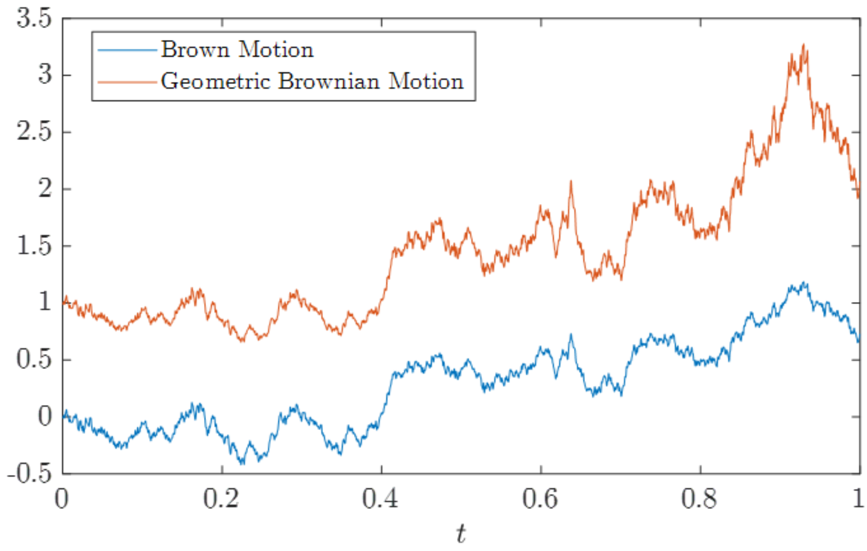

图 7.6. 几何 Brown 运动

例子 7.6.10. 金融市场中，经常假定股票的价格的变化是几何 Brown 运动. 例如，对一个合约期限为 T，交割价格为 K 的欧式看涨期权，标的资产是某种股票，该股票目前的价格为 y，并其价格 $X(t)$ 按照几何 Brown 运动变化，设无风险利率为 r，那么该期权在现时的价格为

$$\begin{aligned} e^{-Tr}\mathbb{E}\left[\max\left(X(T)-K,0\right)\right] &= e^{-rT}\int_0^{\infty}\mathbb{P}\left\{X(T)-K>u\right\}\mathrm{d}u \\ &= e^{-rT}\int_0^{\infty}\mathbb{P}\left\{ye^{B(T)}-K>u\right\}\mathrm{d}u \\ &= e^{-rT}\int_0^{\infty}\mathbb{P}\left\{B(T)>\ln\frac{u+K}{y}\right\}\mathrm{d}u \\ &= e^{-rT}\frac{1}{\sqrt{2\pi}}\int_0^{\infty}\int_{\ln\frac{u+K}{y}}^{\infty}e^{-\frac{x^2}{2T}}\mathrm{d}u \end{aligned}$$

定义 7.6.11. 令

$$X(t)=B(t)+\mu t$$

则称 $X(t)$ 为**有漂移的 Brown 运动**.

注记 7.6.12. 一个感兴趣的量：对有漂移的 Brown 运动，常数 $A,B>0,-B<x<A$，有

$$\mathbb{P}\left\{\text{过程在下降到}-B\text{之前先上升至}A\right\}=\frac{e^{2\mu B}-1}{e^{2\mu B}-e^{-2\mu A}}$$

若 $\mu<0$，令 $B\to\infty$，则有

$$\mathbb{P}\left\{\text{过程迟早上升至}A\right\}=\frac{1}{e^{-2\mu A}}$$

令 $\mu\to 0$，有

$$\mathbb{P}\left\{\text{Brown 运动在下降到}-B\text{之前上升至}A\right\}=\frac{B}{A+B}$$

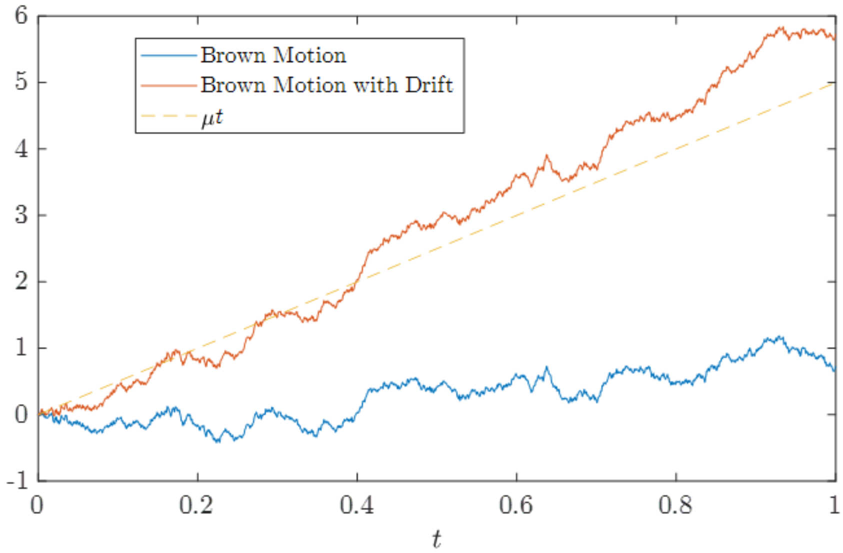

图 7.7. 有漂移的 Brown 运动

7.7 高维 Brown 运动

定义 7.7.1. $B^1(t), B^2(t), \cdots, B^d(t)$ 是相互独立的标准 Brown 运动，则称由它们组成的向量 $\boldsymbol{B}(t) = \left[B^1(t), B^2(t), \cdots, B^d(t)\right]$ 为 **d 维 Brown 运动**.

注记 7.7.2. d 维 Brown 运动的转移概率密度：

$$\frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{1}{2t}\sum_{i=1}^{d}(y_i - x_i)}, x, y \in \mathbb{R}^d$$

定理 7.7.3. d 维 Brown 运动的性质：如果 $\boldsymbol{B}(t)$ 是 d 维 Brown 运动，则

1. $\{\boldsymbol{B}(t), t \geqslant 0\}$ 是鞅；
2. $\left\{|\boldsymbol{B}(t)|^2 - t, t \geqslant 0\right\}$ 是鞅，其中 $|x| = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}, x \in \mathbb{R}^d$；
3. 对任意 $\boldsymbol{u} \in \mathbb{R}^d$，$\left\{e^{\langle \boldsymbol{u}, \boldsymbol{B}(t)\rangle - \frac{1}{2}t|\boldsymbol{u}|^2}, t \geqslant 0\right\}$ 是鞅，称为指数鞅，其中

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{i=1}^{d} x_i y_i, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$$

.

第八章　随机积分与随机微分方程

本章介绍 Itô 积分与随机微分方程，这些内容是理解期权定价理论的基础. 本章要求理解随机游动的积分、Brown 运动的积分、Itô 积分过程及 Itô 积分公式；了解随机微分方程解的存在唯一性定理，掌握常见随机微分方程的解法. 部分内容需要实变函数和测度论的知识，了解即可.

8.1　介绍

8.1.1　关于随机游动的积分

注记 8.1.1. 设 $X_1, X_2, \cdots, X_n$ 为独立同分布的随机变量，且

$$\mathbb{P}\left\{X_i = 1\right\} = \mathbb{P}\left\{X_i = -1\right\} = \frac{1}{2}$$

令 S_n 表示相应的游动，即

$$S_n = X_1 + X_2 + \cdots + X_n$$

形象地来看，X_n 为第 n 次公平赌博的结果，设 $\mathcal{F}_n = \sigma\left(X_1, X_2, \cdots, X_n\right)$，并令 B_n 表示第 n 次赌博的赌注，它只能由第 $n-1$ 次及其之前的信息决定，在第 n 局结果出来之前就以及确定好的，换而言之，它是 $\mathcal{F}_{n-1}$ 可测的随机变量. 于是 n 局结束后，收益 Z_n 为

$$Z_n = \sum_{i=1}^{n} B_i X_i = \sum_{i=1}^{n} B_i\left(S_i - S_{i-1}\right) = \sum_{i=1}^{n} B_i \Delta S_i$$

其中 $\Delta S_i = S_i - S_{i-1}, S_0 = 0$.

我们称 Z_n 为 B_n 关于 S_n 的积分.

容易知道，$\{Z_n\}$ 是关于 $\mathcal{F}_n$ 的鞅，即若 $m < n$ 则

$$\mathbb{E}\left(Z_n \middle| \mathcal{F}_m\right) = Z_m$$

特别地，$\mathbb{E}\left(Z_n\right) = 0$.

此外，如果假定 $\mathbb{E}\left(B_n^2\right) < \infty$，因为

$$Z_n^2 = \sum_{i=1}^{n} B_i^2 X_i^2 + 2 \sum_{1 \leqslant i \leqslant j \leqslant n} B_i B_j X_i X_j$$

而 $\mathrm{Var}\left(X_i^2\right)=1$，如果 $i<j$，则 B_i, X_i, B_j 都是 $\mathcal{F}_{j-1}$ 可测的，且 X_j 独立于 $\mathcal{F}_{j-1}$，那么

$$\mathbb{E}\left(B_iB_jX_iX_j\right)=\mathbb{E}\left[\mathbb{E}\left(B_iB_jX_iX_j\middle|\,\mathcal{F}_{i-1}\right)\right]=\mathbb{E}\left[B_iB_jX_iE\left(X_j\right)\right]=0$$

于是

$$\mathrm{Var}\left(Z_i\right)=\mathbb{E}\left(Z_n^2\right)=\sum_{i=1}^{n}\mathbb{E}\left(B_i^2\right)$$

8.1.2 关于 Brown 运动的积分

注记 8.1.2. 本节的目的是定义关于 Brown 运动的积分 $\int_0^T X\left(t\right)\mathrm{d}B\left(t\right)$ 或 $\int X\mathrm{d}B$.

注记 8.1.3. 简单函数关于 Brown 运动的积分:

设 $\{B\left(t\right)\}$ 是一维的标准 Brown 运动. 考虑一个**非随机的简单过程**$\{X\left(t\right)\}$，即 $X\left(t\right)$ 是一个不依赖于 $B\left(t\right)$ 的简单函数，那么存在 $[0,T]$ 的分割 $0=t_0<t_1<\cdots<t_n=T$ 及常数 $c_0,c_1,\cdots,c_{n-1}$ 使得

$$X\left(t\right)=\begin{cases}c_0, & t=0\\ c_i, & t_i<t\leqslant t_{i+1}, i=0,1,\cdots,n-1\end{cases}$$

与之等价的表达是

$$X\left(t\right)=c_0\mathbb{I}_0\left(t\right)+\sum_{i=0}^{n-1}c_i\mathbb{I}_{(t_i,t_{i+1}]}\left(t\right)$$

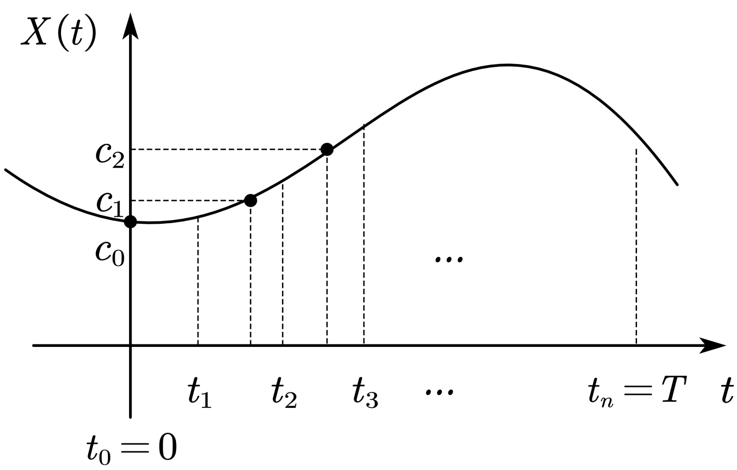

图 8.1. 一般函数积分区间的划分

于是，可以定义其积分为

$$\int_0^T X\left(t\right)\mathrm{d}B\left(t\right)=\sum_{i=1}^{n-1}c_i\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right] \tag{8.1}$$

由 Brown 运动的独立增量性可知，(8.1)是**Gauss 分布的随机变量**，其均值为 0，方差为

$$\begin{aligned}
\operatorname{Var}\left[\int_0^T X(t)\,\mathrm{d}B(t)\right] &= \mathbb{E}\left\{\sum_{i=1}^{n-1} c_i\left[B(t_{i+1})-B(t_i)\right]\right\}^2 \\
&= \mathbb{E}\left\{\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} c_i\left[B(t_{i+1})-B(t_i)\right] c_j\left[B(t_{j+1})-B(t_j)\right]\right\} \\
&= \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} c_i c_j E\left\{\left[B(t_{i+1})-B(t_i)\right]\left[B(t_{j+1})-B(t_j)\right]\right\} \\
&= \sum_{i=1}^{n-1} c_i^2\left(t_{i+1}-t_i\right)
\end{aligned}$$

利用取极限的方法可以将这一定义推广到一般的非随机函数 $X(t)$. 将简单函数中的常数 c_i 用随机变量 ξ_i 来代替，并要求 ξ_i 是 $\mathcal{F}_{t_i}$ 可测的，其中 $\mathcal{F}_t=\sigma(B(u), 0\leqslant u\leqslant t)$. 由 Brown 运动的鞅性，可以得到

$$\mathbb{E}\left\{\xi_i\left[B(t_{i+1})-B(t_i)\right]\middle|\mathcal{F}_i\right\}=\xi_i E\left[B(t_{i+1})-B(t_i)\middle|\mathcal{F}_i\right]=0$$

定义 8.1.4. 设 $\{X(t), 0\leqslant t\leqslant T\}$ 是一个**简单随机过程**，即存在 $[0,T]$ 的分割 $0=t_0<t_1<\cdots<t_n=T$，以及随机变量 $\xi_{-1},\xi_0,\cdots,\xi_{n-1}$ 使得

1. ξ_{-1} 是常数；
2. $\forall i\in\mathbb{N}$，ξ_i 依赖于 $B(t)$ $(t\leqslant t_i)$ 但不依赖于 $B(t)$ $(t>t_i)$；
3. 并且
$$X(t)=\xi_{-1}\mathbb{I}_0(t)+\sum_{i=0}^{n-1}\xi_i\mathbb{I}_{(t_i,t_{i+1}]}(t)$$

此时，定义 **简单随机过程关于 Brown 运动的（Itô）积分** $\int_0^T X\mathrm{d}B$ 为

$$\int_0^T X(t)\,\mathrm{d}B(t)=\sum_{i=0}^{n-1}\xi_i\left[B(t_{i+1})-B(t_i)\right]$$

注记 8.1.5. Itô 积分是一个**随机变量**.

定理 8.1.6. 简单过程的 Itô 积分性质：

1. 线性：如果 $X(t), Y(t)$ 是简单过程，则
$$\int_0^T\left[\alpha X(t)+\beta Y(t)\right]\mathrm{d}B(t)=\alpha\int_0^T X(t)\,\mathrm{d}B(t)+\beta\int_0^T Y(t)\,\mathrm{d}B(t)$$
其中 α,β 为常数.
2.
$$\int_0^T\mathbb{I}_{[a,b]}(t)\,\mathrm{d}B(t)=B(\min\{b,T\})-B(\max\{a,0\})$$
其中 $\mathbb{I}_{[a,b]}(t)$ 是区间 $[a,b]$ 上的示性函数.

3. 零均值性：如果 $\mathbb{E}\left[\xi_i^2\right]<\infty, i=0,1,\cdots,n-1$，则

$$\mathbb{E}\left[\int_0^T X(t)\,\mathrm{d}B(t)\right]=0$$

4. **等距性**：如果 $\mathbb{E}\left[\xi_i^2\right]<\infty, i=0,1,\cdots,n-1$，则

$$\mathbb{E}\left[\int_0^T X(t)\,\mathrm{d}B(t)\right]^2=\int_0^T \mathbb{E}\left[X^2(t)\right]\mathrm{d}t$$

证明. 性质 (1)(2)(3) 由定义即得，下面证明 (4)，由 Cauchy-Schwarz 不等式可知

$$\mathbb{E}\left\{\left|\xi_i\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]\right|\right\}\leqslant\sqrt{\mathbb{E}\left(\xi_i^2\right)\mathbb{E}\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]^2}<\infty$$

于是

$$\begin{aligned}
\operatorname{Var}\left[\int_0^T X(t)\,\mathrm{d}B(t)\right]=&\mathbb{E}\left[\int_0^T X(t)\,\mathrm{d}B(t)\right]^2\\
=&\mathbb{E}\left\{\sum_{i=0}^{n-1}\xi_i\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]\right\}^2\\
=&\mathbb{E}\left\{\sum_{i=0}^{n-1}\xi_i^2\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]^2\right\}\\
&+\mathbb{E}\left\{2\sum_{0\leqslant i<j\leqslant n-1}\xi_i\xi_j\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]\left[B\left(t_{j+1}\right)-B\left(t_j\right)\right]\right\}
\end{aligned}$$

因为 $\forall i\in\mathbb{N}$，ξ_i 依赖于 $B(t)\,(t\leqslant t_i)$ 但不依赖于 $B(t)\,(t>t_i)$，如果 $j>i$ 则有

$$\begin{aligned}
&\mathbb{E}\left\{\xi_i\xi_j\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]\left[B\left(t_{j+1}\right)-B\left(t_j\right)\right]\right\}\\
&=\mathbb{E}\left[\mathbb{E}\left\{\xi_i\xi_j\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]\left[B\left(t_{j+1}\right)-B\left(t_j\right)\right]\mid\mathcal{F}_{t_j}\right\}\right]\\
&=\mathbb{E}\left[\xi_i\xi_j\left[B\left(t_{i+1}\right)-B\left(t_i\right)\right]\mathbb{E}\left\{\left[B\left(t_{j+1}\right)-B\left(t_j\right)\right]\mid\mathcal{F}_{t_j}\right\}\right]\\
&=0
\end{aligned}$$

再利用 Brown 运动的鞅性，可得

$$
\begin{aligned}
\operatorname{Var}\left[\int_0^T X(t)\,\mathrm{d}B(t)\right] &= \mathbb{E}\left\{\sum_{i=0}^{n-1}\xi_i^2\left[B(t_{i+1})-B(t_i)\right]^2\right\}\\
&= \sum_{i=0}^{n-1}\mathbb{E}\left\{\xi_i^2\left[B(t_{i+1})-B(t_i)\right]^2\right\}\\
&= \sum_{i=0}^{n-1}\mathbb{E}\left[\mathbb{E}\left\{\xi_i^2\left[B(t_{i+1})-B(t_i)\right]^2\middle|\mathcal{F}_{t_i}\right\}\right]\\
&= \sum_{i=0}^{n-1}\mathbb{E}\left[\xi_i^2\mathbb{E}\left\{\left[B(t_{i+1})-B(t_i)\right]^2\middle|\mathcal{F}_{t_i}\right\}\right]\\
&= \sum_{i=0}^{n-1}\mathbb{E}\left[\xi_i^2(t_{i+1}-t_i)\right]\\
&= \sum_{i=0}^{n-1}(t_{i+1}-t_i)\,\mathbb{E}\left[\xi_i^2\right]\\
&= \sum_{i=0}^{n-1}(t_{i+1}-t_i)\,\mathbb{E}\left[X(t)\right]^2\\
&= \int_0^T \mathbb{E}\left[X^2(t)\right]\mathrm{d}t
\end{aligned}
$$

□

注记 8.1.7. （该部分教材表述得不够清楚，待完善） 下面用 3 步将随机积分的定义扩展到 V，其中[1]

$$
V := \left\{h: h\text{是定义在}[0,T]\text{上的可测适应过程，满足}\mathbb{E}\left[\int_0^T h^2(s)\,\mathrm{d}s\right]<\infty\right\}
$$

1. 令 $h\in V$ 有界，并且对 $\forall\omega\in\Omega, h(\cdot,\omega)$ 连续，则存在简单过程 $\{\phi_n\}$，其中

$$
\phi_n=\sum_j h(t_j,\omega)\cdot\mathbb{I}_{[t_j,t_{j+1})}(t)\in V
$$

使得当 $n\to\infty$ 时，对 $\forall\omega\in\Omega$ 有

$$
\int_0^T (h-\phi_n)^2\,\mathrm{d}t\to 0
$$

由有界收敛定理知

$$
\mathbb{E}\left[\int_0^T (h-\phi_n)^2\,\mathrm{d}t\right]\to 0
$$

2. 令 $h\in V$ 有界，可以证明：存在 $h_n\in V$ 有界，且对 $\forall\omega\in\Omega,\forall n, h_n(\cdot,\omega)$，使得

$$
\mathbb{E}\left[\int_0^T (h-h_n)^2\,\mathrm{d}t\right]\to 0
$$

[1]回顾适应的定义：设 $\{X(t),t\geqslant 0\}$ 为随机过程，$\{\mathcal{F}_t,t\geqslant 0\}$ 是 σ 代数流，如果 $\forall t, X(t)$ 是 $\mathcal{F}_t$ 可测的，则称 $\{X(t)\}$ 是 $\{\mathcal{F}_t\}$ 适应的.

实际上，不妨设 $|h(t,\omega)| \leqslant M, \forall (t,\omega)$. 再定义

$$h_n(t,\omega) = \int_0^t \psi_n(s-t)\, h(s,\omega)\,\mathrm{d}s$$

其中 ψ_n 是 $\mathbb{R}$ 上的非负连续函数,使得对 $\forall x \in \left(-\frac{1}{n}, 0\right), \psi_n(x) = 0$, 且 $\int_{-\infty}^{+\infty} \psi_n(x)\,\mathrm{d}x = 1$. 那么，对于 $\forall \omega \in \Omega, h_n(\cdot,\omega)$ 连续且 $|h_n(t,\omega)| \leqslant M$. 而 $h \in V$，所以 $h_n \in V$，并且当 $n \to \infty$ 时，对 $\forall \omega \in \Omega$ 有

$$\int_0^T [h_n(s,\omega) - h(s,\omega)]^2\,\mathrm{d}s \to 0$$

再利用有界收敛定理即得所证.

3. 最后，可以证明：对 $\forall f \in V$，存在有界列 $h_n \in V$，使得

$$\mathbb{E}\left[\int_0^T [f(t,\omega) - h_n(t,\omega)]^2\,\mathrm{d}t\right] \to 0$$

事实上，令

$$h_n(t,\omega) = \begin{cases} -n, & f(t,\omega) < -n \\ f(t,\omega), & -n \leqslant f(t,\omega) \leqslant n \\ n, & f(t,\omega) > n \end{cases}$$

再利用控制收敛定理可得所证.

定义 8.1.8. **一般的可测适应随机过程的 Itô 积分**：设 $f \in V(0,T)$，则 f 的 Itô 积分定义为

$$\int_0^T f(t,\omega)\,\mathrm{d}B_t(\omega) = \lim_{n\to\infty} \int_0^T \phi_n(t,\omega)\,\mathrm{d}B_t(\omega) \quad \left(\mathcal{L}^2\text{中极限}\right)$$

其中 $\{\phi_n\}$ 是初等随机过程的序列，使得当 $n \to \infty$ 时，有

$$\mathbb{E}\left[\int_0^T [f(t,\omega) - \phi_n(t,\omega)]^2\,\mathrm{d}t\right] \to 0$$

注记 8.1.9. 实际上，Itô 积分的定义可以推广到更广泛的函数类 V^*

$$V^* = \left\{h : h\text{为定义在}\,[0,T]\,\text{上的可测适应过程，且}\forall T > 0, \int_0^T h^2(s)\,\mathrm{d}s < \infty,\ \text{a.s.}\right\}$$

在实际问题中，我们常常遇到的是仅满足 V^* 中条件的随机过程.

8.2 Itô 积分过程

定义 8.2.1. 假设对任意实数 $T > 0$, 有 $X \in V^*$，那么对 $\forall t \leqslant T$，积分 $\int_0^t X(s)\,\mathrm{d}B(s)$ 是适定的. 因为对任意固定的 t，$\int_0^t X(s)\,\mathrm{d}B(s)$ 是一个随机变量，所以作为上限为 t 的函数，积分

$$Y(t) = \int_0^t X(s)\,\mathrm{d}B(s)$$

定义了一个随机过程 $\{Y(t)\}$，称其为**Itô 积分过程**.

注记 8.2.2. Itô 积分 $Y(t)$ 存在连续的样本路径. 即存在一个连续随机过程 $\{Z(t)\}$ 使得对 $\forall t$, 有 $Y(t)=Z(t)$, a.s..

定理 8.2.3. 设 $X(t)\in V^*$, 并且 $\int_0^\infty \mathbb{E}\left[X^2(s)\right]\mathrm{d}s<\infty$, 则 Itô 积分

$$Y(t)=\int_0^T X(s)\,\mathrm{d}B(s),0\leqslant t\leqslant T$$

是零均值的连续的平方可积鞅, 即 $\sup\limits_{t\leqslant T}\mathbb{E}\left[Y^2(s)\right]<\infty$.

证明. 由 Itô 积分的定义以及 Brown 运动的鞅性, 容易证明

$$\mathbb{E}\left[\int_s^t X(u)\,\mathrm{d}B(u)\middle|\mathcal{F}_s\right]=0,\forall s<t$$

于是

$$\begin{aligned}\mathbb{E}\left[Y(t)|\mathcal{F}_s\right]&=\mathbb{E}\left[\int_0^t X(u)\,\mathrm{d}B(u)\middle|\mathcal{F}_s\right]\\&=\int_0^s X(u)\,\mathrm{d}B(u)+\mathbb{E}\left[\int_s^t X(u)\,\mathrm{d}B(u)\middle|\mathcal{F}_s\right]\\&=\int_0^s X(u)\,\mathrm{d}B(u)\\&=Y(s)\end{aligned}$$

所以 $\{Y(t)\}$ 是鞅, 由 Itô 积分的等距性可以得到二阶矩

$$\mathbb{E}\left[\int_0^t X(s)\,\mathrm{d}B(s)\right]^2=\int_0^t \mathbb{E}\left[X^2(s)\right]\mathrm{d}s<\infty$$

□

推论 8.2.4. 对任意有界的 Borel 可测函数 f, [2]$\int_0^t f\left[B(s)\right]\mathrm{d}B(s)$ 是平方可积鞅.

定理 8.2.5. 如果 $X(t)$ 非随机, 且 $\int_0^T X^2(s)\,\mathrm{d}s<\infty$, 则 $\forall t,Y(t)=\int_0^t X(s)\,\mathrm{d}B(s)$ 是正态分布的随机变量. 即 $\{Y(t),0\leqslant t\leqslant T\}$ 是 Gauss 过程, 其均值为零, 协方差函数为

$$\gamma(t,t+u)=\mathrm{Cov}\left[Y(t),Y(t+u)\right]=\int_0^t X^2(s)\,\mathrm{d}s,u\geqslant 0$$

$\{Y(t)\}$ 也是平方可积鞅.

证明. 因为 $X(t)$ 是非随机的, 所以 $\mathbb{E}\left[X(t)\right]=X(t)$, 那么

$$\int_0^t \mathbb{E}\left[X^2(s)\right]\mathrm{d}s=\int_0^t X^2(s)\,\mathrm{d}s<\infty$$

[2]函数可测的含义是 $\forall a\in\mathbb{R}$ 有 $\{x\in\mathbb{R},f(x)\leqslant a\}\in\mathcal{B}(\mathbb{R})$.

前面已经证明过 $Y(t)$ 的零均值性和鞅性，下面证明协方差函数

$$\begin{aligned}
\gamma(t,t+u) &= Cov\left[Y(t),Y(t+u)\right]\\
&= \mathbb{E}\left[\int_0^t X(s)\,\mathrm{d}B(s)\int_0^{t+u} X(s)\,\mathrm{d}B(s)\right]\\
&= \mathbb{E}\left[\int_0^t X(s)\,\mathrm{d}B(s)\right]^2 + \mathbb{E}\left[\int_0^t X(t)\,\mathrm{d}B(s)\int_t^{t+u} X(s)\,\mathrm{d}B(s)\right]\\
&= \int_0^t \mathbb{E}\left[X^2(s)\right]\mathrm{d}s + \mathbb{E}\left\{\mathbb{E}\left[\int_0^t X(t)\,\mathrm{d}B(s)\int_t^{t+u} X(s)\,\mathrm{d}B(s)\middle|\mathcal{F}_t\right]\right\}\\
&= \int_0^t \mathbb{E}\left[X^2(s)\right]\mathrm{d}s + \mathbb{E}\left\{\int_0^t X(t)\,\mathrm{d}B(s)\mathbb{E}\left[\int_t^{t+u} X(s)\,\mathrm{d}B(s)\middle|\mathcal{F}_t\right]\right\}\\
&= \int_0^t \mathbb{E}\left[X^2(s)\right]\mathrm{d}s
\end{aligned}$$

□

定理 8.2.6. 设 $Y(t)=\int_0^t X(s)\,\mathrm{d}B(s)\ \ (0\leqslant t\leqslant T)$ 是 Itô 积分，则 Y 的二次变差为

$$[Y,Y](t)=\int_0^t X^2(s)\,\mathrm{d}s$$

证明. 因为对一般的过程，可以用简单过程逼近的方法得到，所以这里只证明 $\{X(t)\}$ 为简单过程的情形. 不妨假定 $X(s)$ 在 $[0,T]$ 上仅取两个值，取任意有限多个值的情形可以类似证明. 不失一般性，设 $T=1$，且

$$X(t)=\begin{cases}\xi_0, & 0\leqslant t\leqslant \frac{1}{2}\\ \xi_1, & \frac{1}{2}<t\leqslant 1\end{cases}$$

即

$$X(t)=\xi_0\mathbb{I}_{[0,\frac{1}{2}]}(t)+\xi_1\mathbb{I}_{(\frac{1}{2},1]}(t)$$

于是

$$Y(t)=\int_0^t X(s)\,\mathrm{d}B(s)=\begin{cases}\xi_0 B(t), & 0\leqslant t\leqslant\frac{1}{2}\\ \xi_0 B\left(\frac{1}{2}\right)+\xi_1\left[B(t)-B\left(\frac{1}{2}\right)\right], & \frac{1}{2}<t\leqslant 1\end{cases}$$

因此，对 $[0,t]$ 的任何分割，有

$$Y\left(t_{i+1}^n\right)-Y\left(t_i^n\right)=\begin{cases}\xi_0\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right], & 0\leqslant t_i^n<t_{i+1}^n\leqslant\frac{1}{2}\\ \xi_1\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right], & \frac{1}{2}<t_i^n<t_{i+1}^n\leqslant 1\end{cases}$$

当 $0\leqslant t\leqslant\frac{1}{2}$ 时，有

$$\begin{aligned}
[Y,Y](t) &= \lim_{\delta_n\to0}\sum_{i=0}^{n-1}\left[Y\left(t_{i+1}^n\right)-Y\left(t_i^n\right)\right]^2=\lim_{\delta_n\to0}\sum_{i=0}^{n-1}\xi_0\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2\\
&= \xi_0^2\lim_{\delta_n\to0}\sum_{i=0}^{n-1}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2=\xi_0^2[B,B](t)=\xi_0^2 t=\int_0^t X^2(s)\,\mathrm{d}s
\end{aligned}$$

当 $\frac{1}{2}<t\leqslant 1$ 时，有

$$\begin{aligned}
[Y,Y](t) &= \lim_{\delta_n\to 0}\sum_{i=0}^{n-1}\left[Y\left(t_{i+1}^n\right)-Y\left(t_i^n\right)\right]^2\\
&= \lim_{\delta_n\to 0}\sum_{i=0}^{n-1}\xi_0\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2\\
&= \xi_0^2\lim_{\delta_n\to 0}\sum_{0<t_i\leqslant\frac{1}{2}}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2+\xi_1^2\lim_{\delta_n\to 0}\sum_{\frac{1}{2}<t_i\leqslant 1}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2\\
&= \xi_0^2[B,B]\left(\frac{1}{2}\right)+\xi_1^2[B,B]\left(\left(\frac{1}{2},t\right]\right)\\
&= \frac{1}{2}\xi_0^2+\left(t-\frac{1}{2}\right)\xi_1^2=\int_0^{\frac{1}{2}}X^2(s)\,\mathrm{d}s+\int_{\frac{1}{2}}^t X^2(s)\,\mathrm{d}s\\
&= \int_0^t X^2(s)\,\mathrm{d}s
\end{aligned}$$

□

8.3 Itô 积分公式

注记 8.3.1. Brown 运动二次变差的改写：因为

$$[B,B](t)=t$$

即

$$\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2=t$$

其中 $\{t_i^n\}$ 是 $[0,t]$ 的分割，$\delta_n=\max\limits_{0\leqslant i\leqslant n-1}\left\{t_{i+1}^n-t_i^n\right\}$. 可以将该式子从形式上改写成积分形式：

$$\int_0^t[\mathrm{d}B(s)]^2=t=\int_0^t\mathrm{d}s$$

或

$$[\mathrm{d}B(t)]^2=\mathrm{d}t$$

定理 8.3.2. 设 g 是有界连续函数，$\{t_i^n\}$ 是 $[0,t]$ 的分割，则 $\forall\theta_i^n\in\left(B\left(t_i^n\right),B\left(t_{i+1}^n\right)\right)$，依概率收敛意义下的极限

$$\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}g\left(\theta_i^n\right)\left[B\left(t_{i+1}^n\right)-B\left(t_{i+1}^n\right)\right]^2=\int_0^t g[B(s)]\,\mathrm{d}s$$

证明. 先取 $\theta_i^n=B\left(t_i^n\right)$，由 g 的连续性和积分的定义，有

$$\sum_{i=1}^{n-1}g\left[B\left(t_i^n\right)\right]\left(t_{i+1}^n-t_i^n\right)\to\int_0^t g[B(s)]\,\mathrm{d}s$$

为了证明定理，我们以

$$\sum_{i=0}^{n-1} g\left(\theta_i^n\right)\left[B\left(t_{i+1}^n\right)-B\left(t_{i+1}^n\right)\right]^2$$

作为桥梁.

首先证明

$$\sum_{i=1}^{n-1} g\left[B\left(t_i^n\right)\right]\left(t_{i+1}^n-t_i^n\right)-\sum_{i=0}^{n-1} g\left(\theta_i^n\right)\left[B\left(t_{i+1}^n\right)-B\left(t_{i+1}^n\right)\right]^2 \rightarrow 0 \tag{8.2}$$

在 $\mathcal{L}^2$ 中成立.

记 $\Delta B_i=B\left(t_{i+1}^n\right)-B\left(t_i^n\right), \Delta t_i=t_{i+1}^n-t_i^n$，则由 Brown 运动的独立增量性和迭代期望律，可以得到

$$\begin{aligned}
& \mathbb{E}\left\{\sum_{i=0}^{n-1} g\left[B\left(t_i^n\right)\right]\left[\left(\Delta B_i\right)^2-\Delta t_i\right]\right\}^2 \\
= & \mathbb{E}\left\{\sum_{i=0}^{n-1} g^2\left[B\left(t_i^n\right)\right]\left[\left(\Delta B_i\right)^2-\Delta t_i\right]^2\right\} \\
= & \mathbb{E}\left\{\mathbb{E}\left[\sum_{i=0}^{n-1} g^2\left[B\left(t_i^n\right)\right]\left[\left(\left(\Delta B_i\right)^2-\Delta t_i\right)^2\right] \middle| \mathcal{F}_{t_i}\right]\right\} \\
= & \mathbb{E}\left\{\sum_{i=0}^{n-1} g^2\left[B\left(t_i^n\right)\right] \mathbb{E}\left[\left(\left(\Delta B_i\right)^2-\Delta t_i\right)^2 \middle| \mathcal{F}_{t_i}\right]\right\} \\
= & 2 E\left\{\sum_{i=0}^{n-1} g^2\left[B\left(t_i^n\right)\right]\left(\Delta t_i\right)^2\right\} \leqslant 2 \delta_n E\left\{\sum_{i=0}^{n-1} g^2\left[B\left(t_i^n\right)\right] \Delta t_i\right\} \rightarrow 0
\end{aligned}$$

其中

$$\begin{aligned}
\mathbb{E}\left[\left(\left(\Delta B_i\right)^2-\Delta t_i\right)^2 \middle| \mathcal{F}_{t_i}\right] & =\mathbb{E}\left[\left(\Delta B_i\right)^4-\left(\Delta t_i\right)^2-2\left(\Delta B_i\right)^2 \Delta t_i \middle| \mathcal{F}_{t_i}\right] \\
& =3\left(\Delta t_i\right)^2-\left(\Delta t_i\right)^2-2\left(\Delta t_i\right)^2 \\
& =\left(\Delta t_i\right)^2
\end{aligned}$$

因此，在均方收敛的意义下，有

$$\sum_{i=0}^{n-1} g\left[B\left(t_i^n\right)\right]\left[\left(\Delta B_i\right)^2-\Delta t_i\right] \rightarrow 0$$

这就证明了 $\sum_{i=1}^{n-1} g\left[B\left(t_i^n\right)\right]\left(t_{i+1}^n-t_i^n\right)$ 和 $\sum_{i=0}^{n-1} g\left(\theta_i^n\right)\left[B\left(t_{i+1}^n\right)-B\left(t_{i+1}^n\right)\right]^2$ 有相同的极限，为 $\int_0^t g[B(s)] \mathrm{d} s$，即证得(8.2).

对 $\forall \theta_i^n \in\left(B\left(t_i^n\right), B\left(t_{i+1}^n\right)\right)$，当 $\delta_n \rightarrow 0$ 时，有

$$\sum_{i=0}^{n-1}\left\{g\left(\theta_i^n\right)-g\left[B\left(t_i^n\right)\right]\right\}\left(\Delta B_i\right)^2 \leqslant \max_i\left\{g\left(\theta_i^n\right)-g\left[B\left(t_i^n\right)\right]\right\} \sum_{i=0}^{n-1}\left[B\left(t_{i+1}^n\right)-B\left(t_i^n\right)\right]^2$$

因为 g 是连续函数，B 有连续的样本轨道，则

$$\max_i \{g(\theta_i^n) - g[B(t_i^n)]\} \to 0,\ \text{a.s.}$$

又因为 Brown 运动的二次变差为

$$\sum_{i=0}^{n-1} \left[B\left(t_{i+1}^n\right) - B\left(t_i^n\right)\right]^2 \to t$$

因此，当 $\delta_n \to 0$ 时，有

$$\sum_{i=0}^{n-1} \{g(\theta_i^n) - g[B(t_i^n)]\} (\Delta B_i)^2 \to 0$$

这就证明了 $\sum_{i=1}^{n-1} g(\theta_i^n)(\Delta B_i)^2$ 和 $\sum_{i=0}^{n-1} g(t_i^n)(\Delta B_i)^2$ 有相同的极限 $\int_0^t g[B(s)]\,\mathrm{d}s$（依概率收敛意义下）. □

定理 8.3.3. （Itô 公式）如果 f 是二次可微函数，对任意 t，有

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

证明. 取 $[0,t]$ 的分割 $\{t_i^n\}$，有

$$f[B(t)] = f(0) + \sum_{i=0}^{n-1} \left\{f\left[B\left(t_{i+1}^n\right)\right] - f[B(t_i^n)]\right\} \tag{8.3}$$

对 $f\left[B\left(t_{i+1}^n\right)\right] - f[B(t_i^n)]$ 进行 Taylor 展开，有

$$\begin{aligned} & f\left[B\left(t_{i+1}^n\right)\right] - f[B(t_i^n)] \\ =& f'[B(t_i^n)]\left[B\left(t_{i+1}^n\right) - B(t_i^n)\right] + \frac{1}{2} f''(\theta_i^n)\left[B\left(t_{i+1}^n\right) - B(t_i^n)\right]^2 \end{aligned}$$

其中 $\theta_i^n \in \left(B(t_i^n), B\left(t_{i+1}^n\right)\right)$，将其代入(8.3)得

$$\begin{aligned} f[B(t)] =& f(0) + \sum_{i=0}^{n-1} f'[B(t_i^n)]\left[B\left(t_{i+1}^n\right) - B(t_i^n)\right] \\ &+ \sum_{i=0}^{n-1} \frac{1}{2} f''(\theta_i^n)\left[B\left(t_{i+1}^n\right) - B(t_i^n)\right]^2 \end{aligned}$$

再令 $\delta_n \to 0$，则有

$$\sum_{i=0}^{n-1} f'[B(t_i^n)]\left[B\left(t_{i+1}^n\right) - B(t_i^n)\right] \to \int_0^t f'[B(s)]\,\mathrm{d}B(s)$$

$$\sum_{i=1}^{n-1} f''(\theta_i^n)\left[B\left(t_{i+1}^n\right) - B(t_i^n)\right]^2 \to \int_0^t f''[B(s)]\,\mathrm{d}s$$

那么

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

□

定义 8.3.4. 如果过程 $\{Y(t), 0 \leqslant t \leqslant T\}$ 可以表示为

$$Y(t) = Y(0) + \int_0^t \mu(s)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B(s), 0 \leqslant t \leqslant T$$

其中，过程 $\{\mu(t)\}$ 和 $\{\sigma(t)\}$ 满足

1. $\mu(t)$ 是适应的，并且 $\int_0^T |\mu(t)|\mathrm{d}t < \infty$，a.s. ；
2. $\sigma(t) \in V^*$.

则称 $\{Y(t)\}$ 为**Itô 过程**.

注记 8.3.5. Itô 过程的微分形式:

$$\mathrm{d}Y(t) = \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t), 0 \leqslant t \leqslant T$$

其中，函数 $\mu(t)$ 称为漂移系数，$\sigma(t)$ 称为扩散系数，它们可以易赖与 $Y(t)$ 或 $B(t)$，甚至整个路径 $\{B(s), s \leqslant t\}$，例如 $\mu(t) = \cos\left[\max\limits_{s\leqslant t} B(s) + t\right]$ 依赖于整个路径.

注记 8.3.6. 一个特殊情形：当 $\mu(t)$ 和 $\sigma(t)$ 仅通过 $Y(t)$ 依赖于 t 时，Itô 过程的微分形式为

$$\mathrm{d}Y(t) = \mu[Y(t)]\,\mathrm{d}t + \sigma[Y(t)]\,\mathrm{d}B(t), 0 \leqslant t \leqslant T$$

注记 8.3.7. Itô 公式的微分形式:

$$\mathrm{d}f[B(t)] = f'[B(t)]\,\mathrm{d}B(t) + \frac{1}{2}f''[B(t)]\,\mathrm{d}t$$

例子 8.3.8. 求 $\mathrm{d}\left(e^{B(t)}\right)$.

令 $f(x) = e^x$，由 Itô 公式可得

$$\begin{aligned}
\mathrm{d}\left(e^{B(t)}\right) &= \mathrm{d}f[B(t)] \\
&= f'[B(t)]\,\mathrm{d}B(t) + \frac{1}{2}f''[B(t)]\,\mathrm{d}t \\
&= e^{B(t)}\mathrm{d}B(t) + \frac{1}{2}e^{B(t)}\mathrm{d}t
\end{aligned}$$

设 $X(t) = e^{B(t)}$，即 $\{X(t)\}$ 为几何 Brown 运动，那么

$$\mathrm{d}X(t) = X(t)\,\mathrm{d}B(t) + \frac{1}{2}X(t)\,\mathrm{d}t$$

注记 8.3.9. 在金融应用中，股票的价格 $S(t)$ 用随机微分方程

$$\mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}B(t)$$

来描述，如果 $a(t)$ 表示时刻 t 的股票个股收益，那么在 $[0, T]$ 内的收益为

$$\int_0^T a(t)\,\mathrm{d}S(t) = \mu\int_0^T a(t)\,S(t)\,\mathrm{d}t + \sigma\int_0^T a(t)\,S(t)\,\mathrm{d}B(t)$$

注记 8.3.10. 关于 Itô 过程的 Itô 公式：设 $\{X(t)\}$ 是由

$$\mathrm{d}X(t) = \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t)$$

给出的 Itô 过程，$g(t,x)$ 是 $[0,\infty)\times\mathbb{R}$ 上的二次连续可微函数，则

$$\{Y(t) = g[t, X(t)]\}$$

仍是 Itô 过程，并且

$$\mathrm{d}Y(t) = \frac{\partial g}{\partial t}[t, X(t)]\,\mathrm{d}t + \frac{\partial g}{\partial x}[t, X(t)]\,\mathrm{d}X(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}[t, X(t)]\,[\mathrm{d}X(t)]^2 \tag{8.4}$$

其中 $[\mathrm{d}X(t)]^2 = [\mathrm{d}X(t)]\,[\mathrm{d}X(t)]$，按照表 8.2 的规则（Itô 引理）计算：

表 8.1. Itô 引理

$\times$	$\mathrm{d}B(t)$	$\mathrm{d}t$
$\mathrm{d}B(t)$	$\mathrm{d}t$	0
$\mathrm{d}t$	0	0

那么(8.4)可以改写为

$$\begin{aligned}\mathrm{d}Y(t) =& \frac{\partial g}{\partial t}[t, X(t)]\,\mathrm{d}t + \frac{\partial g}{\partial x}[t, X(t)]\,[\mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t)] \\ &+ \frac{1}{2}\frac{\partial^2 g}{\partial x^2}[t, X(t)]\,[\mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t)]^2 \\ =& \left[\frac{\partial g}{\partial t}x[t, X(t)] + \frac{\partial g}{\partial x}[t, X(t)]\,\mu(t)\right]\mathrm{d}t \\ &+ \frac{\partial g}{\partial x}[t, X(t)]\,\sigma(t)\,\mathrm{d}B(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}[t, X(t)]\,\sigma^2(t)\,\mathrm{d}t\end{aligned}$$

特别地，如果 $g(t,x)$ 仅仅是 x 的函数，即 $g(t,x) = g(x)$，那么

$$\begin{aligned}\mathrm{d}Y(t) &= \frac{\mathrm{d}g}{\mathrm{d}x}[X(t)]\,\mu(t)\,\mathrm{d}t + \frac{\mathrm{d}g}{\mathrm{d}x}[X(t)]\,\sigma(t)\,\mathrm{d}B(t) + \frac{1}{2}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}[X(t)]\,\sigma^2(t)\,\mathrm{d}t \\ &= g'[X(t)]\,\mu(t)\,\mathrm{d}t + g'[X(t)]\,\sigma(t)\,\mathrm{d}B(t) + \frac{1}{2}g''[X(t)]\,\mathrm{d}t\end{aligned}$$

注记 8.3.11. 高维 Itô 公式：设 $\boldsymbol{B}(t) = \left[B^1(t), B^2(t), \cdots, B^d(t)\right]$ 是 d 维 Brown 运动，令

$$\mathrm{d}X_i(t) = \mu_i(t)\,\mathrm{d}t + \sum_{j=1}^{d}\sigma_{ij}(t)\,\mathrm{d}B_j(t), i = 1, 2, \cdots, n$$

是一个 n 维 Itô 过程，设 $\boldsymbol{g} = (g_1, g_2, \cdots, g_m) : [0, +\infty)\times\mathbb{R}^n \to \mathbb{R}^m$ 二次连续可导，则

$$\{\boldsymbol{Y}(t) = \boldsymbol{g}(t, \boldsymbol{X}(t))\}$$

仍是 Itô 过程，即对 $\forall k, 1 \leqslant k \leqslant m$，有

$$\mathrm{d}Y_k(t) = \frac{\partial g_k}{\partial t}[t, \boldsymbol{X}(t)]\,\mathrm{d}t + \sum_j \frac{\partial g_k}{\partial x_j}[t, \boldsymbol{X}(t)]\,\mathrm{d}X_j(t) + \sum_{i,j}\frac{\partial^2 g_k}{\partial x_i \partial x_j}[t, \boldsymbol{X}(t)]\,\mathrm{d}X_i(t)\,\mathrm{d}X_j(t)$$

其中

表 8.2. Itô 引理

$\times$	$\mathrm{d}B_i(t)$	$\mathrm{d}t$
$\mathrm{d}B_j(t)$	$\delta_{ij}\mathrm{d}t$	0
$\mathrm{d}t$	0	0

8.4 随机微分方程

定义 8.4.1. 对一般的 Itô 过程属于**随机微分方程**（SDE），即

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t \tag{8.5}$$

其中，μ, σ 是 $[0, \infty) \times \mathbb{R}$ 上的函数，$\{B_t\}$ 是一维标准 Brown 运动.

注记 8.4.2. 对(8.5)的形式稍作变换，得

$$X_t = X_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s \tag{8.6}$$

这就是随机积分方程.

注记 8.4.3. 对于一般的随机微分方程，通常很难或者根本无法求出显式解.

定义 8.4.4. 设随机过程 $\{X(t), t \geqslant 0\}$ 满足方程(8.6)，则称 $\{X(t), t \geqslant 0\}$ 为随机微分方程(8.5)在**初值为 $X(0) = X_0$ 时的解**.

在随机分析中，有两种类型的解：强解和弱解.

1. 给定漂移系数、扩散系数和随机微分项 $\mathrm{d}B_t$，确定一个随机过程 $\{X_t\}$ 满足(8.6). $\{X_t\}$ 依赖于时间 t 及 Brown 运动过去和现在的值. 这种类型的解称为**强解**.
2. 只需给定漂移系数和扩散系数，确定一个过程 $\left\{\tilde{X}_t\right\}$:

$$\tilde{X}_t = f\left(t, \tilde{B}_t\right)$$

   这里的 Brown 运动和 $\left\{\tilde{X}_t\right\}$ 同时确定. 这种类型的解称为**弱解**.

定理 8.4.5. （强解的存在唯一性定理）设 $[0, T] \times \mathbb{R}$ 上的函数 $\mu(\cdot, \cdot)$ 和 $\sigma(\cdot, \cdot)$ 满足：

1. $\mu(t, x), \sigma(t, x)$ 二元可测，且 $\sqrt{|\mu(t, x)|}, |\sigma(t, x)|$ 平方可积；
2. （Lipschitz 条件）存在常数 $M > 0$，使得对于 $t \in [0, T]$，有

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leqslant M|x - y|, \forall x, y \in \mathbb{R}$$

3. （线性增长条件）存在常数 $K > 0$ 使得

$$|\mu(t, x)| + |\sigma(t, x)| \leqslant K(1 + |x|), \forall t \in [0, T], x \in \mathbb{R}$$

4.（初始条件）随机变量 $X(t_0)$ 关于 $\mathcal{F}_{t_0}$ 可测，且 $\mathbb{E}[X^2(t_0)] < \infty$，

则存在唯一的具有连续路径的随机过程 $\{X_t, t \geqslant t_0\} \in V(t_0, \infty)$ 满足随机微分方程

$$X_t = X_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s, t \geqslant t_0$$

并且该唯一解记为 $\left\{X_t = X_t^{t_0, X_{t_0}}(\omega)\right\}$ 表示解对初始时刻和初始值 X_{t_0} 的依赖.

第九章　随机过程在期权定价中的应用

主要介绍随机过程在金融中的应用之一，Black-Scholes 模型. 首先从基本的金融术语开始，使用随机积分和测度变换的方法对 Black-Scholes 公式进行了推导.

9.1　金融市场的术语与基本假定

注记 9.1.1. 基本概念.

- 套利：开始时无资本，经过资本市场的运作之后，变成非负的随机资金，而且有正的资金的概率为正.
- 无套利原则：市场不允许存在没有初始投资的无风险利润.
- 看涨（跌）期权：在时刻 0 买方与卖方的一个合约，它规定买方有一项**权利**，能在时刻 T（到期日）以价格 K（执行价）从卖方购进（看跌期权对应为卖出）某种有风险的标的资产（一般假设为股票）.
- 看涨期权在时刻 T 的现金流：（股票 T 时刻的市场价格为 S_T）

$$(S_T - K)^+ = \max\{0, S_T - K\} = \begin{cases} S_T - K, & S_T > K \\ 0, & S_T \leqslant K \end{cases}$$

定理 9.1.2. （无套利原则下的看涨与看跌平权关系）假设看涨期权、看跌期权、远期合约在任意时刻 t $(t < T)$ 的价格分别为 $C_t, \mathbb{P}_t, S_t$，在无套利原则下，有

$$C_t - \mathbb{P}_t = S_t - Ke^{-rt}$$

其中 K 为执行价格，r 为无风险利率.

证明. 构建下面组合 A 和组合 B.

- 组合 A：1 个看涨期权、现在 Ke^{-rt} 的现金；
- 组合 B：1 个看跌期权、现在 1 股股票.

在 t 时刻，组合 A 的价值为 $C_t + Ke^{-rt}$，组合 B 的价值为 $\mathbb{P}_t + S_t$，在无套利原则下，组合 A 和 B 的价值应该相等，这就是期权买卖权的平价关系（Pull-Call Parity）. □

注记 9.1.3. 买卖权平价关系的意义：欧式看涨期权和看跌期权中只要知道一个的价格，就可以得到另一个的价格. 特别地，对欧式期权，平权关系还可以写成

$$(S_T - K)^+ - (K - S_T)^+ = S_T - K$$

即表示持有以下两个组合是等价的：

- 组合 A：1 个欧式看涨期权、T 时刻到期的 K 的银行存款；
- 组合 B：1 个欧式看跌期权、T 时刻到期且交割价格为 S_T 的远期合约.

9.2 Black-Scholes 模型

注记 9.2.1. 期权定价的历史：

- 1900 年，Bachelier 首次用随机游走的思想给出股票价格的随机模型；
- 1942 年，Itô 对 Brown 运动引入随机积分，开创随机微分方程理论；
- 1965 年，Samuelson 将随机分析学引入金融学，首次用几何 Brown 运动描述股票价格变动，但他给出的结果依赖于投资者的个人风险偏好；
- 1973 年，Black 和 Scholes 修正了 Samuelson 的结果.

注记 9.2.2. Black-Scholes 模型基本假设：

1. 市场不存在无风险的套利机会；
2. 期权是欧式的，即只能在到期日行权；
3. 市场无摩擦，不存在税收和交易成本，所有证券完全可分割；
4. 无风险利率已知且为常数，投资者可以用无风险利率自由借贷；
5. 股票不发放股利，也不作其他任何形式的利润分配；
6. 对卖空没有限制；
7. 交易时间及价格连续变化.

注记 9.2.3. Black-Scholes 模型基本元素：

1. 风险资产的价格变动：设股票在 t 时刻的价格为 S_t，它的变动满足

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}B_t, t \in [0, T] \tag{9.1}$$

其中，常数 μ 表示风险资产的平均回报率；常数 $\sigma > 0$，表示风险资产的波动率；B_t 是概率测度 $\mathbb{P}$ 下的标准 Brown 运动；T 为期权的到期日. 随机微分方程(9.1)的解是几何 Brown 运动：

$$S_t = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t}$$

证明. 由(9.1)得

$$\frac{\mathrm{d}S_t}{S_t}=\mu\mathrm{d}t+\sigma\mathrm{d}B_t$$

两边同时积分，得

$$\int_0^t\frac{\mathrm{d}S_u}{S_u}=\mu t+\sigma B_t, B_0=0$$

对 $g(t,x)=\ln x$ 用 Itô 公式，得

$$\begin{aligned}\mathrm{d}(\ln S_t)&=\frac{1}{S_t}\mathrm{d}S_t+\frac{1}{2}\left(-\frac{1}{S_t^2}\right)[\mathrm{d}S_t]^2\\&=\frac{1}{S_t}\mathrm{d}S_t+\frac{1}{2}\left(-\frac{1}{S_t^2}\right)\sigma^2S_t^2\mathrm{d}t\\&=\frac{\mathrm{d}S_t}{S_t}-\frac{1}{2}\sigma^2\mathrm{d}t\end{aligned}$$

因此

$$\frac{\mathrm{d}S_t}{S_t}=\mathrm{d}(\ln S_t)+\frac{1}{2}\sigma^2\mathrm{d}t$$

那么

$$\int_0^t\frac{\mathrm{d}S_u}{S_u}=\ln\frac{S_t}{S_0}+\frac{1}{2}\sigma^2=\mu t+\sigma B_t$$

即

$$S_t=S_0e^{\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma B_t}$$

□

2. 无风险资产的价格变动：

$$\mathrm{d}\beta_t=r\beta_t\mathrm{d}t, t\in[0,T]$$

其中，常数 r 为无风险利率. 这是一个常微分方程，容易解得

$$\beta_t=\beta_0e^{rt}$$

3. 在 t 时刻，用数量为 a_t 的风险资产和数量为 b_t 的无风险资产构造投资组合 (a_t,b_t)，该组合的价值为

$$V_t=a_tS_t+b_t\beta_t, t\in[0,T]$$

4. 组合的价值变动：假定投资组合是**自融资的**，即财富的增量仅由组合中 S_t 和 β_t 的变动引起，即[1]

$$\mathrm{d}V_t=a_tS_t+b_t\beta_t$$

[1]一般而言，由 Itô 引理，

$$\mathrm{d}V_t=a_t\mathrm{d}S_t+b_t\mathrm{d}\beta_t+S_t\mathrm{d}a_t+\beta_t\mathrm{d}b_t+\mathrm{d}a_t\mathrm{d}S_t+\mathrm{d}b_t\mathrm{d}\beta_t$$

注记 9.2.4. 1. Black-Scholes 偏微分方程的推导：寻找一个自融资策略 (a_t, b_t)，使得持有该组合相当于持有一个欧式看涨期权，且有

$$V_t = a_t S_t + b_t \beta_t = u\left(T-t, S_t\right), t \in [0,T]$$

其中 $u\left(\tau, x\right)$ 为待求的光滑函数，$\tau = T - t$. 在 T 时刻，根据市场是无套利的假设，投资组合在 T 时刻的价值 V_T 是 T 时刻的现金流 $(S_T - K)^+$，则可以得到一个终端条件：

$$V_T = u\left(0, S_T\right) = \left(S_T - K\right)^+$$

记 $f\left(t,x\right) = u\left(\tau, x\right)$，则

$$V_t = f\left(t, S_t\right)$$

已知 S_t 满足

$$S_t = S_0 + \mu \int_0^t S_s \mathrm{d}s + \sigma \int_0^t S_s \mathrm{d}B_s$$

一方面，对价值过程 $V_t = f\left(t, S_t\right) = u\left(T-t, S_t\right)$ 用 Itô 引理

$$\begin{aligned}
V_t - V_0 =& f\left(t, S_t\right) - f\left(0, S_0\right) \\
=& \int_0^t \mathrm{d}f\left(s, S_s\right) \\
=& \int_0^t \left[\frac{\partial f}{\partial t}\left(s, S_s\right)\mathrm{d}s + \frac{\partial f}{\partial x}\left(s, S_s\right)\mathrm{d}S_s + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\left(s, S_s\right)\left[\mathrm{d}S_s\right]^2\right] \\
=& \int_0^t \left[\frac{\partial f}{\partial t}\left(s, S_s\right)\mathrm{d}s + \frac{\partial f}{\partial x}\left(s, S_s\right)\left[\mu S_s \mathrm{d}s + \sigma S_s \mathrm{d}B_s\right]\right] \\
& + \int_0^t \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\left(s, S_s\right)\left[\mu S_s \mathrm{d}s + \sigma S_s \mathrm{d}B_s\right]^2 \\
=& \int_0^t \left[\frac{\partial f}{\partial t}\left(s, S_s\right)\mathrm{d}s + \mu S_s \frac{\partial f}{\partial x}\left(s, S_s\right)\mathrm{d}s + \frac{1}{2}\sigma^2 S_s^2 \frac{\partial^2 f}{\partial x^2}\left(s, S_s\right)\mathrm{d}s\right] \\
& + \int_0^t \sigma S_s \frac{\partial f}{\partial x}\left(s, S_s\right)\mathrm{d}B_s \\
=& \int_0^t \left[-\frac{\partial u}{\partial t}\left(T-s, S_s\right) + \mu S_s^2 \frac{\partial u}{\partial x}\left(T-s, S_s\right) + \frac{1}{2}\sigma^2 S_s^2 \frac{\partial^2 u}{\partial x^2}\left(T-s, S_s\right)\right]\mathrm{d}s \\
& + \int_0^t \sigma S_s \frac{\partial u}{\partial x}\left(T-s, S_s\right)\mathrm{d}B_s
\end{aligned}$$

另一方面，因为

$$V_t - V_0 = \int_0^t a_s \mathrm{d}S_s + \int_0^t b_s \mathrm{d}\beta_s$$

而

$$\mathrm{d}\beta_t = r\beta_t \mathrm{d}t$$

$$b_t = \frac{V_t - a_t S_t}{\beta_t}$$

所以

$$\begin{aligned}V_t - V_0 &= \int_0^t a_s \mathrm{d}S_s + \int_0^t \frac{V_s - a_s S_s}{\beta_s} r\beta_s \mathrm{d}s \\ &= \int_0^t a_s \mathrm{d}S_s + \int_0^t r\left(V_s - a_s S_s\right)\mathrm{d}s \\ &= \int_0^t a_s\left(\mu S_s \mathrm{d}s + \sigma S_s \mathrm{d}B_s\right) + \int_0^t r\left(V_s - a_s S_s\right)\mathrm{d}s \\ &= \int_0^t \left[rV_s + (\mu - r)\, a_s S_s\right]\mathrm{d}s + \int_0^t a_s \sigma S_s \mathrm{d}B_s\end{aligned}$$

比较 $V_t - V_0$ 的表达式，可知

$$a_t = \frac{\partial u\left(T-t, S_t\right)}{\partial x}$$

和

$$rV_t + (\mu - r)\, a_t S_t = -\frac{\partial u}{\partial t}\left(T-t, S_t\right) + \mu S_t \frac{\partial u}{\partial x}\left(T-t, S_t\right) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 u}{\partial x^2}\left(T-t, S_t\right)$$

于是

$$\begin{aligned}&ru\left(T-t, S_t\right) + (\mu - r)\frac{\partial u\left(T-t, S_t\right)}{\partial x} S_t \\ =&-\frac{\partial u}{\partial t}\left(T-t, S_t\right) + \mu S_t \frac{\partial u}{\partial x}\left(T-t, S_t\right) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 u}{\partial x^2}\left(T-t, S_t\right)\end{aligned}$$

整理得到偏微分方程

$$-ru\left(T-t, S_t\right) + r\frac{\partial u\left(T-t, S_t\right)}{\partial x} S_t - \frac{\partial u}{\partial t}\left(T-t, S_t\right) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 u}{\partial x^2}\left(T-t, S_t\right) = 0$$

其中 $t \in [0, T]$. 或表示为

$$-ru + rx\frac{\partial u}{\partial x} + \frac{\partial u}{\partial \tau} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} = 0, x > 0, t \in [0, T]$$

并且有终端条件 $V_T = u\left(0, S_T\right) = \left(S_T - K\right)^+$，即

$$u(0, x) = (x - K)^+, x > 0$$

注记 9.2.5. 偏微分方程(??)的解为

$$\begin{cases} u(\tau, x) = x\Phi\left[d_1(\tau, x)\right] - Ke^{-rt}\Phi\left[d_2(\tau, x)\right] \\ d_1(\tau, x) = \frac{\ln\left(\frac{x}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}} \\ d_2(\tau, x) = d_1 - \sigma\sqrt{\tau} \end{cases}$$

其中 Φ 是标准正态分布的分布函数. 即

$$\begin{cases} V_0 = u\left(\tau, S_0\right) = S_0\Phi\left[d_1\left(\tau, S_0\right)\right] - Ke^{-rt}\Phi\left[d_2\left(\tau, S_0\right)\right] \\ d_1\left(\tau, S_0\right) = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}} \\ d_2\left(\tau, S_0\right) = d_1 - \sigma\sqrt{\tau} \end{cases}$$

这就是 Black-Scholes 期权定价公式，它与平均回报率 μ 无关，与风险资产的波动率 σ 有关.
使得组合价值运动与期权价值运动等价的自融资策略 (a_t, b_t) 为：

$$\begin{cases} a_t = \frac{\partial u(T-t,S_t)}{\partial x} \\ b_t = \frac{V_t - a_t S_t}{\beta_t} = \frac{u(T-t,S_t)-a_t S_t}{\beta_t} \end{cases}$$

9.3 使用测度变换方法推导 Black-Scholes 公式

注记 9.3.1. 等价鞅测度方法是 Harrison 和 Kreos 再 1979 年提出的，解决期权定价问题的新方法.

定义 9.3.2. 令 $\mathbb{P}$ 和 $\mathbb{Q}$ 是定义在 σ 代数 $\mathcal{F}$ 上的两个概率测度，若存在一个非负函数 f_1，使得

$$\mathbb{Q}(A) = \int_A f_1(\omega)\,\mathrm{d}\mathbb{P}(\omega), A \in \mathcal{F}$$

则称 f_1 为概率测度 $\mathbb{Q}$ 关于概率测度 $\mathbb{P}$ 的**密度**，且称概率测度 $\mathbb{Q}$ 关于概率测度 $\mathbb{P}$**绝对连续**，记为 $\mathbb{Q} \ll \mathbb{P}$.

定义 9.3.3. 若概率测度 $\mathbb{Q}$ 关于概率测度 $\mathbb{P}$ 绝对连续，且概率测度 $\mathbb{P}$ 关于概率测度 $\mathbb{Q}$ 绝对连续，即 $\mathbb{Q} \ll \mathbb{P}$ 且 $\mathbb{P} \ll \mathbb{Q}$，则称 $\mathbb{Q}$和$\mathbb{P}$ 是**等价的概率测度**.

注记 9.3.4. 一般情况下，用 $B = (B_t, t \in [0,T])$ 表示概率测度 $\mathbb{P}$ 下的 Brown 运动，对过程

$$\tilde{B}_t = B_t + qt, t \in [0,T]$$

其中 q 为常数. 当 $q \neq 0$ 时，$\tilde{B}_t$ 不是标准 Brown 运动，但如果将概率测度 $\mathbb{P}$ 转换成一个新的、合适的概率测度 $\mathbb{Q}$，使得 $\tilde{B}_t$ 在新的概率测度 $\mathbb{Q}$ 下是一个标准 Brown 运动. 这就是 Girsanov 定理一个应用.

定理 9.3.5. （Girsanov 定理）令 $\boldsymbol{Y}(t) \in \mathbb{R}^n$ 是一个 Itô 过程，即

$$\mathrm{d}\boldsymbol{Y}_t = \boldsymbol{\rho}_t \mathrm{d}t + \boldsymbol{\theta}_t \mathrm{d}\boldsymbol{B}_t$$

其中，$\boldsymbol{B}_t \in \mathbb{R}^m, \boldsymbol{\rho}_t \in \mathbb{R}^n, \boldsymbol{\theta}_t \in \mathbb{R}^{n\times m}$.

假设存在过程 $\boldsymbol{u}_t$和$\boldsymbol{\alpha}_t$，使得

$$\boldsymbol{\theta}_t \boldsymbol{u}_t = \boldsymbol{\rho}_t - \boldsymbol{\alpha}_t$$

令

$$\boldsymbol{M}_t = e^{-\int_0^t \boldsymbol{u}_s \mathrm{d}\boldsymbol{B}_s - \frac{1}{2}\int_0^t \boldsymbol{u}_s^2 \mathrm{d}s}$$

$$\mathrm{d}\mathbb{Q}(\omega) = \boldsymbol{M}_T(\omega)\,\mathrm{d}\mathbb{P}(\omega)$$

那么，$\boldsymbol{M}_t$ 是关于概率测度 $\mathbb{P}$ 的鞅，$\mathbb{Q}$ 也是概率测度，且过程

$$\tilde{\boldsymbol{B}}_t = \int_0^t \boldsymbol{u}_s \mathrm{d}s + \boldsymbol{B}_t, t \leqslant T$$

是关于概率测度 $\mathbb{Q}$ 的一个标准 Brown 运动，且过程 $\boldsymbol{Y}_t$ 满足方程：

$$\mathrm{d}\boldsymbol{Y}_t = \boldsymbol{\alpha}_t \mathrm{d}t + \boldsymbol{\theta}_t \mathrm{d}\tilde{\boldsymbol{B}}_t$$

注记 9.3.6. 测度变换的目的：消去随机微分方程中的漂移项. 因为

$$\mathrm{d}\boldsymbol{Y}_t = \boldsymbol{\rho}_t \mathrm{d}t + \boldsymbol{\theta}_t \mathrm{d}\boldsymbol{B}_t$$

而

$$\boldsymbol{B}_t = \tilde{\boldsymbol{B}}_t - \int_0^t \boldsymbol{u}_s \mathrm{d}s \Longrightarrow \mathrm{d}\boldsymbol{B}_t = \mathrm{d}\tilde{\boldsymbol{B}}_t - \boldsymbol{u}_t \mathrm{d}t$$

所以

$$\begin{aligned}
\mathrm{d}\boldsymbol{Y}_t &= \boldsymbol{\rho}_t \mathrm{d}t + \boldsymbol{\theta}_t \left(\mathrm{d}\tilde{\boldsymbol{B}}_t - \boldsymbol{u}_t \mathrm{d}t\right) \\
&= \left(\boldsymbol{\rho}_t - \boldsymbol{\theta}_t \boldsymbol{u}_t\right) \mathrm{d}t + \boldsymbol{\theta}_t \mathrm{d}\tilde{\boldsymbol{B}}_t \\
&= \boldsymbol{\alpha}_t \mathrm{d}t + \boldsymbol{\theta}_t \mathrm{d}\tilde{\boldsymbol{B}}_t
\end{aligned}$$

例子 9.3.7. 设 $\boldsymbol{Y}(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} \in \mathbb{R}^2$，且

$$\begin{aligned}
\mathrm{d}Y_1(t) &= 2\mathrm{d}t + \mathrm{d}B_1(t) + \mathrm{d}B_2(t) \\
\mathrm{d}Y_2(t) &= 4\mathrm{d}t + \mathrm{d}B_1(t) - \mathrm{d}B_2(t)
\end{aligned}$$

即

$$\mathrm{d}\boldsymbol{Y}(t) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{bmatrix} \mathrm{d}B_1(t) \\ \mathrm{d}B_2(t) \end{bmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \mathrm{d}\boldsymbol{B}(t)$$

选择 $\alpha_t = 0$，则有

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \Longrightarrow \boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

故令

$$\begin{aligned}
\boldsymbol{M}_t &= e^{-\int_0^t \boldsymbol{u}_s \mathrm{d}\boldsymbol{B}_s - \frac{1}{2}\int_0^t \boldsymbol{u}_s^2 \mathrm{d}s} \\
&= e^{-\int_0^t [u_1 \mathrm{d}B_1(s) + u_2 \mathrm{d}B_2(s)] - \frac{1}{2}\int_0^t \left(u_1^2 + u_2^2\right) \mathrm{d}s} \\
&= e^{-3B_1(t) + B_2(t) - 5t}
\end{aligned}$$

以及

$$\mathrm{d}\mathbb{Q}(\omega) = e^{-3B_1(T)+B_2(T)-5T}\mathrm{d}\mathbb{P}(\omega)$$

$$\mathrm{d}\tilde{\boldsymbol{B}}(t) = \begin{pmatrix} 3 \\ -1 \end{pmatrix}\mathrm{d}t + \mathrm{d}\boldsymbol{B}(t)$$

在概率测度 $\mathbb{Q}$ 下，$\tilde{\boldsymbol{B}}(t)$ 是一个标准 Brown 运动，且

$$\mathrm{d}\boldsymbol{Y}(t) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\mathrm{d}\tilde{\boldsymbol{B}}(t)$$

定理 9.3.8. （一维的 Girsanov 定理）简单来看，Girsanov 定理说明了以下结论成立.

1. 随机过程
$$M_t = e^{-qB_t-\frac{1}{2}q^2t}, t \in [0,T]$$
是概率测度 $\mathbb{P}$ 下，关于 $\mathcal{F}_t = \sigma(B_s, s \leqslant t)$ 的鞅.
2. 式
$$\mathbb{Q}(A) = \int_A M_T(\omega)\,\mathrm{d}\mathbb{P}(\omega), A \in \mathcal{F}$$
定义了一个 $\mathcal{F}$ 上的概率测度 $\mathbb{Q}$，且 $\mathbb{P}$ 与 $\mathbb{Q}$ 等价.
3. 在概率测度 $\mathbb{Q}$ 下，过程
$$\tilde{B}_t = B_t + qt, t \in [0,T]$$
是一个标准 Brown 运动.
4. 过程 $\tilde{B}_t$ 适应于 $\mathcal{F}_t = \sigma(B_s, s \leqslant t)$.

概率测度 $\mathbb{Q}$ 就被称为等价鞅测度.

例子 9.3.9. 线性随机微分方程漂移项的消除.

假设风险资产的价格运动满足

$$\mathrm{d}S_t = \mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}B_t, t \in [0,T]$$

方程的解为

$$S_t = S_0 e^{\left(\mu-\frac{1}{2}\sigma^2\right)t+\sigma B_t}, t \in [0,T]$$

引入

$$\tilde{B}_t = B_t + \frac{\mu}{\sigma}t, t \in [0,T]$$

则有

$$\mathrm{d}S_t = \mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}\left(\tilde{B}_t - \frac{\mu}{\sigma}t\right) = \sigma S_t\mathrm{d}\tilde{B}_t, t \in [0,T]$$

由 Girsanov 定理，在新的概率测度 $\mathbb{Q}$ 下，$\{S_t\}$ 是鞅.

定理 9.3.10. 假设在 Black-Scholes 模型中，存在一个自融资策略 (a_t, b_t) 使得投资组合在 t 时刻的价值为

$$V_t = a_t S_t + b_t \beta_t, t \in [0, T]$$

且此投资组合在到期日 T 时刻的价值等于未定权益在 T 时刻的价值，即

$$V_T = h(S_T)$$

则投资组合在到期日 t 时刻的价值为

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)} h(S_T) \middle| \mathcal{F}_t\right], t \in [0, T]$$

其中 $\mathbb{E}^{\mathbb{Q}}(A|\mathcal{F}_t)$ 表示在新的概率测度 $\mathbb{Q}$ 下随机变量 A 关于 $\mathcal{F}_t = \sigma(B_s, s \leqslant t)$ 的条件期望.

证明. 股票的贴现价格为

$$\tilde{S}_t = e^{-rt} S_t, t \in [0, T]$$

在新的概率测度 $\mathbb{Q}$ 下，$\left\{\tilde{S}_t\right\}$ 是鞅，令

$$f(t, x) = e^{-rt} x$$

由 Itô 公式，可得

$$\begin{aligned}
\mathrm{d}\tilde{S}_t &= -re^{-rt} S_t \mathrm{d}t + e^{-rt} \mathrm{d}S_t \\
&= -re^{-rt} S_t \mathrm{d}t + e^{-rt}(\mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}B_t) \\
&= e^{-rt} S_t [(\mu - r)\mathrm{d}t + \sigma \mathrm{d}B_t] \\
&= \sigma \tilde{S}_t \mathrm{d}\tilde{B}_t
\end{aligned}$$

其中

$$\tilde{B}_t = B_t + \frac{\mu - r}{\sigma} t, t \in [0, T]$$

由 Girsanov 定理可知，存在一个等价鞅测度 $\mathbb{Q}$，使得 $\tilde{B}_t$ 在 $\mathbb{Q}$ 测度下是一个标准 Brown 运动. 可以解得

$$\tilde{S}_t = \tilde{S}_0 e^{-\frac{1}{2}\sigma^2 t + \sigma \tilde{B}_t}$$

类似地，考虑组合的贴现财富过程

$$\tilde{V}_t = e^{-rt} V_t = e^{-rt}(a_t S_t + b_t \beta_t)$$

由 Itô 公式

$$\begin{aligned}
\mathrm{d}\tilde{V}_t &= \mathrm{d}\left(e^{-rt}V_t\right) \\
&= -re^{-rt}V_t\mathrm{d}t + e^{-rt}\mathrm{d}V_t \\
&= -re^{-rt}\left(a_tS_t + b_t\beta_t\right)\mathrm{d}t + e^{-rt}\left(a_t\mathrm{d}S_t + b_t\mathrm{d}\beta_t\right) \\
&= -re^{-rt}\left(a_tS_t + b_t\beta_t\right)\mathrm{d}t + e^{-rt}\left(a_t\mathrm{d}S_t + b_tr\mathrm{d}\beta_t\right) \\
&= -ra_te^{-rt}S_t\mathrm{d}t + e^{-rt}a_t\mathrm{d}S_t \\
&= a_t\left(-re^{-rt}S_t\mathrm{d}t + e^{-rt}\mathrm{d}S_t\right) \\
&= a_t\mathrm{d}\tilde{S}_t \\
&= a_t\sigma\tilde{S}_t\mathrm{d}\tilde{B}_t
\end{aligned}$$

因为

$$\tilde{V}_0 = V_0$$

所以

$$\tilde{V}_t = V_0 + \sigma\int_0^t a_s\tilde{S}_s\mathrm{d}\tilde{B}_s$$

在等价鞅测度 $\mathbb{Q}$ 下，$\tilde{B}$ 为标准 Brown 运动，$\left\{a_t\tilde{S}_t\right\}$ 是 $\mathcal{F}_t$ 适应的. 故 $\left\{\tilde{V}_t\right\}$ 是关于 $\mathcal{F}_t$ 的鞅，那么

$$\tilde{V}_t = \mathbb{E}^{\mathbb{Q}}\left[\tilde{V}_T\middle|\mathcal{F}_t\right], t\in[0,T]$$

而

$$\tilde{V}_T = e^{-rT}h\left(S_T\right)\tilde{V}_t = e^{-rt}V_t$$

故

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}h\left(S_T\right)\middle|\mathcal{F}_t\right], t\in[0,T]$$

得证. □

注记 9.3.11. 用等价鞅测度推导 Black-Scholes 公式：因为

$$\begin{cases}\tilde{S}_T = \tilde{S}_0e^{-\frac{1}{2}\sigma^2T+\sigma\tilde{B}_T} \\ \tilde{S}_t = \tilde{S}_0e^{-\frac{1}{2}\sigma^2t+\sigma\tilde{B}_t}\end{cases} \Longrightarrow \tilde{S}_T = \tilde{S}_te^{-\frac{1}{2}\sigma^2(T-t)+\sigma\left(\tilde{B}_T-\tilde{B}_t\right)}$$

那么

$$S_T = S_te^{-\frac{1}{2}\sigma^2(T-t)+\sigma\left(\tilde{B}_T-\tilde{B}_t\right)}\cdot e^{r(T-t)} = S_te^{\left[r-\frac{1}{2}\sigma^2\right](T-t)+\sigma\left(\tilde{B}_T-\tilde{B}_t\right)}$$

令 $\tau = T-t, t\in[0,T]$，故组合在时刻 t 的价值为

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}h\left(S_T\right)\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}h\left(S_te^{\left[r-\frac{1}{2}\sigma^2\right]\tau+\sigma\left(\tilde{B}_T-\tilde{B}_t\right)}\right)\middle|\mathcal{F}_t\right]$$

在 t 时刻，S_t 是 B_t 的函数，故有 $\sigma(S_t) \subset \mathcal{F}_t$. 而且在概率测度 $\mathbb{Q}$ 下，$\tilde{B}_T - \tilde{B}_t$ 关于 $\mathcal{F}_t$ 独立，且是正态分布 $N(0,\tau)$ 的随机变量. 由条件期望的性质[2]，可得

$$\begin{aligned} V_t &= f(t, S_t) = u(\tau, x) \\ &= e^{-t\tau} \int_{-\infty}^{+\infty} h\left[x e^{\left(r-\frac{1}{2}\sigma^2\right)\tau + \sigma y \sqrt{\tau}}\right] \varphi(y)\,\mathrm{d}y \end{aligned}$$

其中 φ 是标准正态分布的密度函数. 特别地，对于欧式看涨期权，有

$$h(x) = (x-K)^+ = \max(0, x-K)$$

所以

$$\begin{aligned} V_t &= e^{-r\tau} \int_{-\infty}^{+\infty} \max\left[0, x e^{\left(r-\frac{1}{2}\sigma^2\right)\tau + \sigma y\sqrt{\tau}} - K\right] \varphi(y)\,\mathrm{d}y \\ &= \int_{-\infty}^{+\infty} \left(x e^{-\frac{1}{2}\sigma^2\tau + \sigma y \sqrt{\tau}} - K e^{-r\tau}\right) \varphi(y)\,\mathrm{d}y \\ &= \int_{-z_2}^{+\infty} x e^{-\frac{1}{2}\sigma^2\tau + \sigma y\sqrt{\tau}} \varphi(y)\mathrm{d}y - K e^{-r\tau} \int_{-z_2}^{+\infty} \varphi(y)\,\mathrm{d}y \\ &= x\Phi(z_1) - K e^{-r\tau}\Phi(z_2) \end{aligned}$$

其中 Φ 是标准正态分布的分布函数，且

$$z_1 = \frac{\ln\left(\frac{x}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}, z_2 = z_1 - \sigma\sqrt{t}$$

令其中 $x = S_t$ 即得在 t 时刻，看涨期权的价格.

注记 9.3.12. 对于看跌期权，可以采用类似的方法或者看涨看跌期权平价公式得出其价格为

$$V_t = u(\tau, S_t) = K e^{-r\tau}\Phi(-z_2) - S_t\Phi(-z_1)$$

[2]回顾：如果 X 与 $\mathcal{F}$ 相互独立，且随机变量、随机向量或随机过程 G 携带的信息包含在 $\mathcal{F}$ 中，那么对于任意的函数 $h(x,y)$ 有

$$\mathbb{E}\left[h(X,G)|\mathcal{F}\right] = \mathbb{E}\left(\mathbb{E}_X\left[h(X,G)\right]|\mathcal{F}\right)$$

其中 $\mathbb{E}_X[h(X,G)]$ 的含义是固定 G 对 X 取期望.