应用随机过程

李东风

2023-12-21

目录

简介

这是北京大学数学科学学院金融数学系《金融中的随机数学》1 班的讲义，采用的教材为张波、商豪、邓军（2020）《应用随机过程》第 5 版，中国人民大学出版社。

参考书：

- (Ross 2019)：Introduction to Probability Models.
- (Shreve 2004)：Stochastic Calculus for Finance II Continuous Time Models.
- (刘勇 2022)：应用随机分析，内部讲义。
- (严加安 2023) 严加安. 2023. 金融数学引论. 第二版. 科学出版社.

Chapter 1

预备知识

1.1 概率空间

随机试验是概率论的基本概念，试验的结果事先不能准确地预言，但具有如下三个特性:

（1）可以在相同的条件下重复进行；

（2）每次试验的结果不止一个, 但预先知道试验的所有可能的结果；

（3）每次试验前不能确定哪个结果会出现.

- 样本点 (ω)：随机试验的基本结果；
- 样本空间 (Ω)：随机试验所有可能结果组成的集合；
- 基本事件：Ω 中的样本点 ω；
- 必然事件：样本空间 Ω；
- 不可能事件：空集 $\emptyset$；
- 事件：由基本事件组成的 Ω 中的子集 A。

定义 1.1. 设 Ω 是一个样本空间 (或任意一个非空集合)，$\mathscr{F}$ 是 Ω 的某些子集

组成的集合族, 满足:

$$(1)\ \Omega \in \mathscr{F};$$
$$(2)\ 若A \in \mathscr{F}, 则A^c = \Omega \setminus A \in \mathscr{F};$$
$$(3)\ 若A_n \in \mathscr{F}, n = 1, 2, \ldots, 则\bigcup_{n=1}^{\infty} A_n \in \mathscr{F}.$$

则称 $\mathscr{F}$ 为σ **代数**. $(\Omega, \mathscr{F})$ 称为**可测空间**，$\mathscr{F}$ 中的元素称为**随机事件**，简称事件.

σ 代数 $\mathscr{F}$ 的性质：

(1) $\Omega \in \mathscr{F}$, $\emptyset \in \mathscr{F}$;

(2) 对求余运算封闭;

(3) 对有限并和可列并封闭;

(4) 对有限交和可列交封闭;

(5) 对减法封闭.

以 Ω 的某些子集为元素的集合称为 (Ω 上的) 集类. 对于 Ω 上的任一非空集类 $\mathcal{C}$，存在包含 $\mathcal{C}$ 的最小 σ 代数，即

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

称为由 $\mathcal{C}$ 生成的 σ 代数，记为 $\sigma(\mathcal{C})$.

例 1.1. 设 $\mathscr{G}$ 为 σ 代数，则 $\{\Omega, \emptyset\} \subset \mathscr{G}$。

$\{\Omega, \emptyset\}$ 也是一个 σ 代数.

例 1.2. 设 $A \subset \Omega$，$A \neq \Omega$, $A \neq \emptyset$。构造一个 σ 代数。

取 $\sigma(\{A\}) = \{\Omega, \emptyset, A, A^c\}$。这代表了与事件 A 有关的四种情况：一定发生，一定不发生，A 发生，A 不发生。

比如，设概率空间 Ω 为掷一次骰子的结果，则

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

令 A 表示 “掷出偶数点”，则

$$\sigma(\{A\}) = \{\Omega, \emptyset, \{2, 4, 6\}, \{1, 3, 5\}\}.$$

σ 代数表示一个信息集合，知道了上述的 $\sigma(\{A\})$，就能够知道掷出的点数是偶数还是奇数。

定义 1.2. 设 $\Omega = \mathbb{R}$. 由所有半无限区间 $(-\infty, x]$ 生成的 σ 代数 (即包含集族 $\{(-\infty, x], x \in \mathbb{R}\}$ 的最小 σ 代数) 称为 $\mathbb{R}$ 上的 **Borel σ 代数**, 记为 $\mathscr{B}(\mathbb{R})$, 其中的元素称为 Borel 集合.

类似地，可定义 $\mathbb{R}^n$ 上的 Borel σ 代数 $\mathscr{B}(\mathbb{R}^n)$，如

$$\mathscr{B}(\mathbb{R}^n) = \sigma(\{(-\infty, x_1] \times \cdots \times (-\infty, x_n] : (x_1, \ldots, x_n) \in \mathbb{R}^n\}).$$

定义 1.3. 设 $\{A_n, n \geq 1\}$ 为一集合序列. 令

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k;$$

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

分别称其为 $\{A_n\}$ 的**上极限**和**下极限**。上极限有时也记为 $\{A_n, \text{i.o.}\}$

显然有

$$\begin{aligned}\limsup_{n\to\infty} A_n =& \{\omega|\omega\text{属于无穷多个}A_n\}\\ =& \{\omega|\forall n \in \mathbb{N}, \exists k \geq n, \text{使}\omega \in A_k\};\\ \liminf_{n\to\infty} A_n =& \{\omega|\omega\text{至多不属于有限多个}A_n\}\\ =& \{\omega|\exists n \in \mathbb{N}, \forall k \geq n, \text{有}\omega \in A_k\}.\end{aligned}$$

显然 $\liminf_{n\to\infty} A_n$ 的条件更严格，而 $\limsup_{n\to\infty} A_n$ 的条件更宽松，有

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n.$$

定义 1.4. 设 $\{A_n, n \geq 1\}$ 为一集合序列，且 $\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n$，则称 $\{A_n\}$ 的极限存在，并用 $\lim_{n\to\infty} A_n$ 表示，即令

$$\lim_{n\to\infty} A_n = \liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n.$$

特别地，若对每个 n，有 $A_n \subset A_{n+1}$，则称 $\{A_n\}$ 为单调增的（或单调不减的）；若每个 n，有 $A_n \supset A_{n+1}$，则称 $\{A_n\}$ 为单调降的（或单调不增的）. 对单调

增序列 $\{A_n\}$，令 $A=\bigcup_{n=1}^{\infty}A_n$，称 A 为 $\{A_n\}$ 的极限，通常记为 $A_n\uparrow A$；对单调降序列 $\{A_n\}$，令 $A=\bigcap_{n=1}^{\infty}A_n$，称 A 为 $\{A_n\}$ 的极限，记作 $A_n\downarrow A$.

例 1.3. 设有某人在反复地投掷硬币，观察硬币朝上的面是正面或反面. $\Omega=\{$所有由投掷结果正面和反面组成的序列$\}$, $\mathscr{F}=\Omega$的所有子集, 记 A_n 为第 n 次投掷的是“正面”的事件，则

$$\begin{aligned}\limsup_{n\to\infty}A_n=&\{\text{有无限多个投掷结果是正面}\};\\ \liminf_{n\to\infty}A_n=&\{\text{除有限多个外，投掷结果都是正面}\}.\end{aligned}$$

称事件 A,B **互不相容**，若 $A\cap B=\emptyset$。

定义 1.5. 设 $(\Omega,\mathscr{F})$ 是可测空间，$P(\cdot)$ 是定义在 $\mathscr{F}$ 上的实值函数，满足

$$\begin{aligned}&(1)\ P(\Omega)=1;\\ &(2)\ \forall A\in\mathscr{F},\ 0\le P(A)\le 1;\\ &(3)\ \text{对两两互不相容事件}A_1,A_2,\dots,\text{有}\\ &\qquad P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i).\end{aligned}$$

则称 P 是 $(\Omega,\mathscr{F})$ 上的**概率**（或概率测度），$(\Omega,\mathscr{F},P)$ 称为**概率空间**，$\mathcal{F}$ 中的元素称为**事件**，$P(A)$ 称为事件 A 的概率.

概率有如下性质：

(1) 若$A, B \in \mathscr{F}$, 则

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(2) 可减性：若$A, B \in \mathscr{F}$, 且$A \subset B$, 则$P(B - A) = P(B) - P(A)$.

(3) 单调性：若$A, B \in \mathscr{F}$, 且$A \subset B$, 则$P(A) \leq P(B)$.

(4) 次可加性：若$A_n \in \mathscr{F}, n \geq 1$ 则

$$P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n).$$

(5) 从下连续：若$A_n \in \mathscr{F}$ 且$A_n \uparrow A \in \mathscr{F}$ 则

$$P(A) = P(\bigcup_{n=1}^{\infty} A_n) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n);$$

(6) 从上连续：若$A_n \in \mathscr{F}$ 且$A_n \downarrow A \in \mathscr{F}$ 则

$$P(A) = P(\bigcap_{n=1}^{\infty} A_n) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

概率测度是测度的特例。如果集合函数 $\mu : \mathscr{F} \to \mathbb{R}$ 满足概率测度的条件但不要求 $P(\Omega) = 1$，则称 $\mu(\cdot)$ 为可测空间 $(\Omega, \mathscr{F})$ 上的一个**测度**。

最常用的测度是 **Borel 测度**。对区间 $(a, b]$，定义 $\mu((a, b]) = b - a$，这个集合函数 $\mu(\cdot)$ 可以推广到 Borel 代数上定义，就是长度的概念的推广。

如果概率空间 $(\Omega, \mathscr{F}, P)$ 的 P 零测集 (即零概率事件) 的每个子集仍为事件，则称之为**完备的概率空间**. 为了避免 P 零测集的子集不是事件的情形出现，我们把概率测度完备化. 令 $\mathcal{N}$ 代表 Ω 的所有 P 零测集的子集的全体，由 $\{\mathscr{F}, \mathcal{N}\}$ 生成的 σ 代数 (即包含 $\mathscr{F}$ 和 $\mathcal{N}$ 的最小 σ 代数) 称为 $\mathscr{F}$ 的**完备化**，记为 $\overline{\mathscr{F}}$. $\overline{\mathscr{F}}$ 中的每个集合 B 都可以表为 $B = A \cup N$，其中 $A \in \mathscr{F}$, $N \in \mathcal{N}$, 且 $A \cap N = \emptyset$. 定义

$$\bar{P}(B) = \bar{P}(A \cup N) = P(A).$$

则 P 就被扩张到 $\overline{\mathscr{F}}$ 上.

容易验证，$\bar{P}$ 是 $\overline{\mathscr{F}}$ 上的概率测度，集函数 $\bar{P}$ 称为 P 的完备化. 本书总假定 P 是完备的概率测度.

1.2 随机变量与分布函数

1.2.1 随机变量

定义 1.6. 设 $(\Omega, \mathscr{F}, P)$ 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, 如果对任意实数 $x \in \mathbb{R}$, $\{\omega : X(\omega) \leq x\} \in \mathscr{F}$，则称 $X(\omega)$ 是 $\mathscr{F}$ 上的**随机变量**，简称为随机变量.

$$F(x) = P(\omega : X(\omega) \leq x),\ -\infty < x < \infty$$

称为随机变量 X 的**分布函数**.

如果存在函数 $f(x)$, 满足

$$F(x) = \int_{-\infty}^{x} f(t)\, dt,$$

则称 $f(x)$ 为随机变量 X 或其分布函数 $F(x)$ 的**分布密度**. 如果 X 具有分布密度, 则称 X 为**连续型**随机变量; 如果 X 最多以正概率取可数多个值, 则称 X 为**离散型**随机变量.

定义 1.7. 两个随机变量 X 与 Y, 如果满足 $P(\omega \in \Omega : X(\omega) \neq Y(\omega)) = 0$, 则称它们是**等价的**.

对于两个等价的随机变量，我们视为同一个.

定理 1.1. 下列命题等价：

$$\begin{aligned}
&(1)\ X\text{是随机变量};\\
&(2)\ \{\omega : X(\omega) \geq x\} \in \mathscr{F},\ \forall x \in \mathbb{R};\\
&(3)\ \{\omega : X(\omega) > x\} \in \mathscr{F},\ \forall x \in \mathbb{R};\\
&(4)\ \{\omega : X(\omega) < x\} \in \mathscr{F},\ \forall x \in \mathbb{R}.
\end{aligned}$$

证明略。

注：习惯上将 $\{\omega : X(\omega) \geq x\}$ 记为 $\{X \geq x\}$.

例 1.4. 设事件 $A \in \mathscr{F}$，令

$$I_A(\omega) = \begin{cases} 1, & \text{若}\omega \in A, \\ 0, & \text{若}\omega \notin A, \end{cases}$$

则 $I_A(\omega)$ 是随机变量，简记为 I_A 或 $I[A]$，称为 A 的**示性函数**。

证明：对 $x<0$, $\{I_A \le x\} = \emptyset \in \mathscr{F}$; 对 $0 \le x < 1$, $\{I_A \le x\} = \{I_A = 0\} = A^c \in \mathscr{F}$; 对 $x \ge 1$, $\{I_A \le x\} = \Omega \in \mathscr{F}$.

对随机变量 X，令

$$\sigma(X) = \sigma(\{X^{-1}((-\infty, x]),\ x \in \mathbb{R}\}) = \sigma(\{\{\omega : X(\omega) \le x\} :\ x \in \mathbb{R}\}),$$

称为由 X 生成的 σ 代数，可以理解为观测到 X 后所能获得的信息，即对 $\forall A \in \sigma(X)$，在观测到 X 的值后都可以确定 A 是否发生。

$\sigma(X)$ 的一个等价定义是

$$\sigma(X) = \{X^{-1}(B) : B \in \mathscr{B}(\mathbb{R})\},$$

其中 $X^{-1}(B) = \{\omega : X(\omega) \in B\}$。$\Omega \to \mathbb{R}$ 的函数 X 是 $\mathscr{F}$ 随机变量，当且仅当 $\sigma(X) \subset \mathscr{F}$。

分布函数 $F(x)$ 具有如下性质：

(1) 单调增、右连续；

(2) $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$.

满足上述性质的函数 $F(x)$ 称为**分布函数**，必存在概率空间 $(\Omega, \mathscr{F}, P)$ 以及其中的随机变量 X 使得 X 以 $F(x)$ 为分布函数。事实上，取 $\Omega = \mathbb{R}$, $\mathscr{F} = \mathscr{B}(\mathbb{R})$, $P((-\infty, x]) = F(x)$，则对 $\omega \in \mathbb{R}$ 定义 $X(\omega) = \omega$ 即可。所以，任何一个分布函数 $F(x)$ 都定义了 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的一个概率测度 $\mu_F(\cdot)$，满足

$$\mu_F((-\infty, b]) = F(b), \quad \mu_F((a, b]) = F(b) - F(a).$$

例 1.5. 对 $A \in \mathscr{F}$, 考虑 $X = I_A$，易见

$$\sigma(X) = \sigma(\{\emptyset, A, \Omega\}) = \{\emptyset, A, \Omega, A^c\} = \sigma(\{A\}),$$

所以 $\sigma(I_A) = \sigma(\{A\})$ 包含了所有的关于 A 是否发生的信息，观测到 I_A 的值就可以确定事件 A 是否发生。

1.2.2 随机向量

定义 1.8. 若 $X_1, \dots, X_n$ 是 $\mathscr{F}$ 随机变量，称 $X = (X_1, \dots, X_n)$ 为**随机向量**。

定义

$$\sigma(X) = \sigma(X_1, \dots, X_n) = \sigma(\{X_i^{-1}((-\infty, x]) : x \in \mathbb{R}, i = 1, \dots, n\}).$$

用 $\sigma(X_1, \dots, X_n)$ 表示观测到 $X_1, \dots, X_n$ 以后能够确定是否发生的所有事件的集合，或简单理解为知道 $X_1, \dots, X_n$ 的值所能获得的信息。

随机向量 X 也可以看成是 Ω 到 $\mathbb{R}^n$ 的函数，有

$$\sigma(X) = \sigma(\{X^{-1}(B) : B \in \mathscr{B}(\mathbb{R}^n)\}).$$

设 X 和 Y 是两个随机变量，令 $Z = X + iY$（其中 i 是虚数单位），称 Z 为**复值随机变量**，也可以看成是 Ω 到复数域 $\mathbb{C}$ 的可测映射。

定理 1.2. *(1) 若 X, Y 是随机变量，则 $\{X < Y\}$, $\{X \le Y\}$, $\{X = Y\}$ 及 $\{X \ne Y\}$ 都属于 $\mathscr{F}$；*

(2) 若 X, Y 是随机变量，则 $X \pm Y$ 与 XY 亦然；

(3) 若 $\{X_n\}$ 是随机变量序列，则 $\sup_n X_n$，$\inf_n X_n$，$\limsup_{n\to\infty} X_n$ 和 $\liminf_{n\to\infty} X_n$ 都是随机变量.

对于随机向量 $X = (X_1, \dots, X_n)$, 定义它的 (联合) 分布函数定义为

$$F(x_1, \dots, x_n) = P(X_1 \le x_1, \dots, X_n \le x_n).$$

定理 1.3. *若 $F(x_1, \dots, x_n)$ 是联合分布函数，则*

(1) $F(x_1, \dots, x_n)$ 对每个变量都是单调的;

(2) $F(x_1, \dots, x_n)$ 对每个变量都是右连续的;

(3) 对 $i = 1, 2, \dots, n$ 有

$$\lim_{x_i \to -\infty} F(x_1, \dots, x_i, \dots, x_n) = 0,$$
$$\lim_{x_1, x_2, \dots, x_n \to \infty} F(x_1, x_2, \dots, x_n) = 1.$$

如果 $f(x_1, \dots, x_n) = \frac{\partial^n F}{\partial x_1 \dots \partial x_n}$ 对所有的 $(x_1, \dots, x_n) \in \mathbb{R}^n$ 存在, 则称函数 $f(x_1, \dots, x_n)$ 为 $F(x_1, \dots, x_n)$ 或 $X = (X_1, \dots, X_n)$ 的**联合密度函数**, 并且

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \dots, t_n)\, dt_n \cdots dt_1.$$

设 $F(x_1,\dots,x_n)$ 为 $X_1,\dots,X_n$ 的联合分布函数, $1\le k_1<\cdots<k_m\le n$，则 $X_1,\dots,X_m$ 的**边缘分布**$F_{k_1,\dots,k_m}(x_{k_1},\dots,x_{k_m})$ 定义为

$$\begin{aligned}&F_{k_1,\dots,k_m}(x_{k_1},\dots,x_{k_m})\\=&F(\infty,\dots,\infty,x_{k_1},\infty,\dots,\infty,x_{k_2},\infty,\dots,\infty,x_{k_m},\infty,\dots,\infty).\end{aligned}$$

1.2.3 常用分布

常用的两种类型随机变量：

(1) 离散型随机变量 X 的概率分布用分布列描述：

$$p_k=P(X=x_k),\quad k=1,2,\dots,$$

定义

$$f(x_k)=p_k,k=1,2,\dots,$$

称 $f(\cdot)$ 为 X 的概率质量函数（PMF）。其分布函数为

$$F(x)=\sum_{x_k\le x}p_k.$$

(2) 连续型随机变量 X 的概率分布用概率密度 $f(x)$ 描述，其分布函数

$$F(x)=\int_{-\infty}^{x}f(t)\,dt.$$

1.2.3.1 退化分布

若随机变量 X 只取常数 c，即

$$P\{X=c\}=1,$$

则 X 并不随机，但我们把它看作随机变量的退化情况更为方便，因此称之为退化分布，又称单点分布.

1.2.3.2 Bernoulli 分布

在一次试验中，设事件 A 出现的概率为 p, $0\le p\le 1$，不出现的概率为 $1-p$，称 A 出现为成功，不出现为失败，p 为成功概率，若以 X 记事件 A 出现（成功）的次数，即 $X=I_A$，则 X 的可能取值仅为 $0,1$，其对应的概率为

$$P\{X=k\}=p^k(1-p)^{1-k},\ k=0,1.$$

1.2.3.3 二项分布

n 重 Bernoulli 试验表示独立地重复进行 n 次 Bernoulli 试验，设事件 A 在每次试验中出现（成功）的概率均为 $p, 0 \leq p \leq 1$，以 X 记事件 A 出现（成功）的次数，X 的可能取值为 $0, 1, 2, \dots, n$，其对应的概率为

$$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k},\ k = 0, 1, \dots, n,$$

则称之为以 n 和 p 为参数的二项分布，简记为 $X \sim \mathrm{B}(n, p)$.

其中 $\binom{n}{k}$ 是从 n 个不同号码中选取 k 个的不同取法，称为 n 取 k 的**组合数**，计算公式为

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

1.2.3.4 泊松分布

若随机变量 X 可取一切非负整数值，且

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda},\ k = 0, 1, \dots,$$

其中 $\lambda > 0$，则称 X 服从参数为 λ 的泊松分布（Poisson 分布），记为 $X \sim \mathrm{Pois}(\lambda)$.

1.2.3.5 几何分布

在独立重复的 Bernoulli 试验中，设事件 A 在每次试验中出现的概率均为 p, $0 < p < 1$，以 X 记事件 A 首次出现（成功）时的试验次数，X 的可能取值为 $0, 1, 2, \dots,$，其对应的概率分布为

$$P\{X = k\} = p(1-p)^{k-1},\ k = 1, 2, \dots,$$

则称 X 服从几何分布.

1.2.3.6 Pascal 分布

在独立重复的 Bernoulli 试验中，设事件 A 在每次试验中出现的概率均为 p, $0 < p < 1$，以 X 记事件 A 第 r 次出现（成功）时的试验次数，X 的可能取值为 $r, r+1, \dots$，其概率分布为

$$P\{X = k\} = \binom{k-1}{r-1} p^r (1-p)^{k-r},\ k = r, r+1, \dots,$$

称 X 服从 Pascal 分布.

1.2.3.7 负二项分布

在独立重复的 Bernoulli 试验中，设事件 A 在每次试验中出现的概率均为 p, $0<p<1$，以 X 记事件 A 第 r 次出现时已经失败的试验次数，则 X 的可能取值为 $0,1,...$, 其概率分布为

$$P\{X=k\}=\binom{k+r-1}{k}p^r(1-p)^k,\ k=0,1,\dots,$$

称 X 服从负二项分布。这时 $X+r$ 服从 Pascal 分布。

负二项分布通常用于替换泊松分布. 同泊松分布一样，它也在非负整数上取值，但因为它包含两个参数，相比泊松分布其变化更灵活. 泊松分布的方差和均值相等，但负二项分布的方差大于均值，这说明当某类数据集观测到的方差大于均值时，负二项分布要比泊松分布更合适.

1.2.3.8 离散均匀分布

如果随机变量 X 的分布列为

$$p_k=P(X=x_k)=\frac{1}{n},\ k=1,2,\dots,n,$$

称 X 服从 $\{x_1,\dots,x_n\}$ 上的离散均匀分布.

1.2.3.9 均匀分布

如果随机变量 X 有如下密度函数：

$$f(x)=\begin{cases}\frac{1}{b-a}, & \text{若 } a\le x\le b\\ 0, & \text{其他,}\end{cases}$$

其中 $a<b$，则称之为区间 $[a,b]$ 上的均匀分布，记作 $X\sim \mathrm{U}(a,b)$。

注意对连续型分布，单个或者有限个点的密度函数值对于分布函数值没有影响，所以区间 $[a,b]$ 是否包含左右端点不重要。

1.2.3.10 正态分布

如果 X 有如下密度函数：

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\},\ x\in\mathbb{R},$$

则称 X 服从为参数为 μ 和 σ^2 的正态分布，也称为高斯分布，记为 $X \sim \mathrm{N}(\mu, \sigma^2)$.

$\mathrm{N}(0,1)$ 称为标准正态分布，分布密度为

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},\ x \in \mathbb{R},$$

分布函数为 $\Phi(x) = \int_{-\infty}^{x} \phi(u)\, du$。

1.2.3.11 多元正态分布

设 $\mu = (\mu_1, \cdots, \mu_n)^T$，$\Sigma$ 是 n 阶正定对称矩阵，并且其行列式为 $|\Sigma|$. 如果随机向量 X 有如下联合密度函数：

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right\},$$

则称 X 服从为 n 维 (n 元) 正态分布，记为 $X \sim \mathrm{N}(\mu, \Sigma)$ 或 $X \sim \mathrm{N}_n(\mu, \Sigma)$.

多元正态分布推广的定义和性质见节7.1.2。

1.2.3.12 Gamma 分布

如果 X 有如下密度函数：

$$f(x) = \begin{cases} \frac{\lambda^s}{\Gamma(s)} x^{s-1} e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

则称 X 服从以 $s > 0$，$\lambda > 0$ 为参数的 Gamma 分布（伽马分布），其中 Γ 函数定义为

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\, dx,\ s > 0.$$

1.2.3.13 指数分布

如果在 Gamma 分布中令 $s = 1$，即密度函数为

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

则称 X 服从参数为 λ 的指数为指数分布, 记为 $X \sim \mathrm{Exp}(\lambda)$。

X 服从指数分布，当且仅当

$$P(X > x) = e^{-\lambda x},\ \forall x > 0.$$

1.2.3.14 卡方分布

如果在 Gamma 分布中取 $s = \frac{n}{2}$, n 是正整数, $\lambda = \frac{1}{2}$，即密度函数为

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}},\ x > 0,$$

则称 X 服从自由度为 n 的卡方分布，记为 $X \sim \chi^2(n)$.

1.3 数字特征、矩母函数与特征函数

1.3.1 Riemann-Stieltjes 积分

设 $g(x)$, $F(x)$ 为有限区间 $(a,b]$ 上的实值函数，$a = x_0 < x_1 < \cdots < x_n = b$ 为 $(a,b]$ 的一个分割，令

$$\Delta F(x_i) = F(x_i) - F(x_{i-1}),\ \xi_i \in [x_{i-1}, x_i], 1 \le i \le n,$$

$\lambda = \max_{1\le i\le n}(x_i - x_{i-1})$, 如果当 $\lambda \to 0$ 时，极限

$$\lim_{\lambda\to 0} \sum_{i=1}^{n} g(\xi_i)\Delta F(x_i)$$

存在，且与分割的选择以及 $\xi_i \in [x_{i-1}, x_i]$ 的取法无关，则称该极限值为函数 $g(x)$ 关于 $F(x)$ 在 $(a,b]$ 上的 Riemann-Stieltjes 积分，记为

$$\int_a^b g(x)\,dF(x) = \lim_{\lambda\to 0} \sum_{i=1}^{n} g(\xi_i)\Delta F(x_i).$$

当 $F(x) = x$ 时，Riemann-Stieltjes 积分即为 Riemann 积分. Riemann-Stieltjes 积分可以看成是求曲线 $g(x)$ 下方的面积问题的推广，用 $\Delta F(x_i) = F(x_i) - F(x_{i-1})$ 作为小区间 $(x_{k-1}, x_i]$ 的广义长度，以 $g(\xi_i)$ 近似曲边梯形的高度，用矩形面积近似曲边梯形面积。

关于 Riemann-Stieltjes 积分存在的条件，这里不做更进一步的讨论，只给出一个简单的充分条件：若函数 $g(x)$ 连续，$F(x)$ 单调，则 Riemann-Stieltjes 积分存在. 本书中用到的 $g(x)$ 为连续函数，$F(x)$ 为分布函数，因此积分的存在性不成问题.

定义 1.9 (有界变差函数)**.** 设 $f(x)$ 是定义在 $[s,t]$ 的实值函数，作分割 $\Delta : s = t_0 < t_1 < \cdots < t_n = t$，记

$$\nu_\Delta = \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|,$$

令

$$\bigvee_s^t (f) = \sup\{\nu_\Delta : \Delta\text{为}[s,t]\text{的任意分割}\},$$

并称其为 f 在 $[s,t]$ 的全变差；如果 $\bigvee_s^t(f) < \infty$，则称 $f(x)$ 是 $[s,t]$ 上的有界变差函数，记作 $f \in \mathrm{BV}([s,t])$。

$\mathrm{BV}([s,t])$ 构成一个线性空间。

有界变差函数可以写成两个增函数的差，几乎处处可微。

闭区间上连续可微函数必为有界变差。

定理 1.4. *若对任意的 $[a,b]$ 上连续函数 $g(x)$，关于 $F(x)$ 都是在 $[a,b]$ Riemann-Stieltjes 可积的，则 $F(x)$ 是 $[a,b]$ 上的有界变差函数。反之，连续函数关于有界变差函数 Riemann-Stieltjes 可积的。*

为了后面的需要，将积分推广到无限区间上：

$$\int_a^\infty g(x)\,dF(x) = \lim_{b\to+\infty} \int_a^b g(x)\,dF(x),$$
$$\int_{-\infty}^b g(x)\,dF(x) = \lim_{a\to-\infty} \int_a^b g(x)\,dF(x),$$
$$\int_{-\infty}^\infty g(x)\,dF(x) = \lim_{a\to-\infty, b\to+\infty} \int_a^b g(x)\,dF(x).$$

与 Riemannan 积分不同的是

$$\int_{a-}^a dF(x) = \lim_{\delta\to 0+} \int_{a-\delta}^a dF(x) = F(a) - F(a-),$$

当 $F(x)$ 在 $x = a$ 处有跳跃时，上式的值等于 $F(x)$ 在 a 点的跳跃高度. 当 $F(x)$ 是一个阶梯函数时，Riemann-Stieltjes 积分成为一个级数，即设 $F(x)$ 在 $x = x_i$ 处有跳跃高度 p_i, $i = 1, 2, \ldots$，则

$$\int_{-\infty}^\infty g(x)\,dF(x) = \sum_{i=1}^\infty g(x_i) p_i.$$

如果 $F(x)$ 是分布函数，对于离散型和连续型这两种常用情形，有

$$\int_a^b g(x)\,dF(x) = \begin{cases} \sum_{i=1}^{\infty} g(x_i)f(x_i), & \text{离散情形}, \\ \int_a^b g(x)f(x)\,dx, & \text{连续情形}. \end{cases}$$

其中 $f(\cdot)$ 为 $F(x)$ 的概率质量函数或者概率密度函数。

Riemann-Stieltjes 积分的一些基本性质：

(1) 线性性质:

$$\int_a^b [\alpha g_1(x) + \beta g_2(x)]\,dF(x) = \alpha \int_a^b g_1(x)\,dF(x) + \beta \int_a^b g_2(x)\,dF(x).$$

(2) 区间可加性：

$$\int_a^b g(x)\,dF(x) = \int_a^c g(x)\,dF(x) + \int_c^b g(x)\,dF(x),\ c \in (a,b).$$

(3) 广义长度：

$$\int_a^b dF(x) = F(b) - F(a)$$

其中 a, b 均可为有限数或无穷大.

(4) 测度可加性：

$$\int_a^b g(x)d[\alpha F_1(x) + \beta F_2(x)] = \alpha \int_a^b g(x)dF_1(x) + \beta \int_a^b g(x)dF_2(x).$$

(5) 若 $g(x) \geq 0$，$F(x)$ 单调不减, $b > a$, 则 $\int_a^b g(x)\,dF(x) \geq 0$.

1.3.2 数字特征

定义 1.10. 设 X, Y 为随机变量。

(1) 设 X 的分布函数为 $F(x)$，若

$$E(X) = \int_{-\infty}^{\infty} x\,dF(x)$$

存在，称 $E(X)$ 为随机变量 X 的数学期望。$E|X|$ 总存在，等于有限值或 $+\infty$；当 $E|X| < \infty$ 时 $E(X)$ 必存在且为有限值，这时称 X 一阶矩有限。当

$E|X| = +\infty$ 但 $\int_0^\infty x\,dF(x) < \infty$ 或 $\int_{-\infty}^0 |x|\,dF(x) < \infty$ 之一成立时，$E(X)$ 也有定义。如果 $\int_0^\infty x\,dF(x) = \int_{-\infty}^0 |x|\,dF(x) = +\infty$，称 $E(X)$ 不存在。

当 X 为离散型随机变量时，设概率分布列为 $P(X = x_k) = p_k$, $k = 1, 2, \ldots$，则

$$E(X) = \sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} x_k P(X = x_k).$$

（若上述级数收敛）。

当 X 为密度 $f(x)$ 的连续性随机变量时，

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx.$$

（若上述积分存在）。

(2) 对正整数 k，称 $m_k = E(X^k)$ 为 X 的 k 阶**原点矩**. 数学期望是一阶原点矩.

(3) 设 X 为随机变量，若 $E\{[X - E(X)]^2\}$ 存在，则称 $E\{[X - E(X)]^2\}$ 为 X 的方差，记为 $\mathrm{Var}(X)$，即

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} [x - E(X)]^2\,dF(x).$$

(4) 对正整数 k，称 $c_k = E\{[X - E(X)]^k\}$ 为 X 的 k 阶**中心矩**. 方差是二阶中心矩。

(5) 若 $E\{[X - E(X)][Y - E(Y)]\}$ 存在，则称之为 X 与 Y 的协方差，记为 $\mathrm{Cov}(X, Y)$，可知

$$\mathrm{Cov(X,Y)} = E\{[X - E(X)][Y - E(Y)]\} = E(XY) - E(X)E(Y).$$

(6) 对正整数 k, l，称 $E\{[X - E(X)]^k[Y - E(Y)]^l\}$ 为 X, Y 的 $k + l$ 阶**混合中心矩**. 协方差是二阶混合中心矩.

数学期望的性质：设 X, Y 为随机变量，期望存在有限，a, b 为实数。

(1) $E(X)$存在有限 $\iff E|X| < \infty$;

(2) 线性：

$E[aX + bY] = aE(X) + bE(Y)$.

(3) 若$X \geq 0$, a.s.,则

$E(X) \geq 0$.

(4) 若$X \leq Y$,则

$E(X) \leq E(Y)$.

方差的性质：

(1) $\text{Var}(X) = E(X^2) - [E(X)]^2$.

(2) $\text{Var}(a + bX) = b^2\text{Var}(X)$.

(3) $\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i) + 2\sum_{i<j} \text{Cov}(X_i, X_j)$.

定理 1.5. 设 X 为随机变量，分布函数为 $F(x)$，函数 g 为 X 的值域到 $\mathbb{R}$ 的（可测）变换，满足 $\int_{-\infty}^{\infty} |g(x)|\, dF(x) < \infty$，则 $E[g(X)]$ 存在有限且

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, dF(x).$$

对离散分布的 X 有

$$E[g(X)] = \sum_{k=1}^{\infty} g(x_k)P(X = x_k).$$

对连续分布的 X 有

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\, dx.$$

证明略。参见 (Shreve 2004) 定理 1.5.1。

例 1.6. 设 X 服从标准均匀分布 $\text{U}[0, 1]$，求 $E(X^4)$。

解:

$$E(X^4) = \int_0^1 x^4\, dx = \frac{1}{5}.$$

对定义在开区间 I 上的实值函数 $\phi(x)$，如果 $\forall x, y \in I, 0 < \alpha < 1$，都有

$$\phi(\alpha x + (1-\alpha)y) \leq \alpha\phi(x) + (1-\alpha)\phi(y),$$

则称 $\phi(\cdot)$ 为**凸函数**。若 $\phi(x)$ 二阶可微且二阶导数非负，则 $\phi(x)$ 为凸函数。例如：

- $\phi(x) = x^2, -\infty < x < \infty$;
- $\phi(x) = |x|, -\infty < x < \infty$;
- $\phi(x) = x^+ = \max(x, 0), -\infty < x < \infty$;
- $\phi(x) = \frac{1}{x}, x > 0$。

定理 1.6 (Jensen 不等式). *设 $E|X| < \infty$，$\phi(x)$ 是凸函数，则*

$$E[\phi(X)] \geq \phi(E(X)).$$

证明略。典型例子：

$$E[X^2] \geq [E(X)]^2.$$

例 1.7. 设随机变量 $X \geq 0$ 且 $EX < \infty$，证明

$$E[\ln(X)] \leq \ln(E(X)).$$

证明：令 $\phi(x) = -\ln(x), x > 0$，则 $\phi(x)$ 为凸函数。于是

$$E[\phi(X)] \geq \phi(E(X)),$$

得证。

例 1.8. 设随机变量 X 满足 $E|X| < \infty$, 则

$$E[e^X] \geq e^{E(X)}.$$

证明：$\phi(x) = e^x$ 是凸函数。

例 1.9 (匹配问题)**.** 在一次聚会中，n 个人将自己的帽子放到房间中央，混合后每人随机取一个。设随机变量 X 表示取到自己帽子的人数，求 X 的期望和方差.

解：令 X_i 表示第 i 个人取到自己帽子的示性函数，则

$$X = \sum_{i=1}^{n} X_i.$$

由抽签问题的公平性可知每个人抽取到自己帽子的概率相等，显然第一个人抽到自己帽子的概率是 $\frac{1}{n}$。于是

$$\begin{aligned} E(X) =& \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} P(X_i = 1) \\ =& nP(X_1 = 1) = n \times \frac{1}{n} = 1. \end{aligned}$$

易见

$$\begin{aligned} \mathrm{Var}(X_i) =& \mathrm{Var}(X_1) = E(X_1^2) - [E(X_1)]^2 \\ =& E(X_1) - [E(X_1)]^2 = \frac{n-1}{n^2}. \end{aligned}$$

考虑 $E(X_iX_j)$ 的计算（$i < j$）。由抽签问题的公平性，这应该等于 $E(X_1X_2)$。$X_1X_2 = 1$ 当前仅当前两个人抽到自己的帽子，概率为

$$\frac{1}{n} \times \frac{1}{n-1}.$$

于是

$$\mathrm{Cov}(X_i, X_j) = E(X_iX_j) - E(X_i)E(X_j) = \frac{1}{n^2(n-1)}.$$

从而

$$\begin{aligned} \mathrm{Var}(X) =& \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j) \\ =& n \times \frac{n-1}{n^2} + 2 \times \frac{1}{2} n(n-1) \times \frac{1}{n^2(n-1)} = 1. \end{aligned}$$

关于 X 的概率分布见1.6.1。

1.3.3 随机向量的期望和方差阵

设 $X=(X_1,\dots,X_n)^T$ 是随机向量（表示为列向量），定义

$$E(X) = (EX_1,\dots,EX_n)^T.$$

设 M 为 $n\times p$ 矩阵，其中每个元素为随机变量，称 M 为随机矩阵，定义 $E(M)$ 为每个元素的期望所组成的矩阵。

对随机向量 X 定义其协方差阵（方差阵）为

$$\mathrm{Var}(X) = E\left[(X-E(X))(X-E(X))^T\right].$$

设 $\mathrm{Var}(X)=(\sigma_{ij})_{n\times n}$，则

$$\sigma_{ii}=\mathrm{Var}(X_i),\quad \sigma_{ij}=\mathrm{Cov}(X_i,X_j),\ (i\neq j).$$

显然 $\mathrm{Var}(X)$ 是对称阵，且易证明其为非负定阵。

设 X,Y 分别是 n 维和 m 维随机向量，定义其协方差阵为

$$\mathrm{Cov}(X,Y) = E\left[(X-E(X))(Y-E(Y))^T\right],$$

这是一个 $n\times m$ 矩阵，其 (i,j) 元素为 $\mathrm{Cov}(X_i,Y_j)$。易见

$$\mathrm{Cov}(Y,X) = [\mathrm{Cov}(X,Y)]^T.$$

命题 1.1. 设 M,C 为随机矩阵，A,B 为非随机矩阵，则

$$E(C+AMB)=E(C)+AE(M)B.$$

命题 1.2. 设 X 为随机向量，α 为非随机向量，B 为非随机矩阵，令 $Y=\alpha+BX$，则

$$E(Y) = \alpha + BE(X),$$
$$Var(Y) = B\,Var(X)B^T.$$

推论 1.1. 设 X 为 n 维随机向量，$\Sigma = Var(X)$，则 Σ 为非负定矩阵。

证明：只要对任意 $\alpha\in\mathbb{R}^n$ 都有

$$\alpha^T\Sigma\alpha\geq 0.$$

令 $Y=\alpha^T X$，则

$$\mathrm{Var}(Y)=\mathrm{Var}(\alpha^T X)=\alpha^T\Sigma\alpha\geq 0,$$

得证。

命题 1.3. 设 X,Y,Z 为随机向量，α 和 β 为非随机向量，A,B 为非随机矩阵，

$$\begin{aligned}Cov(\alpha+AX,\beta+BY)=&ACov(X,Y)B^T,\\ Cov(X+Z,Y)=&Cov(X,Y)+Cov(Z,Y).\end{aligned}$$

定理 1.7. 设 X 为随机向量，分布函数为 $F(x)$，函数 g 为 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的（可测）变换，则

$$E[g(X)]=\int_{\mathbb{R}^n} g(x)\,dF(x),$$

只要右侧的积分存在。

证明略。

1.3.4 矩母函数

定义 1.11. 设随机变量 X 的分布函数为 $F_X(x)$, 令

$$\phi_X(t)=E[e^{tX}]=\int_{-\infty}^{\infty}e^{tx}\,dF_X(x),$$

若 $\phi_X(t)$ 在某个长度大于 0 的包含原点的区间上取有限值，则称 $\phi_X(t)$ 为 X 的**矩母函数**.

如果 X 是仅取非负值的随机变量，则 $\phi_X(t)$ 在 $(-\infty,0]$ 上有限。

假设对 $\phi(t)$ 求导时，求导运算与求期望运算可以交换次序，有

$$\begin{aligned}\phi'(t)=&E(Xe^{tX})\\ \phi''(t)=&E(X^2e^{tX})\\ &\vdots\\ \phi^{(n)}(t)=&E(X^ne^{tX})\end{aligned}$$

令 $t=0$, 得到

$$\phi^{(n)}(0)=E[X^n],\ n=1,2,...$$

当矩母函数存在时，它唯一地决定分布，因此我们能够用矩母函数刻画随机变量的概率分布. 矩母函数还经常用来讨论独立随机变量和的分布。

例 1.10. 求标准正态分布的矩母函数并求其前 4 阶矩。

解：设 X 服从标准正态分布，则 X 有矩母函数

$$\begin{aligned}
\phi(t) =& E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
=& \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2-2tx)}\, dx \\
=& \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} e^{\frac{1}{2}t^2}\, dx \\
=& e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2}\, dx \\
=& e^{\frac{1}{2}t^2},\ t \in (-\infty, \infty).
\end{aligned}$$

求导得

$$\begin{aligned}
\phi'(t) =& t e^{\frac{t^2}{2}}, \\
E(X) =& \phi'(0) = 0; \\
\phi''(t) =& (1+t^2) e^{\frac{t^2}{2}}, \\
E(X^2) =& \phi''(0) = 1; \\
\phi^{(3)}(t) =& (2t + t(1+t^2)) e^{\frac{t^2}{2}} = (3t + t^3) e^{\frac{t^2}{2}}, \\
E(X^3) =& \phi^{(3)}(0) = 0; \\
\phi^{(3)}(t) =& (3 + 3t^2 + t(3t + t^3)) e^{\frac{t^2}{2}} = (3 + 6t^2 + t^4) e^{\frac{t^2}{2}}, \\
E(X^4) =& \phi^{(4)}(0) = 3.
\end{aligned}$$

一般地，对奇数 k 有 $E(X^k) = 0$，对偶数 k 有 $E(X^k) = (k-1)!! = (k-1)(k-3)\cdots 3 \cdot 1$，见节1.6.2。

例 1.11. 若随机变量 X 有如下密度函数：

$$f(x) = \frac{1}{\pi(1+x^2)},\ -\infty < x < \infty,$$

称 X 服从柯西分布。柯西分布的矩母函数不存在。

证明：对任意 $t \neq 0$，有

$$\begin{aligned} E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\pi(1+x^2)}\, dx \\ &\geq \int_{0}^{\infty} e^{|t|x} \frac{1}{\pi(1+x^2)}\, dx \\ &\geq \int_{0}^{\infty} \frac{1+|t|x}{\pi(1+x^2)}\, dx = +\infty, \end{aligned}$$

所以矩母函数不存在。

随机变量的矩母函数不一定存在，在这种情况下，更方便的是特征函数.

1.3.5 特征函数

定义 1.12. 若随机变量 X 的分布函数为 $F_X(x)$, 则称

$$\psi_X(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx}\, dF_X(x)$$

为 X 的**特征函数**. 特征函数总存在且定义于 $(-\infty, \infty)$。

如果 F_X 有密度 $f(x)$，则 $\psi_X(t)$ 就是 $f(x)$ 的 Fourier 变换：

$$\psi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx.$$

特征函数是一个实变量的复值函数，因为 $|e^{itx}| = 1$，所以它对一切实数 t 都有定义. 可以看成是实部与虚部分别积分的结果：

$$\psi_X(t) = \int_{-\infty}^{\infty} \cos(tx)\, dF_X(x) + i \int_{-\infty}^{\infty} \sin(tx)\, dF_X(x).$$

特征函数有如下性质：

(1) 有界性：$|\psi(t)| \leq 1 = \psi(0)$;

(2) 共轭对称性：$\psi(-t) = \overline{\psi(t)}$;

(3) 一致连续性：

$$|\psi(t+h) - \psi(t)| \leq \int_{-\infty}^{\infty} |e^{ihx} - 1|\, dF(x);$$

(4) 线性变换的作用：设 $Y = aX + b$, 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的特征函数之积;

(6) 非负定性：对于任意的正整数 n，任意实数 $t_1, \ldots, t_n$ 及复数 $\lambda_1, \ldots, \lambda_n$，有

$$\sum_{k=1}^{n}\sum_{j=1}^{n}\psi(t_k - t_j)\lambda_k\overline{\lambda_j} \geq 0.$$

(7) 设随机变量 X 有 n 阶矩存在，则它的特征函数 n 阶导数存在，且当 $k \leq n$ 时，有

$$\psi^{(k)}(0) = i^k E[X^k],$$

其中 i 表示虚数单位。

特别地，特征函数可作如下带皮阿诺型余项的 Taylor 展开：

$$\psi(t) = 1 + itE[X] + \frac{(it)^2}{2!}E[X^2] + \cdots + \frac{(it)^n}{n!}E[X^n] + o(t^n).$$

例 1.12. 求标准正态分布 $\mathrm{N}(0,1)$ 的特征函数.

解： 由定义

$$\psi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{itx}e^{-x^2/2}dx,$$

从而

$$\begin{aligned}
\psi'(t) =& \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ixe^{itx}e^{-x^2/2}\,dx \\
=& \frac{i}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{itx}d(-e^{-x^2/2}) \\
=& -\frac{i}{\sqrt{2\pi}}\left(e^{itx-x^2/2}|_{-\infty}^{\infty} + t\int_{-\infty}^{\infty} e^{itx}e^{-x^2/2}\,dx\right) \\
=& -t\psi(t).
\end{aligned}$$

于是

$$\frac{d}{dt}\log\psi(t) = -t,$$
$$\psi(t) = ce^{-\frac{1}{2}t^2},$$

用 $\psi(0) = 1$ 代入得

$$\psi(t) = e^{-\frac{1}{2}t^2}.$$

例 1.13. 求正态分布 $\mathrm{N}(\mu,\sigma^2)$ 的特征函数.

证明：记标准正态分布的特征函数为 $\psi_Z(t)$，设 $X\sim\mathrm{N}(\mu,\sigma^2)$，则 X 的特征函数为

$$\psi(t)=\int_{-\infty}^{\infty}e^{itx}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\,dx.$$

作变量替换 $x=\mu+\sigma y$，则

$$\begin{aligned}\psi(t)=&\int_{-\infty}^{\infty}e^{it(\mu+\sigma y)}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}y^2}\sigma\,dy\\=&e^{it\mu}\int_{-\infty}^{\infty}e^{i(t\sigma)y}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}\sigma\,dy\\=&e^{it\mu}\psi_Z(t\sigma)=e^{it\mu}e^{-\frac{1}{2}\sigma^2t^2}\\=&\exp\{it\mu-\frac{1}{2}\sigma^2t^2\}.\end{aligned}$$

用例1.12和特征函数性质 (4) 也可以得到上述结果。

例 1.14. 设随机变量 X 服从参数为 λ 的泊松分布，求 X 的特征函数，并由特征函数求 X 的期望和方差.

解：

$$\begin{aligned}\psi(t)=&E[e^{itX}]=\sum_{k=0}^{\infty}e^{itk}\frac{\lambda^k}{k!}e^{-\lambda}\\=&e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^{it})^k}{k!}=e^{-\lambda}e^{\lambda e^{it}}\\=&e^{\lambda(e^{it}-1)}.\\\psi'(t)=&\lambda ie^{it}e^{\lambda(e^{it}-1)},\quad\psi'(0)=\lambda i,\\E(X)=&\frac{1}{i}\psi'(t)=\lambda.\\\psi''(t)=&-\lambda e^{it}e^{\lambda(e^{it}-1)}+\lambda ie^{it}\cdot\lambda ie^{it}e^{\lambda(e^{it}-1)}\\=&-\lambda e^{it}e^{\lambda(e^{it}-1)}-\lambda^2e^{2it}e^{\lambda(e^{it}-1)},\\\psi''(0)=&-\lambda-\lambda^2,\\E(X^2)=&\frac{1}{i^2}\psi''(0)=\lambda+\lambda^2,\\\mathrm{Var}(X)=&E(X^2)-[E(X)]^2=\lambda.\end{aligned}$$

由于特征函数只与分布函数有关，所以称为分布的特征函数. 另一方面, 有下述定理.

定理 1.8 (唯一性定理)**.** *分布函数由其特征函数唯一决定.*

证明略。说明特征函数与分布函数是相互唯一确定的.

若随机向量 $X=(X_1,\dots,X_n)^T$ 的分布函数为 $F(x_1,\dots,x_n)$，与随机变量相仿，类似地定义它的特征函数

$$\begin{aligned}\psi(t_1,\dots,t_n)=&E(e^{it^TX})\\ =&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{i(t_1x_1+\cdots+t_nx_n)}\,dF(x_1,\dots,x_n).\end{aligned}$$

可以类似于一元的场合，建立起 n 元特征函数的理论. 例如，我们在节1.2中定义的多元正态分布是狭义的，这个定义在推导很多性质时并不方便，同时还不能考虑 Σ 不是正定的情形，为此，我们采用下面的定义. 如果随机向量 X 的特征函数为

$$\psi_X(t)=\exp\left\{it^T\mu-\frac{1}{2}t^T\Sigma t\right\},$$

则称 X 服从多元正态分布 $\mathrm{N}(\mu,\Sigma)$. Σ 是非负定 $n\times n$ 矩阵. 当 $\mu=0$，$\Sigma=I_n$ 时，称为标准多元正态分布，记为 $X\sim\mathrm{N}(0,I_n)$，这就是由独立同分布的 n 个标准正态分布随机变量组成的随机向量。

利用特征函数方法不难证明如下的关于正态分布线性组合的性质：

命题 1.4. *若 $X\sim N(0,I_n)$，则 X 的任一线性函数 $Y=A_{m\times n}X+\mu$ 服从 m 维正态分布 $N(\mu,AA^T)$.*

命题 1.5. *若 $Y\sim N(\mu,\Sigma)$，则*

$$AY+b\sim N(A\mu+b,A\Sigma A^T).$$

多元正态分布的更多性质参见节7.1.2。

1.4 收敛性

定义 1.13. (1) 设 $\{X_n,n\geq 1\}$ 是随机变量序列，若存在随机变量 X 使得

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

则称随机变量序列 $\{X_n, n \geq 1\}$ **几乎必然收敛** (或 a.s. 收敛，或以概率 1 收敛) 到 X，记为 $X_n \to X$, a.s. 或 $X_n \xrightarrow{\text{a.s.}} X$.

(2) 设 $\{X_n, n \geq 1\}$ 是随机变量序列，若存在随机变量 X 使得 $\forall \varepsilon > 0$, 有

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

则称随机变量序列 $\{X_n, n \geq 1\}$ **依概率收敛**到 X，记为 $X_n \xrightarrow{P} X$.

(3) 设随机变量序列 $\{X_n\} \subset L^p$, $p \geq 1$, $X \in L^p$，若有

$$\lim_{n\to\infty} E[|X_n - X|^p] = 0,$$

则称随机变量序列 $\{X_n, n \geq 1\}$ 为 p 次**平均收敛**到 X, 或称 $\{X_n\}$ 在 L^p 中**强收敛**到 X. 当 $p = 2$ 时，称为**均方收敛**.

(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**（或依分布收敛）到 $F(x)$，记为 $F_n(x) \xrightarrow{W} F(x)$ 或 $F_n(x) \xrightarrow{d} F(x)$. 设随机变量 X_n, X 的分布函数分别为 $F_n(x)$ 及 $F(x)$, $F_n(x) \xrightarrow{W} F(x)$，则称 $\{X_n\}$ **依分布收敛**到 X，记为 $X_n \xrightarrow{d} X$。

注：L^p 是所有满足 $E|X|^p < \infty$ 的随机变量 X 的集合。

F_n 依分布收敛到 F，当且仅当 F_n 的特征函数 $\psi_n(t)$ 点点收敛到 F 的特征函数 $\psi(t)$。

定理 1.9. *(1) 随机变量序列 $X_n \to X$, a.s. 的充分必要条件是 $\forall \varepsilon > 0$,*

$$\lim_{n\to\infty} P\{\sup_{m\geq n} |X_m - X| \geq \varepsilon\} = 0.$$

(2)} 随机变量序列 $X_n \xrightarrow{P} X$ 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到 X 的子序列.

随机变量序列的这 4 种收敛性之间的关系可以总结为下面的关系图 ($p \geq 1$)：

几乎必然收敛 $\Longrightarrow$ 依概率收敛 $\Longrightarrow$ 依分布收敛;

p次平均收敛 $\Longrightarrow$ 依概率收敛 $\Longrightarrow$ 依分布收敛.

注: 几乎必然收敛与 p 阶矩收敛之间没有蕴含关系.

例 1.15. 取 $\Omega = (0,1]$, $\mathscr{F}$ 为 $(0,1]$ 中全体 Borel 子集所构成 σ 代数, P 为 Lebesgue 测度 (长度的推广), 我们可以构造出两个个随机变量序列, 其中之一是 r 次平均收敛的, 但是不几乎必然收敛; 另外一个则几乎必然收敛, 但不是 r 次平均收敛的.

解: 令

$$Y_{11} = 1;$$

$$Y_{21} = \begin{cases} 1 & \omega \in (0, \frac{1}{2}], \\ 0 & \omega \in (\frac{1}{2}, 1]; \end{cases}$$

$$Y_{22} = \begin{cases} 0 & \omega \in (0, \frac{1}{2}], \\ 1 & \omega \in (\frac{1}{2}, 1]. \end{cases}$$

一般地，将 $(0,1]$ 分成 k 个等长区间，并且令

$$Y_{ki} = \begin{cases} 1 & \omega \in (\frac{i-1}{k}, \frac{i}{k}], \\ 0 & \omega \notin (\frac{i-1}{k}, \frac{i}{k}], \end{cases} \quad i = 1, \dots, k, \ k = 1, 2, \dots$$

定义随机变量序列

$$X_1 = Y_{11}, \ X_2 = Y_{21}, \ X_3 = Y_{22}, \ X_4 = Y_{31}, \ X_5 = Y_{32}, \dots,$$

对任意 $\varepsilon > 0$, 由于

$$E|Y_{ki} - 0|^r = \frac{1}{k} \to 0, \ (k \to \infty)$$

可见 $\{X_n\}$ 为 r 次平均收敛，但是对任意固定的 $\omega \in \Omega$, 任一自然数 k，恰有一个 i，使得 $Y_{ki}(\omega) = 1$, 而对其余的 j 有 $Y_{kj}(\omega) = 0$. 由此知 $\{X_n(\omega)\}$ 中有无穷多个 1 及无穷多个 0，于是 $\{X_n(\omega)\}$ 对每个 $\omega \in \Omega$ 都不收敛.

如果取

$$Z_n = \begin{cases} n^{\frac{1}{r}} & \omega \in (0, \frac{1}{n}] \\ 0 & \omega \notin (0, \frac{1}{n}] \end{cases}$$

$$Z = 0, \ \forall \omega \in \Omega,$$

易见，$Z_n(\omega) \to Z(\omega)$, $\forall \omega \in \Omega$, 所以

$$Z_n \to Z, \ \text{a.s.},$$

但是 $E(|Z_n - Z|^r) = n \cdot \frac{1}{n} = 1 \not\to 0$.

下面我们给出积分号下取极限的三大基本定理.

定理 1.10 (单调收敛定理)**.** 设 X_n 期望存在（不一定有限），$n \geq 1$，则：

(1) 若 $X_n \uparrow X$, *a.s.*, 且 $E(X_1) > -\infty$，则 $E(X)$ 存在，且

$$E(X_n) \uparrow E(X).$$

(2) 若 $X_n \downarrow X$, *a.s.*, 且 $E(X_1) < \infty$，则 $E(X)$ 存在，且

$$E(X_n) \downarrow E(X).$$

.

定理 1.11 (Fatou 引理)**.** 设随机变量序列 $\{X_n\}$ 的期望存在, $n \geq 1$, 则

$$E[\liminf_{n\to\infty} X_n] \leq \liminf_{n\to\infty} E[X_n] \leq \limsup_{n\to\infty} E[X_n] \leq E[\limsup_{n\to\infty} X_n].$$

定理 1.12 (Lebesgue 控制收敛定理)**.** 设 X_n, X 一阶矩有限，$X_n \to X$, *a.s.* 或 $X_n \xrightarrow{P} X$. 若存在一非负随机变量 Y 满足 $E(Y) < \infty$，且使得 $\forall n \geq 1$ 有 $|X_n| \leq Y$, *a.s.*，则 $E(X)$ 有限，且有

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

推论 1.2. 设 X_n 为随机变量序列，$X_n \to X$, *a.s.*, 对某个 $1 \leq p < \infty$，有非负随机变量 Y 满足 $E(Y^p) < \infty$ 且 $|X_n| \leq Y$, *a.s.*，则有

$$E|X_n - X|^p \to 0, \; n \to \infty,$$

即 X_n 在 L^p 意义下收敛到 X。

证明：易见 X_n 和 X 也都存在 p 阶矩。令 $Z_n = |X_n - X|^p$，则 $Z_n \to 0$, a.s., 且

$$Z_n \leq (|X_n| + |X|)^p \leq 2^p Y^p,$$

由控制收敛定理即可知 $EZ_n \to 0$，证毕。

1.5 独立性与条件期望

1.5.1 独立性

定义 1.14. (1) 设 A, B 为两个事件，若

$$P(A\cap B)=P(A)P(B),$$

则称 A 与 B 独立.

(2) 设 $A_1, A_2, \ldots, A_n$ 为 n 个事件，如果对任何 $m\le n$ 及 $1\le i_1<i_2<\cdots<i_m\le n$，有

$$P\left(\bigcap_{j=1}^m A_{i_j}\right)=\prod_{j=1}^m P(A_{i_j}),$$

则称 $A_1, A_2, \ldots, A_n$ 相互独立.

(3) 设 $\{A_i, i\in I\}$ 是一族事件，若对 I 的任意有限子集 $\{i_1, \ldots, i_m\}\neq\emptyset$ 都有

$$P\left(\bigcap_{j=1}^m A_{i_j}\right)=\prod_{j=1}^m P(A_{i_j}),$$

则称 $\{A_i, i\in I\}$ 是相互独立的.

(4) 设 $\{\mathcal{A}_i, i\in I\}$ 是一族事件类，如果对 I 的任意有限子集 $\{i_1, \ldots, i_m\}\neq\emptyset$ 和任意 $A_{i_j}\in\mathcal{A}_{i_j}(j=1,2,\ldots,m)$ 都有

$$P\left(\bigcap_{j=1}^m A_{i_j}\right)=\prod_{j=1}^m P(A_{i_j}),$$

则称 $\{\mathcal{A}_i, i\in I\}$ 是**独立事件类**.

(5) 设 $\{X_i, i\in I\}$ 是 Ω 上一族随机变量，如果 σ 代数族 $\{\sigma(X_i), i\in I\}$ 是独立事件类，则称 $\{X_i, i\in I\}$ **相互独立**.

(6) 设 $\{X_i, i\in I\}$ 是随机变量族的集合，其中的每个 X_i 是一族随机变量，I 为足标集，如果 σ 代数族 $\{\sigma(X_i), i\in I\}$ 是独立事件类，则称 $\{X_i, i\in I\}$ **相互独立**.

注 1：当 $P(B)>0$ 时，A, B 相互独立当且仅当

$$P(A|B)=P(A),$$

即 B 是否发生不影响到 A 发生的概率。

注 2：$A_1, A_2, \dots, A_n$ 两两独立不一定相互独立.

容易证明随机变量 $X_1, \dots, X_n$ 相互独立的充分必要条件是它们的联合分布函数可以分解为

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n).$$

对离散分布的随机变量，相互独立当且仅当

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n),\ \forall x_1, \dots, x_n.$$

对连续型分布的随机向量 $(X_1, \dots, X_n)$，其分量相互独立当且仅当分布密度等于边缘密度的乘积：

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n),\ \forall x_1, \dots, x_n.$$

随机变量 $X_1, \dots, X_n$ 相互独立，当且仅当其联合特征函数等于边缘特征函数的乘积：

$$E[\exp\{i(t_1 X_1 + \cdots + t_n X_n)\}] = E(e^{it_1 X_1}) \cdots E(e^{it_n X_n}),\ \forall (t_1, \dots, t_n) \in \mathbb{R}^n.$$

若随机变量 $X_1, \dots, X_n$ 矩母函数存在，则其相互独立当且仅当联合矩母函数等于边缘矩母函数的乘积：

$$E[\exp\{t_1 X_1 + \cdots + t_n X_n\}] = E(e^{t_1 X_1}) \cdots E(e^{t_n X_n}),\ \forall (t_1, \dots, t_n) \in D,$$

其中 D 为 $\mathbb{R}^n$ 中使得矩母函数有定义的超长方体。

设 A, B 为两个事件，

$$\begin{aligned}
& A, B\text{独立} \\
\iff & A, B^c\text{独立} \\
\iff & A^c, B\text{独立} \\
\iff & A^c, B^c\text{独立} \\
\iff & I_A, I_B\text{独立}.
\end{aligned}$$

定理 1.13. 设 $\{\mathscr{F}_i, i \in I\}$ 是相互独立的 $\mathscr{F}$ 的子 σ 代数族，$\{I_j \subset I : j \in J\}$ 是 I 的互不相交的子集族，则 $\{\sigma(\mathscr{F}_k, k \in I_j) : j \in J\}$ 是相互独立的子 σ 代数族。

证明略，参见 (王寿仁 1997) P.47 系 2.4.7。

推论 1.3. 设随机变量族 $\{X_i : i \in I\}$ 相互独立，$\{I_j \subset I : j \in J\}$ 是 I 的互不相交的子集族，则 $\{(X_k, k \in I_j), j \in J\}$ 相互独立。

推论 1.4. 设 $\{X_n, n \in \mathbb{Z}\}$ 是独立随机变量序列，则其任意的不相交的子集仍为独立的随机变量族。

定理 1.14. 设 $\{X_i, i \in I\}$ 是随机变量族，$\{I_j \subset I : j \in J\}$ 是 I 的互不相交的子集族，$g_j(\cdot), j \in J$ 为可测函数，令 $Y_j = g_j(X_k, k \in I_j)$，则 $\{Y_j, j \in J\}$ 相互独立。

这个定理说明相互独立的随机变量进行不相交的分组后分别作变换，结果仍相互独立。

证明: Y_j 关于 $\sigma(X_k, k \in I_j)$ 可测，从而 $\sigma(Y_j) \subset \sigma(X_k, k \in I_j)$。令 $\mathscr{F}_i = \sigma(X_i)$，由定理1.13可知 $\{\sigma(\mathscr{F}_k, k \in I_j) : j \in J\}$ 是相互独立的子 σ 代数族，由事件类相互独立的定义可知 $\{\sigma(Y_j), j \in J\}$ 相互独立，即 $\{Y_j, j \in J\}$ 相互独立。

定理 1.15. *(1)* 设随机变量 $X_1, \dots, X_n$ 是相互独立且一阶矩有限，则

$$E\left[\prod_{k=1}^n X_k\right] = \prod_{k=1}^n E[X_k].$$

(2) 设随机变量 $X_1, \dots, X_n$ 相互独立且二阶矩有限，则

$$Var\left[\sum_{k=1}^n X_k\right] = \sum_{k=1}^n Var[X_k].$$

定理 1.16 (Borel-Cantelli 第一引理)**.** 设 $\{A_n, n \geq 1\}$ 是一列事件，若

$$\sum_{n=1}^{\infty} P(A_n) < \infty,$$

则

$$P(A_n,\ i.o.) = 0.$$

证明:

$$\begin{aligned}P(A_n,\ \text{i.o.}) =& P(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m)\\ \le& P(\bigcup_{m=n}^{\infty}A_m)\le\sum_{m=n}^{\infty}P(A_m)\to 0\ (n\to\infty).\end{aligned}$$

定理 1.17 (Borel-Cantelli 第二引理). *设 $\{A_n, n\ge 1\}$ 是独立的事件列，若*

$$\sum_{n=1}^{\infty}P(A_n)=+\infty,$$

则

$$P(A_n,\ i.o.)=1.$$

定义 1.15. 设 $\{X_n, n\ge 1\}$ 是随机变量序列，$\mathscr{D}_k=\sigma(X_k, X_{k+1},\dots)$ 是由 $X_k, X_{k+1},\cdots$ 生成的 σ 代数，则 $\{\mathscr{D}_k\}$ 是非增的列，它们的交 $\mathscr{D}=\bigcap_{n\ge1}\mathscr{D}_n$ 称为序列 $\{X_n, n\ge 1\}$ 的**尾 σ 代数**，$\mathscr{D}$ 中的集合称为 $\{X_n, n\ge 1\}$ 的**尾事件**.

定理 1.18 (Kolmogorov 0-1 律). *独立随机变量序列的尾事件的概率或为 0 或为 1.*

1.5.2 独立随机变量和的分布

设随机变量 X_1, X_2 相互独立，F_1, F_2 分别为它们的分布函数. 令 $X=X_1+X_2$，其分布函数记为 $F_X(x)$. 则由独立性, 有

$$\begin{aligned}F_X(x) =& P\{X_1+X_2\le x\} && (1.1)\\ =& \int_{-\infty}^{\infty}P\{X_1+X_2\le x|X_1=t\}dF_1(t)\quad(\text{全期望公式}) && (1.2)\\ =& \int_{-\infty}^{\infty}F_2(x-t)\,dF_1(t). && (1.3)\end{aligned}$$

(1.3)式称作分布函数 F_1, F_2 的卷积，记为 $F_1 * F_2(x)$. 一般地对有界函数 $g(x)$ 和一个单调函数 $F(x)$，都可以定义 F 与 g 的卷积:

$$F * g(x)\triangleq\int_{-\infty}^{\infty}g(x-t)\,dF(t).$$

这里需要注意的是 $F*g$ 的顺序，$g*F$ 可能没有意义. 但是当 F 和 g 都是分布函数时，卷积可以交换顺序，只需注意到卷积中的随机变量 X_1 和 X_2 的地位是对等的即可得到

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

当 F 有密度 f 时，卷积 $F*g$ 就是通常的两个函数的卷积：

$$F * g(x) = \int_{-\infty}^{\infty} g(x-t)\, dF(t) = \int_{-\infty}^{\infty} g(x-t) f(t)\, dt.$$

容易看出，对于分布函数，卷积还满足结合律和分配律. 即对分布函数 F, G, H，有

$$\begin{aligned}(F*G)*H(x) =& F*(G*H)(x),\\ F*(G+H)(x) =& F*G(x) + F*H(x).\end{aligned}$$

于是，更进一步还有，设 $X_k, k=1,2,\dots,n$ 是独立同分布 F 的随机变量，令

$$S_0 = 0,\ S_n = \sum_{k=1}^{n} X_k,\ n = 1, 2, \dots,$$

将 S_n 的分布记作 F_n，则有

$$\begin{aligned}F_0(x) =& \begin{cases} 0, & x<0,\\ 1, & x\ge 0,\end{cases}\\ F_n(x) =& F*F_{n-1}(x), n=1,2,\dots\end{aligned}$$

称 F_n 为 F 的 n 重卷积.

独立随机变量和的分布，经常借助于矩母函数和特征函数研究，这是利用了独立随机变量变量乘积的期望等于期望的乘积。

1.5.3 关于一个随机变量的条件期望

请参考 (刘勇 2022)。

定义 1.16. 设 B 是一个事件，且 $P(B)>0$，则事件 B 发生的条件下事件 A 发生的**条件概率**为

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

当 $P(B) > 0$ 时，A, B 独立当且仅当 $P(A|B) = P(A)$。

定义 1.17. 设 $P(C) > 0$，称事件 A, B 在事件 C 的条件下独立，若

$$P(AB|C) = P(A|C)P(B|C).$$

条件独立不一定独立，独立也不一定条件独立。

定理 1.19 (全概率公式)**.** 设 $\{B_n\}$ 是 Ω 的一个分割，且 $P(B_n) > 0, \forall n$. 对 $A \in \mathscr{F}$，有

$$P(A) = \sum_n P(B_n)P(A|B_n).$$

定理 1.20 (Bayes 公式)**.** 设 $\{B_n\}$ 是 Ω 的一个分割，且 $P(B_n) > 0, \forall n$，如果 $P(A) > 0$，则

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_n P(B_n)P(A|B_n)}, \ \forall k.$$

如果 X 与 Y 是离散型随机变量，设 X 的取值集合为 $\{x_k, k = 1, 2, ...\}$，则给定 $X = x_k$ 时，Y 的条件概率分布定义为：

$$P\{Y = y|X = x_k\} = \frac{P\{X = x_k, Y = y\}}{P\{X = x_k\}}.$$

Y 的条件分布函数定义为：

$$F(y|x_k) = P\{Y \le y|X = x_k\}.$$

Y 的条件期望定义为：

$$E[Y|X = x_k] = \int_{-\infty}^{\infty} y\, dF(y|x_k) = \sum_m y_m P\{Y = y_m|X = x_k\}.$$

如果 X 与 Y 有联合概率密度函数 $f(x, y)$，则对一切使得 $f_X(x) > 0$ 的 x，给定 $X = x$ 时，Y 的条件概率密度函数定义为:

$$f(y|x) = \frac{f(x, y)}{f_X(x)}.$$

Y 的条件分布函数定义为：

$$F(y|x) = P\{Y \le y|X = x\} = \int_{-\infty}^{y} f(u|x)\, du.$$

Y 的条件期望定义为：

$$E[Y|X=x] = \int_{-\infty}^{\infty} y\,dF(y|x) = \int_{-\infty}^{\infty} yf(y|x)\,dy.$$

当 $E[Y|X=x]$ 有定义时，记 $g(x)=E[Y|X=x]$，$g(X)$ 是 X 的函数，记作 $E[Y|X]$。$E[Y|X]$ 是利用自变量 X 的信息对因变量 Y 所作的均方误差最小的预测。

对离散分布的 X，设 X 的取值集合为 $\{x_k, k=1,2,...\}$，这时 $g(x_k) = E[Y|X=x_k]$, $E[Y|X]=g(X)$ 当 $X=x_k$ 时为 $E[Y|X=x_k]$，所以可以写成

$$E[Y|X] = \sum_{k=1}^{\infty} E[Y|X=x_k] I_{\{X=x_k\}}.$$

给定 $X=x$ 条件下的 Y 的条件分布的方差称为条件方差，记为 $\text{Var}(Y|X=x)$：

$$\text{Var}(Y|X=x) = \int_{-\infty}^{\infty} (y - E(Y|X=x))^2\,dF(y|X=x),$$

这是 x 的函数 $h(x)$，记 $\text{Var}(Y|X)=h(X)$。

命题 1.6. 设 $E(Y^2)<\infty$，则

$$Var(Y) = E[Var(Y|X)] + Var[E(Y|X)].$$

证明：

$$
\begin{aligned}
&E[\mathrm{Var}(Y|X)]\\
=&\int_{-\infty}^{\infty} E\left\{[Y-E(Y|X=x)]^2|X=x\right\}\,dF_X(x)\\
=&\int_{-\infty}^{\infty} \left\{E(Y^2|X=x)-[E(Y|X=x)]^2\right\}\,dF_X(x)\\
=&\int_{-\infty}^{\infty} E(Y^2|X=x)\,dF_X(x)-\int_{-\infty}^{\infty}[E(Y|X=x)]^2\,dF_X(x).\\
&\mathrm{Var}[E(Y|X)]\\
=&E\{[E(Y|X)]^2\}-\{E[E(Y|X)]\}^2\\
=&\int_{-\infty}^{\infty}[E(Y|X=x)]^2\,dF_X(x)-[E(Y)]^2,\\
&E[\mathrm{Var}(Y|X)]+\mathrm{Var}[E(Y|X)]\\
=&\int_{-\infty}^{\infty} E(Y^2|X=x)\,dF_X(x)-[E(Y)]^2\\
=&E[E(Y^2|X)]-[E(Y)]^2=E(Y^2)-[E(Y)]^2=\mathrm{Var}(Y).
\end{aligned}
$$

定理 1.21 (全期望公式)**.** 设 X, Y 为随机变量，期望存在，则

$$E[Y]=E\{E[Y|X]\}=\int_{-\infty}^{\infty} E[Y|X=x]\,dF_X(x). \tag{1.4}$$

当 X 为一个离散随机变量时，设其取值集合为 $\{x_k, k=1,2,...\}$，(1.4)式为

$$E[Y]=\sum_{k=1}^{\infty} E[Y|X=x_k]P\{X=x_k\}.$$

当 (X,Y) 为连续型随机向量时，(1.4)式为

$$E[Y]=\int_{-\infty}^{\infty} E[Y|X=x]f_X(x)\,dx.$$

例 1.16 (随机个随机变量之和)**.** 设 $X_1, X_2, ...$ 是一列独立同分布的随机变量；设 N 为取非负整数值的随机变量，且与序列 $X_1, X_2, ...$ 独立. 求 $Y=\sum_{i=1}^{N} X_i$ 的均值和方差.

解：如果 X_1 的矩母函数 $\phi(t)$ 存在，可以求 Y 的矩母函数

$$\begin{aligned}
\phi_Y(t) =& E(e^{tY}) = E[E(e^{tY}|N)] = \sum_{n=0}^{\infty} E(e^{tY}|N=n)P(N=n) \\
=& P(N=0) + \sum_{n=1}^{\infty}\prod_{i=1}^{n} E(e^{tX_i})P(N=n) \\
=& P(N=0) + \sum_{n=1}^{\infty}[\phi(t)]^n P(N=n) \\
=& E[\phi(t)]^N.
\end{aligned}$$

对 $\phi_Y(t)$ 求导得

$$\begin{aligned}
\phi_Y'(t) =& E[N(\phi(t))^{N-1}\phi'(t)], \\
E(Y) =& \phi_Y'(0) = E[NE(X_1)] = E(N)E(X_1). \\
\phi_Y''(t) =& E[N(N-1)(\phi(t))^{N-2}(\phi'(t))^2 + N(\phi(t))^{N-1}\phi''(t)], \\
E(Y^2) =& \phi_Y''(0) = E[N(N-1)[E(X_1)]^2] + E[NE(X_1^2)] \\
=& [E(X_1)]^2 E[N(N-1)] + E(X_1^2)E(N), \\
\mathrm{Var}(Y) =& E(Y^2) - [E(Y)]^2 \\
=& [E(X_1)]^2 E[N(N-1)] + E(X_1^2)E(N) - [E(N)]^2[E(X_1)]^2 \\
=& [E(X_1)]^2 \mathrm{Var}(N) + \mathrm{Var}(X_1)E(N).
\end{aligned}$$

因为 X_1 的矩母函数不一定存在，我们直接求 $E(Y)$ 和 $E(Y^2)$。

$$E(Y) = E[E(Y|N)],$$

其中

$$\begin{aligned}
E(Y|N=n) =& E(\sum_{i=1}^{n} X_i|N=n) \\
=& E(\sum_{i=1}^{n} X_i) \quad (\text{利用独立性}) \\
=& nE(X_1),
\end{aligned}$$

故

$$E(Y) = E(NE(X_1)) = E(X_1)E(N).$$

再来求

$$E(Y^2) = E[E(Y^2|N)],$$

其中

$$
\begin{aligned}
E(Y^2|N=n) &= E[(\sum_{i=1}^{n} X_i)^2|N=n] \\
=&E[(\sum_{i=1}^{n} X_i)^2] = \mathrm{Var}[\sum_{i=1}^{n} X_i] + [E(\sum_{i=1}^{n} X_i)]^2 \\
=&n\mathrm{Var}(X_1) + n^2[E(X_1)]^2,
\end{aligned}
$$

故

$$
\begin{aligned}
E(Y^2) =& E\{N\mathrm{Var}(X_1) + N^2[E(X_1)]^2\} \\
=& E(N)\mathrm{Var}(X_1) + E(N^2)[E(X_1)]^2, \\
\mathrm{Var}(Y) =& E(Y^2) - [E(Y)]^2 \\
=& [E(X_1)]^2\mathrm{Var}(N) + \mathrm{Var}(X_1)E(N).
\end{aligned}
$$

求 $\mathrm{Var}(Y)$，也可以利用

$$
\begin{aligned}
\mathrm{Var}(Y|N=n) &= \mathrm{Var}(\sum_{i=1}^{n} X_i|N=n) \\
=&\mathrm{Var}(\sum_{i=1}^{n} X_i) = n\mathrm{Var}(X_1)
\end{aligned}
$$

得到 $\mathrm{Var}(Y|N) = N\mathrm{Var}(X_1)$，从而得到

$$
\begin{aligned}
\mathrm{Var}(Y) =& E[\mathrm{Var}(Y|N)] + \mathrm{Var}[E(Y|N)] \\
=& E[N\mathrm{Var}(X_1)] + \mathrm{Var}[NE(X_1)] \\
=& [E(X_1)]^2\mathrm{Var}(N) + \mathrm{Var}(X_1)E(N).
\end{aligned}
$$

例 1.17. 一个矿工陷进一个有三个门的矿井。第一个门通向一个隧道，沿此隧道走 2 小时的旅程他可到达安全地；第二个门通向另一个隧道，沿此隧道走 3 小时会使他回到矿井；第三个门通向一个隧道，沿此隧道走 5 小时会使他回到矿井。假定矿工总是等可能地选取任意一个门，用 X 表示矿工到达安全地所需的时间，求 X 的均值及矩母函数。

解：写出 X 的分布列很困难，故无法直接求其均值。在类似的问题中，我们经常使用关于某种初始状态（或选择）取条件的方法。

令 Y 表示矿工第一次选择的门，即 $\{Y=i\}$ 表示第一次选择第 i 个门，由题意知

$$
p\{Y=1\} = P\{Y=2\} = P\{Y=3\} = \frac{1}{3}.
$$

关于 Y 取条件，有

$$
\begin{aligned}
E[X] =& E[E(X|Y)] \\
=& \sum_{i=1}^{3} E[X|Y=i]P(Y=i) \\
=& \frac{1}{3}\{E[X|Y=1]+E[X|Y=2]+E[X|Y=3]\} \\
=& \frac{1}{3}\{2+(3+E[X])+(5+E[X])\}.
\end{aligned}
$$

故有 $E[X]=10$.

矩母函数：

$$
\begin{aligned}
E[e^{tX}] =& E[E(e^{tX}|Y)] \\
=& \frac{1}{3}\{E[e^{tX}|Y=1]+E[e^{tX}|Y=2]+E[e^{tX}|Y=3]\},
\end{aligned}
$$

易知 $E[e^{tX}|Y=1]=e^{2t}$；当 $Y=2$ 时，$X=3+X'$，其中 X' 是回到矿井后再到达安全区所附加的时间，X 与 X' 有相同的分布，故有

$$E[e^{tX}|Y=2]=E[e^{t(3+X')}]=e^{3t}E[e^{tX}],$$

同理可得

$$E[e^{tX}|Y=3]=E[e^{t(5+X')}]=e^{5t}E[e^{tX}].$$

于是得

$$E[e^{tX}]=\frac{e^{2t}}{3-e^{3t}-e^{5t}}.$$

1.5.4 关于 σ 代数的条件期望

最后我们将条件期望推广到一般随机变量及 σ 代数情形.

设 X 是随机变量，B 是事件且 $P(B)>0$，则给定事件 B，随机变量 X 的条件期望定义为

$$E[X|B]=[P(B)]^{-1}E[XI_B].$$

这相当于将 Ω 限制到 B 中得到一个新的概率测度，在这个概率测度下求 X 的期望。

设 $\mathscr{G}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 的子 σ 代数，X 为 $(\Omega, \mathscr{F}, P)$ 中的随机变量。X 与 $\mathscr{G}$ 的关系可以分为如下三种情况：

- X 关于 $\mathscr{G}$ 可测。在已知 $\mathscr{G}$ 的情况下，关于 X 的任何取值论断（严格来讲是 ${ X B$, B 为 Borel 集）都可以确定地给出肯定或者否定的回答。注意这不依赖于概率测度。
- X 与 $\mathscr{G}$ 独立。知道 $\mathscr{G}$ 中的信息对 X 的取值可能性没有任何影响。独立性依赖于概率测度。
- X 既不是关于 $\mathscr{G}$ 可测，与不与 $\mathscr{G}$ 独立。这时可以用 $\mathscr{G}$ 的信息对 X 进行预测，得到条件数学期望 $E(X|\mathscr{G})$。

所谓已知 $\mathscr{G}$ 后所能获得的信息，可以理解为试验结果 $\omega \in \Omega$ 的信息不能完全知道，但 $\forall A \in \mathscr{G}$，$\omega \in A$ 还是 $\omega \notin A$ 都可以给出明确答案。$\mathscr{G}$ 中的集合越多，能给出明确答案的问题就越多；最小的 σ 代数是 $\{\Omega, \emptyset\}$，相当于没有信息。

定义 1.18. 设 Y 是随机变量且 $E|Y| < \infty, \mathscr{G}$ 是 $\mathscr{F}$ 的子 σ 代数，存在唯一的 (几乎必然相等的意义下) 随机变量 Y^*，使得 $E|Y^*| < \infty$，Y^* 关于 $\mathscr{G}$ 可测，且

$$E[Y^* I_B] = E[Y I_B],\ \forall B \in \mathscr{G},$$

称随机变量 Y^* 为 Y 在给定 $\mathscr{G}$ 下的**条件期望**，记为 $Y^* = E[Y|\mathscr{G}]$.

这个定义不如前面的 $E(Y|X)$ 容易理解。我们考虑 $E(Y^2) < \infty$ 的情形，这时称 $Y \in L^2(\Omega, \mathscr{F}, P)$，这是一个 Hilbert 空间，内积为 $\langle X, Y \rangle = E(XY)$。这时必有 $Y^* \in L^2(\Omega, \mathscr{G}, P)$，

$$\begin{aligned}
& Y^* = E(Y|\mathscr{G}) \\
\Longleftrightarrow\ & Y^* \text{关于} \mathscr{G} \text{可测且} E[(Y - Y^*) I_B] = 0,\ \forall B \in \mathscr{G} \\
\Longleftrightarrow\ & Y^* \in L^2(\Omega, \mathscr{G}, P) \text{且} E[(Y - Y^*)\xi] = 0,\ \forall \xi \in L^2(\Omega, \mathscr{G}, P) \\
\Longleftrightarrow\ & Y^* \in L^2(\Omega, \mathscr{G}, P) \text{且} E[(Y - Y^*)^2] \leq E[(Y - \xi)^2],\ \forall \xi \in L^2(\Omega, \mathscr{G}, P).
\end{aligned}$$

即 $E(Y|\mathscr{G})$ 是用关于 $\mathscr{G}$ 可测的随机变量 ξ 对 Y 进行逼近，在均方误差最小准则下的最佳逼近。

如果仅要求 $E|Y| < \infty$，则

$$
\begin{aligned}
& Y^* = E(Y|\mathscr{G}) \\
\Longleftrightarrow\ & Y^* \text{关于}\mathscr{G}\text{可测且} E[(Y - Y^*)I_B] = 0,\ \forall B \in \mathscr{G} \\
\Longleftrightarrow\ & Y^* \text{关于}\mathscr{G}\text{可测且} E[(Y - Y^*)\xi] = 0,\ \forall \text{关于}\mathscr{G}\text{可测的有界随机变量}\xi.
\end{aligned}
$$

定义

$$
\begin{aligned}
E(Y|X) =& E(Y|\sigma(X)), \\
E(Y|X_1, \dots, X_n) =& E(Y|\sigma(X_1, \dots, X_n)), \\
E(Y|X_i, i \in I) =& E(Y|\sigma(X_i, i \in I)),
\end{aligned}
$$

在离散和连续型两种情形下的 $E(Y|X)$ 定义与此定义等价。当 $E(Y^2) < \infty$ 时，$E(Y|X_1, \dots, X_n)$ 是用 $(X_1, \dots, X_n)$ 的函数对 Y 进行逼近时，在均方误差最小准则下的最佳逼近。

设 $\{X_t, t \in T\}$ 是一族随机变量，定义

$$
E(Y|X_t, t \in T) = E[Y|\sigma(\{X_t, t \in T\})].
$$

::: {.example #prob-indce-cesig-cpce} 分析条件期望与条件概率的关系。:::

解：考虑 $E(I_A|I_B)$，这是 I_B 的函数，当 $I_B = 1$ 时，即 B 发生，有

$$
E(I_A|I_B = 1) = E(I_A|B) = P(A|B);
$$

当 $I_B = 0$ 时，即 B^c 发生，有

$$
E(I_A|I_B = 0) = E(I_A|B^c) = P(A|B^c),
$$

于是作为 I_B 的函数，有

$$
\begin{aligned}
E(I_A|I_B) =& E(I_A|I_B = 1)I_B + E(I_A|I_B = 0)I_{B^c} \\
=& P(A|B)I_B + P(A|B^c)I_{B^c}.
\end{aligned}
$$

所以关于随机变量或者 σ 代数的数学期望也是一个随机变量，在这个例子中同时考虑了 B 发生与 B 不发生两种情况。初等概率论中的条件概率或者条件期望是给定某种情况下的条件概率值或者条件期望值，是非随机的。

定理 1.22. 设 X, Y 是一阶矩有限的随机变量，a, b 为实数，条件期望有如下基本性质：

(1) (全期望公式)：

$$E\{E[X|\mathscr{G}]\} = E[X].$$

(2) 若 X 关于 $\mathscr{G}$ 可测，则 $E[X|\mathscr{G}] = X, a.s.$

(3)} 若 X 与 $\mathscr{G}$ 相互独立 *(*即 $\sigma(X)$ 与 $\mathscr{G}$ 相互独立*)*，则有 $E[X|\mathscr{G}] = E[X]$, $a.s.$

(4) 设 $\mathscr{G} = \{\emptyset, \Omega\}$，则 $E[X|\mathscr{G}] = E[X]$, $a.s.$

(5) $E[X|\mathscr{G}] = E[X^+|\mathscr{G}] - E[X^-|\mathscr{G}]$, $a.s.$

*(6) (*保序*)*：若 $X \leq Y$, $a.s.$，则 $E[X|\mathscr{G}] \leq E[Y|\mathscr{G}]$, $a.s.$

(7) （线性）：

$$E[aX + bY|\mathscr{G}] = aE[X|\mathscr{G}] + bE[Y|\mathscr{G}],\ a.s.$$

(8) $|E[X|\mathscr{G}]| \leq E[|X|\,|\mathscr{G}]$, $a.s.$

(9) （单调收敛定理）：设 $0 \leq X_n \uparrow X$, $a.s.$，则 $E[X_n|\mathscr{G}] \uparrow E[X|\mathscr{G}]$, $a.s.$

(10) 设 X 及 XY 的期望存在，且 Y 为 $\mathscr{G}$ 可测，则 $E[XY|\mathscr{G}] = YE[X|\mathscr{G}]$, $a.s.$

(11) 若 $\mathscr{G}_1, \mathscr{G}_2$ 是两个子 σ 代数，使得 $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$, 则

$$E\{E[X|\mathscr{G}_2]\,|\,\mathscr{G}_1\} = E\{E[X|\mathscr{G}_1]\,|\,\mathscr{G}_2\} = E[X|\mathscr{G}_1],\ a.s.$$

(12) 若 X, Y 是两个独立的随机变量（或随机向量），函数 $g(x, y)$ 使得 $E[|g(X, Y)|] < \infty$，令

$$h(y) = E[g(X, y)],$$

则有

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

(13) （条件 *Jensen* 不等式）设 $g(x)$ 为凸函数，则 $E[g(X)|\mathscr{G}] \geq g[E(X|\mathscr{G})]$。

推论 1.5. 设 $E(X^2) < \infty$，则 $E(X^2|\mathscr{G}) \geq [E(X|\mathscr{G})]^2$，$E[E(X|\mathscr{G})]^2 < \infty$。

推论 1.6. 设 $E(Y^2) < \infty$，则对任意关于 $\mathscr{G}$ 可测且满足 $E(\xi^2) < \infty$ 的随机变量 ξ，都有

$$E(Y - E[Y|\mathscr{G}])^2 \leq E(Y - \xi)^2.$$

证明：

$$\begin{aligned}
&E[Y-\xi]^2 = E[(Y - E[Y|\mathscr{G}]) + (E[Y|\mathscr{G}] - \xi)]^2 \\
=&E(Y - E[Y|\mathscr{G}])^2 + E(E[Y|\mathscr{G}] - \xi)^2 \\
&+ 2E\left\{(Y - E[Y|\mathscr{G}])(E[Y|\mathscr{G}] - \xi)\right\},
\end{aligned}$$

其中交叉项

$$\begin{aligned}
&E\left\{(Y - E[Y|\mathscr{G}])(E[Y|\mathscr{G}] - \xi)\right\} \\
=&E\left[E\left\{(Y - E[Y|\mathscr{G}])(E[Y|\mathscr{G}] - \xi) \mid \mathscr{G}\right\}\right] \\
=&E\left[(E[Y|\mathscr{G}] - \xi)\, E\left\{(Y - E[Y|\mathscr{G}]) \mid \mathscr{G}\right\}\right] \\
=&E\left[(E[Y|\mathscr{G}] - \xi)\, 0\right] = 0,
\end{aligned}$$

结论得证。

1.6 补充内容

1.6.1 取帽子问题

求例1.9中 X 的概率分布。

例 1.18. n 个人的帽子混在一起后随机无放回抽取，求至少有一个人拿对自己的帽子的概率，并计算没有人取对自己帽子的概率。

解： 利用 Jordan 公式。把帽子编号 $1, 2, \dots, n$，第 i 个人在第 i 次抽取。设 A_i 表示第 i 个人取对帽子的事件，要求 $P(B) = P(\bigcup_{i=1}^n A_i)$。给定 k 后所有 $P(A_{j_1} A_{j_2} \dots A_{j_k})$ 相等，只要计算 $P(A_1 A_2 \dots A_k)(k = 1, 2, \dots, n)$。帽子取法有 $n!$ 种等可能结果，前 k 人都取对，其他人随意取，有 $(n-k)!$ 种取法。由

Jordan 公式

$$
\begin{aligned}
P(B) &= \sum_{k=1}^{n}(-1)^{k-1}C_n^k P(A_1A_2\dots A_k)\\
&= \sum_{k=1}^{n}(-1)^{k-1}\frac{n!}{k!(n-k)!}\frac{(n-k)!}{n!}\\
&= \sum_{k=1}^{n}(-1)^{k-1}\frac{1}{k!}.
\end{aligned}
$$

于是

$$
\begin{aligned}
&P(\text{无人取对帽子}) = 1 - P(\text{至少一人取对帽子})\\
=&1 - \sum_{k=1}^{n}(-1)^{k-1}\frac{1}{k!}\\
=&\sum_{k=0}^{n}(-1)^{k}\frac{1}{k!}
\end{aligned}
$$

例 1.19. n 个人的帽子混在一起后随机无放回抽取，求恰有 k 个人拿对自己帽子的概率。

解：用 $A_n(k)$ 表示 n 个人依次在编号为 $1 \sim n$ 的帽子中随机选取，恰有 k 人取对的取法数，则

$$
\begin{aligned}
&A_n(k) = C_n^k \cdot A_{n-k}(0),\\
&P(n\text{个人中恰有}k\text{人取对}) = \frac{A_n(k)}{n!}\\
=&\frac{1}{k!}\cdot\frac{A_{n-k}(0)}{(n-k)!}\\
=&\frac{1}{k!}P(n-k\text{个人取}n-k\text{顶帽子都不匹配})\\
=&\frac{1}{k!}\sum_{k=0}^{n-k}(-1)^j\frac{1}{j!} = \sum_{k=0}^{n-k}(-1)^j\frac{1}{k!j!},\ k=0,1,\dots,n.
\end{aligned}
$$

用 X 表示取对帽子的人数，虽然有了上述概率分布，计算 $E(X)$ 和 $\mathrm{Var}(X)$ 仍比较复杂。

1.6.2 正态分布各阶矩

直接用积分来计算标准正态分布的各阶矩。当 k 为奇数时，

$$E(X^k) = \int_{-\infty}^{\infty} x^k \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx$$

是可积的且被积函数为奇函数，所以期望等于 0。

当 k 为偶数时，作积分变量替换 $t = \frac{1}{2}x^2$, $x = 2^{1/2}t^{1/2}$, 则

$$\begin{aligned}
E(X^k) =& 2\int_0^{\infty} x^k \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx \\
=& \frac{2}{\sqrt{2\pi}} \int_0^{\infty} 2^{k/2} t^{k/2} e^t 2^{1/2} \frac{1}{2} t^{-1/2}\, dt \\
=& \frac{1}{\sqrt{\pi}} 2^{k/2} \int_0^{\infty} t^{\frac{k+1}{2}-1} e^t\, dt \\
=& \frac{1}{\sqrt{\pi}} 2^{k/2} \Gamma(\frac{k+1}{2}) = \frac{1}{\sqrt{\pi}} 2^{k/2} \frac{k-1}{2} \Gamma(\frac{k-1}{2}) \\
=& \frac{1}{\sqrt{\pi}} 2^{k/2} \frac{k-1}{2} \frac{k-3}{2} \cdots \frac{3}{2} \frac{1}{2} \Gamma(\frac{1}{2}) \\
=& \frac{1}{\sqrt{\pi}} 2^{k/2} 2^{-k/2} (k-1)(k-3) \cdots 3 \cdot 1 \cdot \sqrt{\pi} \\
=& (k-1)!!.
\end{aligned}$$

这里利用了 $\Gamma(x+1) = x\Gamma(x)$, $\forall x > 0$ 和 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$。

另外，可以计算

$$\begin{aligned}
E|X| =& \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x e^{-\frac{1}{2}x^2}\, dx \\
=& \frac{2}{\sqrt{2\pi}} \int_0^{\infty} 2^{1/2} t^{1/2} e^{-t} 2^{1/2} \frac{1}{2} t^{-1/2}\, dt \\
=& \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-t}\, dt \\
=& \sqrt{\frac{2}{\pi}}.
\end{aligned}$$

1.6.3 随机元和弱收敛

设 X 是概率空间 $(\Omega, \mathscr{F}, P)$ 到可测空间 $(S, \mathscr{B}(S))$ 的可测函数，称为随机元。X 导出了 S 中的测度 $Q(A) = P(X^{-1}(A))$。

设 (S, ρ) 是距离空间，$\mathscr{B}(S)$ 是其中的开集张成的 σ 代数，若 Q_n 和 Q 是 $(S, \mathscr{B}(S))$ 中的概率测度，满足

$$\lim_{n\to\infty} \int_S g(s)\, dQ_n(s) = \int_S g(s)\, dQ(s)$$

对任意定义在 S 上的实值连续有界函数 g 成立，则称 Q_n 弱收敛到 Q。

若 $P(X_n^{-1}(\cdot))$ 弱收敛到 $P(X^{-1}(\cdot))$，则称 X_n 或依分布收敛到 X。X_n 也可以定义在不同的概率空间中。

连续变换可以保持依分布收敛性。

当 $S = \mathbb{R}^n$ 时，依分布收敛当且仅当特征函数点点收敛。

1.6.4 Lebesgue 积分

对概率空间 $(\Omega, \mathscr{F}, P)$ 上的非负随机变量 X，定义取 $0 = y_0 < y_1 < \cdots < y_n$，当 $n \to \infty$ 时 $y_n \to \infty$, 且 $\max_{1\leq k\leq n}(y_k - y_{k-1}) \to 0$, 则

$$\lim_{n\to\infty} \sum_{k=1}^n y_{k-1} P(y_{k-1} < X \leq y_k)$$

存在且不依赖于 $\{y_k\}$ 的选取，定义此极限为

$$\int_\Omega X(\omega) dP(\omega).$$

称这样的积分为 Legesgue 积分。定义

$$E(X) = \int_\Omega X(\omega) dP(\omega).$$

对一般的随机变量 X，令

$$X^+ = \max(X, 0), \quad X^- = \max(-X, 0),$$

则 $X = X^+ - X^-$，若 $\int_\Omega X^+(\omega) dP(\omega)$ 和 $\int_\Omega X^-(\omega) dP(\omega)$ 至少有一个为有限值，则定义

$$E(X) = \int_\Omega X(\omega) dP(\omega) = \int_\Omega X^+(\omega) dP(\omega) - \int_\Omega X^-(\omega) dP(\omega) = E(X^+) - E(X^-).$$

考虑 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$，可以定义关于 $A\in\mathscr{B}(\mathbb{R})$ 的函数 $L(A)$ 满足

(1) $L(A)\geq 0$;

(2) $L([a,b])=b-a,\ -\infty<a<b<\infty$;

(3) 对互不相交的$A_n\in\mathscr{B}(\mathbb{R})$, 有

$$L(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}L(A_n),$$

称 $L(\cdot)$ 为 Lebegue 测度。

对 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$，若函数 $g(x)$ 满足

$$\{x:g(x)\leq y\}\in\mathscr{B}(\mathbb{R}),\ \forall y\in\mathbb{R},$$

称 $g(x)$ 是 Borel 可测函数。对 Borel 可测函数 $g(x)$，可类似定义 $\int_{\mathbb{R}}g(x)dL(x)$，称为 Lebesgue 积分。当 $g(x)$ 黎曼可积时必 Lebesgue 可积，而且积分相等，所以将 Lebesgue 积分也记成

$$\int_{-\infty}^{\infty}g(x)\,dx.$$

区间上的 Lebesgue 积分可类似定义，也可以定义成

$$\int_a^b g(x)\,dx=\int_{\mathbb{R}}g(x)I_{[a,b]}(x)\,dL(x).$$

关于 Lebesgue 积分的性质，可参见数学期望的性质。

1.6.5 测度变换

同一个随机变量在不同的概率测度下的分布不同，期望不同。

例 1.20. 考虑测度空间 $([0,1],\mathscr{B}([0,1]))$。定义随机变量 X 为

$$X(\omega)=\omega,\ \forall\omega\in[0,1].$$

定义概率测度

$$P([a,b])=b-a.$$

这个概率测度对 $[0,1]$ 中的随机结果是等可能看待的。

定义正值随机变量

$$Z(\omega) = 2\omega, \ \forall \omega \in [0,1].$$

易见 $E(Z) = 1$。把 Z 的取值看成是另一个测度 $Q(\cdot)$ 相对于 $P(\cdot)$ 对于 $[0,1]$ 中不同随机结果的概率大小的看法，定义概率测度 $Q(\cdot)$ 为

$$Q([a,b]) = \int_a^b Z(\omega)\, d\omega.$$

则

$$Q([a,b]) = \int_a^b 2\omega\, d\omega = b^2 - a^2 = (b+a)(b-a),$$

当 a, b 在 1 附近时，$Q([a,b]) > P([a,b])$；当 a, b 在 0 附近时，$Q([a,b]) < P([a,b])$，即 $Q(\cdot)$ 给 1 附近的随机结果赋了更大的概率值。

考虑 X 在 Q 下的分布。X 的定义并不随 P, Q 变化，在均匀的 P 测度下，X 是均匀的；但是，在 Q 测度下，1 附近的取值概率更大，就会使得 X 在 1 附近的密度更大。实际上，X 在 Q 下的分布函数和分布密度分别为

$$\begin{aligned}
F^Q(x) =& Q(\{\omega : X(\omega) \leq x\}) = Q(\{\omega : \omega \leq x\}) \\
=& Q([0,x]) = \int_0^x Z(\omega)\, d\omega \\
=& \int_0^x 2\omega\, d\omega = x^2, \ x \in [0,1]. \\
f^Q(x) =& \frac{d}{dx} F^Q(x) = 2x, \ x \in [0,1].
\end{aligned}$$

所以 X 在 $Q(\cdot)$ 下的密度在 1 附近更大。

再来考虑 X 在 Q 下的期望（记作 $E^Q(X)$）。

$$\begin{aligned}
E^Q(X) =& \int_0^1 x f^Q(x)\, dx \\
=& \int_0^1 x\, 2x\, dx = \frac{2}{3}.
\end{aligned}$$

注意

$$E(XZ) = \int_0^1 x\, 2x\, dx,$$

可见

$$E^Q(X) = E(XZ),$$

这是一般结论。

定理 1.23. 设 $(\Omega, \mathscr{F}, P)$ 为概率空间，随机变量 $Z \geq 0$, *a.s.*, 且 $E(Z) = 1$，定义

$$Q(A) = \int_A Z(\omega)\, dP(\omega) = \int_\omega Z(\omega) I_A(\omega)\, dP(\omega) = E(Z I_A), \tag{1.5}$$

则 $Q(\cdot)$ 是 $(\Omega, \mathscr{F})$ 上的概率测度，如果 X 是非负随机变量，则有

$$E^Q(X) = \int_\omega X(\omega)\, dQ(\omega) = E(XZ) = \int_\omega X(\omega) Z(\omega)\, dP(\omega). \tag{1.6}$$

如果 $Z > 0$, *a.s.*, 则对任意非负随机变量 Y 有

$$E(Y) = E^Q(\frac{Y}{Z}). \tag{1.7}$$

这里 $E^Q(X)$ 定义为 $\int_\omega X(\omega)\, dQ(\omega)$。

如果 X（或者 Y）不是非负的，但是(1.6)（或者(1.7)）两边之一存在，则(1.6)（或者(1.7)）仍成立。

证明略，参见 (Shreve 2004) 定理 1.6.1。

定义 1.19. 设 P 和 Q 是测度空间 $(\Omega, \mathscr{F})$ 上的两个概率测度，如果

$$P(A) = 0 \iff Q(A) = 0,\ \forall A \in \mathscr{F},$$

称这两个概率测度等价。

定理 1.24. 设 $(\Omega, \mathscr{F}, P)$ 为概率空间，随机变量 $Z > 0$, *a.s.*, 且 $E(Z) = 1$，定义

$$Q(A) = \int_A Z(\omega)\, dP(\omega) = E(Z I_A), \tag{1.8}$$

则 Q 是与 P 等价的概率测度。

定理 1.25. 设 P 和 Q 是测度空间 $(\Omega, \mathscr{F})$ 上的两个等价的概率测度，则必存在几乎必然为正的随机变量 Z，满足 $E(Z) = 1$，使得

$$Q(A) = \int_A Z(\omega)\, dP(\omega),\ \forall A \in \mathscr{F},$$

称 Z 是 Q 对 P 的 *Radon-Nikodym* 导数，记为

$$\frac{dQ}{dP}(\omega).$$

在金融建模中经常利用测度变换。P 表示真实世界的概率模型，而 Q 表示风险中性世界的概率模型，两者等价，且可以通过正值随机变量 Z 将两个测度联系起来。因为等价性，在两个不同世界其中之一几乎必然成立的事件，在另一个世界也是几乎必然成立的。

例 1.21. 设 X 是 $(\Omega, \mathscr{F}, P)$ 中的标准正态分布随机变量。令

$$Y = X + \theta,$$

则 Y 是 $(\Omega, \mathscr{F}, P)$ 中 $\mathrm{N}(\theta, 1)$ 随机变量。求测度变换 Q 使得 Y 在 $(\Omega, \mathscr{F}, Q)$ 中服从标准正态分布。

解：不妨设 $\theta > 0$。取

$$Z = \exp\{-\theta X - \frac{1}{2}\theta^2\},$$

则 Z 是 $(\Omega, \mathscr{F}, P)$ 中取正值的随机变量，

$$E(Z) = e^{-\frac{1}{2}\theta^2} E e^{-\theta X} = 1.$$

令

$$Q(A) = \int_A Z(\omega)\, dP(\omega),$$

则 Q 是与 P 等价的测度，$Z = \frac{dQ}{dP}$，随机变量 Y 在 $(\Omega, \mathscr{F}, Q)$ 中的矩母函数为

$$\begin{aligned}
E^Q(e^{tY}) =& E(Z e^{t(X+\theta)}) = E \exp\{-\theta X - \frac{1}{2}\theta^2 + tX + t\theta\} \\
=& \exp\{-\frac{1}{2}\theta^2 + t\theta\} E \exp\{(t-\theta)X\} \\
=& \exp\{-\frac{1}{2}\theta^2 + t\theta\} \exp\{\frac{1}{2}(t-\theta)^2\} \\
=& e^{\frac{1}{2}t^2},
\end{aligned}$$

即 Y 在 Q 概率测度下服从标准正态分布。

Chapter 2

随机过程的基本概念和基本类型

2.1 基本概念

定义 2.1. 随机过程是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一族随机变量 $\{X(t), t \in T\}$，其中 t 是参数，它属于某个指标集 T, T 称为参数集.

注：

- 当 $T = \{0, 1, 2, \cdots\}$ 时称之为随机序列或时间序列.
- 参数 t 经常被解释为时间。
- 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, 但仅要求对每个 $t \in T$，$X(t)$ 是随机变量；仅在某些情景下才要求 $X(t, \omega)$ 是二元可测的。
- 参数空间 T 是向量集合时，随机过程 $\{X(t), t \in T\}$ 称为随机场.

$X(t)$ 表示系统在时刻 t 所处的状态. $X(t)$ 的所有可能状态构成的集合为状态空间，记为 S. 一般如果不作说明都认为状态空间是实数集 $\mathbb{R}$ 或 $\mathbb{R}$ 的子集.

随机过程的分类:

（1）依照状态空间可分为连续状态和离散状态；

（2）依照参数集，可分为离散参数过程和连续参数过程。

例 2.1 (随机游动)**.** 一个醉汉在路上行走，以概率 p 前进一步，以概率 $1-p$

后退一步（假定其步长相同）. 以 $X(t)$ 记他在路上的位置，则 $X(t)$ 就是直线上的随机游动.

例 2.2 (布朗运动)**.** 英国植物学家布朗注意到飘浮在液面上的微小粒子不断进行无规则的运动, 这种运动后来称为布朗运动. 它是分子大量随机碰撞的结果. 若记 $(X(t), Y(t))$ 为粒子在平面坐标上的位置，则它是平面上的布朗运动.

例 2.3 (排队模型)**.** 顾客来到服务站要求服务. 当服务站中的服务员都忙碌，即服务员都在为别的顾客服务时，来到的顾客就要排队等候. 顾客的到来、每个顾客所需的服务时间都是随机的，所以如果用 $X(t)$ 表示 t 时刻的队长，用 $Y(t)$ 表示 t 时刻到来的顾客所需的等待时间，则 $\{X(t), t \in T\}$, $\{Y(t), t \in T\}$ 都是随机过程.

2.2 有限维分布与 Kolmogorov 定理

定义 2.2. 对任意有限个 $t_1, \dots, t_n \in T$，定义随机过程的 n **维分布**$F_{t_1,\dots,t_n}(x_1, \dots, x_n)$：

$$F_{t_1,\dots,t_n}(x_1, \dots, x_n) = P(X(t_1) \le x_1, \dots, X(t_n) \le x_n).$$

随机过程的所有的一维分布，二维分布，……，n 维分布等等的全体

$$\{F_{t_1,\dots,t_n}(x_1, \dots, x_n) : t_1, \dots, t_n \in T, n \ge 1\}$$

称为随机过程 $\{X(t), t \in T\}$ 的**有限维分布族**.

注：知道了随机过程的有限维分布族就知道了 $\{X(t), t \in T\}$ 中任意 n 个随机变量的联合分布，也就掌握了这些随机变量之间的相互依赖关系. 但是，对于连续参数的随机过程，如果考虑涉及到一个区间上的随机过程的事件概率，因为涉及到无穷且不可数个随机变量共同的事件概率，仅有限维分布族不能完全确定这些事件的概率，通常需要对随机过程添加一些额外的要求，比如轨道连续，均方连续等。

例 2.4. 本例来自 Karlin and Taylor “A First Course in Stochastic Process” 节 1.3. 设 $U \sim U(0,1)$，随机过程 $\{X_t, t \in [0,1]\}$ 和 $\{Y_t, t \in [0,1]\}$ 定义如下：

$$X_t = \begin{cases} 1, & t = U, \\ 0, & \text{其它}. \end{cases}$$

$$Y_t = 0,\ \forall t \in [0,1].$$

则 $\{X_t\}$ 与 $\{Y_t\}$ 的有限维分布相同。但事件

$$A = \{X_t \leq 0.5,\ t \in [0,1]\}$$

概率为 0，事件

$$B = \{Y_t \leq 0.5,\ t \in [0,1]\}$$

概率为 1.

分布族的性质:

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

$$\begin{aligned}
&F_{t_{j_1},\ldots,t_{j_n}}(x_{j_1},\ldots,x_{j_n})\\
=&P(X(t_{j_1}) \leq x_{j_1},\ldots,X(t_{j_n}) \leq x_{j_n})\\
=&P(X(t_1) \leq x_1,\ldots,X(t_n) \leq x_n)\\
=&F_{t_1,\ldots,t_n}(x_1,\ldots,x_n).
\end{aligned}$$

（2）相容性：对于 $m < n$，有

$$F_{t_1,\ldots,t_m,t_{m+1},\ldots,t_n}(x_1,\ldots,x_m,\infty,\ldots,\infty) = F_{t_1,\ldots,t_m}(x_1,\ldots,x_m).$$

定理 2.1 (Kolmogorov 存在性定理). *设有限维分布函数 $\{F_{t_1,\ldots,t_n}(x_1,\ldots,x_n):\ t_1,\ldots,t_n \in T, n \geq 1\}$ 满足上述的对称性和相容性，则必存在一个随机过程 $\{X(t), t \in T\}$，使*

$$\{F_{t_1,\ldots,t_n}(x_1,\ldots,x_n):\ t_1,\ldots,t_n \in T, n \geq 1\}$$

恰好是 $\{X(t), t \in T\}$ 的有限维分布族.

证明略。

定义 2.3 (高斯过程). 若随机过程 $\{X(t), t \in T\}$ 的所有有限维分布都是多元正态分布，则称 $\{X(t)\}$ 为高斯过程。

注: 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性的有力工具. 但是在实际问题中，要知道随机过程的全部有限维分布是不可能的，因此，人们想到了用随机过程的某些数字特征来刻画随机过程.

定义 2.4. 设 $\{X(t), t \in T\}$ 是一随机过程.

（1）称 $X(t)$ 的期望 $\mu_X(t) = E[X(t)]$ 为过程的**均值函数** (如果存在的话).

（2）如果 $\forall t \in T$, $E[X^2(t)]$ 存在，则称随机过程 $\{X(t), t \in T\}$ 为**二阶矩过程**. 此时，称函数 $\gamma(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))]$, $t_1, t_2 \in T$ 为过程的**协方差函数**；称 $\mathrm{Var}[X(t)] = \gamma(t, t)$ 为过程的**方差函数**.

由 Schwartz 不等式知，二阶矩过程的协方差函数存在。

有些教材还定义如下的相关函数：

$$R(t_1, t_2) = E[X(t_1)X(t_2)] = \gamma(t_1, t_2) + \mu_X(t_1)\mu_X(t_2).$$

协方差函数 $\gamma(s, t)$ 满足如下的性质：

（1）对称性：$\gamma(s, t) = \gamma(t, s)$；

（2）非负定性：对任意正整数 n 和任意 $t_1, \dots, t_n \in T$, 任意实数 $a_1, \dots, a_n$，都有

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(t_i, t_j) \geq 0.$$

例 2.5. $X(t) = X_0 + tV$, $a \leq t \leq b$, 其中 X_0 和 V 是相互独立且服从 $N(0, 1)$ 分布的随机变量. 求均值函数和协方差函数。

解答: 由多元正态分布的性质可知 $X(t)$ 服从正态分布，且 $X(t_1), \dots, X(t_n)$ 也是 n 维正态分布. 所以只要知道它的一阶矩和二阶矩就完全确定了它的分布. 每一条轨道是一条直线，截距和斜率都是随机变量的值。

$$\begin{aligned}
\mu_X(t) &= E[X(t)] = E(X_0 + tV) = EX_0 + tEV = 0, \\
\gamma(t_1, t_2) &= E[X(t_1)X(t_2)] = E[(X_0 + t_1 V)(X_0 + t_2 V)] \\
&= E[X_0^2] + t_1 t_2 E[V^2] = 1 + t_1 t_2.
\end{aligned}$$

高斯过程的有限维分布完全由其均值函数和协方差函数确定；反过来，有如下存在性定理：

定理 2.2 (高斯过程存在定理). *设 $\mu(t)$，$t \in T$ 为实值函数，$\gamma(s, t)$，$s, t \in T$ 为二元实值函数，满足对称性与非负定性条件，则存在高斯过程 $\{X(t), t \in T\}$ 使得 $\{X(t)\}$ 以 $\mu(t)$ 为均值函数，以 $\gamma(s, t)$ 为协方差函数。*

证明略，参见 (谢衷洁 1990) P.5。

定义 2.5. 设 $\{X(t), t \geq 0\}$ 和 $\{Y(t), t \geq 0\}$ 是两个随机过程。若对任意的正整数 n, m 和任意的 $0 \leq t_1 < t_2 < \cdots < t_n$ 和 $0 \leq s_1 < s_2 < \cdots < s_m$，随机向量 $(X(t_1), X(t_2), \dots, X(t_n))$ 与随机向量 $(X(s_1), X(s_2), \dots, X(s_m))$ 相互独立，则称这两个随机过程相互独立。

命题 2.1. 设 $\{X(t), t \geq 0\}$ 和 $\{Y(t), t \geq 0\}$ 是两个随机过程。若对任意的正整数 n 和任意的 $0 \leq t_1 < t_2 < \cdots < t_n$，随机向量 $(X(t_1), X(t_2), \dots, X(t_n))$ 与随机向量 $(Y(t_1), Y(t_2), \dots, Y(t_n))$ 相互独立，则这两个随机过程相互独立。

证明: 对任意的正整数 n, m 和任意的 $0 \leq t_1 < t_2 < \cdots < t_n$ 和 $0 \leq s_1 < s_2 < \cdots < s_m$，设 $\{t_1, t_2, \dots, t_n\} \cup \{s_1, s_2, \dots, s_m\}$ 的各个元素从小到大排列为

$$0 \leq u_1 < u_2 < \cdots < u_k,$$

其中 $k \leq n + m$。则随机向量 $(X(u_1), X(u_2), \dots, X(u_k))$ 与随机向量 $(Y(u_1), Y(u_2), \dots, Y(u_k))$。两个随机向量相互独立，其中的子集也相互独立，故结论成立。

2.3 随机过程的基本类型

2.3.1 平稳过程

定义 2.6. 如果随机过程 $\{X(t), t \in T\}$ 对任意的 $t_1, \dots, t_n \in T$ 和任意的 h(使得 $t_i + h \in T$)，$(X(t_1 + h), \dots, X(t_n + h))$ 与 $(X(t_1), \dots, X(t_n))$ 具有相同的联合分布，记为

$$(X(t_1 + h), \dots, X(t_n + h)) \stackrel{d}{=} (X(t_1), \dots, X(t_n)),$$

则称 $\{X(t), t \in T\}$ 为**严平稳**的.

定义 2.7. 如果 $X(t)$ 是二阶矩过程，并且均值函数 $E[X(t)] = \mu$（不依赖于 t）, 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：

- 对于宽平稳过程，由于 $\gamma(s,t)=\gamma(0,t-s)$, $s,t\in\mathbb{R}$，可记为 $\gamma(t-s)$;
- $\gamma(\tau)$ 为偶函数，且 $\gamma(0)=\mathrm{Var}(X(t))$, $|\gamma(\tau)|\le\gamma(0)$;
- $\gamma(\tau)$ 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

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

当参数 t 仅取整数值 $0,\pm1,\pm2,...$ 或 $0,1,2,...$ 时，称宽平稳过程为宽平稳序列. 在时间序列分析中经常简称宽平稳序列为平稳列。

例 2.6 (白噪声序列)**.** 设 $\{X_n, n=0,1,...\}$ 为一列两两互不相关的随机变量序列，满足 $EX_n=0$, $n=0,1,2,...$, 且

$$E(X_mX_n)=\begin{cases}0 & \text{当}m\neq n,\\ \sigma^2 & \text{当}m=n,\end{cases}$$

则称 $\{X_n\}$ 是白噪声序列，记为 $\mathrm{WN}(0,\sigma^2)$。白噪声序列 $\{X_n\}$ 是宽平稳的. 这是因为协方差函数 $\mathrm{Cov}(X_n,X_m)=E(X_nX_m)$ 只与 $m-n$ 有关.

例 2.7 (线性序列)**.** 设 $\{\varepsilon(n), n\in\mathbb{Z}\}$ 为白噪声列 $\mathrm{WN}(0,\sigma^2)$，实数列 $\{a_j, j\in\mathbb{Z}\}$ 满足 $\sum_{j=-\infty}^{\infty}|a_j|<\infty$，定义

$$X(t)=\sum_{j=-\infty}^{\infty}a_j\varepsilon(t-j),\ t\in\mathbb{Z},$$

则 $\{X(t)\}$ 定义的级数 a.s. 收敛，称 $\{X(t)\}$ 为线性序列。如果条件放宽到 $\{a_j\}$ 平方可和，则定义仍成立，其中的级数是 L^2 收敛。

例 2.8 (AR 模型)**.** 设 $\{\varepsilon_n, n\in\mathbb{Z}\}$ 为白噪声列 $\mathrm{WN}(0,\sigma^2)$，实数 a 满足 $|a|<1$，则存在平稳列 $\{X(t)\}$ 使得

$$X(t)=c+aX(t-1)+\varepsilon_t.$$

称 $\{X(t)\}$ 服从一阶自回归 (AR) 模型。

例 2.9 (MA 模型)**.** 设 $\{\varepsilon_n, n\in\mathbb{Z}\}$ 为白噪声列 $\mathrm{WN}(0,\sigma^2)$，实数 b 满足 $|b|<1$，定义 $\{X(t)\}$ 为

$$X(t)=\mu+\varepsilon_t+b\varepsilon_{t-1},\ t\in\mathbb{Z}.$$

称 $\{X(t)\}$ 服从一阶滑动平均 (MA) 模型。

对于平稳序列，如果我们要从实际数据建立平稳序列的模型，就需要估计其均值和自协方差函数。均值 $\mu = E[X(t)]$ 和自协方差函数 $\gamma(k) = E[(X(t) - \mu)(X(t+k) - \mu)] = \text{Cov}(X(t), X(t+k))$ 都是针对整个概率空间 Ω 的积分，如

$$\mu = E[X(t)] = \int_{\Omega} X(t, \omega) \, dP(\omega).$$

理论上，这需要观察许多条独立的轨道，对这些不同轨道的观测值进行平均。但是，实际中我们一般仅能观测到一条轨道，需要从这一条轨道的有限个观测值估计平稳序列的均值和自协方差函数。如果这样的估计是相合的，就成这个平稳列具有某种“遍历性”。

随机过程中“严平稳遍历”的定义很复杂，参见王梓坤《随机过程通论》第 197–204 页（北京师范大学出版社，1996）。这里给出一个严平稳遍历的充分条件，和严平稳遍历的性质。

定理 2.3. 如果 $\{X(t)\}$ 为线性序列,且其中的白噪声序列独立同分布,则 $\{X(t)\}$ 为严平稳遍历序列。

定理 2.4. 如果 $\{X(t)\}$ 为严平稳遍历序列，则

（*1*）强大数律成立：如果 $E|X(1)| < \infty$，则

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} X(t) = E(X(1)), \ \ a.s.;$$

（*2*）对任意多元函数 $\phi(x_1, \dots, x_m)$, $Y(t) = \phi(X(t), X(t-1), \dots, X(t-m+1))$ 也是严平稳遍历序列。

这两个定理的证明略，定理参见 (何书元 2003) 节 1.5。

对于非严平稳遍历的情形，也有一些略放松的条件可以使得从一条轨道估计的均值函数和自协方差函数相合。

定理 2.5. （*1*）设 $\{X(t), -\infty < t < \infty\}$ 是平稳过程，则当且仅当

$$\lim_{B\to\infty} \frac{1}{B} \int_0^{2B} (1 - \frac{\tau}{2B}) \gamma(\tau) \, d\tau = 0$$

时

$$E\left[\frac{1}{2B} \int_{-B}^{B} X(t) \, dt - E(X(1))\right]^2 \to 0, \ B \to \infty.$$

(2) 设 $\{X_n, n \in \mathbb{Z}\}$ 是平稳序列，其均值函数为 μ，协方差函数为 $\gamma(\tau)$，则当且仅当

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

时

$$E\left[\frac{1}{N}\sum_{t=1}^{N} X(t) - E(X(1))\right]^2 \to 0,\ N \to \infty.$$

证明略，见教材。

推论 2.1. 若 $\int_{-\infty}^{\infty} |\gamma(\tau)|\,d\tau < \infty$，则定理2.5中（1）的条件满足。

证明略，见教材。

推论 2.2. 对于平稳序列，若 $\gamma(\tau) \to 0(\tau \to \infty)$，则定理2.5中（2）的条件满足。

证明略，见教材。

定理 2.6. 设 $\{X(t), t \in \mathbb{R}\}$ 为四阶矩有限的宽平稳过程，均值为零，则当且仅当

$$\lim_{B\to\infty} \frac{1}{B} \int_0^{2B} \left(1 - \frac{s}{2B}\right) [Q(s) - \gamma^2(\tau)]\, ds = 0,\ \forall \tau \in \mathbb{R}$$

时

$$E\left[\frac{1}{2B}\int_{-B}^{B} X(t)X(t+\tau)\,dt - \gamma(\tau)\right]^2 \to 0,\ B \to \infty.$$

其中 $Q(s) = E[X(t+\tau+s)X(t+s)X(t+\tau)X(t)]$。

证明略，参见教材。

2.3.2 独立增量和平稳增量

定义 2.8. 对随机过程 $\{X(t), t \geq 0\}$，如果对任意正整数 n 和 $t_1, t_2, \dots, t_n$，其中 $0 \leq t_1 < t_2 < \cdots < t_n$，随机变量 $X(t_1), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ 是相互独立的，则称 $X(t)$ 为**独立增量过程**. 如果对任意的 $t_1 < t_2$ 和 $h > 0$，$X(t_1+h) - X(t_1)$ 与 $X(t_2+h) - X(t_2)$ 分布相同，则称 $\{X(t), t \in T\}$ 为是**平稳增量过程**. 兼有独立增量和平稳增量的过程称为**平稳独立增量过程**.

对于平稳独立增量过程，设 $X(0) = 0$，$0 \leq t_1 < t_2 < \cdots < t_n$，$h > 0$，则 $(X(t_1), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ 与 $(X(t_1+h), X(t_2+h) - X(t_1+h), \dots, X(t_n+h) - X(t_{n-1}+h)$ 有相同的联合分布。

随机过程 $\{X(t), t \geq 0\}$ 的特征函数为 $\psi_{X(t)}(u) = E[e^{iuX(t)}]$，我们有如下定理:

定理 2.7. 设 $\{X(t), t \geq 0\}$ 是一个独立增量过程，$X(0) = 0$，则 $X(t)$ 具有平稳增量的充分必要条件是：其特征函数具有可乘性，即

$$\psi_{X(t+s)}(u) = \psi_{X(t)}(u)\psi_{X(s)}(u).$$

证明：必要性：注意到

$$X(t+s) = X(s) + [X(t+s) - X(s)]$$

可得。

充分性：由独立增量性有

$$\begin{aligned}
&\psi_{X(t)}(u)\psi_{X(s)}(u) = \psi_{X(t+s)}(u) \\
=&E[e^{iuX(t+s)}] = E[e^{iu\{[X(t+s)-X(s)]+X(s)\}}] \\
=&E\{e^{iu[X(t+s)-X(s)]}\}E[e^{iuX(s)}] \\
=&\psi_{X(t+s)-X(s)}(u)\psi_{X(s)}(u).
\end{aligned}$$

从两边约掉 $\psi_{X(s)}(u)$，得

$$\psi_{X(t)}(u) = \psi_{X(t+s)-X(s)}(u).$$

这说明 $X(t+s) - X(s)$ 与 $X(t) - X(0)$ 同分布，从而有平稳增量。

平稳独立增量过程的均值函数是时间 t 的线性函数。本书后面将要介绍的泊松过程和布朗运动都是这类过程。这两类过程是随机过程理论中的两块最重要的基石。

设 $\{X(t), t \geq 0\}$ 是平稳独立增量过程，$X(0) = 0$, $\mu(t) = E[X(t)]$，$b = \mu(1)$。则

$$\mu(2) = E\{X(2) - X(0)\} = E\{[X(2) - X(1)] + [X(1) - X(0)]\} = 2b.$$

归纳可知 $\mu(k) = kb$, $k = 0, 1, 2, ...$。同理可知对有理数 p/q 有 $\mu(\frac{p}{q}) = b\frac{p}{q}$。若假定 $\mu(t)$ 连续，则有 $\mu(t) = bt$。

Chapter 3

泊松过程

3.1 泊松过程定义

定义 3.1. 随机过程 $\{N(t), t \geq 0\}$ 称为**计数过程**或**点过程**，如果 $N(t)$ 表示从时刻 0 到 t 某一特定事件 A 发生的次数，它具备以下两个特点：

（1）$N(t) \geq 0$ 且取值为整数;

（2）$s < t$ 时，$N(s) \leq N(t)$ 且 $N(t) - N(s)$ 表示 $(s, t]$ 时间内事件 A 发生的次数。

一般都设 $N(0) = 0$。

定义 3.2. 计数过程 $\{N(t), t \geq 0\}$ 称为参数为 $\lambda(\lambda > 0)$ 的泊松过程，如果

（1）$N(0) = 0$;

（2）过程有独立增量;

（3）对任意的 $s, t \geq 0$,

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

注 1：泊松过程是独立平稳增量的计数过程.

注 2：由于 $E[N(t)] = \lambda t$, 于是可认为 λ 是单位时间内发生的事件的平均次数，故一般称 λ 是泊松过程的强度或速率.

例 3.1 (泊松过程在排队论中的应用)**.** 在随机服务系统中排队现象的研究中，经常用到泊松过程模型. 例如, 到达电话总机的呼叫数目，到达某服务设施的顾客数，都可以用泊松过程来描述。以某火车站售票处为例，设从早上 8:00 开始，此售票处连续售票，乘客依 10 人/小时的平均速率到达，则从 9:00 到 10:00 这 1 小时内最多有 5 名乘客来此购票的概率是多少？从 10:00-11:00 没有人来买票的概率是多少？

解：我们用一个泊松过程来描述. 设 8:00 为 0 时刻, 则 9:00 为 1 时刻, 参数 $\lambda = 10$. 由泊松过程的平稳增量性知

$$\begin{aligned} P(N(2)-N(1)\le 5) =& \sum_{n=0}^{5} e^{-10\cdot 1}\frac{(10\cdot 1)^n}{n!}, \\ P(N(3)-N(2)=0) =& e^{-10}\cdot\frac{(10)^0}{0!} = e^{-10}. \end{aligned}$$

例 3.2 (事故发生次数与保险公司接到的索赔数)**.** 若以 $N(t)$ 表示某场所在 $(0,t]$ 时间内发生不幸事故的数目，则泊松过程就是 $\{N(t), t\ge 0\}$ 的一种很好近似. 例如，保险公司接到赔偿请求的次数（设一次事故就导致一次索赔）都是可以应用泊松过程的模型。我们考虑一种最简单情况，设保险公司每次的赔付都是 1，每月平均接到索赔要求 4 次，则一年中它要付出的金额平均为多少？

解：设一年开始为 0 时刻，1 月末为时刻 1，2 月末为时刻 2，……，则年末为时刻 12。

$$P(N(12)-N(0)=n) = \frac{(4\times 12)^n}{n!}e^{-4\times 12}.$$

均值

$$E[N(12)-N(0)] = 4\times 12 = 48.$$

为什么实际中有这么多的现象可以用泊松过程来反映呢？其根据是小概率事件原理. 我们在概率论的学习中已经知道，Bernoulli 试验中，每次试验成功的概率很小而试验的次数很多时，二项分布会逼近 Poisson 分布. 这一想法很自然地推广到随机过程情况. 比如上面提到的事故发生的例子，在很短的时间内发生事故的概率是很小的，但假如考虑很多个这样很短的时间的连接，事故的发生将会有一个大致稳定的速率，这很类似于 Bernoulli 试验以及二项分布逼近 Poisson 分布时的假定.

泊松过程的另一等价定义:

定义 3.3. 设 $\{N(t), t \geq 0\}$ 是一个计数过程，若满足

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

$(2)'$ 过程有平稳独立增量;

$(3)'$ 存在$\lambda > 0$, 当$h \downarrow 0$时

$$P(N(t+h) - N(t) = 1) = \lambda h + o(h);$$

$(4)'$ 当$h \downarrow 0$时

$$P(N(t+h) - N(t) \geq 2) = o(h).$$

则称 $\{N(t), t \geq 0\}$ 为泊松过程.

事实上，把 $[0, t]$ 划分为 n 个等间隔的时间区间，则由条件 $(4)'$ 可知，当 $n \to \infty$ 时，在每个小区间内事件发生两次或两次以上的概率可以近似为 0，因此，事件发生一次的概率 $p \approx \lambda \frac{t}{n}$(显然 p 会很小)，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. 其中事件发生一次即为试验成功，不发生即为失败，再由条件 $(2)'$ 给出的平稳独立增量性，$N(t)$ 就相当于 n 次独立 Bernoulli 试验中试验成功的总次数，由 Poisson 分布的二项分布逼近可知 $N(t)$ 将服从参数为 λt 的 Poisson 分布.

泊松过程两定义等价的严格的数学证明:

定理 3.1. *满足上述条件 $(1)' - (4)'$ 的计数过程 $\{N(t), t \geq 0\}$ 是泊松过程，反过来泊松过程一定满足这 4 个条件.*

证明：先证明泊松过程满足条件 $(1)' \sim (4)'$，只需验证条件 $(3)', (4)'$ 成立. 由定义中第三个条件可得

$$P(N(t+h) - N(t) = 1) = e^{-\lambda h} \frac{\lambda h}{1!},$$

而

$$\lim_{h \to 0} \frac{e^{-\lambda h} \lambda h - \lambda h}{h} = \lim_{h \to 0} \lambda (e^{-\lambda h} - 1) = 0,$$

按 $o(\cdot)$ 定义即有 $e^{-\lambda h}\lambda h = \lambda h + o(h)$，从而 $(3)'$ 成立。对于 $(4)'$，

$$\begin{aligned}
&\lim_{h\to 0}\frac{P(N(t+h)-N(t)\geq 2)}{h}\\
=&\lim_{h\to 0}\frac{1-P[N(t+h)-N(t)=0]-P[N(t+h)-N(t)=1]}{h}\\
=&\lim_{h\to 0}\frac{1-e^{-\lambda h}-e^{-\lambda h}\lambda h}{h}\\
=&\lim_{h\to 0}\frac{1-e^{-\lambda h}}{h}-\lim_{h\to 0}e^{-\lambda h}\lambda\\
=&\lim_{h\to 0}\frac{e^{-\lambda h}\lambda}{1}-\lambda=0,
\end{aligned}$$

即 $(4)'$ 成立。

反过来，设计数过程 $\{N(t), t\geq 0\}$ 满足条件 $(1)'\sim(4)'$, 要证明它是泊松过程. 可以看到, 其实只需验证 $N(t)$ 服从参数为 λt 的泊松分布即可. 记

$$P_n(t) = P(N(t)=n),\ n=0,1,2\ldots$$

则对 $h\to 0$，由 $(3)',(4)'$ 可知

$$P_0(h) = 1-\lambda h + o(h).$$

于是对 $t+h$ 有

$$\begin{aligned}
P_0(t+h) =& P(N(t)=0,\, N(t+h)-N(t)=0)\\
=& P(N(t)=0)\,P(N(t+h)-N(t)=0)\\
=& P_0(t)P_0(h) = P_0(t)[1-\lambda h+o(h)],\\
P_0'(t) =& \lim_{h\to 0}\frac{P_0(t+h)-P_0(t)}{h} = -\lambda P_0(t),\\
\frac{d}{dt}\log P_0(t) =& -\lambda,
\end{aligned}$$

配合 $P_0(0)=1$ 解得 $P_0(t)=e^{-\lambda t}$，$t\geq 0$。注意 $P_0'(t)$ 是按（右）导数定义计算的，所以 $P_0(t)$ 可微。

归纳地，设

$$P_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$$

对 $k=0,1,\ldots,n-1$ 成立，则

$$\begin{aligned}
P_n(t+h) =& P(N(t)=n,\, N(t+h)-N(t)=0) \\
& + P(N(t)=n-1,\, N(t+h)-N(t)=1) \\
& + P(N(t)\le n-2,\, N(t+h)-N(t)\ge 2) \\
=& P_n(t)P_0(h)+P_{n-1}(t)P_1(h) \\
& + P(N(t)\le n-2)P(N(t+h)-N(t)\ge 2) \\
=& P_n(t)[1-\lambda h+o(h)]+P_{n-1}(t)[\lambda h+o(h)]+o(h), \\
P_n'(t) =& \frac{P_n(t+h)-P_n(t)}{h} \\
=& P_n(t)(-\lambda)+P_{n-1}(t)\lambda,
\end{aligned}$$

以归纳法假定的 $P_{n-1}(t)$ 代入，变换得

$$\begin{aligned}
P_n'(t)e^{\lambda t}+P_n(t)\lambda e^{\lambda t} =& \lambda\frac{(\lambda t)^{n-1}}{(n-1)!}, \\
\frac{d}{dt}[P_n(t)e^{\lambda t}] =& \lambda\frac{(\lambda t)^{n-1}}{(n-1)!}, \\
P_n(t)e^{\lambda t} =& c_1+\frac{(\lambda t)^n}{n!}, \\
P_n(t) =& c_1e^{-\lambda t}+e^{-\lambda t}\frac{(\lambda t)^n}{n!},
\end{aligned}$$

利用 $P_n(0)=P(N(0)=n)=0$ 得 $c_1=0$，从而

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

得证。

定理 3.2. *事件 A 的发生形成强度为 λ 的泊松过程 $\{N(t),t\ge 0\}$. 如果每次事件发生时以概率 p 能够被记录下来，并以 $M(t)$ 表示到 t 时刻被记录下来的事件总数，则 $\{M(t),t\ge 0\}$ 是一个强度为 λp 的泊松过程.*

事实上，由于每次事件发生时，对它的记录和不记录都与其他的事件能否被记录独立，而且事件到来服从泊松过程. 所以 $M(t)$ 也是具有平稳独立增量的（这样证明平稳独立增量不严格，严格证明可借助条件期望，参见 (何书元 2008) P.49 定理 4.3 或3.4.1），故只需验证 $M(t)$ 服从均值为 λpt 的泊松分布.

即证明对 $t>0$, 有

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

由于

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

例 3.3. 若每条蚕的产卵数服从泊松分布，强度为 λ，而每个卵变为成虫的概率为 p，且各个卵是否变为成虫彼此间没有关系，求每条蚕养活 k 只小蚕的概率.

解: 由定理3.2的证明方法可知小蚕数服从强度为 λp 的泊松分布，故所求概率为

$$\frac{(\lambda pt)^k}{k!}e^{-\lambda pt}.$$

例 3.4. 观察资料表明，天空中星体数服从泊松分布，其参数为 λV，这里 V 是被观察区域的体积. 若每个星球上有生命存在的概率为 p，则在体积为 V 的宇宙空间中有生命存在的星球数服从参数为 λpV 的泊松分布.

3.2 与泊松过程相联系的若干分布

首先给出泊松过程的有关记号，泊松过程 $\{N(t),t\geq 0\}$ 的一条样本路径一般是跳跃高度为 1 的阶梯函数. $T_n,n=1,2,3\ldots$ 表示第 n 次事件发生的时刻，规定 $T_0=0$. X_n, $n=1,2,\ldots$ 表示第 n 次与第 $n-1$ 次事件发生的时间间隔.

3.2.1 X_n 和 T_n 的分布

定理 3.3. $X_n, n = 1, 2, ...$ 服从参数为 λ 的指数分布，且相互独立.

证明：首先考虑 X_1 的分布，注意到事件 $\{X_1 > t\}$ 等价于事件 $\{N(t) = 0\}$，即 $(0, t]$ 内没有事件发生. 因此

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

说明 X_1 服从指数分布 $\mathrm{Exp}(\lambda)$。

再来看 X_2。

$$\begin{aligned} P(X_2 > t | X_1 = s) =& P(N(s+t) - N(s) = 0 | X_1 = s) \\ =& P(N(s+t) - N(s) = 0) \quad (\text{独立增量性}) \\ =& e^{-\lambda t}. \end{aligned}$$

所以 X_2 与 X_1 独立, 且都服从参数为 λ 的指数分布. 重复同样的推导，可得定理结论.

注 1：定理3.3的结果应该是预料之中的，由于泊松过程有平稳独立增量，过程在任何时刻都“重新开始”，换言之，这恰好就是“无记忆”的体现，与指数分布的“无记忆性”是对应的. 关于指数分布和泊松过程的无记忆性，参见例4.5关于余寿的讨论。

注 2：更严格的证明见 (林元烈 2002) P.39 定理 2.2.1，或 (钱敏平 et al. 2011)。可以先用计算超矩形上概率的方法求 $(T_1, T_2, ..., T_n)$ 的联合密度，由此求 $(X_1, ..., X_n)$ 的联合密度。

定理 3.4. T_n, $n = 1, 2, 3, ...$ 服从参数为 n 和 λ 的 Γ 分布.

证明：由于 $T_n = \sum_{i=1}^n X_i$, 而由定理3.3知道，X_i 是相互独立的且有同指数分布，同时指数分布是 Γ 分布的一种特殊情形（$n = 1$）。由 Γ 分布可加性，易得 T_n 服从参数为 n 和 λ 的 Γ 分布，但我们在这里用另外的方法导出. 注意到

$$N(t) \geq n \iff \quad T_n \leq t$$

即第 n 次事件发生在时刻 t 或之前相当于到时刻 t 已经发生的事件数目至少是 n. 因此

$$P(T_n \leq t) = P(N(t) \geq n) = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}.$$

对第上式两端求导可得 T_n 的密度函数

$$\begin{aligned}f(t)=&-\sum_{j=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^j}{j!}+\sum_{j=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!}\\=&\lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}=\frac{\lambda^n}{\Gamma(n)}t^{n-1}e^{-\lambda t},\ t>0.\end{aligned}$$

利用 X_i 的性质，可以给出泊松过程又一定义方法:

定义 3.4. 计数过程 $\{N(t),t\geq 0\}$ 是参数为 λ 的泊松过程, 如果每次事件发生的时间间隔 $X_1,X_2,...$ 相互独立，且服从同一参数为 λ 的指数分布.

定义的等价性证明见 (林元烈 2002) P.39 定理 2.2.1。

定义3.4提供了对泊松过程进行计算机模拟的方便途径：只需产生 n 个同指数分布的随机数，将其作为 X_i, $i=1,2,...$ 即可得到泊松过程的一条样本路径.

例 3.5. 设从早上 8:00 开始有无穷多的人排队等候服务，只有一名服务员，且每个人接受服务的时间是独立的并服从均值为 20 分钟的指数分布，则到中午 12:00 为止平均有多少人已经离去，已有 9 个人接受服务的概率是多少？

解: 由所设条件可知，离去的人数 $\{N(t)\}$ 是强度为 3 的泊松过程（这里以小时为单位）. 设 8:00 为零时刻，则

$$P(N(4)-N(0)=n)=e^{-12}\frac{(12)^n}{n!}.$$

其均值为 12，即到 12:00 为止，离去的人平均是 12 名. 而有 9 个人接受过服务的概率是

$$P(N(4)=9)=e^{-12}\frac{(12)^9}{9!}.$$

例 3.6. 假定某天文台观测到的流星流是一个泊松过程，根据以往资料统计为每小时平均观察到 3 颗流星. 试求：在上午 8 点到 12 点期间，该天文台没有观察到流星的概率.

解: 设早晨 8 时为 0 时刻，以 $N(t)$ 表示 0 时到 t 时观测到的流星数，则 $\{N(t)\}$ 是强度为 3 的泊松过程，则有

$$N(4)-N(0)\sim \text{Pois}(3\times 4).$$

故在上午 8 点到 12 点期间，该天文台没有观察到流星的概率为

$$P\{N(4)-N(0)=0\}=e^{-12}.$$

3.2.2 事件发生时刻的条件分布

定理 3.5. 在已知 $[0,t]$ 内 A 只发生一次的前提下，A 发生的时刻在 $[0,t]$ 上是均匀分布.

证明：对于 $s\le t$,

$$\begin{aligned}
P(T_1\le s|N(t)=1)&=\frac{P(T_1\le s,N(t)=1)}{P(N(t)=1)}\\
&=\frac{P(\text{A 发生在}s\text{时刻之前 }(s,t]\text{内 A 没有发生})}{P(N(t)=1)}\\
&=\frac{P(N(s)=1)\cdot P(N(t)-N(s)=0)}{P(N(t)=1)}\\
&=\frac{\lambda s e^{-\lambda s}\cdot e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}}=\frac{s}{t}.
\end{aligned}$$

这个定理的逆命题也成立。设事件发生时间间隔 X_n 独立同分布，$\{N(t)\}$ 是由 $\{X_n\}$ 定义的计数过程，设 $P(X_n=0)=0$，则若上述定理的条件分布成立，则 $\{X(t)\}$ 必为泊松过程。见 (林元烈 2002) P.49 定理 2.4.3。

定理 3.6. 在已知 $N(t)=n$ 的条件下，事件发生的 n 个时刻 T_1,T_2, ..., T_n 的联合分布密度是

$$f(t_1,t_2,\dots,t_n)=\frac{n!}{t^n},\ 0<t_1<t_2<\dots<t_n.$$

即 n 个独立同均匀分布 $U(0,\,t)$ 的随机变量的次序统计量的联合分布。

证明：设 $0<t_1<t_2<\dots<t_n<t_{n+1}=t$. 取 h_i 充分小使得 $t_i+h_i<t_{i+1}$, $i=1,2,\dots,n$,

$$\begin{aligned}
&P(t_i<T_i\le t_i+h_i,i=1,2,\dots,n|N(t)=n)\\
=&\frac{P(N(t_i+h_i)-N(t_i)=1,N(t_{i+1})-N(t_i+h_i)=0,1\le i\le n,N(t_1)=0)}{P(N(t)=n)}\\
=&\frac{\lambda h_1 e^{-\lambda h_1}\cdots\lambda h_n e^{-\lambda h_n}e^{-\lambda(t-h_1-h_2-\dots-h_n)}}{e^{-\lambda t}(\lambda t)^n/n!}\\
=&\frac{n!}{t^n}h_1\cdots h_n.
\end{aligned}$$

故由联合密度定义，给定 $N(t)=n$ 时，$(T_1,\ldots,T_n)$ 的 n 维条件分布密度函数

$$\begin{aligned}
&f(t_1,\cdots,t_n)\\
=&\lim_{\substack{h_i\to 0\\ 1\le i\le n}}\frac{P(t_i<T_i\le t_i+h_i,1\le i\le n|N(t)=n)}{h_1h_2\cdots h_n}\\
=&\frac{n!}{t^n},\quad 0<t_1<t_2<\cdots<t_n.
\end{aligned}$$

注: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 $T_1,T_2,\ldots,T_n$（不排序）可看做相互独立的随机变量，且都服从 $[0,t]$ 上的均匀分布. 这也可以作为模拟泊松过程轨道的一种方法。

例 3.7. 乘客按照强度为 λ 的泊松过程来到某火车站，火车在时刻 t 启程，计算在 $(0,t]$ 内到达的乘客等待时间的总和的期望值，即求

$$E\left(\sum_{i=1}^{N(t)}(t-T_i)\right),$$

其中 T_i 是第 i 个乘客来到的时刻.

解: 在 $N(t)$ 给定条件下，取条件期望

$$\begin{aligned}
&E\left[\sum_{i=1}^{N(t)}(t-T_i)|N(t)=n\right]\\
=&E\left[\sum_{i=1}^{n}(t-T_i)|N(t)=n\right]\\
=&nt-E\left[\sum_{i=1}^{n}T_i|N(t)=n\right].
\end{aligned}$$

记 $U_1,U_2,\cdots,U_n$ 为 n 个独立的服从 $(0,t]$ 上的均匀分布的随机变量，由定理3.6

$$E\left[\sum_{i=1}^{n}T_i|N(t)=n\right]=E\left[\sum_{i=1}^{n}U_i\right]=\frac{nt}{2}.$$

从而

$$E\left[\sum_{i=1}^{n}(t-T_i)|N(t)=n\right]=nt-\frac{nt}{2}=\frac{nt}{2}.$$

所以

$$
\begin{aligned}
&E\left[\sum_{i=1}^{N(t)}(t-T_i)\right]\\
=&E\left[E\left(\sum_{i=1}^{N(t)}(t-T_i)|N(t)\right)\right]\\
=&\frac{t}{2}E[N(t)]=\frac{\lambda t^2}{2}.
\end{aligned}
$$

例 3.8. 考虑定理3.2中每次事件发生时被记录到的概率随时间发生变化时的情况，设事件 A 在 s 时刻发生被记录到的概率是 $p(s)$, 若以 $M(t)$ 表示到 t 时刻被记录的事件数，那么它还是泊松过程吗？试给出 $M(t)$ 的分布.

解: 易看出 $M(t)$ 已不能形成一个泊松过程，因为虽然它仍然具有独立增量性，但由于 $p(s)$ 的影响，它已不再有平稳增量性. 但可以证明，对 $\forall t$, $M(t)$ 依然是泊松分布，参数与 t 和 $p(s)$ 有关. 实际上，$M(t)$ 的均值为 $\lambda t\tilde{p}(t)$，其中

$$
\tilde{p}(t)=\frac{1}{t}\int_0^t p(s)\,ds.
$$

事实上，若对 $N(t)$ 给定的条件下取条件期望, 则有

$$
\begin{aligned}
&P(M(t)=m)\\
=&\sum_{k=0}^{\infty}P(M(t)=m|N(t)=m+k)P(N(t)=m+k)\\
=&\sum_{k=0}^{\infty}P(\text{已知}[0,t]\text{中发生了}m+k\text{次事件的条件下只有}m\text{件被记录})\\
&\times P(N(t)=m+k).
\end{aligned}
$$

在发生的 $m+k$ 次事件中，每一次的发生时间可以认为是独立同 U(0,t) 分布的。

对其中的某一次事件，其发生时间 $X\sim \mathrm{U}(0,t)$，用 $Y=1$ 表示此事件被记录，

$Y=0$ 表示不被记录，被记录的概率用全期望公式计算为

$$
\begin{aligned}
\tilde{p}(t)=&P(Y=1)=E(Y)=E[E(Y|X)]\\
=&\int_0^t E(Y|X=x)\frac{1}{t}\,ds\\
=&\int_0^t P(Y=1|X=x)\frac{1}{t}\,ds\\
=&\frac{1}{t}\int_0^t p(s)\,ds.
\end{aligned}
$$

这 $m+k$ 个事件是否被记录是相互独立的，而且每个事件的发生时间不计事件次序可以认为是独立的均匀分布 U(0,t)，所以被记录的概率是相等的，都等于 $\tilde{p}(t)$，于是 $m+k$ 个事件中被记录的个数服从二项分布 B($m+k,\tilde{p}(t)$)，于是

$$
\begin{aligned}
&P(M(t)=m)\\
=&\sum_{k=0}^{\infty}\binom{m+k}{m}[\tilde{p}(t)]^m[1-\tilde{p}(t)]^k\frac{(\lambda t)^{m+k}}{(m+k)!}e^{-\lambda t}\\
=&\sum_{k=0}^{\infty}\frac{(m+k)!}{m!k!}[\tilde{p}(t)]^m[1-\tilde{p}(t)]^k\frac{(\lambda t)^{m+k}}{(m+k)!}e^{-\lambda t}\\
=&\frac{[\lambda t\tilde{p}(t)]^m}{m!}e^{-\lambda t}\sum_{k=0}^{\infty}\frac{[\lambda t(1-\tilde{p}(t))]^k}{k!}\\
=&\frac{[\lambda t\tilde{p}(t)]^m}{m!}e^{-\lambda t}e^{\lambda t(1-\tilde{p}(t))}\\
=&\frac{[\lambda t\tilde{p}(t)]^m}{m!}e^{-\lambda t\tilde{p}(t)},\ m=0,1,...
\end{aligned}
$$

3.3 泊松过程的推广

3.3.1 非齐次泊松过程

当泊松过程的强度 λ 不再是常数，而与时间 t 有关时，泊松过程被推广为非齐次（非时齐）泊松过程. 一般来说, 非齐次泊松过程是不具备平稳增量的 (见例3.8). 在实际中，非齐次泊松过程也是比较常用的. 例如在考虑设备的故障率时，由于设备使用年限的变化, 出故障的可能性会随之变化；放射性物质的衰变速度，会因各种外部条件的变化而随之不同；昆虫产卵的平均数量随年龄和季

节而变化等. 在这样的情况下，再用齐次泊松过程来描述就不合适了，于是改用非齐次的泊松过程来处理.

定义 3.5. 计数过程 $\{N(t), t \geq 0\}$ 称做强度函数为 $\lambda(t) > 0$ $(t \geq 0)$ 的非齐次泊松过程，如果

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

(2) 过程有独立增量;

(3) $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$;

(4) $P(N(t+h) - N(t) \geq 2) = o(h)$.

类似于泊松过程，非齐次泊松过程有如下的等价定义.

定义 3.6. 计数过程 $\{N(t), t \geq 0\}$ 称为强度函数为 $\lambda(t) > 0 (t \geq 0)$ 的非齐次泊松过程, 若

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

(2) 过程有独立增量;

(3) 对任意实数$t \geq 0, s \geq 0, N(t+s) - N(t)$服从参数为

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

的泊松分布.

注：$m(t) = \int_0^t \lambda(s)\, ds$. 对齐次（时齐）的泊松过程，$m(t) = \lambda \cdot t$，不论时齐与否都有 $m(t) = E[N(t)]$，即 $m(t)$ 是 $(0, t]$ 时间段发生的事件个数的期望值。

泊松过程与非齐次泊松过程之间转换关系:

定理 3.7. *设 $\lambda(t) > 0$, $\forall t \geq 0$，$m(t) = \int_0^t \lambda(u)\, du$。若 $\{N(t), t \geq 0\}$ 是一个强度函数为 $\lambda(t)$ 的非齐次泊松过程. 令 $N^*(t) = N(m^{-1}(t))$，则 $\{N^*(t)\}$ 是一个强度参数为 1 的泊松过程。反之，若 $\{N^*(t)\}$ 是一个强度参数为 1 的泊松过程，则 $N(t) = N^*(m(t))$ 是一个强度函数为 $\lambda(t)$ 的非齐次泊松过程。*

证明: 用定义3.2验证 $\{N^*(t)\}$ 是一个强度参数为 1 的泊松过程. 显然 $N^*(0) = 0$, 且具有独立增量，只要证明 $N^*(t+s) - N^*(s)$ 服从参数为 t 的泊松分布。事

实上

$$\begin{aligned}
&N^*(t+s) - N^*(s) \\
=&N(m^{-1}(t+s)) - N(m^{-1}(s)) \\
\sim&\text{Pois}[m(m^{-1}(t+s)) - m(m^{-1}(s))] \\
=&\text{Pois}(t).
\end{aligned}$$

反过来，如果 $\{N^*(t)\}$ 是一个强度参数为 1 的泊松过程，$N(t) = N^*(m(t))$，则 $N(0) = 0$，$\{N(t)\}$ 仍有独立增量，

$$\begin{aligned}
&N(t+s) - N(s) \\
=&N^*(m(t+s)) - N^*(m(s)) \\
\sim&\text{Pois}[m(t+s) - m(s)],
\end{aligned}$$

由定义3.6可知 $\{N(t)\}$ 是强度函数为 $\lambda(t)$ 的非齐次泊松过程。

注 1: 用此定理可以简化非齐次泊松过程的问题到泊松过程中进行讨论. 也从一个参数为 1 的齐次泊松过程构造一个强度函数为 $\lambda(t)$ 的非齐次泊松过程.

注 2：若 $m(t)$ 非严格单调增，比如在某一区间 $\lambda(t) = 0$ 的情形，可以取

$$m^{-1}(u) = \inf\{t : t > 0, m(t) \geq u\},\ u > 0,$$

这时 $u \leq m(t) \iff t \geq m^{-1}(u)$。

例 3.9. 设某设备的使用期限为 10 年，在前 5 年内它平均 2.5 需要维修一次，后 5 年平均 2 年需维修一次。试求它在使用期内只维修过一次的概率.

解：用非齐次泊松过程考虑，强度函数

$$\lambda(t) = \begin{cases} \frac{1}{2.5}, & 0 \leq t \leq 5, \\ \frac{1}{2}, & 5 < t \leq 10, \end{cases}$$

有

$$m(10) = \int_0^{10} \lambda(t)dt = \int_0^5 \frac{1}{2.5}dt + \int_5^{10} \frac{1}{2}dt = 4.5$$

因此

$$P(N(10) - N(0) = 1) = e^{-4.5}\frac{(4.5)^1}{1!} = \frac{9}{2}e^{-\frac{9}{2}}.$$

3.3.2 复合泊松过程

定义 3.7. 称随机过程 $\{X(t), t \geq 0\}$ 为复合泊松过程, 如果对于 $t \geq 0$，$X(t)$ 可以表示为

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

其中 $\{N(t), t \geq 0\}$ 是一个泊松过程，Y_i, $i = 1, 2, ...$ 是一族独立同分布的随机变量，并且与 $\{N(t), t \geq 0\}$ 也是独立的.

注：复合泊松过程不一定是计数过程，但是当 $Y_i \equiv c$, $i = 1, 2, ...$, c 为常数时, 可化为泊松过程. 定理3.2可以看成是 $Y_i \sim \mathrm{b}(1, p)$ 的复合泊松过程，但恰好构成了参数为 λp 的泊松过程。

例 3.10. 保险公司接到的索赔次数服从一个泊松过程 $\{N(t)\}$，每次要求赔付的金额 Y_i 都相互独立，且有同分布 F, 每次的索赔数额与它发生的时刻无关，则 $[0, t]$ 时间区间内保险公司需要赔付的总金额 $\{X(t)\}$ 就是一个复合泊松过程，

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

例 3.11 (顾客成批到达的排队系统)**.** 设顾客到达某服务系统的时间 $S_1, S_2, ...$ 形成一强度为 λ 的泊松过程，在每个时刻 S_n 可以同时有多名顾客到达. Y_n 表示在时刻 S_n 到达的顾客人数，假定 Y_n, $n = 1, 2, ...$ 相互独立，并且与 $\{S_n\}$ 也独立，则在 $[0, t]$ 时间区间内到达服务系统的顾客总人数也可用一复合泊松过程来描述.

例 3.12. 假设顾客按照参数为 λ 的泊松过程进入一个商店，又假设各顾客所花费的金额形成一族独立同分布的随机变量。以 $X(t)$ 记到时间 t 为止顾客在此商店所花费的总额，易知 $\{X(t), t \geq 0\}$ 是一个复合泊松过程。

定理 3.8. 设 $\{X(t) = \sum_{i=1}^{N(t)} Y_i, t \geq 0\}$ 是一复合泊松过程，泊松过程 $\{N(t), t \geq 0\}$ 的强度为 λ, 则

$$\begin{aligned}
&(1)\ X(t)\text{有独立增量;}\\
&(2)\ E[X(t)] = \lambda t \cdot E(Y_1);\\
&\quad\ Var[X(t)] = \lambda t \cdot E(Y_1^2).
\end{aligned}$$

证明：（1）令 $0 \le t_0 < t_1 < t_2 < \cdots < t_n$，则

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

由过程的独立增量性及各 $Y_i (i = 1, 2, \dots, n)$ 之间的独立性不难得出 $X(t)$ 的独立增量性.

（2）用全期望公式，

$$E[X(t)] = E[E(X(t)|N(t))],$$

其中

$$E(X(t)|N(t) = n) = E\left(\sum_{i=1}^{n} Y_i | N(t) = n\right)$$
$$= E\left(\sum_{i=1}^{n} Y_i\right) = nE(Y_1),$$

故

$$E[X(t)] = E[N(t)E(Y_1)] = E[N(t)]E(Y_1).$$

对于 $\mathrm{Var}(X(t))$，要利用条件方差公式

$$\mathrm{Var(X(t))} = E\left\{\mathrm{Var}\left[X(t)|N(t)\right]\right\} + \mathrm{Var}\left\{E\left[X(t)|N(t)\right]\right\},$$

其中

$$\begin{aligned}
&\mathrm{Var}\left[X(t)|N(t) = n\right] \\
=&E\left\{\left[X(t) - E(X(t)|N(t) = n)\right]^2 | N(t) = n\right\} \\
=&E\left\{\left[\sum_{i=1}^{n} Y_i - nE(Y_1)\right]^2 | N(t) = n\right\} \\
=&E\left\{\left[\sum_{i=1}^{n} Y_i - nE(Y_1)\right]^2\right\} \\
=&E\left\{\left[\sum_{i=1}^{n} (Y_i - EY_i)\right]^2\right\} \\
=&\mathrm{Var}\left[\sum_{i=1}^{n} (Y_i - EY_i)\right] = n\mathrm{Var}(Y_1),
\end{aligned}$$

于是

$$\begin{aligned}\mathrm{Var}(\mathrm{X(t)}) =&E[N(t)\mathrm{Var}(Y_1)] + \mathrm{Var}(N(t)E(Y_1))\\ =&\lambda t\mathrm{Var}(Y_1) + \lambda t[E(Y_1)]^2 = \lambda t E(Y_1^2).\end{aligned}$$

例 3.13. 在保险中的索赔模型中，设保险公司接到的索赔要求是强度为每个月两次的泊松过程。每次赔付服从均值为 10000 元的正态分布，则一年中保险公司平均的赔付额是多少?

解：由定理3.8

$$E[X(12)] = 2\times 12\times 10000 = 240000(\text{元}).$$

例 3.14. 设顾客以每分钟 6 人的平均速率进入某商场，这一过程可以用泊松过程来描述. 又设进入该商场的每位顾客买东西的概率为 0.9，且每位顾客是否买东西互不影响，也与进入该商场的顾客数无关，求一天 (12 小时) 在该商场买东西的顾客数的分布与均值.

解：由定理3.2，这是一个复合泊松过程也是泊松过程，速率参数为 0.9×6，所以一天 (12 小时，即 720 分钟) 在该商场买东西的顾客数的分布为均值 $0.9\times 6\times 720 = 3888$ 的泊松分布，

注：这个例子中若以 Z_i 表示进入该商场的第 i 位顾客在该商场所花的钱数 (单位: 元)，且有 $Z_i \sim B(200, 0.5)$，则

$$\tilde{X}(t) = \sum_{i=1}^{N_1(t)} Z_i$$

表示在时间 $(0, t]$ 内该商场的营业额，则该商场一天的平均营业额为

$$E[\tilde{X}(720)] = E[Z_1]E[N(720)] = (200\times 0.5)\times(6\times 720) = 432000(\text{元}).$$

3.3.3 条件泊松过程

条件泊松过程，有些像是统计学中的混合分布。设随机变量 I 取值于 $\{1, 2\}$，$\alpha = P(I = 1)$, 随机变量 $X_1 \sim f_1(x)$, $X_2 \sim f_2(x)$，随机变量

$$Y = \begin{cases} X_1, & \text{当} I = 1,\\ X_2, & \text{当} I = 2,\end{cases}$$

则 Y 的分布密度为

$$f(x) = \alpha f_1(x) + (1-\alpha) f_2(x),$$

称这样的分布为混合分布。

泊松过程描述的是一个有着“风险”参数 λ 的个体发生某一事件的频率，如果我们考虑一个总体，其中的个体存在差异，比如发生事故的倾向性因人而异，这时我们可以把 $N(t)$ 所服从的泊松分布解释为给定 λ 时，$N(t)$ 的条件分布.

定义 3.8. 设随机变量 $\Lambda > 0$，在 $\Lambda = \lambda$ 的条件下，计数过程 $\{N(t), t \geq 0\}$ 是参数为 λ 的泊松过程. 则称 $\{N(t), t \geq 0\}$ 为条件泊松过程.

命题 3.1. 设 Λ 的分布是 $G(\cdot)$，那么在长度为 t 的时间区间内发生 n 次事件的概率为

$$P(N(t+s) - N(s) = n) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} \, dG(\lambda).$$

证明： 利用全期望公式。记 I_A 为事件 A 的示性函数，则 $E(I_A) = P(A)$，

$$\begin{aligned}
& P(N(t+s) - N(s) = n) \\
=& E\left[I_{\{N(t+s)-N(s)=n\}}\right] \\
=& E\left\{E\left[I_{\{N(t+s)-N(s)=n\}} | \Lambda\right]\right\} \\
=& \int_0^\infty E\left[I_{\{N(t+s)-N(s)=n\}} | \Lambda = \lambda\right] dG(\lambda) \\
=& \int_0^\infty P\left[N(t+s) - N(s) = n | \Lambda = \lambda\right] dG(\lambda) \\
=& \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} \, dG(\lambda).
\end{aligned}$$

定理 3.9. 设 $\{N(t), t \geq 0\}$ 是条件泊松过程，且 $E(\Lambda^2) < \infty$，则

$$\begin{aligned}
&(1)\ E[N(t)] = tE(\Lambda); \\
&(2)\ Var[N(t)] = t^2 Var(\Lambda) + tE(\Lambda).
\end{aligned}$$

证明：（1）

$$E[N(t)] = E[E(N(t)|\Lambda)] = E(t\Lambda) = tE(\Lambda).$$

（2）

$$\begin{aligned}\mathrm{Var}(N(t)) =& E[N^2(t)] - [E(N(t))]^2\\ =& E[\Lambda t + \Lambda^2 t^2] - [tE(\Lambda)]^2\\ =& tE(\Lambda) + t^2 E(\Lambda^2) - t^2[E(\Lambda)]^2\\ =& tE(\Lambda) + t^2 \mathrm{Var}(\Lambda).\end{aligned}$$

例 3.15. 设意外事故的发生频率受某种未知因素影响有两种可能 λ_1, λ_2，且 $P(\Lambda = \lambda_1) = p$, $P(\Lambda = \lambda_2) = 1 - p = q$, $0 < p < 1$ 为已知. 已知到时刻 t 已发生了 n 次事故. 求下一次事故在 $t + s$ 之前不会到来的概率. 另外，这个发生频率为 λ_1 的后验概率是多少？

解: 关于第一个问题，有

$$\begin{aligned}& P((t, t+s]\text{内没有事故发生}|N(t) = n)\\ =& \frac{P((t, t+s]\text{内没有事故发生},\ N(t) = n)}{P(N(t) = n)}\\ =& \frac{P(N(t+s) - N(t) = 0,\ N(t) = n)}{P(N(t) = n)}\\ =& \frac{\sum_{i=1}^{2} P(\Lambda = \lambda_i)\ P(N(t+s) - N(t) = 0,\ N(t) = n|\Lambda = \lambda_i)}{\sum_{i=1}^{2} P(\Lambda = \lambda_i)\ P(N(t) = n|\Lambda = \lambda_i)}\\ =& \frac{\sum_{i=1}^{2} P(\Lambda = \lambda_i)\ \frac{(\lambda_i t)^n}{n!} e^{-\lambda_i t} e^{-\lambda_i s}}{\sum_{i=1}^{2} P(\Lambda = \lambda_i)\ \frac{(\lambda_i t)^n}{n!} e^{-\lambda_i t}}\\ =& \frac{p\lambda_1^n e^{-\lambda_1(t+s)} + (1-p)\lambda_2^n e^{-\lambda_2(t+s)}}{p\lambda_1^n e^{-\lambda_1 t} + (1-p)\lambda_2^n e^{-\lambda_2 t}}.\end{aligned}$$

关于第二个问题，有

$$\begin{aligned}& P(\Lambda = \lambda_1|N(t) = n)\\ =& \frac{P(N(t) = n, \Lambda = \lambda_1)}{P(N(t) = n)}\\ =& \frac{P(\Lambda = \lambda_1)P(N(t) = n|\Lambda = \lambda_1)}{P(\Lambda = \lambda_1)P(N(t) = n|\Lambda = \lambda_1) + P(\Lambda = \lambda_2)P(N(t) = n|\Lambda = \lambda_2)}\\ =& \frac{p\frac{(\lambda_1 t)^n}{n!} e^{-\lambda_1 t}}{p\frac{(\lambda_1 t)^n}{n!} e^{-\lambda_1 t} + (1-p)\frac{(\lambda_2 t)^n}{n!} e^{-\lambda_2 t}}\\ =& \frac{p\lambda_1^n e^{-\lambda_1 t}}{p\lambda_1^n e^{-\lambda_1 t} + (1-p)\lambda_2^n e^{-\lambda_2 t}}.\end{aligned}$$

3.4 补充

3.4.1 泊松过程合并与分解

定理 3.10. 设 $\{N_1(t), t \geq 0\}$ 是参数为 λ 的泊松过程，$\{N_2(t), t \geq 0\}$ 是参数为 μ 的泊松过程，且这两个泊松过程独立，令 $N(t) = N_1(t) + N_2(t)$，则 $\{N(t), t \geq 0\}$ 是参数为 $\lambda + \mu$ 的泊松过程。

证明：易见 $N(0) = 0$，独立增量性和平稳增量性从 N_1 和 N_2 的平稳独立增量性可得，而 $N_1(t) + N_2(t)$ 服从泊松 $\text{Pois}(\lambda + \mu)$ 分布，结论成立。

上述定理可以推广到多个独立的泊松过程的和。

关于泊松过程的分解，先给出定理3.2的严格证明。参见 (何书元 2008) P.49 定理 4.3。

用 Y_i 表示第 i 个事件是否被记录的两点分布随机变量，$\{Y_i\}$ 相互独立且与 $\{N(t)\}$ 相互独立，

$$M(t) = \sum_{j=1}^{N(t)} Y_j,\ t \geq 0.$$

来证明 $\{M(t), t \geq 0\}$ 是参数为 $p\lambda$ 的泊松过程。

对于 $0 \leq s < t$，

$$M(t) - M(s) = \sum_{j=N(s)+1}^{N(t)} Y_j$$

是 $(s, t]$ 内事件的到达次数。由独立性，对 $k \geq l$ 可计算如下的条件概率：

$$\begin{aligned}
& P(M(t) - M(s) = n \,|\, N(s) = l, N(t) = k) \\
=& P(\sum_{j=l+1}^{k} Y_j = n \,|\, N(s) = l, N(t) = k) \\
=& P(\sum_{j=l+1}^{k} Y_j = n) \\
=& \binom{k-l}{n} p^n q^{k-l-n} \\
\triangleq& g(k-l, n).
\end{aligned}$$

$g(k,n)$ 是 $\mathrm{B}(k,p)$ 随机变量等于 n 的概率。

于是，

$$E[I_{\{M(t)-M(s)=n\}} \,|\, N(s), N(t)] = g(N(t)-N(s), n).$$

两边取期望得

$$P(M(t)-M(s)=n) = E[g(N(t)-N(s), n)].$$

显然 $M(0)=0$，由(3.4.1)和 $\{N(t)\}$ 的平稳增量性可知 $\{M(t)\}$ 为平稳增量过程。

来证明独立增量性。对任意正整数 m 和 $0=t_0 \le t_1 < t_2 < \cdots < t_m$ 和整数 $0=n_0 \le n_1 \le n_2 \le \cdots \le n_m$，定义

$$N = (N(t_1), N(t_2), \dots, N(t_m)),\ n = (n_1, n_2, \dots, n_m).$$

由(3.4.1)可知，在 $N=n$ 条件下，以下各个随机变量

$$M(t_j) - M(t_{j-1}) = \sum_{i=n_{j-1}}^{n_j} Y_i,\ j=1,2,\dots,m$$

相互独立，并且与 N 独立。于是

$$\begin{aligned}
& P(M(t_j)-M(t_{j-1})=k_j,\ 1\le j\le m \,|\, N=n) \\
=& P\left(\sum_{i=n_{j-1}}^{n_j} Y_i = k_j,\ 1\le j\le m \,|\, N=n\right) \\
=& P\left(\sum_{i=n_{j-1}}^{n_j} Y_i = k_j,\ 1\le j\le m\right) \\
=& \prod_{j=1}^{m} P\left(\sum_{i=n_{j-1}}^{n_j} Y_i = k_j\right) \\
=& \prod_{j=1}^{m} g(n_j - n_{j-1}, k_j).
\end{aligned}$$

于是

$$\begin{aligned}
& E[I_{\{M(t_j)-M(t_{j-1})=k_j,\ 1\le j\le m\}} \,|\, N] \\
=& \prod_{j=1}^{m} g(N(t_j)-N(t_{j-1}), k_j).
\end{aligned}$$

对上式求无条件期望。由 $\{N(t)\}$ 的独立增量性有

$$\begin{aligned}&P(M(t_j)-M(t_{j-1})=k_j,\ 1\le j)\\=&\prod_{j=1}^{m}E[g(N(t_j)-N(t_{j-1}),k_j)].\end{aligned}$$

这说明增量 $M(t_1)-M(t_0), M(t_2)-M(t_1),\dots,M(t_n)-M(t_{n-1})$ 相互独立，独立增量性得证。

最后只要证明 $M(t)$ 服从参数为 $p\lambda t$ 的泊松分布。由(3.4.1)，

$$\begin{aligned}&P(M(t)=n)\\=&E[g(N(t),n)]\\=&\sum_{k=n}^{\infty}g(k,n)P(N(t)=k)\\=&\sum_{k=n}^{\infty}\binom{k}{n}p^n q^{k-n}\frac{\lambda^k}{k!}e^{-\lambda}\\=&\frac{(p\lambda)^n}{n!}e^{-p\lambda},\ n=0,1,\dots.\end{aligned}$$

证毕。

定理 3.11. 设 $\{N(t)\}$ 为参数 λ 的泊松过程，$\{Y_j\}$ 独立同 $b(1,p)$ 分布，且与 $\{N(t)\}$ 独立。令

$$N_1(t)=\sum_{j=1}^{N(t)}Y_j,\ N_2(t)=\sum_{j=1}^{N(t)}(1-Y_j),$$

则 $\{N_1(t)\}$ 是参数为 $p\lambda$ 的泊松过程，$\{N_2(t)\}$ 是参数为 $(1-p)\lambda$ 的泊松过程，且 $\{N_1(t)\}$ 和 $\{N_2(t)\}$ 相互独立。

证明：前面已证明 $\{N_1(t)\}$ 是参数为 $p\lambda$ 的泊松过程，$\{N_2(t)\}$ 同理可证。下面用命题2.1证明这两个过程相互独立。

对任意正整数 n 和 $0\le t_1<t_2<\dots<t_n$，对于整数

$$\begin{aligned}&0=k_0\le k_1\le k_2\le\dots\le k_n,\\&0=m_0\le m_1\le m_2\le\dots\le m_n,\end{aligned}$$

记 $n_j = k_j + m_j$, $n = (n_1, n_2, \dots, n_n)$。如下的 n 个随机变量

$$\xi_j = \sum_{i=n_{j-1}+1}^{n_j} Y_i,\ j = 1, 2, \dots$$

相互独立，并且与泊松过程 $\{N(t)\}$ 独立。注意 $\xi_j \sim \mathrm{B}(n_j - n_{j-1}, p)$，于是

$$\begin{aligned}
&P(N_1(t_j) = k_j, N_2(t_j) = m_j, 1 \le j \le n) \\
=&P(N_1(t_j) = k_j, N(t_j) = n_j, 1 \le j \le n) \\
=&P(N_1(t_j) - N_1(t_{j-1}) = k_j - k_{j-1}, N(t_j) = n_j, 1 \le j \le n) \\
=&P(\xi_j = k_j - k_{j-1}, N(t_j) - N(t_{j-1}) = n_j - n_{j-1}, 1 \le j \le n) \\
=&P(\xi_j = k_j - k_{j-1}, 1 \le j \le n) \\
&P(N(t_j) - N(t_{j-1}) = n_j - n_{j-1}, 1 \le j \le n) \\
=&\prod_{j=1}^n P(\xi_j = k_j - k_{j-1}) P(N(t_j) - N(t_{j-1}) = n_j - n_{j-1}) \\
=&\prod_{j=1}^n \frac{(n_j - n_{j-1})!}{(k_j - k_{j-1})!(m_j - m_{j-1})!} p^{k_j - k_{j-1}} q^{m_j - m_{j-1}} \\
&\cdot \frac{[\lambda(t_j - t_{j-1})]^{n_j - n_{j-1}}}{(n_j - n_{j-1})!} e^{-\lambda(t_j - t_{j-1})} \\
=&\prod_{j=1}^n \frac{[p\lambda(t_j - t_{j-1})]^{k_j - k_{j-1}}}{(k_j - k_{j-1})!} \\
&\prod_{j=1}^n \frac{[q\lambda(t_j - t_{j-1})]^{m_j - m_{j-1}}}{(m_j - m_{j-1})!}.
\end{aligned}$$

所以两个随机过程独立。

上述泊松过程分解定理可以推广到分解成多个相互独立的泊松过程：

推论 3.1. 设 $\{N(t)\}$ 为参数 λ 的泊松过程，$p_j > 0$, $j = 1, 2, \dots, m$ 且 $p_1 + p_2 + \cdots + p_m = 1$。则存在 m 个相互独立的泊松过程 $\{N_j(t), t \ge 0\}$，使得

$$N(t) = \sum_{j=1}^m N_j(t).$$

3.4.2 剩余寿命和年龄

对齐次泊松过程 $\{N(t)\}$，T_n 表示第 n 次事件的发生时刻，$N(t)$ 表示截止到 t 为止已发生的事件个数，则 $T_{N(t)}$ 是 $(0,t]$ 时间段内最后一次事件的时间；$T_{N(t)+1}$ 是 (t,∞) 时间段内第一次事件的时间，$t \in [T_{N(t)}, T_{N(t)}+1)$。将每次事件的发生看成是更换新的设备，并且 $t=0$ 时刚刚更换过设备，令

$$\begin{aligned} W(t) =& T_{N(t)+1} - t, \\ V(t) =& t - T_{N(t)}, \end{aligned}$$

则 $W(t)$ 是 t 时刻正在使用的设备的剩余寿命，$V(t)$ 是 t 时刻正在使用的设备的年龄（已经使用的时间），$X_n = V(t) + W(t)$ 是第 $n-1$ 次事件到第 n 次事件之间的间隔时间，服从参数为 λ 的指数分布且各 X_n 相互独立。讨论 $W(t)$ 和 $V(t)$ 的分布。

命题 3.2. （*1*）$W(t)$ 与 X_n 同分布。

（*2*）$V(t)$ 服从截尾的指数分布：

$$P(V(t) \le x) = \begin{cases} 1 - e^{-\lambda x}, & 0 \le x < t, \\ 1, & x \ge t. \end{cases}$$

证明：（1）

$$W(t) > x \iff N(t+x) - N(t) = 0,$$

故

$$\begin{aligned} & P(W(t) > x) = P(N(t+x) - N(t) = 0) \\ =& P(N(x) = 0) = e^{-\lambda x},\ x > 0 \end{aligned}$$

而 $x < t$ 时

$$V(t) > x \iff N(t) - N(t-x) = 0,$$

$x \ge t$ 时必然有 $V(t) \le x$，所以可得 $V(t)$ 的分布。注意 $V(t)$ 是一个混合分布，分布函数在 $x = t$ 处有一个跳跃点。

定理 3.12. 对计数过程 $\{N(t), t \ge 0\}$，$N(0) = 0$，X_n 表示事件之间的间隔时间，各 X_n 相互独立同分布，分布函数为 $F(x)$，$F(0) = 0$，若对任意 $t > 0$，余寿 $W(t)$ 与 X_1 同分布，则 $\{N(t), t \ge 0\}$ 是泊松过程。

见 (林元烈 2002) P.46 定理 2.3.3。

Chapter 4

更新过程

4.1 更新过程的定义及若干分布

4.1.1 定义

定义 4.1. 设 $\{X_n, n = 1, 2, ...\}$ 是一列独立同分布的非负随机变量, 分布函数为 $F(x)$。为了避免平凡的情况，设 $F(0) = P(X_n = 0) < 1$，记 $\mu = E[X_n]$, 则 $0 < \mu \leq +\infty$. 令

$$T_n = \sum_{i=1}^{n} X_i,\ n \geq 1,\ T_0 = 0.$$

我们把由

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

定义的计数过程称为**更新过程**.

更新过程的一个典型例子是机器零件的更换. 在 0 时刻，安装上一个新零件并开始运行，设此零件在 X_1 时刻损坏，马上用一个新的来替换 (假定替换不需要时间)，则第二个零件在 X_1 时刻开始运行，设它在经过 X_2 时间后损坏，同样马上换第三个，……. 很自然可以认为这些零件的使用寿命是独立同分布的，那么到 t 时刻为止所更换的零件数目就构成一个更新过程.

注：在更新过程中我们将事件发生一次叫做一次更新，X_n 就是第 $n-1$ 次和第 n 次更新相距的时间，T_n 是第 n 次更新发生的时刻，而 $N(t)$ 就是 t 时刻

之前发生的总的更新次数.

更新过程 $\{N(t), t \geq 0\}$ 是泊松过程的推广，其事件（更新）发生的时间间隔仍为独立同分布但不要求是指数分布。

注：从更新过程定义看出，如果将某次更新时刻 T_k 作为新的时间原点，则此后的计数过程仍是相同的更新过程。即：

$$N^*(t) = N(T_k + t) - N(T_k),\ t \geq 0,$$

是与 $\{N(t), t \geq 0\}$ 分布相同的更新过程。若已知 $T_k = t_0$，则 $N^*(t) = N(t_0 + t) - N(t_0)$ 是与 $\{N(t), t \geq 0\}$ 分布相同的更新过程，且与 $\{N(t), t \in [0, t_0]\}$ 独立。

更新过程不再是平稳独立增量过程。举例说明。

例 4.1. 设 $X_i \sim U(1, 2)$，则两次更新之间的时间间隔至少为 1，至多为 2。这时

$$\begin{aligned} P(N(1) - N(0) = 1) = 0, \\ P(N(2) - N(1) = 1) = 1, \end{aligned}$$

所以不是平稳增量的。

考虑 $N(1.1) - N(1)$ 与 $N(2) - N(1.1)$ 是否独立。当 $N(1.1) - N(1) = 1$ 时，第一次更新 T_1 发生于 $(1, 1.1]$, $T_2 = T_1 + X_2 \in (2, 3.1]$，所以

$$P(N(2) - N(1.1) = 1 \,|\, N(1.1) - N(1) = 1) = 0;$$

当 $N(1.1) - N(1) = 0$ 时，第一次更新 T_1 发生于 $(1.1, 2]$,

$$P(N(2) - N(1.1) = 1 \,|\, N(1.1) - N(1) = 1) = 1,$$

这说明 $N(1.1) - N(1)$ 与 $N(2) - N(1.1)$ 不独立，即非独立增量。

4.1.2 性质

命题 4.1.

$$N(t) \geq n \Longleftrightarrow T_n \leq t.$$

下面的命题说明有限时间内不会有无穷次更新发生，从而 $N(t)$ 是普通的取非负整数值的随机变量。

命题 4.2.

$$P(N(t) < \infty) = 1.$$

证明：令 $\mu = EX_1$，因为 $X_1 \geq 0$ 且 $P(X_1 = 0) < 1$ 所以 $0 < \mu \leq +\infty$。

如果 $0 < \mu < \infty$，由强大数律可知

$$\frac{1}{n}T_n = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu, \text{ a.s.}$$

对任意给定的 $t > 0$，以概率 1 地对每一条轨道，$T_n \to \infty$, 存在 n_0 使得 $n \geq n_0$ 时 $T_n > t$, 即 t 之前仅有有限个更新发生，即以概率 1 地成立 $N(t) < \infty$。

如果 $\mu = +\infty$，必存在 $t_0 > 0$ 使得 $P(X_n \leq t_0) > 0$，令 $Y_n = X_n I[X_n \leq t_0]$，则 $Y_n \leq X_n$ 且 $0 < EY_n \leq t_0 < \infty$，以 $\{Y_i\}$ 为更新间隔时间的更新过程的更新次数以概率 1 为有限值，而以 $\{X_i\}$ 为更新间隔时间的更新过程比 $\{Y_i\}$ 对应的更新过程的每次更新发生更晚，所以更新次数以概率 1 为有限值。

前面的命题说明 $N(t)$ 是一个取非负整数值的随机变量。考虑 $N(t)$ 的分布。更新过程 $N(t)$ 分布第一个公式：

命题 4.3.

$$P(N(t) = n) = F_n(t) - F_{n+1}(t),\ t \geq 0, n = 0, 1, 2, ...$$

其中 $F_n(\cdot)$ 是 T_n 的分布函数，为 $F(\cdot)$ 的 n 重卷积，$F(\cdot)$ 是 X_1 的分布函数。

证明：

$$N(t) \geq n \iff T_n \leq t,$$

所以

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

更新过程边缘分布的第二个公式（来自 (Ross 2019) P.434）：

命题 4.4.

$$P(N(t)=n)=\int_0^t \overline{F}(t-y)\,dF_n(y)=\int_0^t \overline{F}(t-y)f_n(y)\,dy.$$

其中 $\overline{F}(x)=1-F(x)$，$f_n(\cdot)$ 是 T_n 的分布密度函数（如果存在）。

证明: 以 T_n 为条件求 $I[N(t)=n]$ 的条件期望。

$$\begin{aligned}
&P(N(t)=n)\\
=&E\left\{E\left(I[N(t)=n]|T_n\right)\right\}\\
=&\int_0^\infty P(N(t)=n|T_n=y)\,dF_n(y)\\
=&\int_0^t P(N(t)=n|T_n=y)\,dF_n(y)\\
&(\text{因为}y>t\text{即}T_n>t\text{时}N(t)<n)\\
=&\int_0^t P(X_{n+1}>t-y|T_n=y)\,dF_n(y),
\end{aligned}$$

最后一式是因为在 $T_n=y(y\in(0,t])$ 条件下，$N(t)=n$ 当且仅当第 $n+1$ 次更新不发生在 $(y,t]$ 内，即 $X_{n+1}>t-y$。于是

$$\begin{aligned}
&P(N(t)=n)\\
=&\int_0^t P(X_{n+1}>t-y)\,dF_n(y)\\
&(\text{因为}X_{n+1}\text{与}T_n\text{独立})\\
=&\int_0^t \overline{F}(t-y)\,dF_n(y).
\end{aligned}$$

$N(t)<+\infty$, a.s. 并不能保证 $EN(t)<\infty$。下面的定理表明 $EN(t)<\infty$ 且给出了一个计算公式。

定理 4.1. 级数

$$M(t)=\sum_{n=1}^\infty F_n(t),\ t\ge 0$$

收敛，$M(t),t\ge 0$ 是单调不减函数，且

$$M(t)=EN(t),\ t\ge 0,$$

分段来证明这个定理。

引理 4.1. 设 $\{a_n\}$ 为非负数列，则级数 $\sum_{n=1}^{\infty} a_n$ 收敛或者等于 $+\infty$，且结果与求和次序无关。

证明见 §4.5.1。

推论 4.1. 正项的二重级数可以按任意次序计算，可以交换次序计算。

引理 4.2. 设 Y 为取非负整数值的随机变量，则

$$EY = \sum_{n=1}^{\infty} P(Y \geq n).$$

证明

$$\begin{aligned}
&\sum_{n=1}^{\infty} P(Y \geq n) = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P(Y = k) \\
=&\sum_{k=1}^{\infty} \sum_{n=1}^{k} P(Y = k) = \sum_{k=1}^{\infty} kP(Y = k) \\
=&\sum_{k=0}^{\infty} kP(Y = k) = EY.
\end{aligned}$$

更新时间的序列为 $T_n, n = 1, 2, ...,$ 相应的分布函数为 $F_n(t), t \geq 0$。

引理 4.3.

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

证明：因为 $N(t)$ 是取非负整数值的随机变量，所以

$$\begin{aligned}
EN(t) =& \sum_{n=1}^{\infty} P(N(t) \geq n) \\
=& \sum_{n=1}^{\infty} P(T_n \leq t) \\
=& \sum_{n=1}^{\infty} F_n(t).
\end{aligned}$$

这里用到了 $N(t) \geq n \Longleftrightarrow T_n \leq t$。

下面证明级数 $\sum_{n=1}^{\infty} F_n(t)$ 收敛，而且通项是以负指数速度衰减的。此级数定义的函数显然是单调增的。

引理 4.4. 对任意 $t \geq 0$，存在 $0 < \alpha < 1$ 使得

$$F_n(t) = O(\alpha^n),$$

从而

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

证明见4.5.2，这就证明了定理4.1。

例 4.2. 考虑一个时间离散的更新过程 $\{N_j, j = 1, 2, ...\}$, 在每个时刻独立地做 Bernoulli 试验，设成功的概率为 p，失败的概率为 $q = 1 - p$. 以试验成功做为事件（更新），求此过程的更新函数 $M(k)$。

解：首先, 易知更新的时间间隔 $\{X_i\}$ 为独立的同几何分布

$$P(X_i = k) = q^{k-1}p,\ k = 1, 2, ...$$

则第 r 次成功（更新）发生的时刻 $T_r = \sum_{i=1}^{r} X_i$ 服从 Pascal 分布

$$P(T_r = k) = \binom{k-1}{r-1} q^{k-r} p^r,\ k = r, r+1, ...$$

因此

$$\begin{aligned}
&P(N_n = r) \\
=&P(N_n \geq r) - P(N_n \geq r+1) \\
=&P(T_r \leq n) - P(T_{r+1} \leq n) \\
=&\sum_{k=r}^{n} \binom{k-1}{r-1} q^{k-r} p^r - \sum_{k=r+1}^{n} \binom{k-1}{r} q^{k-r-1} p^{r+1} \\
=&\binom{n}{r} p^r q^{n-r},\ r = 0, 1, \dots, n.
\end{aligned}$$

上面的最后一个等式可以直接证明，或者注意到 $\{N_n = r\}$ 就是 n 次重复独立试验中成功 r 次，其分布为二项分布 $\mathrm{B}(n, p)$。所以更新函数为

$$M(n) = E[N(n)] = \sum_{r=0}^{n} rP(N_n = r) = np.$$

4.2 更新方程及其应用

随机过程研究中经常以某个早期发生事件为条件导出方程，更新方程就是其中的一个常见形式。

4.2.1 更新方程

定义 4.2. 在 $M(t)$ 的导数存在的条件下，其导数 $M'(t)$ 称为**更新密度**，记为 $m(t)$. 由 $M(t) = \sum_{n=1}^{\infty} F_n(t)$ 两边求导得

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

其中 $f_n(t)$ 是 $F_n(t)$ 的密度函数.

$M(t)$ 是到 t 时刻为止的平均更新次数，$m(t)$ 就是瞬时的单位时间平均更新次数。

定理 4.2. *$M(t)$ 和 $m(t)$ 分别满足积分方程*

$$M(t) = F(t) + \int_0^t M(t-s)\, dF(s) \tag{4.1}$$

$$m(t) = f(t) + \int_0^t m(t-s) f(s)\, ds. \tag{4.2}$$

其中 $f(t) = F'(t)$.

证明： 只要证明第一式，第二式由第一式两边取导数可得. 由定义可得

$$\begin{aligned}
M(t) = \sum_{n=1}^{\infty} F_n(t) &= F(t) + \sum_{n=2}^{\infty} F_n(t) \\
&= F(t) + \sum_{n=2}^{\infty} \int_0^t F_{n-1}(t-s)\, dF(s) \\
&= F(t) + \int_0^t \sum_{n=2}^{\infty} F_{n-1}(t-s)\, dF(s) \quad (\text{由单调收敛定理}) \\
&= F(t) + \int_0^t M(t-s)\, dF(s) \\
&= F(t) + (F * M)(t).
\end{aligned}$$

定义 4.3 (更新方程)**.** 称如下形式的积分方程为**更新方程**

$$K(t) = H(t) + \int_0^t K(t-s)\, dF(s),$$

其中 $H(t)$, $F(t)$ 为已知，且当 $t<0$ 时 $H(t)$, $F(t)$ 均为 0. 当 $H(t)$ 在任意闭区间有界时, 称上述方程为**适定 (Proper) 更新方程**, 简称为更新方程.

更新方程可以用卷积记号写成

$$K(t) = H(t) + (F * K)(t),\ t \geq 0,$$

其中函数 $K(\cdot)$ 为未知的函数。H 和 F 已知，一般 F 为单调增有界函数。

命题 4.5. *假设 $B(t)$, $B_1(t)$, $B_2(t)$ 是单调增加的右连续函数, 且 $B(0)=0$. $C(t)$, $C_1(t)$, $C_2(t)$ 为光滑有界函数（注：这些条件可以保证卷积有定义），则有*

(1) $\max\limits_{0\le t\le T} |(B*C)(t)| \le \max\limits_{0\le t\le T} |C(t)|\cdot B(T)$;

(2) $(B*C_1)(t) + (B*C_2)(t) = (B*(C_1+C_2))(t)$;

(3) $(B_1*C)(t) + (B_2*C)(t) = ((B_1+B_2)*C)(t)$;

(4) $(B_1*(B_2*C))(t) = ((B_1*B_2)*C)(t)$.

证明略。

定理 4.3. *设更新方程中 $H(t)$ 为在任意闭区间有界的函数，$F(t)$ 为 $[0,\infty)$ 上的分布函数，则方程在有限区间内存在唯一的闭区间有界解*

$$K(t) = H(t) + \int_0^t H(t-s)dM(s), \tag{4.3}$$

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

更新方程的解又可以用卷积记号写成

$$K(t) = H(t) + (M*H)(t),\ t \geq 0.$$

证明： 我们先证(4.3) 满足闭区间有界性条件。

由 $M(t)$ 是更新函数, 由定理4.1可知 $M(t)$ 在任意闭区间有界且单调不减, 再由 $H(t)$ 闭区间有界，对任意 $T>0$，有

$$\begin{aligned}\sup_{0\le t\le T}|K(t)| &\le |H(t)|+\int_0^T \sup_{0\le s\le T}|H(T-s)|\,dM(s)\\ &\le \sup_{0\le t\le T}|H(t)|(1+M(T))<\infty.\end{aligned}$$

所以在任意闭区间上 $K(t)$ 是有界的.

再来证明(4.3)确实是更新方程的解，这只要证明 $(F*K)(t)=(M*H)(t)$。事实上，

$$\begin{aligned}(F*K)(t)=&(F*(H+M*H))(t)\\ =&(F*H)(t)+(F*(M*H))(t)\\ =&(F*H)(t)+((F*M)*H)(t)\\ =&((F+F*M)*H)(t)\\ &\quad(\text{由卷积性质 }(3))\\ =&(M*H)(t),\end{aligned}$$

最后一个等式利用了定理4.2。所以，(4.3)确实是更新方程的解。

最后证明惟一性，只需证明更新方程的任何解都有(4.3)式的形式. 设 $\tilde{K}(t)$ 是更新方程的解，即

$$\tilde{K}=H+F*\tilde{K}.$$

并且满足闭区间有界性条件。$\forall T>0$，设

$$|K(t)|\le B, |\tilde{K}(t)|\le B,\ \forall t\in[0,T].$$

连续代换 $\tilde{K}(t)$, 有

$$\begin{aligned}\tilde{K}(t) =&H(t) + (F * \tilde{K})(t)\\ =&H(t) + (F * (H + F * \tilde{K}))(t)\\ =&H(t) + (F * H)(t) + (F * (F * \tilde{K}))(t)\\ =&H(t) + (F * H)(t) + (F_2 * \tilde{K})(t)\\ =&H(t) + (F * H)(t) + (F_2 * (H + F * \tilde{K}))(t)\\ =&H(t) + (F * H)(t) + (F_2 * H)(t) + (F_3 * \tilde{K})(t)\\ =&\cdots\cdots\\ =&H(t) + \left[\left(\sum_{k=1}^{n-1} F_k\right) * H\right](t) + (F_n * \tilde{K})(t).\end{aligned}$$

同理有

$$K(t) = H(t) + \left[\left(\sum_{k=1}^{n-1} F_k\right) * H\right](t) + (F_n * K)(t),$$

于是

$$K(t) - \tilde{K}(t) = (F_n * (K - \tilde{K}))(t).$$

$\forall t \in [0, T]$，由 $M(T) = \sum_{n=1}^{\infty} F_n(T)$ 收敛可知

$$\lim_{n\to\infty} \sup_{t\in[0,T]} F_n(t) \leq \lim_{n\to\infty} F_n(T) = 0.$$

而

$$\begin{aligned}(F_n * |K - \tilde{K}|)(t) =& \int_0^t |K - \tilde{K}|(t - s)\, dF_n(s)\\ \leq& 2B \int_0^t dF_n(s) = 2BF_n(t) \to 0\ (n \to \infty),\end{aligned}$$

所以对 $0 \leq t \leq T$，

$$\begin{aligned}|K(t) - \tilde{K}(t)| =& \left|F_n * (K - \tilde{K})(t)\right|\\ \leq& (F_n * |K - \tilde{K}|)(t) \to 0,\end{aligned}$$

即 $K(t) = \tilde{K}(t)$，由 T 和 t 的任意性可知

$$\tilde{K}(t) = K(t),\ t \geq 0,$$

这就证明了闭区间有界解的唯一性。

例 4.3 (Wald 等式)**.** 对更新过程设 $EX_1 < \infty$, $i = 1, 2, ...$, 证明

$$E(T_{N(t)+1}) = E(X_1 + X_2 + \cdots + X_{N(t)+1}) = EX_1 E(N(t) + 1).$$

证明：$N(t) + 1$ 是 (t, ∞) 区间内的首次更新的序号，$T_{N(t)+1}$ 是 (t, ∞) 区间内的首次更新时刻，或 t 之后（不含）的首次更新时刻。

对第一次更新的时刻 X_1 取条件，令 $X_1 = x$，条件期望

$$E(T_{N(t)+1}|X_1 = x) = \begin{cases} x & \text{若} x > t, \\ x + E(T_{N(t-x)+1}) & \text{若} x \leq t. \end{cases}$$

事实上,如果 $X_1 = x > t$,则 $[0, t]$ 内没有更新,$N(t)+1 = 1$, $T_{N(t)+1} = X_1 = x$。如果 $X_1 = x \leq t$，则 $[0, t]$ 内至少有一次更新，$N(t) + 1 \geq 2$，更新过程可以从 x 时刻重新开始，所以 t 时刻之后的首次更新时刻的条件期望等于以 x 为新的时间原点时的 $t - x$ 之后首次更新的时间的期望。

记 $K(t) = E(T_{N(t+1)})$, 则

$$\begin{aligned} K(t) =& E[T_{N(t)+1}] = E\left[E(T_{N(t)+1}|X_1)\right] \\ =& \int_0^\infty E(T_{N(t)+1}|X_1 = x)\, dF(x) \\ =& \int_t^\infty x\, dF(x) + \int_0^t [x + K(t - x)]\, dF(x) \\ =& \int_0^\infty x\, dF(x) + \int_0^t K(t - x)\, dF(x) \\ =& EX_1 + \int_0^t K(t - x)\, dF(x). \end{aligned}$$

这是更新方程，由定理定理4.3知

$$\begin{aligned} K(t) =& EX_1 + \int_0^t (EX_1)\, dM(x) \\ =& E(X_1) \cdot (1 + M(t)) \\ =& E(X_1) \cdot (E[N(t) + 1]). \end{aligned}$$

注：教材中给出的利用独立性的证明是错误的，因为在 $N(t) = 1$ 条件下 X_1 的分布有变化。事实上，$N(t) + 1$ 是 $\{X_n, n = 1, 2, ...\}$ 的一个停时，可以利用停时证明 Wald 等式，见6.3。

例 4.4 (更新过程的余寿)**.** 设 $\{N(t), t \geq 0\}$ 为更新过程，X_i 表示第 i 次更新与上次更新的间隔时间 $(i = 1, 2, ...)$，比如，X_i 是某产品（如电池）的寿命。设 $X_1 \sim F(x)$。令 $T_n = \sum_{i=1}^{n} X_i$ 为第 n 次更新的时间，也是第 n 个产品的失效时间。令

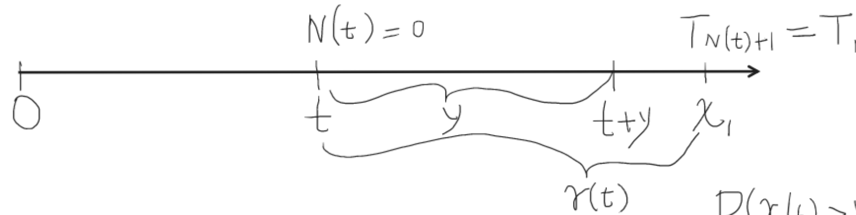
$$r(t) = T_{N(t)+1} - t$$

为 t 时刻的剩余寿命，证明 $r(t)$ 的分布为

$$P(r(t) > y) = \bar{F}(t + y) + \int_0^t \bar{F}(t + y - x)\, dM(x),\ y > 0. \tag{4.4}$$

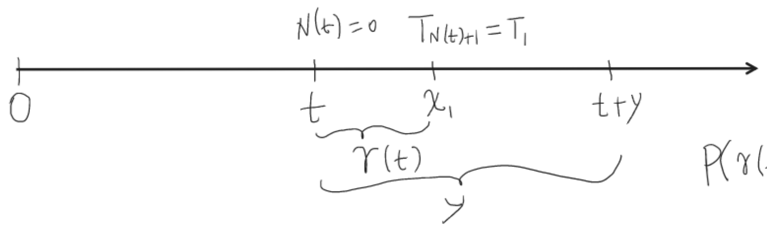
余寿示意图

情形1： $X_1 > t+y$



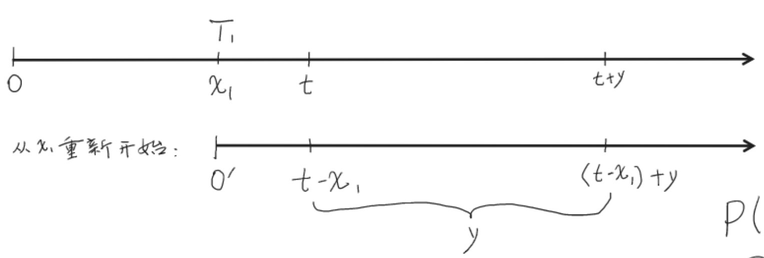
$$P(\gamma(t) > y \mid X_1 = x_1) = 1$$

情形2： $X_1 \in (t, t+y]$



$$P(\gamma(t) > y \mid X_1 = x_1) = 0$$

情形3： $X_1 \in [0, t]$



$$P(\gamma(t) > y \mid X_1 = x_1)$$
$$= P(\gamma(t - x_1) > y)$$

证明: 令

$$\bar{R}_y(t) = P(r(t) > y),\ y > 0.$$

对第一次更新时间 $T_1 = X_1$ 取条件，设 $X_1 = x$，若 $x > t + y$，则第一个产品在时刻 $t + y$ 还在工作，从 t 时刻开始计算的余寿 $> y$，于是这时

$$P(r(t) > y | X_1 = x) = 1;$$

若 $t < x \le t + y$，则第一个产品的更新发生在 $(t, t + y]$ 之间，这时 $N(t) = 0$，$T_{N(t)+1} = T_1 = X_1 = x \le t + y$，从 t 时刻开始计算的余寿 $\le y$，所以这时

$$P(r(t) > y | X_1 = x) = 0;$$

若 $0 < x \le t$，则 $N(t) \ge 1$，更新过程从 x 时刻重新开始，这时有

$$P(r(t) > y | X_1 = x) = \bar{R}_y(t - x).$$

总之有

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

由全期望公式，

$$\begin{aligned}
\bar{R}_y(t) =& E\left\{ E\left(I[r(t) > y] | X_1 \right) \right\} \\
=& \int_0^\infty E\left(I[r(t) > y] | X_1 = x \right) dF(x) \\
=& \int_0^\infty P\left(r(t) > y | X_1 = x \right) dF(x) \\
=& \int_{t+y}^\infty 1 \cdot dF(x) \\
& + \int_t^{t+y} 0 \cdot dF(x) \\
& + \int_0^t \bar{R}_y(t - x) dF(x) \\
=& 1 - F(t + y) + \int_0^t \bar{R}_y(t - x) dF(x),
\end{aligned}$$

记 $\bar{F}(x) = 1 - F(x)$，由定理4.3得

$$\bar{R}_y(t) = \bar{F}(t+y) + \int_0^t \bar{F}(t+y-x)\,dM(x).$$

例 4.5 (泊松过程的余寿)**.** 设 $\{N(t), t \geq 0\}$ 为速率 λ 的泊松过程，$r(t)$ 为 t 时刻的剩余寿命，证明 $r(t)$ 服从参数为 λ 的指数分布，这称为泊松过程的无记忆性。

证明: 由例4.4得

$$P(r(t) > y) = \bar{F}(t+y) + \int_0^t \bar{F}(t+y-x)\,dM(x),\ y > 0.$$

对泊松过程，X_1 服从参数 λ 的指数分布，

$$\begin{aligned} \bar{F}(x) =& e^{-\lambda x}, x > 0, \\ M(t) =& EN(t) = \lambda t, \end{aligned}$$

所以有

$$\begin{aligned} & P(r(t) > y) \\ =& e^{-\lambda(t+y)} + \int_0^t e^{-\lambda(t+y-x)} \lambda\,dx \\ =& e^{-\lambda(t+y)} \left[1 + \int_0^t de^{\lambda x} \right] \\ =& e^{-\lambda y},\ y > 0, \end{aligned}$$

即泊松过程时刻 t 处的余寿分布为参数为 λ 的泊松分布，与所处时刻和已经发生的更新次数都没有关系。

4.2.2 更新方程在人口学中的一个应用

考虑一个确定性的人口模型：

- $B(t)$ 表示在时刻 t 女婴的出生率, 即在 $[t, t+dt]$ 时间区间内有 $B(t)dt$ 个女婴出生. 已知过去的 $B(t)(t \leq 0)$, 要预测未来的 $B(t)(t > 0)$。
- 假定生存函数为 $S(x)$(指一个新生女婴能够活到年龄 x 的概率）已知；

- 假定生育的年龄强度为 $\beta(x)(x>0)$(指年龄为 x 的母亲生育女婴的速率，即 $\beta(x)dt$ 为这个母亲在长度为 dt 的时间区间内生下的女婴数) 已知.

考虑对给定 $t>0$ 时刻，年龄段 $[x,x+dx]$ 的女性，她们出生于 $[t-x-dx,t-x]$ 时间段，该时间段内有 $B(t-x)dx$ 个女婴出生，但到 t 时刻存活的人数为 $B(t-x)S(x)dx$，这些人在 $[t,t+dt]$ 时间段将生育女婴 $B(t-x)S(x)\,dx\cdot\beta(x)\,dt$ 个。将 t 时刻的所有年龄的育龄妇女生出的女婴个数求和，即求积分，得到在 $[t,t+dt]$ 时间段内出生的女婴个数为

$$B(t)\,dt=\int_0^{\infty}B(t-x)S(x)\beta(x)\,dx\,dt. \tag{4.5}$$

根据过去与未来的生育情况，将上述积分分成两段

$$B(t)\,dt=\int_t^{\infty}B(t-x)S(x)\beta(x)\,dx\,dt \tag{4.6}$$

$$+\int_0^{t}B(t-x)S(x)\beta(x)\,dx\,dt. \tag{4.7}$$

(4.7)的第一个积分中 $t-x<0$，$B(t-x)$ 已知；第二个积分中 $t-x\geq 0$, $B(t-x)$ 未知。消去(4.7)中的 dt，记 $f(x)=S(x)\beta(x)$, $H(t)=\int_t^{\infty}B(t-x)S(x)\beta(x)\,dx$, 则

$$B(t)=H(t)+\int_0^{t}B(t-x)f(x)\,dx.$$

这是一个更新方程，对 $H(t)$ 积分作变量替换 $x=y+t$, 得

$$H(t)=\int_0^{\infty}B(-y)S(y+t)\beta(y+t)\,dy.$$

注意 $H(t)dt$ 是由年龄为 t 或更大的女性在时间 $[t,t+dt]$ 之间生育的女婴数，此外，每一个新生的女婴将期待在年龄 x 与 $x+dx$ 之间生育 $f(x)dx$ 个女婴. 于是，每一新生女婴在死亡或生存到年龄 x 之前（不论哪种情况先发生）将期待生育 $F(x)=\int_0^x f(s)ds$ 个女婴，从而在她一生中将期待生育 $F(\infty)$ 个女婴.

若 $F(\infty)>1$，可以解得 $B(t)\sim Ce^{Rt}$, $t\to\infty$(解的过程略), 其中 C 为常数，$R>0$ 满足方程

$$\int_0^{\infty}e^{-Ry}S(y)\beta(y)\,dy=1.$$

即出生速率（以及具有此速率的人群）将以渐近指数速度爆炸性增长. 若 $F(\infty)<1$, 则 $B(t)\sim Ce^{-Rt}$, $B(t)$ 渐近负指数地趋于 0，也就是说人群最终要消亡，只有当 $F(\infty)=1$ 时, 出生速率将最终趋于一个有限的正数.

4.3 更新定理

$N(t)$ 的分布虽然可以通过 $F_n(t)$ 计算，但是很困难，我们更关心极限情况下平均更新速率这样的问题。在第3章中我们已经知道，强度为 λ 的泊松过程的两次事件发生的时间间隔 X_n 服从参数为 λ 的指数分布，且其更新函数

$$M(t) = E[N(t)] = \lambda t,$$

于是

$$\lim_{t\to\infty} \frac{M(t)}{t} = \lambda = \frac{1}{E[X_n]}.$$

那么，对于一般的更新过程是否还有这种性质呢？有如下的极限 (elementary renewal theorem)。

定理 4.4 (Feller 初等更新定理)**.** 记 $\mu = EX_n$，则

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

其中 $\mu = +\infty$ 时 $\frac{1}{\mu} = 0$。

证明：首先假设 $\mu < \infty$。注意到 $T_{N(t)+1}$ 是 t 时刻之后第一次更新的时间，所以

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

两边取期望, 利用 Wald 等式（见例4.3）得

$$\mu[M(t) + 1] > t.$$

于是

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

从而推出

$$\liminf_{t\to\infty} \frac{M(t)}{t} \geq \frac{1}{\mu}. \tag{4.9}$$

另一方面，固定常数 $B > 0$, 令

$$X_n^c = \begin{cases} X_n, & X_n \leq B \\ B, & X_n > B, \end{cases}$$

则

$$X_n^c \le B, \quad X_n^c \le X_n,$$

$X_n^c, n = 1, 2, ...$ 确定了一个新的更新过程，记

$$T_n^c = \sum_{i=1}^n X_i^c,$$

$$N(t)^c = \sup\{n : T_n^c \le t\}.$$

由于 $X_n^c \le B$, 则

$$T_{N^c(t)+1}^c = T_{N^c(t)}^c + X_{N^c(t)+1}^c \le t + B.$$

再利用 Wald 等式得

$$[M^c(t) + 1] \cdot \mu_B \le t + B,$$

其中 $\mu_B = E(X_n^c)$, $M^c(t) = E[N^c(t)]$, 由上式得

$$\frac{M^c(t)}{t+B} \le \frac{1}{\mu_B} - \frac{1}{t+B},$$

从而

$$\limsup_{t\to\infty} \frac{M^c(t)}{t} \le \frac{1}{\mu_B}.$$

又因为 $X_n^c \le X_n$, 得 $T_n^c \le T_n$, $N^c(t) \ge N(t)$, $M^c(t) \ge M(t)$, 于是

$$\limsup_{t\to\infty} \frac{M(t)}{t} \le \frac{1}{\mu_B}.$$

令 $B \to \infty$, 有 $X_n^c \to X_n$, $E(X_n^c) \to EX_n$（控制收敛定理）, 即

$$\mu_B \to \mu, \ (B \to \infty),$$

所以

$$\limsup_{t\to\infty} \frac{M(t)}{t} \le \frac{1}{\mu}. \tag{4.10}$$

利用由(4.9)和(4.10)式可得 $\mu < \infty$ 时

$$\lim_{t\to\infty} \frac{M(t)}{t} = \frac{1}{\mu}.$$

当 $\mu = \infty$ 时, 仍有

$$\limsup_{t\to\infty} \frac{M(t)}{t} \le \frac{1}{\mu_B},$$

当 $B\to\infty$ 时, $\mu_B\to\infty$，所以 $\limsup_{t\to\infty}\frac{M(t)}{t}=0$，即 $\lim_{t\to\infty}\frac{M(t)}{t}=0$，所以 $\mu=+\infty$ 时定理结论成立.

注: 设 $X\geq 0$, $EX=+\infty$，$X_B=XI_{\{X\leq B\}}+BI_{\{X>B\}}$，则 $E(X_B)\to\infty(B\to\infty)$。事实上，易见 $E(X_B)$ 是 B 的单调增函数，取 $B=n$，由单调收敛定理有

$$\lim_{n\to\infty}E(X_n)\geq\lim_{n\to\infty}E[XI_{\{X\leq n\}}]=E(X)=+\infty,$$

所以 $\lim_{B\to\infty}E(X_B)=+\infty$。

命题 4.6.

$$\lim_{n\to\infty}\frac{S_n}{n}=E(X_1),\ \ a.s.$$

由强大数律直接可得此结论。

命题 4.7.

$$\lim_{n\to\infty}\frac{N(t)}{t}=\frac{1}{E(X_1)},\ \ a.s.$$

证明：易见 $\lim_{t\to\infty}N(t)=\infty$，a.s.，又已证明 $\lim_{n\to\infty}\frac{S_n}{n}=E(X_1)$, 由复合极限性质知

$$\lim_{t\to\infty}\frac{S_{N(t)}}{N(t)}=E(X_1).$$

由

$$S_{N(t)}\leq t<S_{N(t)+1},$$

得

$$\frac{S_{N(t)}}{N(t)}\leq\frac{t}{N(t)}\leq\frac{S_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)},$$

于是可知

$$\lim_{t\to\infty}\frac{t}{N(t)}=E(X_1),\ \ \text{a.s.},$$

从而

$$\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{E(X_1)},\ \ \text{a.s.}$$

例 4.6. 某控制器用一节电池供电，设电池寿命 $X_i(i = 1,2,...)$ 服从均值为 45 小时的正态分布，电池失效时需要去仓库领取，领取新电池的时间 $Y_i(i = 1,2,...)$ 服从期望为 0.5 小时的均匀分布. 求长时间工作时，控制器更换电池的速率.

解: 以 $N(t)$ 表示到时刻 t 为止更换电池的次数，平均更新时间为 $E[X_i+Y_i] = 45.5$ 小时.

由初等更新定理（Feller），更换电池的速率为

$$\lim_{t\to\infty}\frac{M(t)}{t} = \frac{1}{45.5}(\text{次/小时}).$$

例 4.7. 考虑离散时间的更新过程 $N(n), n = 0,1,2,...$，在每个时间点独立地做 Bernoulli 试验，设试验成功的概率为 p，失败的概率为 $q = 1-p$，以试验成功作为更新事件，以 $M(n)$ 记此过程的更新函数，求其更新率 $\lim_{n\to\infty}\frac{M(n)}{n}$.

解: 相邻两次试验成功的时间间隔 X_i 的分布为几何分布，$P\{X_i = k\} = pq^{k-1}$，$k = 1,2,...$，

平均更新时间为几何分布的期望值：

$$\begin{aligned}E[X_1] =& \sum_{k=1}^{\infty} kpq^{k-1} = p\sum_{k=1}^{\infty}(q^k)' \\ =& p\left(\sum_{k=1}^{+\infty} q^k\right)' = p(\frac{q}{1-q})' = \frac{1}{p}.\end{aligned}$$

由初等更新定理（Feller）有

$$\lim_{n\to\infty}\frac{M(n)}{n} = \frac{1}{E[X_1]} = p.$$

这个例子不用更新定理也能解决。实际上，$N(n) \sim \mathrm{B}(n,p)$，所以 $M(n) = E[N(n)] = np$, $\frac{M(n)}{n} = p$。

例 4.8. 某电话交换台的电话呼叫次数服从平均 1 分钟 λ 次的 Poisson 过程，通话时间 $Y_1, Y_2, \cdots$ 是相互独立且服从同一分布的随机变量序列，满足 $E[Y_1] < \infty$，

假定通话时电话打不进来，用 $N(t)$ 表示到时刻 t 为止电话打来的次数，试证

$$\lim_{n\to\infty} \frac{E[N(t)]}{t} = \frac{\lambda}{1+\lambda E[Y_1]}.$$

解：

简化地考虑，每次通话结束，平均等待有电话打入的时间是 $1/\lambda$，平均通话时间是 $E(Y_1)$，所以通话结束之间的间隔 X_1 的期望是 $\frac{1}{\lambda}+E(Y_1)$。在第一次通话结束之后，由泊松过程的无记忆性，可以将此刻作为新的电话开始打入时刻，所以第一次通话结束到第二次通话结束之间的间隔 X_2 与 X_1 同分布且独立，同理可知各 $X_1, X_2, \ldots$ 独立同分布，构成更新过程的间隔时间，于是

$$\lim_{n\to\infty} \frac{E[N(t)]}{t} = \frac{1}{\frac{1}{\lambda}+E(Y_1)} = \frac{\lambda}{1+\lambda E[Y_1]}.$$

下面更严格第证明通话结束时刻的序列为更新过程的发生时间. 以 X_i 表示从第 $i-1$ 次通话结束到下一次 i 次电话打进来并开始通话的间隔时间，X_1 是第一个电话打入的时间。用 Z_i 表示所有电话（接通与未接通）的间隔时间，则 $\{Z_i\}$ 相互独立服从 $\text{Exp}(\lambda)$，定义了一个速率参数为 λ 的泊松过程 $\{\tilde{N}(t), t \geq 0\}$，记为 $\text{PP}(\lambda)$，记第 n 个电话的到来时刻为 $\tilde{T}_n = \sum_{i=1}^n Z_i$。记 $W_i = X_i + Y_i$，来证明 $W_i, i=1,2,\ldots$ 独立同分布且 $X_i \sim \text{Exp}(\lambda)$。

第一个电话在 X_1 时刻到来并开始通话，要到 X_1+Y_1 时刻才能继续接听新的电话，在 X_1 到 X_1+Y_1 之间到来的电话被忽略。以 $X_1=x_1, Y_1=y_1$ 为条件，记 $t_1=x_1+y_1$，则 $\tilde{N}(t_1)\geq 1$，第 $\tilde{N}(t_1)+1$ 个电话在 $\tilde{T}_{\tilde{N}(t_1)+1}$ 时刻到来并开始通话，是第二个被接通的电话，所以

$$X_2 = \tilde{T}_{\tilde{N}(t_1)+1} - t_1$$

是 $\text{PP}(\lambda)$ 在 t_1 时刻的余寿，服从 $\text{Exp}(\lambda)$ 且不依赖于 t_1，因此 X_2 与 (X_1, Y_1) 独立，从而 W_2 与 W_1 独立。类似可证明 W_n 与 $W_1, \ldots, W_{n-1}$ 独立且 X_n 服从 $\text{Exp}(\lambda)$，这样，$\{W_n = X_n + Y_n\}$ 定义了一个更新过程，平均更新时间为

$$\mu = E[X_i + Y_i] = \frac{1}{\lambda} + E[Y_1],$$

例 4.9. 设有一个单服务员银行，顾客到达可看作速率为 λ 的 Poisson 过程，服务员为每一位顾客服务的时间是随机变量，分布为均值为 $\frac{1}{\mu}$ 指数分布. 顾客到达门口只能在服务员空闲时才准进来，服务员忙时顾客则离开不返回. 试求：

（1）顾客进银行的速率；

（2）服务员工作的时间所占的比例.

解：这与例4.8模型相同。用 X_i 表示第 $i-1$ 个顾客结束服务后下一个顾客到来的间隔，由泊松过程的无记忆性，X_i 相互独立且服从 $\mathrm{Exp}(\lambda)$ 分布。

（1）两位进入银行的顾客的平均间隔时间为 $\frac{1}{\lambda}+\frac{1}{\mu}$.

反过来，顾客进银行的速率为 $\frac{\lambda\mu}{\lambda+\mu}$.

（2）服务员工作的时间所占比例

$$\frac{1}{\mu}/(\frac{1}{\lambda}+\frac{1}{\mu})=\frac{\lambda}{\lambda+\mu}.$$

$\mu<\infty$ 时，定理4.4可以看做当 $t\to\infty$ 时，$M(t)\sim\frac{t}{\mu}$（$\sim$ 记号含义是 $\lim_{t\to\infty}\frac{M(t)}{\frac{t}{\mu}}=1$）. 很自然, 我们会猜想当 $t\to\infty$ 时, 对于任意一个 h，是否有

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

由这一猜想, 引出了另一个重要的更新定理——Blackwell 更新定理. 为了描述这一定理，我们先引入一个新的概念.

定义 4.4 (格点分布)**.** 若存在 $d\geq 0$, 使得

$$\sum_{n=0}^{\infty}P(X=nd)=1,$$

称随机变量 X 服从**格点分布** (或称随机变量 X 的分布 F 是格点分布)，称满足上述条件的最大的 d 为此格点分布的**周期**.

注：从上述定义只能知道格点分布在不是 d 的整数倍处取值的概率为 0，但并不一定所有 $nd,n=0,1,2,...$ 都有正概率.

定理 4.5 (Blackwell 更新定理)**.** 记 $\mu=EX_1$，

（1）若 F 不是格点分布，则对一切 $a \geq 0$,

$$\lim_{t\to\infty}[M(t+a)-M(t)] = \frac{a}{\mu}.$$

（2）若 F 是格点分布，周期为 $d>0$, 则

$$\lim_{\to\infty} P(\text{在}nd\text{处发生更新}) = \frac{d}{\mu}.$$

证明略去。

注：Blackwell 定理指出, 在远离原点的某长度为 a 的区间内，更新次数的期望是 $\frac{a}{\mu}$, 直观上这是容易理解的，因为 $\frac{1}{\mu}$ 本来就可以看做长时间后更新过程发生的平均速率 (见习题 4.3). 但当 F 是格点时，T_n 的分布也是格点的，以 d 为周期，由于更新只能发生在 d 的整数倍处, (1) 就不能成立了，因为更新次数的多少是依赖于区间上形如 nd 的点的数目，而同样长度的区间内含有此类点的数目是可以不同的.

初等更新定理是 Blackwell 定理的特殊情形. 事实上, 令 $b_n = M(n)-M(n-1)$, 当 F 是非格点时, 由 (1) 知

$$b_n \to \frac{1}{\mu},\ n\to\infty,$$

由极限的性质得

$$\frac{M(n)}{n} = \frac{b_1+b_2+\cdots+b_n}{n} \to \frac{1}{\mu},\ n\to\infty.$$

而对任何实数 t,

$$\frac{[t]}{t}\cdot\frac{M([t])}{[t]} \leq \frac{M(t)}{t} \leq \frac{[t]+1}{t}\cdot\frac{M([t]+1)}{[t]+1},$$

这里 $[t]$ 表示 t 的整数部分, 即不超过 t 的最大整数. 令 $t\to\infty$ 可得

$$\frac{M(t)}{t} \to \frac{1}{\mu}.$$

这就是 Feller 定理. 当 F 是格点分布时，请读者自己论证.

定理 4.6 (关键更新定理)**.** 记 $\mu = EX_1$, 设函数 $h(t), t\in[0,\infty)$ 满足

$$\text{(i) } h(t)\text{非负不增};$$

$$\text{(ii) } \int_0^\infty h(t)\,dt < \infty.$$

$H(t)$ 是更新方程

$$H(t) = h(t) + \int_0^t H(t-x)\, dF(x)$$

的解. 那么

（*1*）若 F 不是格点分布, 有

$$\lim_{t\to\infty} H(t) = \begin{cases} \frac{1}{\mu}\int_0^\infty h(x)\, dx, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$$

（*2*）若 F 是格点分布，对于 $0 \le c < d$, 有

$$\lim_{n\to\infty} H(c+nd) = \begin{cases} \frac{d}{\mu}\sum_{n=0}^\infty h(c+nd), & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$$

关键更新定理与 Blackwell 定理的等价性（不严格的讨论）：只考虑 F 不是格点的情况，在关键更新定理中，取满足 (i),(ii) 两个条件的 $h(t) = I_{[0,a)}(t)$. 代入更新方程的解, 则有

$$\begin{aligned} H(t) =& h(t) + \int_0^t h(t-x) dM(x) \\ =& \int_{t-a}^t dM(x) = M(t) - M(t-a). \end{aligned}$$

又

$$\lim_{t\to\infty} H(t) = \frac{1}{\mu}\int_0^\infty h(x)dx = \frac{1}{\mu}\int_0^a dx = \frac{a}{\mu},$$

从而有

$$\lim_{t\to\infty}(M(t) - M(t-a)) = \frac{a}{\mu}.$$

反过来，由 Blackwell 定理知道, 对任意的 $a > 0$,

$$\lim_{t\to\infty} \frac{M(t+a) - M(t)}{a} = \frac{1}{\mu}.$$

故

$$\lim_{a\to 0}\lim_{t\to\infty} \frac{M(t+a) - M(t)}{a} = \frac{1}{\mu}.$$

假设极限次序可交换, 将有

$$\lim_{t\to\infty} \frac{dM(t)}{dt} = \frac{1}{\mu}.$$

即当 t 很大时, 有 $dM(t) \sim \frac{1}{\mu}dt$. 又 $\int_0^\infty h(x)dx < \infty$, 则当 $t \to \infty$ 时，$h(t)$ 将快速趋于 0, 因此当 t 变得很大时, 对于 $h(t-x)$ 主要考虑当 $t-x$ 比较小也就是 x 比较大的情况. 则有

$$\int_0^t h(t-x)dM(x) \approx \int_0^t h(t-x)dx \cdot \frac{1}{\mu} = \frac{1}{\mu}\int_0^t h(x)dx.$$

这正是关键更新定理的结论.

例 4.10 (剩余寿命与年龄的极限分布)**.** 以 $r(t) = T_{N(t)+1} - t$ 表示更新过程在时刻 t 的剩余寿命，$s(t) = t - T_{N(t)}$ 为 t 时刻的年龄. 求 $r(t)$ 和 $s(t)$ 的极限分布.

解：对给定的 $y > 0$，记 $H(t) = P(r(t) > y)$，由例4.4有

$$H(t) = \bar{F}(t+y) + \int_0^t H(t-x)\,dF(x),$$

这是更新方程，其解为

$$P(r(t) > y) = \bar{F}(t+y) + \int_0^t \bar{F}(t+y-x)\,dM(x),\ y > 0.$$

设 $\mu = EX_1 < \infty$, 则

$$\mu = \int_0^\infty [1-F(x)]dx < \infty.$$

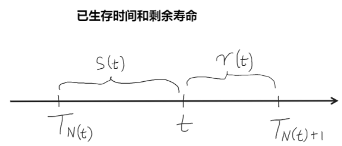
所以

$$\int_0^\infty [1-F(t+y)]dt = \int_y^\infty [1-F(x)]dx < \infty.$$

即 $h(t) = 1 - F(t+y)$ 满足关键更新定理的条件，于是

$$\lim_{t\to\infty} P(r(t) > y) = \lim_{t\to\infty} H(t) \tag{4.11}$$

$$= \frac{1}{\mu}\int_0^\infty h(x)\,dx = \frac{1}{\mu}\int_y^\infty [1-F(x)]\,dx, \quad y > 0. \tag{4.12}$$



年龄 $s(t)$ 的分布可由(4.12)式导出. 注意到

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

从而

$$\begin{aligned}\lim_{t\to\infty} P(r(t)>x, s(t)>y) =& \lim_{t\to\infty} P(r(t-y)>x+y)\\ =&\frac{1}{\mu}\int_{x+y}^{\infty}[1-F(z)]dz.\end{aligned}$$

特别地

$$\begin{aligned}\lim_{t\to\infty} P(s(t)>y) =& \lim_{t\to\infty} P(s(t)>y, r(t)>0)\\ =&\frac{1}{\mu}\int_{y}^{\infty}[1-F(z)]\,dz.\end{aligned}$$

可见更新过程在 $t\to\infty$ 时的寿命和余寿有相同的渐近分布。

注 1：在应用关键更新定理时，与更新定理的应用类似，常常如4.10这样先关于某次更新 (一般对第一次或 t 时刻前的最后一次) 取条件而得到一个更新方程，再利用关键更新定理.

注 2：例4.10中得到的年龄与剩余寿命的极限分布是相同的，如何理解这一结论？

4.4 更新过程的推广

4.4.1 延迟更新过程

定义 4.5. 更新过程要求时间间隔 $X_1, X_2, ...$ 是分布函数为 F 的独立同分布序列. 如果放宽对 X_1 的要求，允许它服从别的分布 G, 则称由 $X_1, X_2, ...$ 确定的计数过程是**延迟更新过程**.

例如，设 X_i 表示第 i 个可替换零件的运行时间，假设开始观察时第一个零件已经运行了一段时间，则到它损坏的使用时间 X_1，就应认为与新零件能使用的时间 $X_2, X_3, \ldots$ 不同分布（除非是指数分布寿命）. 但是要注意的是 $X_1, X_2, \ldots$ 之间的独立性仍然是必需的.

可以证明对于延迟更新过程，节4.3中的三个更新定理依然成立，但要注意这时 $M(t) = \sum_{n=1}^{\infty} F_n(t)$ 中的 $F_n(t)$ 相应用 $F_1(t) = G(t)$, $F_n(t) = G * F_{n-1}(t)(n \geq 2)$ 来替换。直观上这也是比较显然的，因为我们并不认为第一次更新发生的时刻会对无穷远处的情况有很大的影响.

4.4.2 更新回报过程

定义 4.6. 设

$$R(t) = \sum_{i=1}^{N(t)} R_i,$$

其中 $\{N(t), t \geq 0\}$ 是一个更新过程, R_n, $n = 1, 2, \ldots$ 独立同分布且与 $\{N(t), t \geq 0\}$ 独立, 则称 $R(t)$ 是一个**更新回报过程**。

注：这一名称是由于我们可以把 R_n 看做第 n 次更新发生时带给我们的回报，$R(t)$ 看成是截止到 t 时刻获得的总回报。更一般的更新回报过程，可以允许 R_n 依赖于 X_n(即回报的多少与等待的时间有关)，只要求随机向量列 (X_n, R_n) 独立同分布.

命题 4.8. 设 $E|R_1| < \infty$，$E|X_1| < \infty$，则

$$E[R(t)] = E(R_1) \cdot E[N(t)].$$

如果 $E(R_1^2) < \infty$，$Var[N(t)] < \infty$，则

$$Var[R(t)] = Var(R_1)E[N(t)] + (ER_1)^2 \, Var[N(t)].$$

证明若 $E|R_1| < \infty$，则

$$\begin{aligned}
E[R(t)] =& \sum_{n=0}^{\infty} E\left(\sum_{i=1}^{n} Y_i \middle| N(t)=n \right) P(N(t)=n) \\
=& \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{n} Y_i \right) \ P(N(t)=n) \\
=& E(Y_1) \sum_{n=1}^{\infty} nP(N(t)=n) \\
=& E(Y_1)E[N(t)].
\end{aligned}$$

当 $\mathrm{Var}(N(t)) < \infty$ 时，

$$\begin{aligned}
\mathrm{Var}[R(t)] =& E\left[\mathrm{Var}\left(R(t)|N(t)\right)\right] + \mathrm{Var}\left[E\left(R(t)|N(t)\right)\right] \\
=& E\left[\mathrm{Var}\left(\sum_{i=1}^{N(t)} R_i \middle| N(t) \right) \right] + \mathrm{Var}\left[E\left(\sum_{i=1}^{N(t)} R_i \middle| N(t) \right) \right] \\
=& E\left[N(t)\mathrm{Var}(R_1)\right] + \mathrm{Var}\left[N(t) \cdot E(R_1)\right] \\
=& \mathrm{Var}(R_1)E[N(t)] + (ER_1)^2 \mathrm{Var}[N(t)].
\end{aligned}$$

注：更新回报过程是复合泊松过程（3.3.2）的推广。对复合泊松过程，$E[N(t)] = \lambda t$, $E(X_1) = \frac{1}{\lambda}$，$E[R(t)] = E(Y_1)\lambda t$，有

$$\frac{E[R(t)]}{t} = E(Y_1)\lambda = \frac{E(Y_1)}{E(X_1)}.$$

更一般的更新回报过程可以考虑极限情形。

定理 4.7 (更新回报定理)**.** *若更新间隔 $X_1, X_2, \ldots$ 满足 $EX_1 < \infty$, 每次得到的回报 $\{R_i\}$ 满足 $ER_1 < \infty$. 则*

$$\begin{aligned}
&(1)\ \lim_{t\to\infty} \frac{R(t)}{t} = \frac{ER_1}{EX_1}\ \ a.s.; \\
&(2)\ \lim_{t\to\infty} \frac{E[R(t)]}{t} = \frac{ER_1}{EX_1}.
\end{aligned}$$

证明：（1）先证明

$$\lim_{t\to\infty} N(t) = +\infty,\ \text{a.s.}$$

因为 $E(X_1) < \infty$ 所以 $X_1 < \infty$, a.s.，于是

$$
\begin{aligned}
&P(\lim_{t\to\infty} N(t) < \infty) \\
=&P(\text{存在}T\text{使得}t \geq T\text{时}N(t)\text{为常数}) \\
\leq&P(\text{存在}K\text{使得}X_K = +\infty) \\
\leq&\sum_{n=1}^{\infty} P(X_n = +\infty) = 0.
\end{aligned}
$$

于是 a.s. 地

$$
\lim_{t\to\infty} \frac{N(t)}{t} = \lim_{t\to\infty} \frac{N(t)}{T_{N(t)} + [t - T_{N(t)}]},
$$

其中

$$
\lim_{t\to\infty} \frac{T_{N(t)}}{N(t)} = \lim_{t\to\infty} \frac{\sum_{i=1}^{N(t)} X_i}{N(t)} = E(X_1), \text{ a.s.},
$$

$$
t - T_{N(t)} \leq X_{N(t)+1}
$$

故（这里不够严格）

$$
\lim_{t\to\infty} \frac{t - T_{N(t)}}{N(t)} = 0, \text{ a.s.},
$$

所以

$$
\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{E(X_1)}, \text{ a.s.}
$$

进一步有

$$
\begin{aligned}
&\lim_{t\to\infty} \frac{R(t)}{t} = \lim_{t\to\infty} \frac{R(t)}{N(t)} \lim_{t\to\infty} \frac{N(t)}{t} \\
=&E(R_1) \cdot \frac{1}{E(X_1)}, \text{ a.s.}
\end{aligned}
$$

（2）由命题4.8有

$$
E[R(t)] = E(R_1)E[N(t)] = E(R_1)M(t),
$$

由 Feller 初等更新定理有

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

注：后面所说的“长期的平均回报率”，按定理的第 (2) 条定义较好。

推论 4.2. 对更新过程 $N(t)$，有

$$\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{E(X_1)},\ a.s.$$

教材例题 4.4.1（产品保修策略）叙述较混乱，改为如下例题。例题来自 Ross: Introduction to Probablity Models 12th Ed P.452 例 7.14。

例 4.11 (产品替换策略)**.** 考虑汽车用户甲的替换策略。设汽车的寿命 Z 服从连续分布，分布函数和分布密度分别为 H 和 h，当现有汽车失效或者使用年限达到 T 时甲就废弃旧车购买新车，购买新车的费用是 C_1 元，如果汽车失效时购买新车还要有一个额外的 C_2 元的费用，废弃的旧车不产生回报。问：甲的长期的单位时间购置费用为多少？如果汽车寿命服从 0 到 10 的均匀分布，$C_1=3$, $C_2=\frac{1}{2}$，如何取 T 使得长期单位时间购置费用最低？

解答：将每次购置新车的时间点看作一次更新发生，构成更新过程，每次更新时支付的费用作为一个随机变量 Y_i，这些 Y_i 独立同分布，到时刻 t 为止购置新车花费的费用是一个更新回报过程 $\{R(t),t\geq 0\}$。长期的单位时间购置费用定义为

$$\lim_{t\to\infty}\frac{ER(t)}{t}.$$

由定理4.7上式等于

$$\frac{EY_1}{EX_1},$$

其中 X_i 表示两次更新之间的间隔。由题意

$$X_1=\begin{cases}T, & Z>T\\ Z, & Z\leq T,\end{cases}$$

其中 Z 表示第一辆汽车的寿命，则

$$\begin{aligned}E(X_1)=&E\{TI[Z>T]\}+E\{ZI[Z\leq T]\}\\ =&\int_T^\infty Th(z)\,dz+\int_0^T zh(z)\,dz\\ =&T[1-H(T)]+\int_0^T zh(z)\,dz.\end{aligned}$$

而费用

$$Y_1=\begin{cases}C_1, & Z>T,\\ C_1+C_2, & Z\leq T,\end{cases}$$

所以

$$
\begin{aligned}
E(Y_1) =& C_1 P(Z > T) + (C_1 + C_2) P(Z \leq T) \\
=& C_1 + C_2 H(T).
\end{aligned}
$$

于是，长期的单位时间购置费用为

$$
\mu_C = \frac{EY_1}{EX_1} = \frac{C_1 + C_2 H(T)}{T[1 - H(T)] + \int_0^T z h(z)\, dz}.
$$

若 $Z \sim \text{U}(0, 10)$, $C_1 = 3$, $C_2 = \frac{1}{2}$，取 $T \leq 10$，则 $H(T) = T/10$,

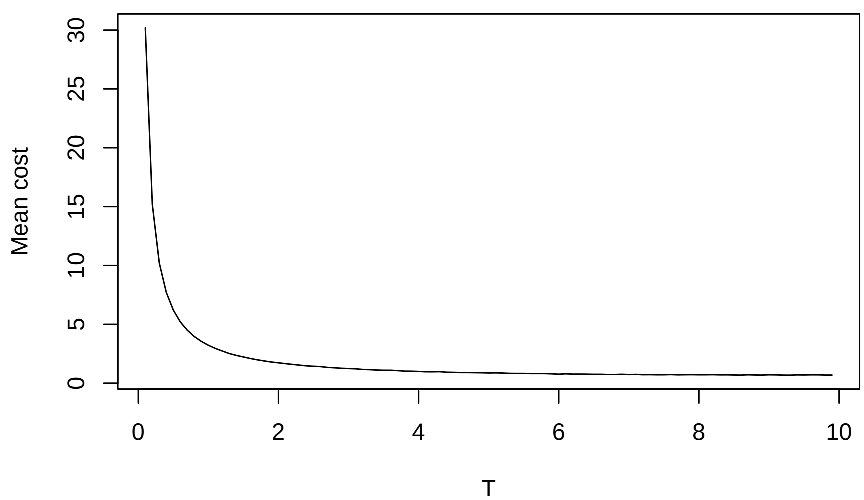
$$
\mu_C = \frac{3 + \frac{T}{20}}{T(1 - \frac{T}{10}) + \frac{T^2}{20}},
$$

令关于 T 的一阶导数等于零，可解得 $T \approx 9.282$。

可以用随机模拟方法验证结果。这里只模拟一条轨道，实际计算到是 $\frac{R(t)}{t}$ 在 t 很大时的值（取了 $t = 10^6$）。

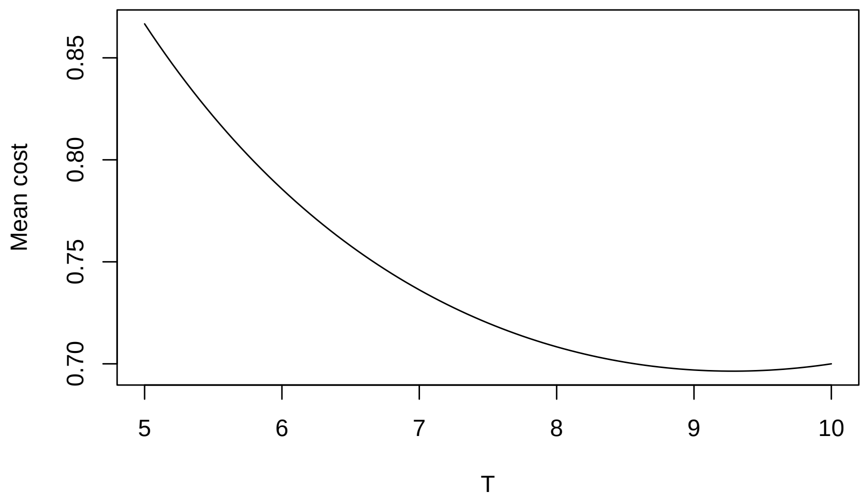
```
simone <- function(Tc=5.0, end_time=1e6){
  c1 = 3
  c2 = 0.5

  n = 0
  time = 0
  costs = 0
  while(time < end_time){
    n = n+1
    # 一辆汽车实际寿命
    life1 = runif(1, 0, 10)
    # 使用时间
    use1 = min(c(Tc, life1))
    cost1 = ifelse(life1 < Tc, c1+c2, c1)
    costs = costs + cost1
    time = time + use1
  }
```

```
  costs / end_time
}
## simone(T=5.0)
set.seed(101)
tv <- seq(0.1, 9.9, by=0.1)
cv <- sapply(tv, simone, end_time=1e4)
plot(tv, cv, type="l",
     xlab = "T", ylab="Mean cost")
```



模拟产生的平均费用关于截止时间 T 的曲线好像是单调减函数，这实际是因为曲线在右端斜率过小了。用理论值在右端放大作图：

```
f <- function(x) (3+x/20)/(x*(1-x/10) + x^2/20)
curve(f(x), 5, 10,
      xlab="T", ylab="Mean cost")
```



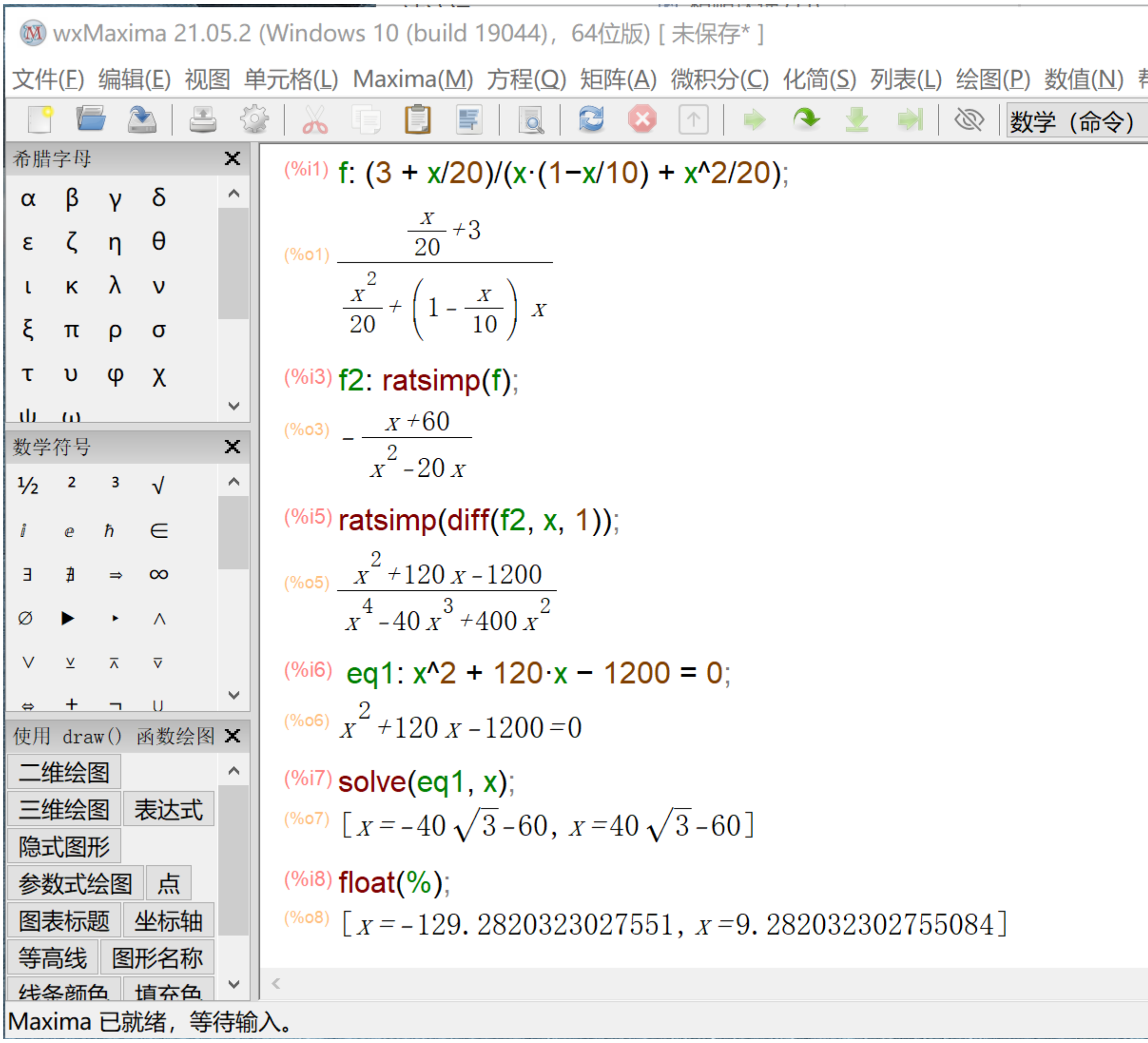

```r
optimize(f, c(0, 10))
```

```
## $minimum
## [1] 9.282025
##
## $objective
## [1] 0.6964102
```

```r
df <- function(x, eps=1e-6) (f(x + eps) - f(x))/eps
df(9.282)
```

```
## [1] -4.453105e-07
```

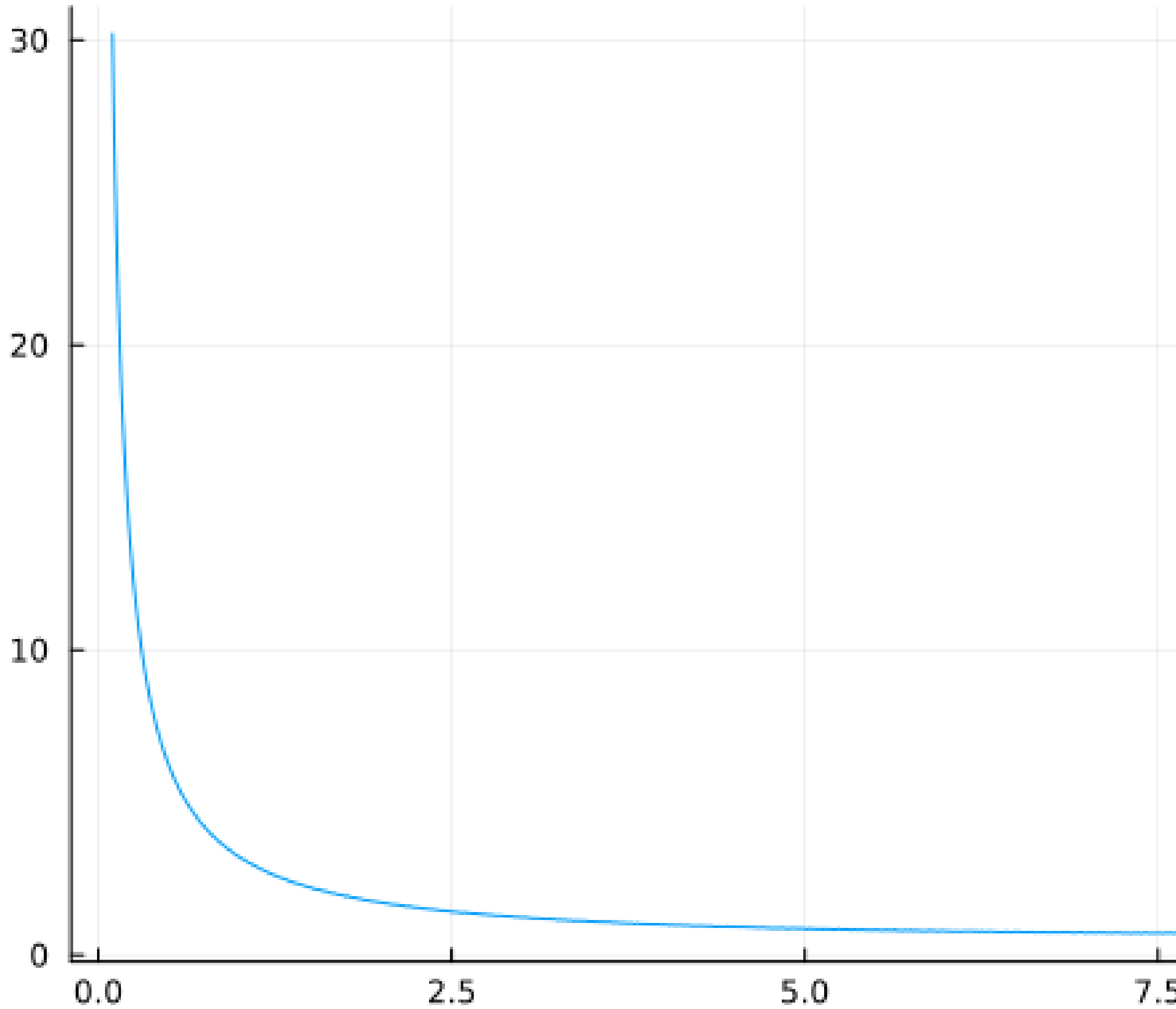
借助于符号计算软件 wxMaxima(参见https://www.math.pku.edu.cn/teachers/lidf/docs/statcomp/html/_statcompbook/maxima.html) 求导并求解最大值：



使用 Julia 编程可以提高速度，前面的 R 程序如果令模拟时间达到 10^6 则计算时间太长，无法满足绘制曲线的要求。Julia 程序的编译运行则可以使用很长的模拟时间。

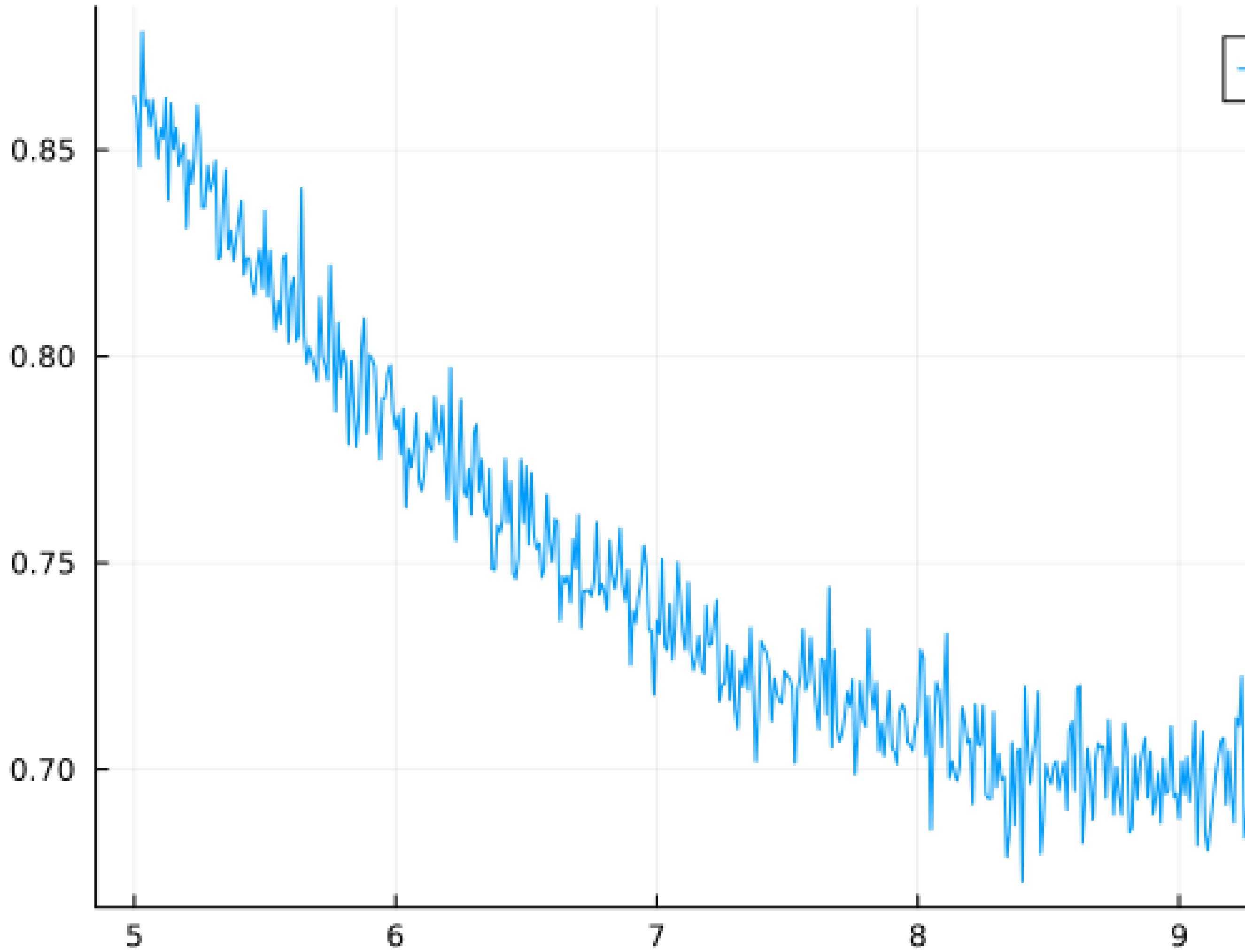
```
# 更新过程中汽车替换例子模拟
using Random, Distributions
using Statistics, StatsBase
function simone(T=5.0;
```

```julia
    end_time = 1e4, max_life=10.0, c1=3.0, c2=0.5)
    n = 0
    time = 0.0
    costs = 0.0
    while time < end_time
        n += 1
        life1 = rand()*max_life
        use1 = life1 < T ? life1 : T
        cost1 = life1 < T ? c1+c2 : c1
        costs += cost1
        time += use1
    end # while

    return costs / end_time
end # function simone
tv = range(0.1, 9.9, step=0.01)
cv = [simone(T; end_time=1e4) for T in tv];

using Plots
gr()
plot(tv, cv)
#Plots.savefig("renewal-carrepl-011.png")
```



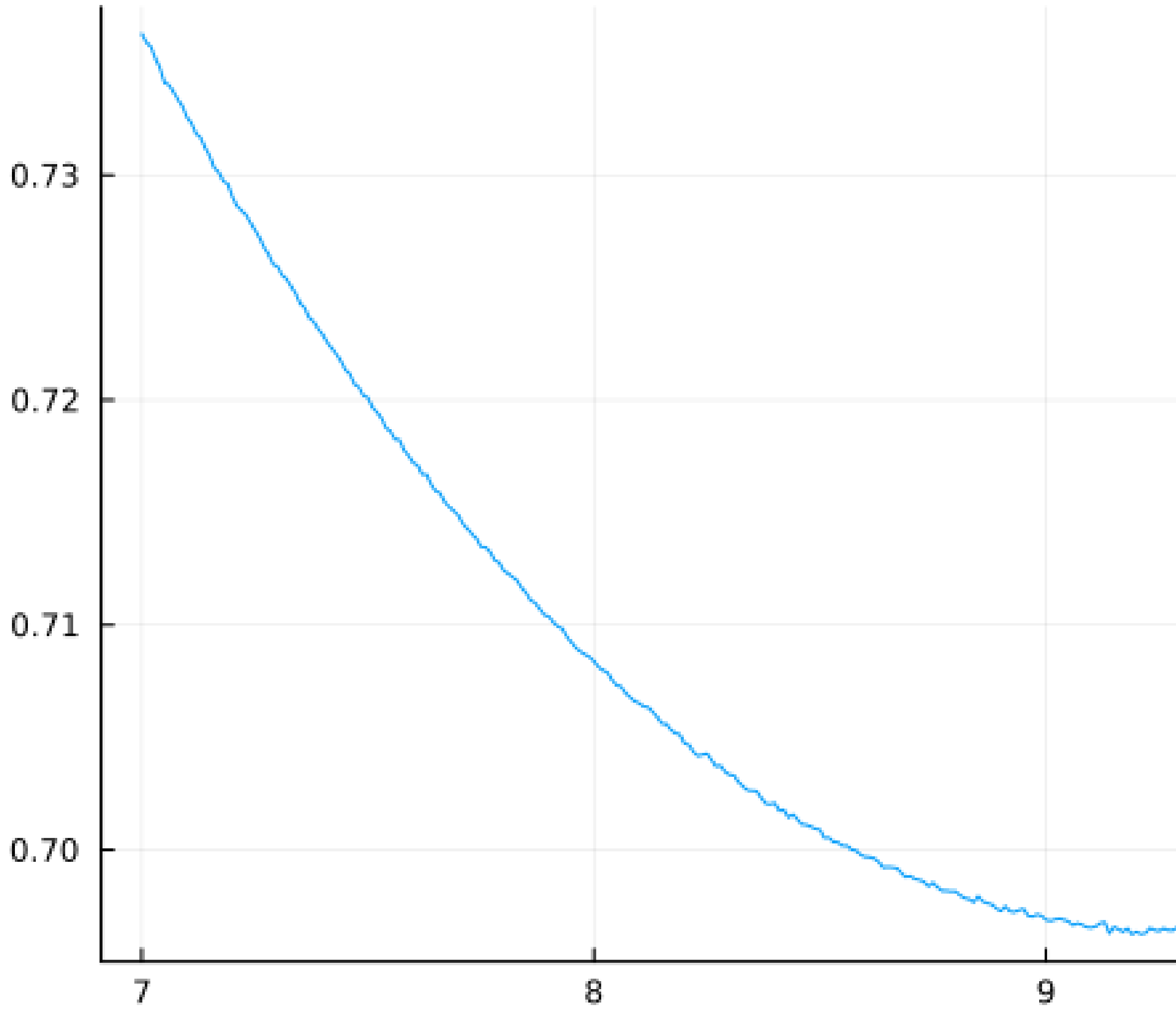
放大上图的 $T > 7.0$ 局部来看：

```
sele = tv .>= 5.0
plot(tv[sele], cv[sele])
#Plots.savefig("renewal-carrepl-012.png")
```



将这一部分重新模拟计算，模拟时间 10^8：

```
tv = range(7.0, 10.0, step=0.01)
cv = [simone(T; end_time=1e8) for T in tv];
plot(tv, cv)
Plots.savefig("renewal-carrepl-013.png")
```



4.4.3 交替更新过程

定义 4.7. 在更新过程中，我们考虑系统只有一个状态的情况, 比如机器一直是开的（即更换零件不需时间）. 而实际中，零件损坏之后，会有一个拆卸、更换的过程，这段时间机器是“关”的. 这里我们就来考虑有”开”、“关”两种状态的更新过程, 我们将这种过程称做**交替更新过程**.

注 1：设系统最初是开的，持续开的时间是 Z_1，而后关闭, 时间为 Y_1, 之后再打开，时间为 Z_2, 又关闭, 时间 Y_2, ……，交替进行, 每当系统被打开称做一次更新.

注 2：我们假设随机向量列 $\{(Z_n, Y_n),\ n = 1, 2, \ldots\}$ 是独立同分布的，从而 $\{Z_n\}, \{Y_n\}$ 都是独立同分布的，Z_i, Y_j 在 $i \neq j$ 时独立，但 Z_i, Y_i 允许不独立.

定理 4.8. 设 H 是 Z_n 的分布，G 是 Y_n 的分布，F 是 $Z_n + Y_n$ 的分布. 并记 $P(t) = P(t\text{时刻系统是开的})$，设 $E(Y_n + Z_n) < \infty$，且 F 不是格点的, 则

$$\lim_{t\to\infty} P(t) = \frac{E(Z_1)}{E(Z_1) + E(Y_1)}.$$

证明：对第一次更新的时刻 $X_1 = Z_1 + Y_1$ 取条件, 得

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

于是

$$\begin{aligned} P(t) =& E\{I[t\text{时刻系统开着}]\} \\ =& E\{E(I[t\text{时刻系统开着}] \mid X_1)\} \\ =& \int_0^\infty P(t\text{时刻系统开着}|X_1 = x)\, dF(x) \\ =& \int_t^\infty P(Z_1 > t|X_1 = x)\, dF(x) + \int_0^t P(t - x)\, dF(x) \end{aligned}$$

来证明 $x \geq t$ 时 $\int_t^\infty P(Z_1 > t|X_1 = x)\, dF(x) = \bar{H}(t)$。

$$\begin{aligned} &\bar{H}(t) = P(Z_1 > t) = P(Z_1 > t,\ X_1 > t) \\ =& EI_{\{Z_1 > t, X_1 > t\}} = E\{E(I_{\{Z_1 > t, X_1 > t\}}|X_1)\} \\ =& \int_t^\infty E(I_{\{Z_1 > t\}}|X_1 = x)\, dF(x) \\ =& \int_t^\infty P(Z_1 > t|X_1 = x)\, dF(x). \end{aligned}$$

从而

$$P(t) = \bar{H}(t) + \int_0^t P(t - x)\, dF(x). \tag{4.13}$$

(4.13)是更新方程，其解为

$$P(t) = \bar{H}(t) + \int_0^t \bar{H}(t-x)\, dM(x).$$

又 $\int_0^\infty \bar{H}(t)dt = EZ_1 < \infty$，且 $\bar{H}(t)$ 非负不增，由关键更新定理得

$$\lim_{t\to\infty} P(t) = \frac{\int_0^\infty \bar{H}(t)\, dt}{E(Y_1 + Z_1)} = \frac{EZ_1}{EZ_1 + EY_1}.$$

例 4.12. 例题来自 Ross Introduction to Probablity Models 12th Ed P.439 例 7.7。

设有 n 个人博弈，每次随机选择一个人胜出，选第 i 个人的概率为 $p_i(i = 1,2,\dots,n)$，当连续 k 次都为同一个人胜出时博弈完成一局，此局以该连续胜出者为胜方。比如，$k = 2$ 时各次结果若为 $1,2,4,3,5,2,1,3,3$ 则最后第 3 个人胜出，一局用时 9。问：第 i 个人最终胜出的概率是多少，每局平均需要的时间是多少？

解答先考虑简单的只有两个人的情况，第一个人每次胜出的概率为 $p = p_1$，进行独立重复试验直到第一个人连续 k 次胜出（不考虑第二个人连续 k 次胜出问题）。

用 T 表示第一个人连续 k 次胜出时的时间，记 Z 为首次出现非第一个人胜出的时刻，即前 $Z-1$ 次都是第一个人胜出，用 $ET = E(E(T|Z))$ 求 ET。如果 $Z \geq k+1$，则开始的 k 次博弈为连续的第一个人胜出，这时 $T = k$; 对 $Z = 1,2,\dots,k$，前面的 $Z-1$ 次都是第一个人胜出，第 Z 次为第二个人胜出，从第 $Z+1$ 次第一个人连续 k 次胜出的概率变成与重新开始概率相同，这样得

$$\begin{aligned}
E(T) =& E[E(T|Z)] \\
=& \sum_{j=1}^{\infty} E(T|Z=j)P(Z=j) \\
=& \sum_{j=1}^{k} E(T|Z=j)P(Z=j) + kP(Z \geq k+1) \\
=& \sum_{j=1}^{k} [j + ET]p^{j-1}(1-p) + kp^k,
\end{aligned}$$

求解 ET 得

$$ET = k + \frac{1-p}{p^k}\sum_{j=1}^{k} jp^{j-1}.$$

化简得

$$\begin{aligned}
ET =& \frac{kp^k + (1-p)\sum_{j=1}^{k} jp^{j-1}}{p^k} \\
=& \frac{kp^k + \sum_{j=1}^{k} jp^{j-1} - \sum_{j=1}^{k} jp^j}{p^k} \\
=& \frac{kp^k + \sum_{m=0}^{k-1}(m+1)p^m - \sum_{j=1}^{k} jp^j}{p^k} \\
=& \frac{kp^k + p^0 + \sum_{m=1}^{k-1} p^m - kp^k}{p^k} \\
=& \frac{\sum_{m=0}^{k-1} p^m}{p^k} \\
=& \frac{1-p^k}{p^k(1-p)}.
\end{aligned}$$

回到原始的问题，设有 n 个人博弈，每次有任何一个人连续 k 次胜出时重新开始博弈，这样实际上每次博弈都是一个独立重复试验，与当前在哪一局没有关系。在这样的独立重复试验中，对给定 i，将第 i 人连续 k 次胜出作为一次更新，这样关于第 i 个人构成更新过程，用 N_i 表示第 i 人完成一局的时间，由上面的推导可知

$$EN_i = \frac{1-p_i^k}{p^k(1-p_i)}.$$

由更新定理，第 i 个人胜出的速率（平均每次博弈胜出比例）为

$$\lim_{t\to\infty}\frac{M(t)}{t} = \frac{1}{EN_i} = \frac{p_i^k(1-p_i)}{1-p_i^k}.$$

其中 $N(t)$ 是 t 次博弈中第 i 人胜出的局数。

因为每个人都可能先胜出 k 次博弈，每个人先胜出的概率与其胜出的速率成正比（这里是数学直观不是严格证明），所以长期来看第 i 个人胜出的概率为

$$\begin{aligned}
&\frac{\text{第}i\text{个人胜出速率}}{\sum_{j=1}^{n}\text{第}j\text{个人胜出速率}} \\
=&\frac{\frac{p_i^k(1-p_i)}{1-p_i^k}}{\sum_{j=1}^{n}\frac{p_j^k(1-p_j)}{1-p_j^k}}.
\end{aligned}$$

注意到每局决出胜负的时间是所有 n 个人先胜出 k 次博弈的时间的最小值，所以将有人胜出作为一次更新，更新的速率是 n 个速率之和（这需要严格证明）：

$$\sum_{j=1}^n \frac{p_j^k (1-p_j)}{1-p_j^k},$$

以 T 为有人先胜出 k 次博弈的时间，则由更新定理

$$\sum_{j=1}^n \frac{p_j^k (1-p_j)}{1-p_j^k} = \frac{1}{ET},$$

即

$$ET = \left(\sum_{j=1}^n \frac{p_j^k (1-p_j)}{1-p_j^k} \right)^{-1}.$$

因为上面的推导不够严格，用随机模拟方法来验证。取 $n=3$, $(p_1, p_2, p_3) = (1/2, 1/3, 1/6)$, $k=2$, 每个人的胜出速率为：

```
k <- 2
n <- 3
pv <- 1 / c(2, 3, 6)
rates <- pv^k*(1-pv)/(1 - pv^k)
rates
```

```
## [1] 0.16666667 0.08333333 0.02380952
```

结束一局时每个人的胜出概率为：

```
rates / sum(rates)
```

```
## [1] 0.60869565 0.30434783 0.08695652
```

平均每局时间：

```
1 / sum(rates)
```

```
## [1] 3.652174
```

模拟一条长度为 N 的轨道：

```
simren_one <- function(N = 10000, k = 2){
  n <- 3
```

```r
pv <- 1 / c(2, 3, 6)

# 每次博弈的胜出者
xv <- sample(1:n, size=N, replace=TRUE, prob = pv)
nr <- 0 # 总局数
wv <- numeric(N) # 每局的胜出人代号
tv <- numeric(N) # 每局的时间长度

# 下面循环找到每一局，条件是最后的 k 个为相同胜者
istart <- 1
i <- k
##print(xv)
while(i <= N){
  if(all(xv[(i-k+1):i] == xv[i])){
    ## 找到 k 个连续，结束一局，开始新的一局
    nr <- nr + 1
    wv[nr] <- xv[i]
    tv[nr] <- i - istart + 1
    istart <- i+1
    i <- i+k
  } else {
    i <- i+1
  }
}
wv <- wv[1:nr]
tv <- tv[1:nr]

## 模拟估计的胜出概率与理论概率:
print(rbind(prop.table(table(wv)),
            rates / sum(rates)))

## 模拟估计的每局时间长度与理论平均长度
print(c(mean(tv), 1 / sum(rates)))
```

```
  invisible(list(winners=wv, lengths=tv) )
}
simren_one(N=1000000)
## 模拟估计的胜出概率与理论概率:
##              1         2          3
## [1,] 0.6092890 0.3043518 0.08635923
## [2,] 0.6086957 0.3043478 0.08695652
## 模拟估计的每局时间长度与理论平均长度:
## [1] 3.658511 3.652174
```

可见模拟估计的值与理论值比较接近，但是需要模拟很长时间（这里时间长度为一百万次博弈）。

使用 Julia 语言编程：

```
# 多人连胜游戏。每次某人连胜 k 轮时结束一局比赛。
# 输入:
#   pw: 每人的获胜概率
#   k: 连胜多少轮算一局结束
#   N: 模拟重复的轮数
using Random, Distributions
using Statistics, StatsBase
function simone(pw::Vector; k=2, N=1000_000)
    n = length(pw)
    rates = pw .^ k .* (1 .- pw) ./ (1 .- pw .^ k)
    # 每人理论获胜概率
    probs = rates ./ sum(rates)
    # 平均每局时间的理论值
    mean_time = 1 / sum(rates)

    # 一次性生成 N 轮的输赢
    xv = StatsBase.sample(1:n, StatsBase.Weights(pw),
        N; replace=true)
    winners = Int[]
```

```
    times = Int[]

    # 下面循环找到每一局，条件是最后的 k 个为相同胜者
    istart = 1
    i = k
    while i <= N
        if all(xv[i-k+1:i-1] .== xv[i])
            ## 找到 k 个连续，结束一局，开始新的一局
            push!(winners, xv[i])
            push!(times, i - istart + 1)
            istart = i+1
            i += k
        else
            i += 1
        end
    end # while i

    # 获胜者频率估计
    freqs = counts(winners, 1:n)
    probs_est = freqs ./ sum(freqs)
    # 平均每局时间估计
    mean_time_est = mean(times)

    println(" 理论概率与模拟估计概率: ")
    display(round.([probs probs_est], digits=6))
    println(" 理论平均时间与模拟估计值: ")
    show(round.([mean_time, mean_time_est], digits=4))

    return
end

function test(N=1000)
    n = 3
```

```
    pw = [1/2, 1/3, 1/6]
    k = 2
    simone(pw, k=k, N=N)
end
Random.seed!(101)
test(1000_000)
## 理论概率与模拟估计概率:
## 3×2 Matrix{Float64}:
## 0.608696  0.608658
## 0.304348  0.303514
## 0.086957  0.087829
## 理论平均时间与模拟估计值:
## [3.6522, 3.6565]
```

4.5 附录：一些证明

4.5.1 引理4.1证明

令 $S_n = \sum_{j=1}^n a_j$，则 S_n 为非负单调增数列。如果 S_n 有上界，则必有上确界，为 S_n 的极限。如果 S_n 无上界，即级数等于 $+\infty$。

设 $j_i, i = 1, 2, \dots$ 是 $1, 2, \dots$ 的排列，即 j_i 是 $\{1, 2, \dots\}$ 上的一个一一映射。考虑级数 $\sum_{i=1}^\infty a_{j_i}$。分两种情况讨论。

如果 $\sum_{i=1}^\infty a_i = A < \infty$，则 $A > 0$(若 $A = 0$ 即所有 $a_i = 0$，结论显然)，对任意正整数 n 有

$$\sum_{i=1}^n a_{j_i} \leq A,$$

所以级数 $\sum_{i=1}^\infty a_{j_i} \leq A < \infty$。$\forall \epsilon > 0$，$\exists N$ 使得 $n \geq N$ 时

$$\sum_{i=1}^n a_i > A - \epsilon.$$

$\exists N_1 \geq N$ 使得

$$\{j_i, i = 1, 2, \dots, N_1\} \supset \{1, 2, \dots, N\},$$

于是 $n \geq N_1$ 时

$$\sum_{i=1}^{n} a_{j_i} \geq \sum_{i=1}^{N} a_i \geq A - \epsilon,$$

进一步有

$$\sum_{i=1}^{\infty} a_{j_i} \geq A - \epsilon,$$

于是

$$A - \epsilon \leq \sum_{i=1}^{\infty} a_{j_i} \leq A,$$

令 $\epsilon \to 0$ 即得

$$\sum_{i=1}^{\infty} a_{j_i} = A = \sum_{i=1}^{\infty} a_i.$$

如果 $\sum_{i=1}^{\infty} a_i = +\infty$，则 $\forall M > 0$，存在 N 使得 $n \geq N$ 时

$$\sum_{i=1}^{n} a_i > M.$$

于是存在 $N_1 \geq N$ 使得

$$\{j_i, i = 1, 2, \dots, N_1\} \supset \{1, 2, \dots, N\},$$

从而 $n \geq N_1$ 时

$$\sum_{i=1}^{n} a_{j_i} \geq \sum_{i=1}^{N} a_i > M,$$

于是 $\sum_{i=1}^{\infty} a_{j_i} = +\infty$。

4.5.2 引理4.4证明

此引理证明较长，分为若干段。

(1) 考虑间隔时间 X_1 的分布函数 $F(x)$。简记 X_1 为 X。因为 $F(0) < 1$，所以 $P(X > 0) > 0$，而

$$\begin{aligned}
P(X > 0) =& P\left(\bigcup_{\ell=1}^{\infty} \{X > \frac{1}{\ell}\}\right) \\
=& \lim_{\ell \to \infty} P(X > \frac{1}{\ell}) > 0,
\end{aligned}$$

存在 $\ell > 1$ 使得 $P(X > \frac{1}{\ell}) > 0$，记 $b = \frac{1}{\ell}$，则 $0 < b < 1$，且

$$F(b) = P(X \leq \frac{1}{\ell}) < 1. \tag{4.14}$$

(2) 得到上述的 $0 < b < 1$ 使得 $F(b) < 1$ 以后，考虑任意固定的 t，下面的讨论都假定 t 为某个固定值，对此 t 必存在正整数 k 使得 $kb > t$，这时有

$$\{T_k \leq t\} \subset \{T_k \leq kb\} \subset \{X_1 > b, \dots, X_k > b\}^c,$$

(k 个更新时间之和小于等于 kb 则不能每个间隔时间都大于 b)，所以

$$\begin{aligned}
F_k(t) =& P(T_k \leq t) \\
\leq& 1 - P(X_1 > b, \dots, X_k > b) \\
=& 1 - [1 - F(b)]^k \\
=& 1 - c_1,
\end{aligned}$$

其中 $c_1 = [1 - F(b)]^k > 0$, 令 $c_2 = 1 - c_1$, 则 $0 \leq c_2 < 1$ 使得

$$F_k(t) \leq c_2. \tag{4.15}$$

(3) 将更新的发生分成每 k 个一组，考察第 mk 次更新时间 T_{mk} 的分布 $F_{mk}(t)$, $m = 1, 2, \dots$。因为

$$\{T_{mk} \leq t\} \subset \{T_k - T_0 \leq t, T_{2k} - T_k \leq t, \dots, T_{mk} - T_{(m-1)k} \leq t\},$$

这是由于总更新时间小于等于 t 则更新分成 m 段后每一段的更新时间也小于等于 t。其中 $T_0 = 0$。

由更新过程的间隔时间独立同分布性质，可得

$$P(T_{mk} \leq t) \leq [P(T_k \leq t)]^m = [F_k(t)]^m \leq c_2^m.$$

(4) 对正整数 n，必存在非负整数 m 和 $0 \leq j \leq k - 1$ 使得

$$n = mk + j,$$

实际上 m 是 n 除以 k 的整除商，j 是余数。这时，因为

$$\{T_n \leq t\} \subset \{T_{mk} \leq t\},$$

所以

$$\begin{aligned}F_n(t) =&P(T_n \le t)\\ \le&P(T_{mk} \le t) \le c_2^m.\end{aligned}$$

如果 $c_2 = 0$，则 $F_n(t) = 0$，定理结论成立，下面假设 $0 < c_2 < 1$。

利用

$$\frac{n}{k} - 1 < m \le \frac{n}{k}$$

可得

$$c_2^m \le c_2^{\frac{n}{k}-1} = c_2^{-1}(c_2^{\frac{1}{k}})^n,$$

令

$$\alpha = c_2^{\frac{1}{k}}, C = c_2^{-1} > 0,$$

则 $0 < \alpha < 1$，使得

$$F_n(t) \le C\alpha^n,$$

引理得证。

Chapter 5

马氏链

5.1 基本概念

有一类随机过程，它具备所谓的 “无后效性”(马氏性)，即要确定过程将来的状态，知道它此刻的情况就足够了，并不需要对它以往状况的认识，这类过程称为 Markov 过程. 我们将介绍 Markov 过程中最简单的两种类型：离散时间的马氏链 (简称马氏链) 及连续时间的马氏链.

5.1.1 定义和例子

定义 5.1. 随机过程 $\{X_n, n = 0, 1, 2, \ldots\}$ 称为**马氏链**，若它只取有限或可列个值 (若不另外说明，以非负整数集 $\{0, 1, 2, \ldots\}$ 来表示)，并且对任意的 $n \geq 0$，及任意状态 $i, j, i_0, i_1, \ldots, i_{n-1}$，有

$$P\{X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i\} = P\{X_{n+1} = j \mid X_n = i\} \tag{5.1}$$

其中 $X_n = i$ 表示过程在时刻 n 处于状态 i，称 $\{0, 1, 2, \ldots\}$ 为该过程的**状态空间**，记为 S. 式(5.1)刻画了马氏链的特性，称为**马氏性**。

条件独立性：设 A, B, C 为随机事件，若 $P(BC) > 0$, 则

$$P(A|BC) = P(A|B)$$

等价于

$$P(AC|B) = P(A|B)P(C|B),$$

这时称给定 B 的条件下 A 和 C 相互独立。集合函数 $P_B(A) = P(A|B)$（$A \in \mathscr{F}$）是一个概率测度，也有乘积公式、全概率公式等性质。

对随机变量 X, Y, Z，若

$$P(X \le x, Y \le y|Z = z) = P(X \le x|Z = z)P(Y \le y|Z = z)$$

则称 X 和 Y 在给定 Z 的条件下独立。X, Y, Z 都服从离散分布时，条件独立性等价于

$$P(X = x, Y = y|Z = z) = P(X = x|Z = z)P(Y = y|Z = z),$$

也等价于

$$P(X = x|Y = y, Z = z) = P(X = x|Z = z).$$

条件(5.1)等价于 X_{n+1} 和 $X_0, \dots, X_{n-1}$ 在给定 X_n 的条件下条件独立。

定义 5.2. 称(5.1)式中的条件概率 $P\{X_{n+1} = j|X_n = i\}$ 为马氏链 $\{X_n, n = 0, 1, 2, \dots\}$ 的**一步转移概率**，简称为转移概率.

一般情况下，转移概率与状态 i, j 和时刻 n 有关.

定义 5.3. 当马氏链的一步转移概率 $P\{X_{n+1} = j|X_n = i\}$ 只与状态 i, j 有关，而与 n 无关时，称此马氏链为**时齐马氏链**，记 $P\{X_{n+1} = j|X_n = i\} = p_{ij}$，代表处于状态 i 的过程下一步转移到状态 j 的概率. 若 $P\{X_{n+1} = j|X_n = i\}$ 与 n 有关，就称该马氏链为非时齐的.

在本书中，我们只讨论时齐马氏链，并且简称为马氏链.

当马氏链的状态为有限时，称为有限链，否则称为无限链. 但无论状态有限还是无限，我们都可以将 $p_{ij}(i, j \in S)$ 排成一个矩阵的形式。令

$$P = (p_{ij}) = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

称 P 为**转移概率矩阵**，一般简称为转移矩阵. 由于概率是非负的，且过程必须转移到某个状态，所以容易看出 $p_{ij}(i,j\in S)$ 有性质

$$\begin{aligned}&(1)p_{ij}\geq 0,\ i,j\in S;\\&(2)\sum_{j\in S}p_{ij}=1,\ \forall i\in S.\end{aligned}\tag{5.2}$$

定义 5.4. 称矩阵为**随机矩阵**，若矩阵元素具有(5.2)) 式中两条性质.

易见随机矩阵每一行元素的和都为 1.

例 5.1 (一个简单的疾病、死亡模型，Fix-Neyman)**.** 考虑一个包含两个健康状态 S_1,S_2 以及两个死亡状态 S_3,S_4 (即由不同原因引起的死亡) 的模型. 若个体病愈，则认为它处于状态 S_1，若患病，说处于 S_2，个体可以从 S_1, S_2 进入 S_3 和 S_4，易见这是一个马氏链的模型，转移矩阵为

$$P=\begin{pmatrix}p_{11}&p_{12}&p_{13}&p_{14}\\p_{21}&p_{22}&p_{23}&p_{24}\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

例 5.2 (赌徒的破产或称带吸收壁的随机游动)**.** 系统的状态是 $0\sim n$，反映赌博者在赌博期间拥有的钱数，当他输光或拥有钱数为 n 时，赌博停止，否则他将持续赌博. 每次以概率 p 赢得 1，以概率 $q=1-p$ 输掉 1. 这个系统的转移矩阵为

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

例 5.3 (带反射壁的随机游动)**.** 设上例中当赌博者输光时将获得赞助 1 让他继续赌下去，就如同一个在直线上做随机游动的球在到达左侧 0 点处就立刻反弹

回 1 一样，这就是一个一侧带有反射壁的随机游动. 此时转移矩阵为

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

同样可考虑两侧均有反射壁的情况.

例 5.4 (自由随机游动)**.** 设一个球在全直线上做无限制的随机游动，它的状态为 $0, \pm1, \pm2, \ldots.$ 它仍是一个马氏链，转移矩阵为

$$P=\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdots & q & 0 & p & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdots & 0 & q & 0 & p & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \ddots & \ddots & \ddots & & & & & \\ \cdots & \cdot & \cdot & \cdot & q & 0 & p & 0 & \cdot & \cdot & \cdot \\ \cdots & \cdot & \cdot & \cdot & 0 & q & 0 & p & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

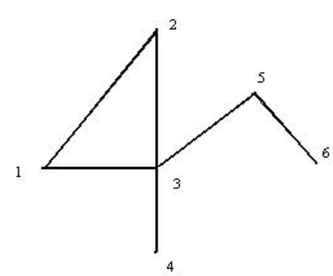


图 5.1: 图上的简单随机游动

例 5.5 (图上的简单随机游动)**.** 设有一蚂蚁在如图5.1上爬行，当两个结点相临时，蚂蚁将爬向它临近的一点，并且爬向任何一个邻居的概率是相同的. 则此马

氏链的转移矩阵为

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

例 5.6 (Wright-Fisher 遗传模型)**.** 遗传的要素是染色体. 遗传性质的携带者称为基因，它们位于染色体上. 基因控制着生物的特征，它们是成对出现的. 控制同一特征的不同基因称为等位基因，记这对等位基因为 A 和 a，分别称为显性的与隐性的. 在一个总体中基因 A 和 a 出现的频率称为基因频率，分别记为 p 和 $1-p$.

设总体中的个体数为 $2N$，每个个体的基因按 A 型基因的基因频率的大小，在下一代中转移成为 A 型基因. 因此，繁殖出的第二代的基因型是由试验次数为 $2N$ 的 Bernoulli 试验所确定的，即如果在第 n 代母体中 A 型基因出现了 i 次，而 a 型基因出现了 $2N-i$ 次，则下一代出现 A 型基因的概率为 $p_i = \frac{i}{2N}$，而出现 a 型基因的概率为 $1-p_i$.

记 X_n 为第 n 代中携带 A 型基因的个体数，则易知 $\{X_n\}$ 是一个状态空间为 $S=\{0,1,2,\dots,2N\}$ 的时齐马氏链，其转移概率矩阵为 $P=(p_{ij})$，其中

$$\begin{aligned} p_{ij} =& P\{X_{n+1}=j|X_n=i\} = C_{2N}^j p_i^j (1-p_i)^{2N-j} \\ =& C_{2N}^j \left(\frac{i}{2N}\right)^j \left(1-\frac{i}{2N}\right)^{2N-j} \end{aligned}$$

下面我们再给出几个所谓“嵌入马氏链”的例子，在这些情况下模型的马氏性不是明显的.

例 5.7 (M/G/1 排队系统)**.** 假设顾客依照参数为 λ 的 Poisson 过程来到一个只有一名服务员的服务站，若服务员空闲则顾客就能立刻得到服务, 否则排队等待直至轮到该顾客. 设每名顾客接受服务的时间是独立的随机变量, 有共同的分布 G，而且与来到过程独立. 这个系统称为 $M/G/1$ 排队系统，字母 M 代表顾客来到的间隔服从指数分布，G 代表服务时间的分布，数字 1 表示只有 1 名服务员. 考虑其中嵌入的马氏链。

解:

若以 $X(t)$ 表示时刻 t 系统中的顾客人数，则 $\{X(t), t \geq 0\}$ 是不具备马氏性的，因为若已知 t 时刻系统中的人数，要预测未来，虽然可以不用关心从最近的一位顾客到来又过去了多长时间 (因过程无记忆，所以这段时间不影响下一位顾客的到来)，但要注意此刻在服务中的顾客已经接受了多长时间的服务 (因为 G 不是指数的，不具备 "无记忆性"，所以已经服务过的时间将影响到他何时离去).

我们可以这样考虑，令 X_n 表示第 n 位顾客走后剩下的顾客数，$n \geq 1$. 再令 Y_n 记第 $n+1$ 位顾客接受服务期间到来的顾客数，则

$$X_{n+1} = \begin{cases} X_n - 1 + Y_n, & \text{若} X_n > 0, \\ Y_n, & \text{若} X_n = 0. \end{cases} \tag{5.3}$$

记 T_n 为第 n 位顾客到来的时刻，$N(t)$ 为截止到 t 时刻顾客到来的人数，则 $\{N(t)\}$ 与 T_n 是参数为 λ 的泊松过程以及相应的事件到来时刻。

如果 $X_n = 0$，这表示第 n 位顾客接受完服务离开时没有其他顾客等候，服务员变空闲，直到 T_{n+1} 时第 $n+1$ 位顾客到来，这位顾客是当时的唯一顾客，马上接受服务，在第 $n+1$ 位接受服务期间又到来 Y_n 位顾客排队等候，所以在 $X_n = 0$ 条件下第 $n+1$ 位离开时系统中有 Y_n 位顾客。

如果 $X_n > 0$，则第 n 位顾客接受完服务离开时排队的队首顾客是第 $n+1$ 位到来的顾客，这位顾客马上接受服务，还有 $X_n - 1$ 位顾客排队等待，在第 $n+1$ 位顾客接受服务期间又到来 Y_n 位顾客加入排队，所以在 $X_n > 0$ 条件下在第 $n+1$ 位顾客离开时队伍中共有 $X_n - 1 + Y_n$ 人。

举例说明。设 $n = 5$，X_5 表示第 5 位顾客走后剩余的顾客数，为一个非负整数值。X_6 表示第 6 位顾客走后剩余的顾客数，这依赖于 X_5 的值和第 5 位顾客走后、第 6 位顾客走之前新到来的顾客人数 Y_5。分为两种情况，若 $X_5 = 0$，则第 6 位顾客到来时直接接受服务，这期间又来了 Y_5 位顾客，第 6 位顾客离开时还剩 Y_5 位顾客；若 $X_5 > 0$，则第 5 位顾客离开时第 6 位顾客是排在队首的顾客，这时除了第 6 位顾客还有 $X_5 - 1 \geq 0$ 位顾客，在第 6 位顾客接受服务期间又来了 Y_5 位顾客，第 6 位离开时还剩下原来队伍中的 $X_5 - 1$ 位顾客和新来的 Y_5 位顾客。

可以证明 Y_n 的分布为：

$$P\{Y_n = j\} = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} \, dG(x), \ j = 0, 1, 2, ... \tag{5.4}$$

设第 $n+1$ 位顾客接受服务的时间长度为 Z, $Z \sim G(\cdot)$，在 $Z=x$ 条件下，接受服务期间到达的顾客数 Y_n 服从泊松 $\mathrm{P}(\lambda x)$ 分布，于是

$$P(Y_n = j|Z = x) = e^{-\lambda x}\frac{(\lambda x)^j}{j!},$$

由全期望公式，

$$\begin{aligned}
P(Y_n = j) =& E[P(Y_n = j|Z)] \\
=& \int_0^\infty P(Y_n = j|Z = x)\, dG(x) \\
=& \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^j}{j!}\, dG(x).
\end{aligned}$$

这给出 $\{Y_n\}$ 的分布，此分布并不依赖 $X_0, X_1, \dots, X_n$ 的值，所以 $P(X_{n+1} = j|X_n = i)$ 与给定 $X_0, X_1, \dots, X_n$ 的条件概率相同，X_n 马氏性成立。所以由(5.3)、(5.4)式得 $\{X_n, n = 1, 2, \dots\}$ 是马氏链，转移概率为

$$\begin{aligned}
p_{0j} =& P(X_{n+1} = j|X_n = 0) = P(Y_n = j|X_n = 0) = P(Y_n = j) \\
=& \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^j}{j!} dG(x), \quad j \geq 0 \\
p_{ij} =& P(X_{n+1} = j|X_n = i) \qquad (i \geq 1) \\
=& P(Y_n + i - 1 = j|X_n = i) \\
=& P(Y_n = j - i + 1) \qquad (i \geq 1, j \geq i - 1) \\
=& \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^{j-i+1}}{(j-i+1)!} dG(x),\ j \geq i-1, i \geq 1, \\
p_{ij} =& 0,\ \text{其他}
\end{aligned}$$

例 5.8 (订货问题)**.** 考虑定货问题. 设某商店使用 (s, S) 定货策略，每天早上检查某商品的剩余量，设为 x, 则定购额为

$$\begin{cases} 0, & \text{若} x \geq s, \\ S - x, & \text{若} x < s. \end{cases} \tag{5.5}$$

设定货和进货不需要时间，每天的需求量 Y_n 独立同分布且 $P\{Y_n = j\} = a_j, j = 0, 1, 2, \dots$. 现在我们要从上述问题中寻找一个马氏链.

解: 令 X_n 为第 n 天结束时的存货量，因为要考虑需求量 Y_{n+1} 大于存货量 X_n 的情况，所以第 $n+1$ 天结束时的存货量为

$$X_{n+1} = \begin{cases} \max(0, X_n - Y_{n+1}), & X_n \geq s, \\ \max(0, S - Y_{n+1}), & X_n < s. \end{cases}$$

对 $X_n = i \geq s$，有

$$X_{n+1} = \begin{cases} i - Y_{n+1}, & 0 \leq Y_{n+1} < i \\ 0, & Y_{n+1} \geq i, \end{cases}$$

所以对 $1 \leq j \leq i$,

$$\begin{aligned} p_{ij} =& P(X_{n+1} = j | X_n = i) \\ =& P(Y_{n+1} = i - j) = a_{i-j}, \end{aligned}$$

其中 $0 \leq i - j < i$。而对 $j = 0$ 有

$$\begin{aligned} p_{i0} =& P(i - Y_{n+1} \leq 0) = P(Y_{n+1} \geq i) \\ =& 1 - \sum_{k=0}^{i-1} a_k = 1 - \sum_{j=1}^{i} p_{ij}. \end{aligned}$$

当 $0 \leq i < s$ 时，

$$X_{n+1} = \begin{cases} S - Y_{n+1}, & 0 \leq Y_{n+1} < S \\ 0, & Y_{n+1} \geq S, \end{cases}$$

所以对 $1 \leq j \leq S$,

$$\begin{aligned} p_{ij} =& P(X_{n+1} = j | X_n = i) \\ =& P(Y_{n+1} = S - j) = a_{S-j}, \end{aligned}$$

其中 $0 \leq S - j < S$。而 $j = 0$ 时

$$\begin{aligned} p_{i0} =& P(S - Y_{n+1} \leq 0) = P(Y_{n+1} \geq S) \\ =& 1 - \sum_{k=0}^{S-1} a_k = 1 - \sum_{j=1}^{S-1} p_{ij}. \end{aligned}$$

总之有

$$p_{ij}=\begin{cases}a_{i-j}, & j=1,2,\dots,i,s\le i\le S\\ 1-\sum_{k=0}^{i-1}a_k, & j=0,s\le i\le S\\ a_{S-j}, & j=1,2,\dots,S,0\le i<s\\ 1-\sum_{k=0}^{S-1}a_k, & j=0,0\le i<s.\end{cases}$$

例 5.9. 以 S_n 表示保险公司在时刻 n 的盈余，这里的时间以适当的单位来计算 (如天, 月等). 初始盈余 $S_0=x$ 显然为已知，但未来的盈余 $S_1,S_2,\dots$ 却必须视为随机变量，增量 S_n-S_{n-1} 解释为时刻 $n-1$ 和时刻 n 之间获得的盈利，取整数值 (可以为负). 假定 $X_1,X_2,\dots$ 是不包含利息的盈利且独立同分布，概率质量函数为 $p(x)$，则

$$S_n=S_{n-1}(1+\gamma)+X_n,$$

其中 γ 为固定的利率, $\{S_n\}$ 是一马氏链，转移概率为

$$p_{xy}=p(y-(1+\gamma)x).$$

5.1.2 多步转移概率与 C-K 方程

命题 5.1. 设 $\{X_n,n=0,1,\dots\}$ 是时齐离散时间离散状态马氏链，状态空间为 S，则对任意 $n\ge 0$, $m\ge 1$ 和状态 $x_0,\dots,x_{n-1},i,j\in S$，有

$$\begin{aligned}&P(X_{n+m}=j|X_n=i,X_{n-1}=x_{n-1},\dots,X_0=x_0)\\ =&P(X_{n+m}=j|X_n=i)\\ =&p_{ij}^{(m)},\end{aligned}$$

与 n 无关。

当 $m=1$ 时就是时齐马氏性的定义。

证明：对 m 用数学归纳法。已知 $m=1$ 时成立。设已知结论对 $1,\dots,m$ 成立，

来证明结论对 $m+1$ 成立。

$$
\begin{aligned}
&P(X_{n+m+1}=j|X_n=i,X_{n-1}=x_{n-1},\dots,X_0=x_0)\\
=&\sum_{k\in S}P(X_{n+m+1}=j,X_{n+m}=k|X_n=i,X_{n-1}=x_{n-1},\dots,X_0=x_0)\\
=&\sum_{k\in S}P(X_{n+m+1}=j|X_{n+m}=k,X_n=i,X_{n-1}=x_{n-1},\dots,X_0=x_0)\\
&\quad P(X_{n+m}=k|X_n=i,X_{n-1}=x_{n-1},\dots,X_0=x_0)\\
=&\sum_{k\in S}P(X_{n+m+1}=j|X_{n+m}=k)P(X_{n+m}=k|X_n=i)\\
=&\sum_{k\in S}p_{kj}p_{ik}^{(m)}
\end{aligned}
$$

与 n 和 $x_{n-1},\dots,x_0$ 无关，所以在 $X_n=i$ 条件下 X_{n+m+1} 与 $X_{n-1},\dots,X_0$ 条件独立，有

$$
\begin{aligned}
&P(X_{n+m+1}=j|X_n=i,X_{n-1}=x_{n-1},\dots,X_0=x_0)\\
=&P(X_{n+m+1}=j|X_n=i)=p_{ij}^{(m+1)}
\end{aligned}
$$

与 n 无关。定理证毕。

定义 5.5. 称条件概率

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

为马氏链的n **步转移概率**，相应地称 $P^{(n)}=(p_{ij}^{(n)})$ 为n **步转移矩阵**.

当 $n=1$ 时, $p_{ij}^{(1)}=p_{ij}$, $P^{(1)}=P$，此外规定

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

显然，n 步转移概率 $p_{ij}^{(n)}$ 指的就是系统从状态 i 经过 n 步后转移到 j 的概率，它对中间的 $n-1$ 步转移经过的状态无要求.

下面的定理给出了 $p_{ij}^{(n)}$ 和 p_{ij} 的关系.

定理 5.1 (Chapman-Kolmogorov 方程，简称 C-K 方程). 对一切 $n,m\geq 0$,

$i,j \in S$ 有

$$(1)\ p_{ij}^{(m+n)} = \sum_{k\in S} p_{ik}^{(m)} p_{kj}^{(n)}; \tag{5.6}$$

$$(2)\ P^{(n)} = P^n. \tag{5.7}$$

证明：

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

(2) 是 (1) 的矩阵形式，利用矩阵乘法易得.

推论 5.1. 对于正整数 $n,m,k,n_1,n_2,...,n_k$ 和状态 i,j,l，总有

$$\begin{aligned}
&(1)\ p_{ij}^{(n+m)} \ge p_{il}^{(n)}p_{lj}^{(m)};\\
&(2)\ p_{ii}^{(n+k+m)} \ge p_{ij}^{(n)}p_{jl}^{(k)}p_{li}^{(m)};\\
&(3)\ p_{ii}^{(n_1+n_2+\cdots+n_k)} \ge p_{ii}^{(n_1)}p_{ii}^{(n_2)}...p_{ii}^{(n_k)};\\
&(1)\ p_{ii}^{(nk)} \ge (p_{ii}^{(n)})^k.
\end{aligned}$$

例 5.10. 设例5.2中，$n=3$, $p=q=\frac{1}{2}$. 赌博者从 2 元赌金开始赌博，求解他经过 4 次赌博之后输光的概率.

解：记 $p_{ij}=P(X_{n+1}=j|X_n=i)$, 则对 $0<i<3$，$p_{i,i-1}=p_{i,i+1}=\frac{1}{2}$；对

$p0,0=1$, $p_{3,3}=1$。转移概率矩阵为

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

依题意，$X_0=2$，求 $p_{20}^{(4)}$。这只要求 P^4 的第 3 行第 1 列元素。手工计算，可以计算 P^2 再计算 $(P^2)^2$。

如果不要求结果用有理数表示，可以用如下 R 程序：

```
P <- rbind(
  c(1,0,0,0),
  c(0.5, 0, 0.5, 0),
  c(0, 0.5, 0, 0.5),
  c(0,0,0,1))
P4 <- P %*% P %*% P %*% P; P4
```

```
##        [,1]   [,2]   [,3]   [,4]
## [1,] 1.0000 0.0000 0.0000 0.0000
## [2,] 0.6250 0.0625 0.0000 0.3125
## [3,] 0.3125 0.0000 0.0625 0.6250
## [4,] 0.0000 0.0000 0.0000 1.0000
```

MASS 包的 `fractions()` 函数近似为有理数显示：

```
MASS::fractions(P4)
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    1    0    0    0
## [2,]  5/8 1/16    0 5/16
## [3,] 5/16    0 1/16  5/8
## [4,]    0    0    0    1
```

Julia 语言直接支持有理数：

```
P = [1 0 0 0;
1//2 0 1//2 0;
0 1//2 0 1//2;
0 0 0 1]
P^4
## 4×4 Matrix{Rational{Int64}}:
## 1//1   0//1   0//1   0//1
## 5//8   1//16  0//1   5//16
## 5//16  0//1   1//16  5//8
## 0//1   0//1   0//1   1//1
```

所以结果为 $\frac{5}{16}$。

例 5.11. 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 $\{X_n, n=0,1,2,\cdots\}$ 为时齐马氏链，求在甲获得 1 分的情况下，不超过两局可结束比赛的概率.

解：这里 $-2, 2$ 是吸收态。$\{X_n, n=0,1,2,\cdots\}$ 的一步转移概率矩阵为

$$
P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ q & r & p & 0 & 0 \\ 0 & q & r & p & 0 \\ 0 & 0 & q & r & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}
$$

两步转移概率矩阵为

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

故在甲获得 1 分的情况下，不超过两局可结束比赛的概率为

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

例 5.12. 质点在数轴上的点集 $\{-2,-1,0,1,2\}$ 上做随机游动. 质点到达点-2 后，以概率 1 停留在原处；到达点 2 后，以概率 1 向左移动一点；到达其它点后，分别以概率 $\frac{1}{3}$ 向左、向右移动一点，以概率 $\frac{1}{3}$ 停留在原处. 试求在已知该质点处于状态 0 的条件下，经三步转移后仍处于状态 0 的概率.

解： 一步转移概率矩阵为

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

二步转移概率矩阵为

$$P^{(2)} = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ \frac{1}{9} & \frac{2}{9} & \frac{3}{9} & \frac{2}{9} & \frac{1}{9} \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$

三步转移概率矩阵为

$$P^{(3)} = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ \frac{1}{9} & \frac{2}{9} & \frac{3}{9} & \frac{2}{9} & \frac{1}{9} \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \cdot \begin{pmatrix} * & * & 0 & * & * \\ * & * & \frac{1}{3} & * & * \\ * & * & \frac{1}{3} & * & * \\ * & * & \frac{1}{3} & * & * \\ * & * & 0 & * & * \end{pmatrix} = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & \frac{7}{27} & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$

即在已知处于状态 0 的条件下，经三步转移后仍处于状态 0 的概率为 $\frac{7}{27}$. (转移矩阵中的 $*$ 部分与最终要计算的结果无关，故省略计算)

也可以借助于用 R 和 Julia 计算矩阵乘方。

R 版本：

```r
P <- matrix(c(
  1, 0, 0, 0, 0,
  1/3, 1/3, 1/3, 0, 0,
  0, 1/3, 1/3, 1/3, 0,
  0, 0, 1/3, 1/3, 1/3,
  0, 0, 0, 1, 0),
  byrow=TRUE, ncol=5)
P %*% P %*% P |> MASS::fractions()
```

```
##       [,1]  [,2]  [,3]  [,4]  [,5]
## [1,]     1     0     0     0     0
## [2,] 14/27  4/27  5/27   1/9  1/27
## [3,]  5/27  5/27  7/27  8/27  2/27
## [4,]  1/27   1/9  8/27 10/27  5/27
## [5,]     0   1/9   2/9   5/9   1/9
```

结果 $p_{0,0}^{(3)} = \frac{7}{27}$。

利用 Julia 计算：

```julia
P = [
1 0 0 0 0;
1//3 1//3 1//3 0 0 ;
0 1//3 1//3 1//3 0 ;
0 0 1//3 1//3 1//3 ;
0 0 0 1 0]
P^3
## 5×5 Matrix{Rational{Int64}}:
##   1//1   0//1   0//1    0//1   0//1
##  14//27  4//27  5//27   1//9   1//27
##   5//27  5//27  7//27   8//27  2//27
##   1//27  1//9   8//27  10//27  5//27
##   0//1   1//9   2//9    5//9   1//9
```

定理 5.2. 设时齐马氏链 $\{X_n, n \geq 0\}$ 的状态转移矩阵为 P，X_0 的分布为

$P(X_0=j)=p_j$，$j\in S$，则 $\{X_n\}$ 有限维分布完全被 $\{p_j\}$ 和 P 决定，有

$$P(X_0=i_0,X_1=i_1,\dots,X_n=i_n)=p_{i_0}p_{i_0i_1}p_{i_1i_2}\cdots p_{i_{n-1}i_n}.$$

证明：对时刻 $(0,1,\dots,n)$，由概率的乘法公式和马氏性有

$$\begin{aligned}&P(X_0=i_0,X_1=i_1,\dots,X_n=i_n)\\=&P(X_n=i_n|X_0=i_0,X_1=i_1,\dots,X_{n-1}=i_{n-1})\\&\cdots P(X_1=i_1|X_0=i_0)P(X_0=i_0)\\=&p_{i_0}p_{i_0i_1}p_{i_1i_2}\cdots p_{i_{n-1}i_n},\end{aligned}$$

仅依赖于初始分布和转移概率矩阵。

推论 5.2. 设马氏链中 X_0 分布列为 $\pi_0=(p_0,p_1,\dots)$，则 X_n 的分布列为

$$\pi_n=\pi_0P^n.$$

证明：

$$\begin{aligned}P(X_1=j)=&\sum_{i\in S}P(X_1=j|X_0=i)P(X_0=i)\\=&\sum_{i\in S}p_ip_{ij},\end{aligned}$$

记 π_1 表示 X_1 概率分布列，可以写成

$$\pi_1=\pi_0P.$$

同理

$$\pi_n=\pi_{n-1}P.$$

于是递推可得结论。

例 5.13 (广告效益的推算)**.** 某种啤酒 A 的广告改变了广告方式，经调查发现买 A 种啤酒及另外三种啤酒 B,C,D 的顾客每两个月的平均转换率如下 (设市场中只有这四种啤酒):

$$\begin{array}{cccccc}A&\to&A(0.95)&B(0.02)&C(0.02)&D(0.01)\\B&\to&A(0.30)&B(0.60)&C(0.06)&D(0.04)\\C&\to&A(0.02)&B(0.01)&C(0.07)&D(0.00)\\D&\to&A(0.20)&B(0.20)&C(0.10)&D(0.50)\end{array}$$

假设目前购买 A,B,C,D 四种啤酒的顾客的分布为 $(25\%, 30\%, 35\%, 10\%)$，试求半年后啤酒 A 的市场份额.

解：令 P 为转移矩阵，则显然有

$$P = \begin{pmatrix} 0.95 & 0.02 & 0.02 & 0.01 \\ 0.30 & 0.60 & 0.06 & 0.04 \\ 0.20 & 0.10 & 0.70 & 0.00 \\ 0.20 & 0.20 & 0.10 & 0.50 \end{pmatrix}$$

令

$$\mu = (\mu_1, \mu_2, \mu_3, \mu_4) = (0.25, 0.30, 0.35, 0.10),$$

计算经过半年后顾客在这四种啤酒上的转移概率 P^3,

$$P^2 = \begin{pmatrix} 0.9145 & 0.035 & 0.0352 & 0.0153 \\ 0.485 & 0.38 & 0.088 & 0.047 \\ 0.36 & 0.134 & 0.5 & 0.006 \\ 0.37 & 0.234 & 0.136 & 0.26 \end{pmatrix},$$

$$P^3 = \begin{pmatrix} 0.8894 & 0.0458 & 0.0466 & 0.01820 \\ 0.60175 & 0.2559 & 0.0988 & 0.04355 \\ 0.4834 & 0.1388 & 0.36584 & 0.01196 \\ 0.5009 & 0.2134 & 0.14264 & 0.14306 \end{pmatrix}.$$

我们关心啤酒 A 半年后的市场占有率，即从 A,B,C,D 四种啤酒经 3 次转移后转到 A 的概率，求得 A 的市场占有率变为

$$\begin{aligned} P(X_3 = 1) =& \sum_{i=1}^{4} P(X_3 = 1 | X_0 = i) P(X_0 = i) \\ =& (0.25, 0.3, 0.35, 0.10) \begin{pmatrix} 0.8894 \\ 0.60175 \\ 0.4834 \\ 0.5009 \end{pmatrix} \approx 0.624 \end{aligned}$$

可见啤酒 A 的市场份额由原来的 25% 增至 62.4%，新的广告方式很有效.

5.2 状态的分类及性质

本节我们首先来讨论一下马氏链各个状态之间的关系，并以这些关系将状态分类，最后来研究它们的性质.

定义 5.6 (互通)**.** 称状态 i **可达**状态 $j(i,j\in S)$，若存在 $n\geq 0$ 使得 $p_{ij}^{(n)}>0$，记为 $i\to j$. 若同时有 $j\to i$，则称 i 与 j **互通**，记为 $i\leftrightarrow j$.

注：定义中之所以用 $n\geq 0$ 而不是 $n>0$，是为了包含 $i=j$ 时 $p_{ii}^{(0)}=1>0$ 的情况，从而状态 i 与自身是互通的。

定理 5.3. *互通是一种等价关系，即满足：*

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

证明：从互通的定义可知 1、2 是显然的, 只证 3.

由互通定义可知需证 $i\to k$ 且 $k\to j$. 首先，由 $i\to j$, $j\to k$ 知道存在 $m,n\geq 0$，使得 $p_{ij}^{(m)}>0, p_{jk}^{(n)}>0$. 再由 C-K 方程知道

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

故 $i\to k$. 同理可证 $k\to i$，即有 $i\leftrightarrow k$.

我们把任何两个互通状态归为一类，由上述定理可知，同在一类的状态应该都是互通的，并且任何一个状态不能同时属于两个不同的类.

定义 5.7. 若马氏链只存在一个类，就称它是**不可约的** (irreducible)；否则称为可约的 (reducible).

例 5.14. 我们来看例5.1中疾病死亡模型的四个状态之间的关系. 为清楚起见，经常以图5.2所示的转移图来表示马氏链的状态变化. 由转移矩阵容易看出：S_1 和 S_2 互通，S_1 和 S_2 可达 S_3 和 S_4，但 S_3 和 S_4 不可达除本身以外的其它状态。状态可分为三类 $\{S_1,S_2\}$, $\{S_3\}$ 和 $\{S_4\}$. 称 S_3 和 S_4 为**吸收态**，这样的状态 i 满足

$$p_{ii}^{(n)}=1,\ n=0,1,2,...$$

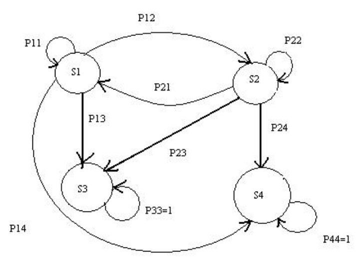


图 5.2: 图上的简单随机游动

读者可用类似的方法来说明赌徒输光问题 (例5.2) 中任何两个状态 $i, j(0 < i, j < n)$ 都互通，并可将所有状态分为三类：$\{0\}$, $\{1, 2, \cdots, n-1\}$, $\{n\}$.

下面我们给出状态的一些性质，然后证明同在一类的状态具有相同的性质.

定义 5.8 (周期). 若集合 $\{n : n \geq 1, p_{ii}^{(n)} > 0\}$ 非空，则称它的最大公约数 $d = d(i)$ 为状态 i 的**周期**. 若 $d > 1$，称 i **是周期的**；若 $d = 1$，称 i 是**非周期的**. 并特别规定上述集合为空集时，称 i 的周期为无穷大.

注 1：周期无穷大的状态 i，是在这个状态下一步必然离开而且永不返回的状态，是吸收态的一个反面。

注 2：由定义5.8知道，即使 i 有周期 d，但并不是对所有的 $n, p_{ii}^{(nd)}$ 都大于 0. 例如，设集合 $\{n : n \geq 1, p_{ii}^{(n)} > 0\}$ 为 $\{3, 9, 18, 21, ...\}$，则最大公约数 $d = 3$，即 3 是 i 的周期，显然，$n = 6, 12, 15$ 都不属于此集合，即 $p_{ii}^{(6)} = 0, p_{ii}^{(12)} = 0$, $p_{ii}^{(15)} = 0$. 类似地，若非周期，即 $d = 1$，也不一定有 $p_{ii} > 0$，如果 $p_{ii} = 0$，$p_{ii}^{(2)} > 0$，$p_{ii}^{(3)} > 0$，则最大公约数也是 1。但是可以证明，当 n 充分大之后一定有 $p_{ii}^{(dn)} > 0$.

例 5.15. 考察如图5.3所示的马氏链.

由状态 1 出发再回到状态 1 的可能步长为 $T = \{4, 6, 8, 10, ...\}$，它的最大公约数是 2，虽然从状态 1 出发 2 步并不能回到状态 1，我们仍然称 2 是状态 1 的周期.

定理 5.4. *若状态 i, j 同属一类，则 $d(i) = d(j)$.*

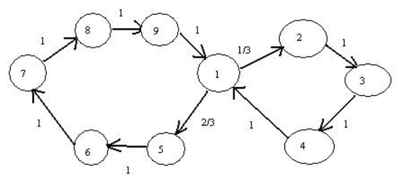


图 5.3: 状态分类

证明：由类的定义知 $i \leftrightarrow j$，即存在 $m, n \geq 0$，使 $p_{ij}^{(m)} > 0, p_{ji}^{(n)} > 0$，则

$$p_{ii}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{ki}^{(n)} \geq p_{ij}^{(m)} p_{ji}^{(n)} > 0.$$

对所有使得 $p_{jj}^{(s)} > 0$ 的 s, 有 $p_{ii}^{(n+s+m)} \geq p_{ij}^{(m)} p_{jj}^{(s)} p_{ji}^{(n)} > 0$. 显然 $d(i)$ 应同时整除 $n+m$ 和 $n+m+s$，则它必定整除 s. 而 $d(j)$ 是 j 的周期，所以也有 $d(i)$ 整除 $d(j)$. 反过来也可证明 $d(j)$ 整除 $d(i)$，于是 $d(i) = d(j)$.

定义 5.9 (首达概率). 对于任何状态 i, j，以 $f_{ij}^{(n)}$ 记从 i 出发经 n 步后首次到达 j 的概率，记

$$\begin{aligned} f_{ij}^{(n)} =& P\{X_n = j, X_k \neq j, k = 1, 2, \dots, n-1 | X_0 = i\},\ n \geq 1, \\ f_{ij}^{(0)} =& \delta_{i-j} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \end{aligned}$$

令

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

注：f_{ij} 是从 i 出发经过有限步可达状态 j 的概率。$i \to j$ 当且仅当 $f_{ij} > 0$。

定义 5.10 (常返状态). 若 $f_{jj} = 1$，称状态 j 为**常返状态** (recurrent)；若 $f_{jj} < 1$，称状态 j 为**非常返状态**或**瞬过状态**.

注：对常返状态 i，从 i 出发以概率 1 在有限步内回到状态 i。

注：对非常返状态 i，有概率 $p = 1 - f_{ii} > 0$ 使得从 i 出发不再返回 i。每次回到 i 后从 i 出发看成一个 Bernoulli 试验，不返回视作成功，返回视作失败，

记 Y 为这样的试验的次数，则 Y 服从几何分布，是取有限值的随机变量，而每次这样的试验，在成功（一去不返）前，都是有限次状态转移，所以从瞬态 i 出发在有限次状态转移之后必定一去不返。

对于常返状态 i，令 T 表示从 i 出发首次返回 i 的时间，即 $X_0 = i$ 条件下

$$T = \inf\{n \geq 1 : X_n = i\},$$

则 T 为一个取有限值的随机变量，$P(T = n) = f_{ii}^{(n)}, n = 1, 2, ...$，期望值为

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

即 μ_i 表示的是由 i 出发再返回到 i 所需的平均步数 (时间).

定义 5.11 (正常返状态)**.** 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**. 特别地，若 i 为正常返状态，且是非周期的，则称之为**遍历状态**.

显然对于吸收态，$f_{ii}^{(1)} = p_{ii} = 1$, $f_{ii}^{(n)} = 0 (n \geq 2)$，$f_{ii} = 1$，有 $\mu_i = 1$，是正常返状态。

例 5.16. 设马氏链的状态空间为 $S = \{1, 2, 3, 4\}$，其一步转移概率矩阵为

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

试将状态进行分类.

解： 由一步转移概率矩阵 P，对一切 $n \geq 1$，$f_{44}^{(n)} = 0$，从而

$$f_{44} = \sum_{n=1}^{\infty} f_{44}^{(n)} = 0 < 1,$$

故状态 4 是非常返态. 这就是周期等于正无穷的情况：从状态 4 出发以概率 1 不再返回。

又

$$
\begin{aligned}
f_{33}^{(1)} =& \frac{2}{3},\\
f_{33}^{(n)} =& 0\ (n \geq 2),\\
f_{33} =& \sum_{n=1}^{\infty} f_{33}^{(n)} = \frac{2}{3} < 1,
\end{aligned}
$$

故状态 3 是非常返态.

但状态 1 与 2 是常返态，因为

$$
\begin{aligned}
f_{11} =& f_{11}^{(1)} + f_{11}^{(2)} = \frac{1}{2} + \frac{1}{2} = 1,\\
f_{22} =& \sum_{n=1}^{\infty} f_{22}^{(n)} = 0 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots = 1.
\end{aligned}
$$

事实上，从状态 1 出发，以概率 $\frac{1}{2}$ 在第一步回到 1，以概率 $\frac{1}{2}$ 先去状态 2 然后从状态 2 在第二步返回状态 1。

从状态 2 出发，必然在第一步去状态 1，但从状态 1 又可以以概率 $\frac{1}{2}$ 在第二步回到状态 2，也可能以概率 $(\frac{1}{2})^{n-2} \cdot \frac{1}{2}$ 在状态 1 停留 $n-2$ 步后才返回状态 2，这样，$f_{22}^{(1)} = 0$，$f_{22}^{(2)} = \frac{1}{2}$，$f_{22}^{(n)} = (\frac{1}{2})^{n-1}$。

计算得

$$
\begin{aligned}
\mu_1 =& \sum_{n=1}^{\infty} n f_{11}^{(n)} = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2} < \infty,\\
\mu_2 =& \sum_{n=1}^{\infty} n f_{22}^{(n)} = 1 \cdot 0 + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{2^2} + \cdots + n \cdot \frac{1}{2^{n-1}} = 3 < \infty,
\end{aligned}
$$

故状态 1 与 2 都是正常返状态，又因其周期都是 1，故它们都是遍历态.

计算中用了

$$
\sum_{n=2}^{\infty} n x^{n-1} = \frac{d}{dx} \sum_{n=2}^{\infty} x^n = \frac{d}{dx} x^2 \sum_{n=0}^{\infty} x^n = \frac{d}{dx} \frac{x^2}{1-x} = \frac{x(2-x)}{(1-x)^2}.
$$

我们可以证明，对于同属一类的状态 i, j, 它们同为常返状态或非常返状态，并且当它们是常返状态时，又同为正常返状态和零常返状态.

引理 5.1. 对任意状态 i,j 及 $1\le n<\infty$，有

$$p_{ij}^{(n)}=\sum_{l=1}^{n}f_{ij}^{(l)}p_{jj}^{(n-l)}. \tag{5.8}$$

证明：令

$$\begin{aligned}
p_{ij}^{(n)} =&P(X_n=j|X_0=i)\\
=&P\left(\bigcup_{l=1}^{n}\{X_k\ne j,1\le k\le l-1,X_l=j,X_n=j\}|X_0=i\right)\\
=&\sum_{l=1}^{n}P\left(X_k\ne j,1\le k\le l-1,X_l=j,X_n=j|X_0=i\right)\\
=&\sum_{l=1}^{n}P\left(X_k\ne j,1\le k\le l-1,X_l=j|X_0=i\right)\\
&P(X_n=j|X_k\ne j,1\le k\le l-1,X_l=j,X_0=i)\\
=&\sum_{l=1}^{n}f_{ij}^{(l)}P(X_n=j|X_l=j)\quad(\text{利用马氏性})\\
=&\sum_{l=1}^{n}f_{ij}^{(l)}p_{jj}^{(n-l)}.
\end{aligned}$$

以下我们首先引入常返性的另一个判定方法.

定理 5.5. 状态 i 为常返的当且仅当 $\sum_{n=0}^{\infty}p_{ii}^{(n)}=\infty$；状态 i 为非常返的当且仅当 $\sum_{n=0}^{\infty}p_{ii}^{(n)}<\infty$，且这时有

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

证明：常返时，$f_{ii}=1$，每次从 i 出发必然在有限步回到 i，然后重新开始，以回到 i 作为一次更新，这是更新过程，于是从 $X_0=i$ 出发系统有无穷多次处于状态 i。用 I_n 表示 $X_n=i$，则常返时

$$P(\sum_{n=1}^{\infty}I_n=+\infty|X_0=i)=1,$$

于是

$$E(\sum_{n=1}^{\infty}I_n|X_0=i)=+\infty,$$

但

$$E(\sum_{n=1}^{\infty} I_n|X_0=i) = \sum_{n=1}^{\infty} E(I_n|X_0=i)$$
$$= \sum_{n=1}^{\infty} P(X_n=i|X_0=i) = \sum_{n=1}^{\infty} p_{ii}^{(n)},$$

所以常返时 $\sum_{n=1}^{\infty} p_{ii}^{(n)} = +\infty$。

如果非常返，$f_{ii} < 1$，有 $1-f_{ii} > 0$ 的概率一去不返，以从 i 出发是否一去不返为一次成败型试验，令 Y 表示一去不返前返回的次数加一，则 Y 为几何分布的随机变量，$EY = \frac{1}{1-f_{ii}}$。

$$\sum_{n=1}^{\infty} I_n = Y - 1,$$

于是

$$E(\sum_{n=1}^{\infty} I_n|X_0=i) = \sum_{n=1}^{\infty} p_{ii}^{(n)} = E(Y) - 1 = \frac{1}{1-f_{ii}} - 1 < \infty$$

从而

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

定理得证。

注：教材的证明有一些问题。

推论 5.3. *若 j 为非常返状态，则对任意 $i \in S$，*

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty,$$
$$\lim_{n\to\infty} p_{ij}^{(n)} = 0.$$

证明：由引理5.1,

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

关于 n 求和得

$$\begin{aligned}\sum_{n=1}^{N} p_{ij}^{(n)} &= \sum_{n=1}^{N}\sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} \\ &= \sum_{l=1}^{N} f_{ij}^{(l)} \sum_{n=l}^{N} p_{jj}^{(n-l)} \\ &= \sum_{l=1}^{N} f_{ij}^{(l)} \sum_{m=0}^{N-l} p_{jj}^{(m)} \\ &\le \sum_{m=0}^{\infty} p_{jj}^{(m)} \sum_{l=1}^{\infty} f_{ij}^{(l)} \\ &= \sum_{m=0}^{\infty} p_{jj}^{(m)} f_{ij} < \infty.\end{aligned}$$

推论 5.4. 若 j 为常返状态，则对任意 $i \in S$，若 $i \to j$，则

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty,$$

若 $i \not\to j$，则

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = 0.$$

证明：当 i 不可达 j 时，每个 $p_{ij}^{(n)} = 0$，所以第二条结论成立。当 $i \to j$ 时，存在 $m > 0$ 使得 $p_{ij}^{(m)} > 0$，由 C-K 方程，

$$p_{ij}^{(m+n)} \ge p_{ij}^{(m)} p_{jj}^{(n)},$$

于是

$$\sum_{n=1}^{\infty} p_{ij}^{(m+n)} \ge p_{ij}^{(m)} \sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty.$$

命题 5.2. 若 i 与 j 互通且都为常返状态，则 $f_{ij} = 1$。

证明：如果 $f_{ij} < 1$，则设 $p_{ji}^{(n)} > 0$，有正概率 $p_{ji}^{(n)}(1 - f_{ij})$ 使得从 j 出发不能返回 j，与 j 常返矛盾。

定理 5.6. 常返性是一个类性质.

证明：只要证明若 $i \leftrightarrow j$，则 i, j 同为常返或非常返状态.

由 $i \leftrightarrow j$ 知，存在 $n, m \geq 0$，使得 $p_{ij}^{(n)} > 0, p_{ji}^{(m)} > 0$，由 C-K 方程总有

$$\begin{aligned} p_{ii}^{(n+m+l)} &\geq p_{ij}^{(n)} p_{jj}^{(l)} p_{ji}^{(m)} \\ p_{jj}^{(n+m+l)} &\geq p_{ji}^{(m)} p_{ii}^{(l)} p_{ij}^{(n)}, \end{aligned}$$

求和得到

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

可见，$\sum_{l=0}^{\infty} p_{jj}^{(l)}$, $\sum_{l=0}^{\infty} p_{ii}^{(l)}$ 相互控制，同为无穷或有限，从而 i, j 同为常返或非常返状态.

其次我们还可以证明，当 i, j 同为常返状态时，它们同为正常返状态或零常返状态. 证明将在下一节给出.

推论 5.5. 若常返状态 i 可达状态 j，则 j 也是常返状态。

证明：这时 j 必然可达 i，否则设 $p_{ij}^{(n)} > 0$，则从 i 出发有正概率 $p_{ij}^{(n)}$ 达到 j 后不再回到 i，与常返性矛盾。i, j 互通，则常返性相同，证毕。

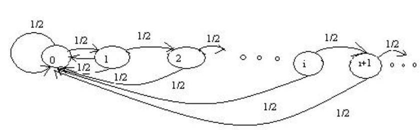


图 5.4: 每个状态返回 0

例 5.17. 设马氏链的状态空间 $S = \{0, 1, 2, \dots\}$, 转移概率为 $p_{00} = \frac{1}{2}$, $p_{i,i+1} = \frac{1}{2}$, $p_{i0} = \frac{1}{2}$, $i \in S$. 由图5.4易知, $f_{00}^{(1)} = \frac{1}{2}$, $f_{00}^{(2)} = \frac{1}{2} \cdot \frac{1}{2}$, $f_{00}^{(3)} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$, ……,

$f_{00}^{(n)} = \frac{1}{2^n}$，故

$$f_{00} = \sum_{n=1}^{\infty} \frac{1}{2^n} = 1,$$

$$\mu_0 = \sum_{n=1}^{\infty} n 2^{-n} < \infty.$$

可见状态 0 是正常返状态，显然它是非周期的，故 0 是遍历态. 对其他状态 $i > 0$，由 $i \leftrightarrow 0$, 故 i 也是遍历的.

例 5.18. 考虑直线上无限制的随机游动的常返性。状态空间为 $S = \{0, \pm 1, \pm 2, ...\}$，转移概率为 $p_{i,i+1} = 1 - p_{i,i-1} = p$, $i \in S$.

解：对于状态 0，可知 $p_{00}^{(2n+1)} = 0$, $n = 1, 2, ...$, 即从 0 出发奇数次不可能返回到 0. 而

$$p_{00}^{(2n)} = \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{n!n!} [p(1-p)]^n.$$

即经过偶数次回到 0 当且仅当它向左、右移动距离相同.

由 Stirling 公式知, 当 n 充分大时，$n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}$，则

$$p_{00}^{(2n)} \sim \frac{[4p(1-p)]^n}{\sqrt{n\pi}}.$$

而 $p(1-p) \leq \frac{1}{4}$ 且 $p(1-p) = \frac{1}{4} \iff p = \frac{1}{2}$. 于是 $p = \frac{1}{2}$ 时, $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$, 否则 $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$, 即当 $p \neq \frac{1}{2}$ 时状态 0 是瞬过状态，$p = \frac{1}{2}$ 时是常返状态. 显然，过程的各个状态都是相通的，故以此可得其他状态的常返性.

5.3 极限定理及平稳分布

5.3.1 极限定理

对于一个系统来说，考虑它的长期的性质是很必要的，本节我们将研究马氏链的极限情况和平稳马氏链的有关性质. 首先来看两个例子.

例 5.19. 设马氏链的转移矩阵为

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

考虑 $P^{(n)}$ 当 $n \to \infty$ 时的情况.

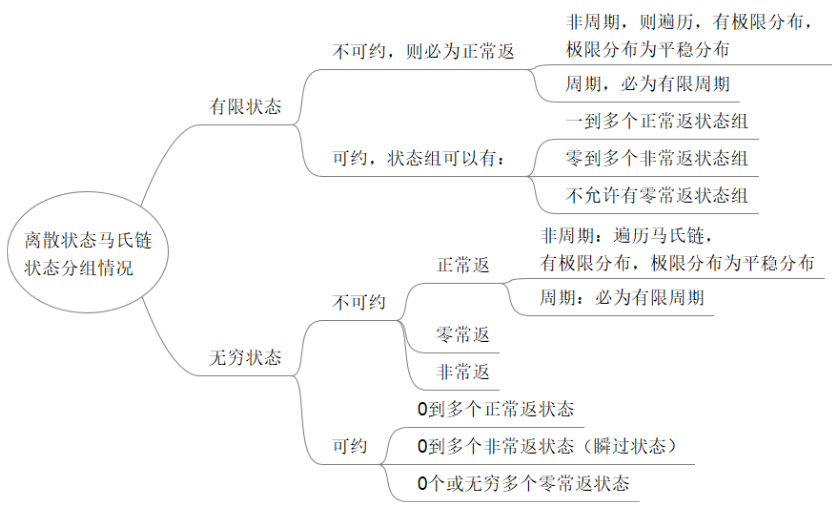


图 5.5: 离散状态马氏链状态分类

解： 由 $P^{(n)} = P^n$ 知，只需计算 P 的 n 重乘积的极限. 令

$$Q = \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 1-p-q \end{pmatrix}$$

则

$$Q^{-1} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ -\frac{1}{p+q} & \frac{1}{p+q} \end{pmatrix}, \qquad P = QDQ^{-1},$$

这是 P 的相似变换，其中 D 为对角阵，这使得矩阵的幂的计算化简，有

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

由于 $|1-p-q| < 1$，P^n 的极限为

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix},$$

可见此马氏链的 n 步转移概率有一个稳定的极限.

例 5.20. 在例5.18中令 $p=\frac{1}{3}$，系统为不可约非常返马氏链，有

$$\lim_{n\to\infty} p_{00}^{(2n)} = \lim_{n\to\infty} \frac{(4\times\frac{1}{3}\times\frac{2}{3})^n}{\sqrt{n\pi}} = 0.$$

令 $p=\frac{1}{2}$，系统为不可约零常返马氏链，有

$$\lim_{n\to\infty} p_{00}^{(2n)} = \lim_{n\to\infty} \frac{(4\times\frac{1}{2}\times\frac{1}{2})^n}{\sqrt{n\pi}} = 0.$$

即从零出发经过无穷次的转移之后，系统在某一规定时刻回到 0 的概率趋于 0.

我们容易证明例5.19中所有状态是正常返状态，而例5.20中当 $p=\frac{1}{3}$ 时状态 0 是非常返状态，当 $p=\frac{1}{2}$ 时，0 是零常返状态. 那么两个例子给出的是不是一般结论呢？答案是肯定的，我们不加证明地引入马氏链的一个基本极限定理.

定理 5.7. *若状态 i 是周期为 d 的常返状态，则*

$$\lim_{n\to\infty} p_{ii}^{(nd)} = \frac{d}{\mu_i}. \tag{5.9}$$

这样，若 i 为正常返状态，极限为正值；若 i 为零常返状态，极限等于 0.

另外，如果 i 是非常返状态，由推论5.3，

$$\lim_{n\to\infty} p_{ji}^{(n)} = 0,\ \forall j\in S.$$

证明见 (林元烈 2002) P.96 定理 3.3.6。

注 1：μ_i 表示常返状态 i 的平均返回时间。

推论 5.6. *设 i 为常返状态，则*

$$i\text{为零常返状态} \iff \lim_{n\to\infty} p_{ii}^{(n)} = 0.$$

证明：若 i 为零常返状态，则 $\mu_i=\infty$, 从而 $\lim_{n\to\infty} p_{ii}^{(nd)}=0$. 而当 m 不是 d 的整数倍时, $p_{ii}^{(m)}=0$，故 $\lim_{n\to\infty} p_{ii}^{(n)}=0$.

另一方面，若 $\lim_{n\to\infty} p_{ii}^{(n)}=0$，如果 i 为正常返状态，则 $\mu_i<\infty$，由定理5.7知 $\lim_{n\to\infty} p_{ii}^{(nd)}>0$，矛盾.

推论 5.7. 两个互通的常返状态或者同为正常返，或者同为零常返，从而同一类的状态或者同为正常返，或者同为零常返，或者同为非常返。

证明：设 $i \leftrightarrow j$ 为常返状态且 i 为零常返状态，则

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

有互通性，必存在 $n \geq 1, l \geq 1$ 使得 $p_{ij}^{(n)} > 0, p_{ji}^{(l)} > 0$，从而由 C-K 方程可知

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

令 $m \to \infty$, 对上式取极限，知

$$\lim_{m\to\infty} p_{jj}^{(m)} = 0$$

从而 j 也为零常返状态. 反之由 j 为零常返状态也可推得 i 为零常返状态，从而证明了 i, j 同为零常返状态或正常返状态.

下面我们要利用定理5.7来讨论 $p_{ij}^{(n)}$ 的极限性质. 一般说来, 我们讨论两个问题. 一是极限 $\lim_{n\to\infty} p_{ij}^{(n)}$ 是否存在, 二是其极限是否与 i 有关. 首先有

定理 5.8. *(1)* 若 j 为非常返状态或零常返状态，则 $\forall i \in S$ 都有

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

(2) 若马氏链为遍历链（不可约、正常返、非周期），则对任意 $i, j \in S$，有

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j}.$$

证明：(1) 由引理5.1，得

$$p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}. \tag{5.10}$$

对 $N < n$，有

$$\sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} \leq \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}. \tag{5.11}$$

于是

$$\begin{aligned}\overline{\lim_{n\to\infty}}\, p_{ij}^{(n)} \le& \sum_{l=1}^{N} f_{ij}^{(l)} \overline{\lim_{n\to\infty}}\, p_{jj}^{(n-l)} + \sum_{l=N+1}^{\infty} f_{ij}^{(l)} \\ =&0 + \sum_{l=N+1}^{\infty} f_{ij}^{(l)}.\end{aligned}$$

由 $\sum_{l=1}^{\infty} f_{ij}^{(l)} \le 1$ 可知 $\lim_{N\to\infty} \sum_{l=N+1}^{\infty} f_{ij}^{(l)} = 0$，于是 $\overline{\lim}_{n\to\infty} p_{ij}^{(n)} \le 0$，$\lim_{n\to\infty} p_{ij}^{(n)} = 0$。

（2）由(5.10)，取 $N < n$, 有

$$\begin{aligned}\underline{\lim_{n\to\infty}}\, p_{ij}^{(n)} \ge& \sum_{l=1}^{N} f_{ij}^{(l)} \underline{\lim_{n\to\infty}}\, p_{jj}^{(n-l)} \\ =& \sum_{l=1}^{N} f_{ij}^{(l)} \frac{1}{\mu_j},\end{aligned}$$

令 $N \to \infty$ 有

$$\underline{\lim_{n\to\infty}}\, p_{ij}^{(n)} \ge \frac{1}{\mu_j} \sum_{l=1}^{\infty} f_{ij}^{(l)} = \frac{1}{\mu_j} f_{ij} = \frac{1}{\mu_j}.$$

这里用到 $i \to j$ 且 j 常返则 $f_{ij} = 1$。

另一方面，由(5.10)，

$$\begin{aligned}\overline{\lim_{n\to\infty}}\, p_{ij}^{(n)} \le& \sum_{l=1}^{N} f_{ij}^{(l)} \overline{\lim_{n\to\infty}}\, p_{jj}^{(n-l)} + \sum_{l=N+1}^{\infty} f_{ij}^{(l)} \\ =& \frac{1}{\mu_j} \sum_{l=1}^{N} f_{ij}^{(l)} + \sum_{l=N+1}^{\infty} f_{ij}^{(l)},\end{aligned}$$

令 $N \to \infty$ 得

$$\overline{\lim_{n\to\infty}}\, p_{ij}^{(n)} \le \frac{1}{\mu_j} \sum_{l=1}^{\infty} f_{ij}^{(l)} + 0 = \frac{1}{\mu_j} f_{ij} = \frac{1}{\mu_j},$$

故 $\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j}$。

注：结论 (2) 中 $\lim_{n\to\infty} p_{ij}^{(n)}$ 可以理解为系统从状态 i 出发，当 n 充分大时系统状态 X_n 位于状态 j 的比例。5.3.4给出了针对逐条轨道的结论。

推论 5.8. *有限状态的马氏链，不可能全为非常返状态，也不可能有零常返状态，从而不可约的有限马氏链是正常返的.*

证明： 设状态空间 $S = \{1, 2, \dots, N\}$. 若全部 N 个状态非常返，$\forall i, j \in S$ 有 $p_{ij}^{(n)} \to 0$, 注意

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

令 $n \to \infty$，等式左侧极限为 0，右侧为 1，矛盾。

若 S 中有零常返状态，设为 i。令 $C = \{j : i \to j\}$, 由推论5.5可知 C 中的状态也是常返状态，再由推论5.7可知 C 中的状态都是零常返状态，从而

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

注意

$$\sum_{j\in C} p_{ij}^{(n)} = 1,$$

在上式中令 $n \to \infty$，则左侧极限为 0，右侧为 1，矛盾。

推论 5.9. *若 Markov 链有一个零常返状态，则必有无限个零常返状态.*

证明： 设 i 为零常返状态，$C = \{j : i \to j\}$, 则 C 中的状态也是零常返状态，有 $\lim_{n\to\infty} p_{ij}^{(n)} = 0$。注意到

$$\sum_{j\in C} p_{ij}^{(n)} = 1,$$

如果 C 为有限集合，则上式两边令 $n \to \infty$，左侧极限为 0，右侧为 1，矛盾，所以 C 必为无限集合，于是有无穷多个零常返状态。

定理5.8没有讨论当 j 是正常返，但非遍历时 $\lim_n p_{ij}^{(n)}$ 的情况。这种情况下，极限不一定存在，存在时可能与 i 有关。下面的定理侧面描述了这种情况：

定理 5.9. *设 j 为常返状态，则对于任意的 $i \in S$，有*

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} p_{ij}^{(k)} = \frac{f_{ij}}{\mu_j}.$$

证明见 (林元烈 2002) P.94 定理 3.3.5。

注 1：左边的极限描述了从状态 i 出发，在长期内平均处于状态 j 的概率，这个概率等于从 i 可达 j 的概率，除以从 j 返回 j 的平均间隔。5.3.4给出了针对逐条轨道的结论。

如果 $i = j$，这个平均概率就等于 $1/\mu_j$。

注 2：可以将所有状态分为瞬态组 T（不要求所有瞬态互通）和若干个常返状态组 $C_k, k = 1, 2, \ldots$，其中 C_k 中的状态互通且为常返状态，而 $C_k \cap C_j = \emptyset (k \neq j)$。系统可以一直从瞬态出发一直在瞬态中运动（这要求有无穷多个瞬态），还可以从某个 C_k 出发一直在 C_k 中运动，也可以从某个瞬态出发在有限步后进入某个 C_k，然后保持在 C_k 中运动。

下面的定理给出了 j 正常返但周期大于 1 的 $\lim_{n\to\infty} p_{ij}^{(n)}$ 结果。需要按周期中不同相位分别计算极限：

定理 5.10. *若 j 为正常返状态，周期为 d，则对任意 $i \in S$，$0 \leq r \leq d-1$，有*

$$\lim_{n\to\infty} p_{ij}^{(nd+r)} = f_{ij}(r) \frac{d}{\mu_j}, \tag{5.12}$$

其中

$$f_{ij}(r) = \sum_{m=0}^{\infty} f_{ij}^{(md+r)},$$

当 $d = 1$ 时 (这时状态 j 遍历) 即

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

见 (林元烈 2002) P.108 定理 3.5.2，证明参考 (何声武 1999)。

注：定理结论说明，若 j 的周期 $d \geq 2$，则一般 $\lim_n p_{ij}^{(n)}$ 不存在，可以将 n 按照除以 d 的余数分为 d 个子序列，每个子序列有单独的极限。

例 5.21. 设马氏链的状态空间为 $S = \{1, 2, 3, 4, 5\}$，转移矩阵为

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

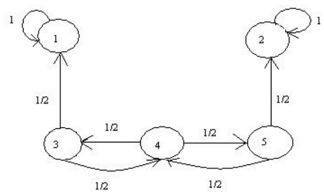


图 5.6: 图上的简单随机游动

试确定常返状态，瞬过状态，并对常返状态 i 确定其平均回转时间 μ_i.

解: 这是有限链，不能有零常返状态。

画出转移图5.6. 显然，状态 $1, 2$ 的周期为 1，状态 $3, 4, 5$ 的周期为 2.

$$P^n \overset{n\to\infty}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \end{pmatrix}$$

从而 $1, 2$ 是吸收态，为正常返，$\mu_1 = \mu_2 = 1$. $3, 4, 5$ 为瞬过状态。此马氏链可分为三类，即 $\{1\}$,$\{2\}$ 和 $\{3, 4, 5\}$.

5.3.2 平稳分布

前面我们只讨论了马氏链的转移概率 p_{ij} 的有关问题，下面我们将就它的初始分布的问题给出一些结论. 首先是关于马氏链的平稳分布和极限分布的概念.

定义 5.12. 对于马氏链，概率分布 $\{p_j, j \in S\}$ 称为**平稳分布**，若

$$p_j = \sum_{i \in S} p_i p_{ij},\ \forall j \in S.$$

平稳分布又称**不变分布**。

记 $\pi = (p_1, p_2, ...)^T$，则平稳分布 π 必须满足

$$\pi^T P = \pi^T, \quad \pi^T 1 = 1.$$

若马氏链的初始分布 $P\{X_0=j\}=p_j$ 是平稳分布，则 X_1 的分布将是

$$\begin{aligned}P\{X_1=j\}&=\sum_{i\in S}P\{X_1=j|X_0=i\}\cdot P\{X_0=i\}\\&=\sum_{i\in S}p_{ij}p_i=p_j,\ \forall j\in S.\end{aligned}$$

这与 X_0 的分布是相同的，依次递推可知 $X_n, n=0,1,2,3,...$ 将有相同的分布，这也是为什么称 $\{p_i, i\in S\}$ 为平稳分布的原因. 易见

$$\pi^T P^n=\pi^T,\ n\geq 1.$$

定理 5.11. *设马氏链 $\{X_n, n\geq 0\}$ 有平稳分布 π，且 $X_0\sim\pi$，则 $\{X_n\}$ 是严平稳的随机过程。*

证明：这时每一个 X_n 的边缘分布都是 π，于是由定理5.2，

$$P(X_0=i_0,X_1=i_1,...,X_n=i_n)=\pi_{i_0}p_{i_0i_1}p_{i_1i_2}...p_{i_{n-1}i_n},$$

同理有 $m\geq 0$ 时

$$P(X_m=i_0,X_{m+1}=i_1,...,X_{m+n}=i_n)=\pi_{i_0}p_{i_0i_1}p_{i_1i_2}...p_{i_{n-1}i_n},$$

因此 $\{X_n, n\geq 0\}$ 是严平稳的。

定义 5.13. 称马氏链是遍历的，如果所有状态不可约、非周期、正常返. 对于遍历的马氏链, 极限

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

称为马氏链的**极限分布**。

定义中极限 $\pi_j=\frac{1}{\mu_j}$，是利用了定理5.8. 此定理说明，对于遍历链，不论从那个状态开始，当 n 充分大时，对任意 n，X_n 都有正概率处于状态 j，j 也是任意状态；X_n 处于状态 j 的概率为正值 $\frac{1}{\mu_j}$。这给出了“遍历”这一术语的解释，即系统在长时间运行中可以在任意时间到达任意状态。

下面的定理说明对于遍历的马氏链，极限分布就是平稳分布并且还是唯一的平稳分布.

定理 5.12. *(1) 对于遍历马氏链，$\pi_j=\lim_{n\to\infty}p_{ij}^{(n)}=\frac{1}{\mu_j}>0(j\in S)$ 是平稳分布且是唯一的平稳分布；*

(2) 若马氏链所有状态都是非常返或零常返的，则平稳分布不存在.

证明：（1）对遍历的马氏链，由定理5.8知

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j} > 0,$$

记为 π_j. 于是极限分布存在（极限分布存在则必唯一）。

先证明 $\{\pi_j, j \in S\}$ 是平稳分布，然后再来证明它是唯一的平稳分布。这里仅给出有限链时的证明，状态个数无限的链在证明时涉及到级数与极限交换次序，需要较复杂的讨论，见5.7.2。

由于

$$\sum_{j\in S} p_{ij}^{(n)} = 1,$$

则有

$$\lim_{n\to\infty} \sum_{j\in S} p_{ij}^{(n)} = 1. \tag{5.13}$$

当 S 为有限状态时极限与求和可交换。于是有

$$\sum_{j\in S} \pi_j = 1.$$

利用 C-K 方程得

$$p_{ij}^{(n+1)} = \sum_{k\in S} p_{ik}^{(n)} p_{kj},$$

两边取极限，若 S 为有限状态则极限与求和可交换，于是

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

即 $\pi_j = \sum_{k\in S} \pi_k p_{kj}$, 从而 $\{\pi_j, j \in S\}$ 是平稳分布.

再证 $\{\pi_j, j \in S\}$ 是唯一的平稳分布. 假设另外还有一个平稳分布 $\{\tilde{\pi}_j, j \in S\}$, 则由

$$\tilde{\pi}_j = \sum_{k\in S} \tilde{\pi}_k p_{kj}$$

归纳得到

$$\tilde{\pi}_j = \sum_{k\in S} \tilde{\pi}_k p_{kj}^{(n)}, \quad n = 1, 2, \dots, \tag{5.14}$$

令 $n \to \infty$，若 S 为有限状态则极限与求和可交换，有

$$\begin{aligned}\tilde{\pi}_j &= \sum_{i \in S} \tilde{\pi}_i \lim_{n \to \infty} p_{ij}^{(n)} \\ &= \sum_{i \in S} \tilde{\pi}_i \cdot \pi_j = \pi_j \cdot \sum_{i \in S} \tilde{\pi}_i = \pi_j,\ \forall j \in S.\end{aligned}$$

即平稳分布唯一.

(2) 所有状态都是零常返或瞬态时，假设存在一个平稳分布 $\{\pi_j, j \in S\}$, 则由(5.14)有

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

令 $n \to \infty$，若 S 为有限状态则极限与求和可交换，因 $p_{ij}^{(n)} \to 0$, 推出 $\pi_j = 0, j \in S$, 这与 $\sum_{j \in S} \pi_j = 1$ 矛盾。于是对于所有状态都非常返或零常返的马氏链不存在平稳分布.

注 1：若马氏链为有限、不可约、非周期，必为正常返从而为有限状态遍历马氏链，平稳分布存在唯一且等于极限分布。这时平稳方程存在唯一解且等于极限分布，只要求解平稳方程。这同时也给出了求 $\mu_j = \frac{1}{\pi_j}$ 的方法。

注 2：定理5.12没有给出马氏链存在正常返状态，但非遍历链的结果。实际上，只要存在正常返状态，平稳分布就存在，但不一定唯一；如果只有非常返和零常返状态，平稳分布一定不存在。下面的定理给出了一般的结果。

定理 5.13. *(1) 当且仅当所有状态都是非常返和零常返时，马氏链没有平稳分布。*

(2) 平稳分布存在唯一的充分必要条件是马氏链存在正常返状态且正常返状态都是互相连通的。

(3) 有限状态马氏链总存在平稳分布（不一定唯一）。

(4) 有限的不可约、非周期马氏链存在唯一的平稳分布。

见 (林元烈 2002) P.111 定理 3.5.7。

注 1：结论 (1) 说明只要有正常返状态就一定存在平稳分布但不一定唯一，否则一定不存在平稳分布。存在正常返状态是存在平稳分布的充分必要条件，与是否不可约和是否非周期无关。

注 2：结论 (2) 说明如果正常返状态按互通性分成两个或两个以上的组，平稳分布就有多个。

注 3：第 (4) 条是因为这时马氏链遍历，由定理5.12第一条即可得结论。

注 4: 结论 (4) 是因为有限状态、不可约、非周期必然遍历。类似结论对无限状态马氏链是否成立？有如下结果。

定理 5.14. 不可约、非周期马氏链为遍历链（即状态都是正常返的），当且仅当它存在平稳分布。这时平稳分布等于极限分布。

见 (林元烈 2002) P.113 定理 3.5.8。所以不可约、非周期马氏链或者是遍历的，存在唯一的平稳分布同时也是极限分布；或者是非常返或者零常返的，这时不存在平稳分布，也不存在极限分布。

定理 5.15. 设马氏链 $\{X_n, n \geq 0\}$ 不可约、非周期，且有平稳分布 π，则

$$\lim_{n\to\infty} \sum_{j\in S} |p_{ij}^{(n)} - \pi_j| = 0,\ \forall i \in S.$$

参见 (钱敏平 et al. 2011) P.19 定理 1.5.5。注意这时马氏链为遍历链，π 是唯一的平稳分布和极限分布。定理说明 $p_{ij}^{(n)}$ 当 $n \to \infty$ 时关于 j 一致地趋于 π_j。

推论 5.10. 若不可约马氏链存在平稳分布 π，则

$$\lim_{n\to\infty} \sum_{j\in S} \left| \frac{1}{n} \sum_{m=0}^{n} p_{ij}^{(m)} - \pi_j \right| = 0,\ \forall i \in S.$$

参见 (钱敏平 et al. 2011) P.21 推论 1.5.6。这时马氏链为正常返，但不一定是非周期的。结论说明 $p_{ij}^{(n)}$ 的局部平均序列 $\frac{1}{n}\sum_{m=0}^{n} p_{ij}^{(m)}$ 当 $n \to \infty$ 时关于 j 一致地趋于 π_j。因为没有要求非周期，所以 $p_{ij}^{(n)}$ 有可能等于 0，所以结论退化到了考虑 $\frac{1}{n}\sum_{m=0}^{n} p_{ij}^{(m)}$，这可以理解为从状态 i 出发，时间 $1, 2, \dots, n$ 中的状态平均处于状态 j 的概率。5.3.4给出了针对逐条轨道的结论。

例 5.22 (用 R 和 Julia 计算平稳分布)**.** 当马氏链为有限、不可约、非周期时，平稳分布存在唯一，只要求解平稳方程。用编程方法求解平稳方程。

平稳分布 π 满足

$$\pi^T P = \pi^T, \quad 1^T \pi = 1,$$

这是一个超定方程：

$$\begin{pmatrix} P^T - I \\ 1^T \end{pmatrix} \pi = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

可以用最小二乘法求解 π。

R 的输入 P 求解 π 的函数：

```
solve.sta <- function(P, rational=FALSE){
  n <- nrow(P)
  y <- c(rep(0,n), 1)
  X <- rbind(t(P) - diag(n), 1)
  p <- solve(crossprod(X), crossprod(X, y)) |> c()
  if(rational){
    MASS::fractions(p)
  } else {
    p
  }
}
```

Julia 语言：

```
using LinearAlgebra
function solve_sta(P, rat=false)
  P = float(P)
  n = size(P,1)
  X = [P' - Int[i==j ? 1 : 0 for i=1:n, j=1:n];
       ones(1, n)]
  y = [ones(n); 1]
  p = X \ y
  if rat
    return Rational.(round.(p; digits=6))
  else
    return p
  end
end
```

例 5.23. 设马氏链的转移阵为

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

求极限分布。

解：易见马氏链为有限状态，不可约，非周期，所以是遍历马氏链，平稳存在唯一且等于极限分布。平稳方程为

$$\begin{cases} \pi_1 &=& 0.5\pi_1 &+& 0.5\pi_2 \\ \pi_2 &=& 0.5\pi_1 &+& 0.5\pi_3 \\ \pi_3 &=& 0.5\pi_2 &+& 0.5\pi_3 \end{cases}$$

求解得

$$\pi=(\pi_1,\pi_2,\pi_3)=(\frac{1}{3},\frac{1}{3},\frac{1}{3}).$$

于是极限分布为

$$\lim_{n\to\infty} p_{ij}^{(n)}=\lim_{n\to\infty} P\{X_n=j|X_0=i\}=\frac{1}{3}.$$

即 0 时刻从 i 出发在很久的时间之后马氏链处于状态 $1,2,3$ 的概率均为 $\frac{1}{3}$，即 X_n 的极限分布为离散均匀分布.

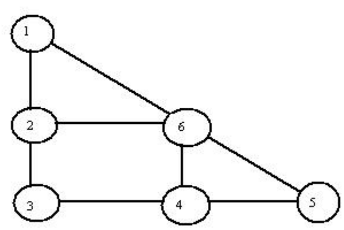


图 5.7: 6 个车站

例 5.24. 设有 6 个车站, 车站中间的公路连接情况如图5.7所示. 汽车每天可以从一个站驶向与之直接相临的车站，并在夜晚到达车站留宿，次日凌晨重复相同的活动. 设每天凌晨汽车开往临近的任何一个车站都是等可能的，试说明很

长时间后，各站每晚留宿的汽车比例趋于稳定. 求出这个比例以便正确地设置各站的服务规模.

解：以 $\{X_n, n = 0, 1, ...\}$ 记第 n 天某辆汽车留宿的车站号. 这是一个马氏链，转移概率矩阵为

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

从图形易见所有状态互通，其中车站 1 可以 2 步或者 3 步返回，所以周期为 1，因此这是遍历马氏链，平稳分布存在唯一且等于极限分布。解平稳方程

$$\begin{cases} \pi P & = & \pi \\ \sum_{i=1}^{6} \pi_i & = & 1 \end{cases}$$

其中 $\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6)$, 可得

$$\pi = (\frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{1}{4}).$$

所以在这个系统运行很长时间后，无论开始汽车从哪一个车站出发它在任一个车站留宿的概率都是固定的，从而所有的汽车也将以一个稳定的比例在各车站留宿.

可以用 R 程序求解上述方程：

```r
P <- matrix(c(
  0, 1/2, 0, 0, 0, 1/2,
  1/3, 0, 1/3, 0, 0, 1/3,
  0, 1/2, 0, 1/2, 0, 0,
  0, 0, 1/3, 0, 1/3, 1/3,
  0, 0, 0, 1/2, 0, 1/2,
  1/4, 1/4, 0, 1/4, 1/4, 0),
  byrow=TRUE, ncol=6)
solve.sta(P, rational=TRUE)
## [1]  1/8 3/16  1/8 3/16  1/8  1/4
```

Julia 语言：

```
P = [
  0 1//2 0 0 0 1//2;
  1//3 0 1//3 0 0 1//3;
  0 1//2 0 1//2 0 0 ;
  0 0 1//3 0 1//3 1//3;
  0 0 0 1//2 0 1//2;
  1//4 1//4 0 1//4 1//4 0]
solve_sta(P, true) |> show
## Rational{Int64}[
##   1//8, 3//16, 1//8, 3//16, 1//8, 1//4]
```

例 5.25. 设甲袋中有 k 个白球和 1 个黑球，乙袋中有 $k+1$ 个白球，每次从两袋中各任取一球，交换后放入对方的袋中. 证明经过 n 次交换后，黑球仍在甲袋中的概率 p_n 满足

$$\lim_{n\to\infty} p_n = \frac{1}{2}.$$

解：以 X_n 表示第 n 次取球后甲袋中的黑球数，则 $\{X_n, n=0,1,2,\dots\}$ 是状态空间为 $S=\{0,1\}$ 的时齐马氏链，一步转移概率矩阵为

$$P = \begin{pmatrix} \frac{k}{k+1} & \frac{1}{k+1} \\ \frac{1}{k+1} & \frac{k}{k+1} \end{pmatrix}.$$

这是不可约、非周期、有限状态马氏链，是遍历的。求解其平稳方程

$$\begin{cases} \pi_0 &= \frac{k}{k+1}\pi_0 + \frac{1}{k+1}\pi_1 \\ \pi_1 &= \frac{1}{k+1}\pi_0 + \frac{k}{k+1}\pi_1 \end{cases}$$

以及 $\pi_0+\pi_1=1$，解得

$$\pi = (\pi_0, \pi_1) = (\frac{1}{2}, \frac{1}{2}),$$

故经过 n 次交换后，黑球仍在甲袋中的概率 p_n 的极限为

$$\lim_{n\to\infty} p_n = \lim_{n\to\infty} P\{X_n = 1\} = \pi_1 = \frac{1}{2}.$$

例 5.26. 我国某种商品在国外销售情况共有连续 24 个季度的数据 (其中 1 表示畅销，2 表示滞销)：

$$1, 1, 2, 1, 2, 2, 1, 1, 1, 2, 1, 2, 1, 1, 2, 2, 1, 1, 2, 1, 2, 1, 1, 1$$

如果该商品销售情况近似满足时齐性与马氏性.

(1) 试确定销售状态的一步转移概率矩阵.

(2) 如果现在是畅销，试预测这之后的第四个季度的销售状况.

(3) 如果影响销售的所有因素不变，试预测长期的销售状况.

解：以 X_n 表示第 n 季度该种商品在国外的销售情况，则 $\{X_n, n = 1, 2, ...\}$ 是一状态空间为 $\{1, 2\}$ 的时齐马氏链. 这是有限状态、不可约、非周期马氏链，是遍历的。

(1) 由 $1 \to 1$ 有 7 次；由 $1 \to 2$ 有 7 次，得 $p_{11} = p_{12} = \frac{7}{14} = \frac{1}{2}$.

由 $2 \to 1$ 有 7 次；由 $2 \to 2$ 有 2 次，得 $p_{21} = \frac{7}{9}, p_{22} = \frac{2}{9}$.

一步转移概率矩阵为

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{7}{9} & \frac{2}{9} \end{pmatrix}.$$

(2)

$$P^{(4)} = P^4 = \begin{pmatrix} 0.611 & 0.389 \\ 0.605 & 0.395 \end{pmatrix}$$

从而

$$p_{11}^{(4)} = 0.611 > p_{12}^{(4)} = 0.389,$$

即如果现在是畅销，这之后第四个季度该种商品将以概率 0.611 畅销.

(3) 由平稳方程

$$(\pi_1, \pi_2) = (\pi_1, \pi_2)P$$

及规范性条件 $\pi_1 + \pi_2 = 1$ 得

$$\begin{cases} \pi_1 & = \frac{1}{2}\pi_1 + \frac{7}{9}\pi_2 \\ \pi_2 & = \frac{1}{2}\pi_1 + \frac{2}{9}\pi_2 \\ \pi_1 + \pi_2 & = 1 \end{cases}$$

解得

$$\pi = (\pi_1, \pi_2) = (\frac{14}{23}, \frac{9}{23}) = (0.61, 0.39).$$

即长期下去，该种商品将以 $\frac{14}{23} = 61\%$ 的概率畅销. 从观测数据的统计看，畅销比例为 $15/24 = 62.5\%$。

不可约的马氏链的极限概率有如下总结：

$$\begin{cases} \text{常返} \begin{cases} \text{正常返}(\Leftrightarrow \text{有平稳分布}) \begin{cases} \text{非周期 (遍历)：} p_{ij}^{(n)} \to \pi_j \\ \text{周期：} \frac{1}{n}\sum_{k=0}^{n} p_{ij}^{(k)} \to \pi_j \end{cases} \\ \text{零常返：} p_{ij}^{(n)} \to 0 \end{cases} \\ \text{非常返：} p_{ij}^{(n)} \to 0 \end{cases}$$

5.3.3 平稳可逆分布

设马氏链 $\{X_n, n \geq 0\}$ 转移概率矩阵为 $P = (p_{ij})$，若存在概率分布 $\pi = \{\pi_i\}$ 使得

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

则称 π 为此马氏链的一个**平稳可逆分布**或**可逆分布**。

平稳可逆分布必为平稳分布：

$$\begin{aligned} \pi_i =& \pi_i \sum_{j \in S} p_{ij} = \sum_{j \in S} \pi_i p_{ij} \\ =& \sum_{j \in S} \pi_j p_{ji}, \end{aligned}$$

即 π 是平稳分布。

所以，为了构造一个具有平稳分布的马氏链，只要取 P 满足不可约性和平稳可逆条件(5.3.3)，(5.3.3)也称为细致平衡 (detailed balanced) 条件。这时所有状态是正常返的。MCMC 算法的设计就利用了这种性质。

如果 X_0 服从平稳可逆分布，则将 $\{X_n\}$ 的时间方向反转后仍是以 P 为转移概率矩阵的马氏链。

5.3.4 强大数律

定理 5.16 (遍历定理)**.** *设 $\{X_n, n \geq 0\}$ 是不可约、正常返马氏链，μ_j 是从 j 出发返回 j 的平均时间，则 $\{X_n\}$ 有平稳分布 $\{\pi_i = \frac{1}{\mu_j}\}$，且在 X_0 服从任何初始分布的条件下，都有*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_{\{X_n = i\}} = \pi_i,\ a.s.$$

即平稳分布也是马氏链处于状态 i 的长期比例；反过来，若 $\{X_n, n \geq 0\}$ 是不可约马氏链但平稳分布不存在，则状态必为非常返或零常返。

参见 (Ross 2019) P. 218 定理 4.1, (钱敏平 et al. 2011) P.23 定理 1.6.1 和定理 1.6.3。

定理 5.17. *设 $\{X_n, n \geq 0\}$ 是不可约、正常返马氏链，则 $\{X_n\}$ 有平稳分布 $\{\pi_i\}$，设随机变量 Y 分布为 $\{\pi_i\}$，函数 $g(x)$ 定义在状态空间 S 上且为有界函数，则*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g(X_n) = Eg(Y),\ a.s.$$

见 (Ross 2019) P.230 命题 4.6，(钱敏平 et al. 2011) P.26 推论 1.6.6。

证明：记 $a_i(N)$ 为 $\{X_n\}$ 在 $n = 1, 2, \dots, N$ 中等于状态 i 的次数，则由定理5.16可知

$$\lim_{N\to\infty} \frac{a_i(N)}{N} = \pi_i,\ \text{a.s.}$$

而

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^{N} g(X_n) = & \frac{1}{N} \sum_{i=0}^{\infty} a_i(N) g(i) \\ = & \sum_{i=0}^{\infty} \frac{a_i(N)}{N} g(i), \end{aligned}$$

设 $|g(i)| \leq C, \forall i \in S$，则

$$\sum_{i=0}^{\infty} \left| \frac{a_i(N)}{N} g(i) \right| \leq C \frac{\sum_{i=0}^{\infty} a_i(N)}{N} = C,$$

由控制收敛定义可得

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}g(X_n)=\sum_{i=0}^{\infty}\pi_i g(i)=Eg(Y).$$

推论 5.11. 若 $\{X_n\}$ 是不可约有限状态马氏链，则 $\{X_n\}$ 存在平稳分布 $\{\pi_i\}$，对于定义在状态空间 S 上的任意函数 $g(x)$，都有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}g(X_n)=\sum_{i\in S}g(i)\pi_i,\ a.s.$$

证明：这时 $\{X_n\}$ 必为正常返状态，且 $g(x)$ 为有界函数，由定理5.17可知结论成立。

定理 5.18. 设 $\{X_n, n\geq 0\}$ 是不可约、正常返马氏链，则有 $\{X_n\}$ 平稳分布 $\{\pi_i\}$，设随机变量 Y 分布为 $\{\pi_i\}$，函数 $g(x)$ 定义在状态空间 S 上，满足 $E|g(Y)|<\infty$，则

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}g(X_n)=Eg(Y),\ a.s.$$

见 (何书元 2008) P.194 定理 8.2(3)。

推论 5.12. 如果 $\{X_n\}$ 是遍历马氏链，则 $\{X_n\}$ 存在唯一的平稳分布 $\{\pi_i\}$（也是极限分布），设定义在状态空间 S 上的函数 $g(\cdot)$ 满足 $E|g(Y)|<\infty$，则

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}g(X_n)=Eg(Y),\ a.s.$$

这个推论说明，可以用模拟马氏链的方法来估计与平稳分布 $\{\pi\}$ 有关的数字特征。设随机向量 Y 服从分布 $\{\pi\}$，$E[g(Y)]$ 很难计算，直接生成 Y 的简单随机样本也很困难，就可以设计马氏链 $\{X_n\}$，使得 $\{X_n\}$ 遍历且以 $\{\pi\}$ 为极限分布和平稳分布，则可以用 $\frac{1}{N}\sum_{n=1}^{N}g(X_n)$ 估计 $E[g(Y)]$。虽然只要 $\{X_n\}$ 不可约、正常返就可以，但是最好加上非周期条件，使得从任意初始状态 i 出发到任意状态 j 的概率极限都等于 π_j。对于不可约、非周期马氏链，若转移概率满足细致平衡条件，则存在平稳分布，从而是遍历链，所以在 MCMC 算法中，Metropolis-Hastings 算法就利用了这样的设计。

5.4 马氏链的应用

5.4.1 群体消失模型 (离散时间分支过程)

考虑一个能产生同类后代的个体组成的群体. 每一个体生命结束时以概率 $p_j(j = 0,1,2,3,...)$ 产生了 j 个新的后代，与别的个体产生的后代个数相互独立. 初始的个体数以 X_0 表示，称为第零代的总数；第零代的后代构成第一代，其总数记为 X_1，第一代的每个个体以同样的分布产生第二代，……，一般地，以 X_n 记第 n 代的总数. 这里的 n 代表第几代而非具体时间。此马氏链 $\{X_n, n = 0,1,2,...\}$ 称为**离散时间分支过程**.

现在假设群体是从单个祖先开始的，即 $X_0 = 1$，则有

$$X_{n+1} = \sum_{i=1}^{X_n} Z_{n,i},\ n = 0,1,...$$

其中 $Z_{n,i}$ 表示第 n 代的第 i 个成员的后代的个数.

首先来考虑第 n 代的平均个体数 $E[X_n]$，设

$$\mu = \sum_{i=0}^{\infty} ip_i$$

是每个个体的后代个数的均值，对 X_n 取条件期望，有

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

可以看出，若 $\mu < 1$，则平均个体数单调下降趋于 0. 若 $\mu = 1$ 时，各代平均个体数相同. 当 $\mu > 1$ 时，平均个体数按指数阶上升至无穷.

下面就来考虑群体最终会消亡的概率 π_0. 对第一代个体数取条件，则

$$\pi_0 = P\{\text{群体消亡}\} \tag{5.15}$$

$$= \sum_{j=0}^{\infty} P\{\text{群体消亡}|X_1 = j\} \cdot p_j \tag{5.16}$$

$$= \sum_{j=0}^{\infty} \pi_0^j p_j \tag{5.17}$$

上面的第二个等式是因为群体最终灭绝是以第 1 代为祖先的 j 个家族全部消亡，而各家族已经假定为独立的，每一家族灭绝的概率均为 π_0.

很自然我们会假设：家族消亡与 μ 有关，在此我们给出一个定理，以证明 $\pi_0 = 1$ 的充要条件是 $\mu \le 1$。不考虑 $p_0 = 1$ 和 $p_0 = 0$ 的平凡情况，即家族在第零代后就消失或永不消失.

定理 5.19. 设 $0 < p_0 < 1$，则 $\pi_0 = 1 \iff \mu \le 1$.

证明： 定义

$$F(x) = \sum_{j=0}^{\infty} p_j x^j,\ 0 \le x \le 1,$$

显然 $F(0) = 0,\ F(1) = 1$，$F(x)$ 是严格单调增函数。由(5.17)式可知 π_0 满足

$$F(\pi_0) = \pi_0,$$

即 π_0 是直线 $y = x$ 和曲线 $y = F(x)$ 交点的横坐标。显然 (1,1) 是一个交点.

当 $p_0 + p_1 = 1$ 时，$y = F(x) = p_0 + p_1 x$ 是一条直线，斜率 p_1 在 $(0, 1)$ 之间，与 $y = x$ 只能有一个交点 $(1, 1)$，即这时必有 $\pi_0 = 1$，最终消亡概率为 1。

当 $p_0 + p_1 < 1$ 时，由于

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

可见这时 $F(x)$ 是严格单调增的严格凸函数，严格凸函数与任一直线都至多有 2 个交点。曲线 $y = F(x)$ 与直线 $y = x$ 除了交点 $(1, 1)$ 以外，只可能有 0 或者 1 个交点。

于是，当 $F(x)$ 非直线时，可按与 $y = x$ 的交点是 1 个还是两个分为如下两种情况：

(1) 只有 $(1, 1)$ 一个交点的情况。这时群体最终消亡概率等于 1。

(2) 存在一个 $0 < s < 1$ 使得 $F(s) = s$, 那么此时 π_0 应取 s 还是 1 呢？可以证明，π_0 必定取值为 s，为此只需证明 π_0 是方程 $x = F(x)$ 的最小解，利用这种情况下最小解为 s 就可得 $\pi_0 = s < 1$.

设 π 是方程 $x = F(x)$ 的任意一个解，利用归纳法证明 $\pi \geq P(X_n = 0)$。当 $n = 1$ 时，

$$\pi = \sum_{j=0}^{\infty} \pi^j p_j \geq \pi^0 p_0 = p_0 = P\{X_1 = 0\}.$$

归纳地，假设 $\pi \geq P\{X_n = 0\}$ 成立, 则

$$\begin{aligned} P\{X_{n+1} = 0\} &= \sum_{j=0}^{\infty} P\{X_{n+1} = 0 | X_1 = j\} \cdot p_j \\ &= \sum_{j=0}^{\infty} (P\{X_n = 0\})^j \cdot p_j \\ &\leq \sum_{j=0}^{\infty} \pi^j p_j \\ &= \pi. \end{aligned}$$

从而对一切 n，

$$\pi \geq P\{X_n = 0\}.$$

令 $A_n = \{X_n = 0\}$，则 $A_n \subset A_{n+1}$，于是

$$\begin{aligned} \pi_0 &= P(\{\text{群体最终灭绝}\}) \\ &= P\left(\bigcup_{n=1}^{\infty} \{X_n = 0\}\right) \\ &= \lim_{n\to\infty} P(X_n = 0) \\ &\leq \pi, \end{aligned}$$

即 π_0 是方程 $x = F(x)$, $x \in [0, 1]$ 的最小解，这就证明了在这种情况下 π_0 取值应为 s，最终消亡概率小于 1。

这样，可以根据 $F(x)$ 的不同表现将分布 $\{p_j\}$ 分为如下三种情况，每一种情况下有对应的 μ 值：

第一种情况，若 $p_0 + p_1 = 1$，已证明 $\pi_0 = 1$，而显然

$$\mu = \sum_{j=0}^{\infty} j p_j = 0 \times p_0 + 1 \times p_1 = p_1 < 1.$$

第二种情况，若 $p_0 + p_1 < 1$ 而且方程 $x = F(x)$ 只有一个根，有 $\pi_0 = 1$。因为 $F(0) > 0$, $F(1) = 1$, 由 $F(x)$ 和 $y = x$ 的连续性可知曲线 $y = F(x)$ 在直

线 $y=x$ 上方，即 $F(x)>x, x\in(0,1)$。这时必有此时

$$
\begin{aligned}
F'(1)=&\lim_{x\to 1-0}\frac{F(x)-F(1)}{x-1}=\lim_{x\to 1-0}\frac{1-F(x)}{1-x}\\
\le&\lim_{x\to 1-0}\frac{1-x}{1-x}\quad(\text{因为}F(x)\ge x)\\
=&1.
\end{aligned}
$$

而

$$
F'(1)=\sum_{j=0}^{\infty}jp_j=\mu,
$$

从而 $\mu\le 1$.

第三种情况下，由 $F(x)$ 得严格凸性可知对 $x\in(s,1)$ 有 $F(x)<x$, 类似可得 $F'(1)=\mu>1$。

综合三种情况就可以推出 $\pi_0=1 \iff \mu\le 1$.

在实际应用中，考虑一个群体的真实增长时，分支过程的假定在群体达到无限之前就不成立了 (比如独立同分布性). 但另一方面，利用分支过程研究消亡现象是有意义的，因为一般灭绝常常发生在过程的早期.

5.4.2 人口结构变化的马氏链模型

这个应用在原教材中论述不够充分，略去。

5.5 连续时间马氏链

5.5.1 概念

前面几节讨论的是时间和状态空间都是离散的马氏过程，本节我们将介绍另外一种情况的马氏过程，它的状态空间仍然是离散的，但时间是连续变化的，称为连续时间马氏链. 我们会给出它的一些性质，一个重要的方程 (Kolmogrov 方程) 和一个重要的应用 (生灭过程).

定义 5.14. 设过程 $\{X(t),t\ge 0\}$ 的状态空间 S 为离散空间，为方便书写设 $S=\{0,1,2,...\}$ 或其子集. 若对一切 $s,t\ge 0$ 及 $i,j\in S$, 有

$P\{X(t+s)=j|X(s)=i,X(u)=x(u),0\le u<s\}=P\{X(t+s)=j|X(s)=i\}$

成立，则称 $\{X(t), t \geq 0\}$ 是一个**连续时间马氏链**.

连续时间马氏链定义中是 $\{x(u), 0 \leq u < s\}$ 和 $x(s) = i$ 条件下对未来的单个时间点 $t+s$ 的条件分布仅依赖于 $X(s) = i$，实际上，对 $\sigma(\{x(u) : u > s\})$ 中的事件的条件概率也仅依赖于 $X(s) = i$。

条件概率 $P\{X(t+s) = j | X(s) = i\}$ 记作 $p_{ij}(s,t)$，表示过程在时刻 s 处于状态 i，经 t 时间后转移到 j 的转移概率，并称 $P(s,t) = (p_{ij}(s,t))$ 为相应的转移概率矩阵.

定义 5.15. 称连续时间马氏链是**时齐的**，若 $p_{ij}(s,t)$ 与 s 无关. 简记 $p_{ij}(s,t) = p_{ij}(t)$，相应地记 $P(t) = (p_{ij}(t))$.

我们只讨论时齐的连续时间马氏链，并且简称为连续时间马氏链 (在不引起混淆的情况下有时也称为马氏链).

对于连续时间马氏链来说, 除了要考虑在某一时刻它将处于什么状态外, 还关心它在离开这个状态之前会停留多长的时间，从连续时间的马氏性来看，因为已知 t 时刻的状态后，未来的发展与 t 时刻之前的轨迹无关，当然也与到 t 时刻为止已经在当前状态停留的时间长度无关，所以系统在某一状态的 “停留时间” 具备 “无记忆性” 的特征，应该服从指数分布，下面我们给出一个具体的解释.

定理 5.20. *设 $\{X(t), t \geq 0\}$ 是连续时间马氏链, 假定在时刻 0 过程刚刚到达 $i (i \in S)$. 以 τ_i 记过程在离开 i 之前在 i 停留的时间，则 τ_i 服从指数分布.*

证明：我们只需证明对 $s, t \geq 0$, 有

$$P\{\tau_i > s+t | \tau_i > s\} = P\{\tau_i > t\}, \tag{5.18}$$

即无记忆性.

注意到

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

则

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

这里实际上利用了更强的马氏性：$\mathscr{F}_s$ 下关于 $\sigma(\{X(u): u>s\})$ 可测随机变量的条件期望等于对 $X(s)$ 的条件期望。

由上述定理，实际上我们得到了另外一个构造连续时间马氏链的方法，它是具有如下两条性质的随机过程.

(1) 在转移到下一个状态之前处于状态 i 的时间服从速率参数为 μ_i 的指数分布；

(2) 在过程离开状态 i 时，将以概率 p_{ij} 到达 j，且 $\sum_{j\in S} p_{ij} = 1$.

注: 当 $\mu_i = \infty$ 时，它在状态 i 停留的平均时间为 0，即一旦进入马上离开，称这样的状态为瞬过的. 瞬过状态的实际应用价值不大，我们假设所考虑的连续时间马氏链中不存在瞬过状态，即设

$$0 \le \mu_i < +\infty, \ \forall i.$$

若 $\mu_i = 0$，称 i 为吸收态，即一旦进入，将停留的平均时间为无限长. 由此我们看出，连续时间马氏链是一个做下面的运动的随机过程：它以一个马氏链的方式在各个状态之间转移，在两次转移之间以指数分布停留. 从上面已经知道，这个指数分布与过程现在所处的状态有关，但一定要与下面将转移到的状态独立 (试用马氏性说明原因).

定义 5.16. 称一个连续时间马氏链是**正则的**，若以概率 1 在任意有限长的时间内转移的次数是有限的.

如果非正则，就存在一个 $\tau<\infty$，在 $t\to\tau-$ 时转移间隔趋于零，在 $[0,\tau]$ 内发生无穷次转移，不容易定义 $t>\tau$ 时的马氏链状态。某些教材称正则条件为**非爆炸**的。

由正则性可得连续性条件

$$\lim_{t\to 0} p_{ij}(t) = \delta_{ij}.$$

以下我们总假定所考虑的马氏链都满足正则性条件. 下面是几个连续时间马氏链的典型例子.

例 5.27 (Poisson 过程)**.** 参数为 λ 的 Poisson 过程 $\{N(t), t \geq 0\}$，取值于 $\{0, 1, 2, ...\}$. 由第3章知道，它在任一个状态 i 停留的时间服从指数分布，并且在离开 i 时以概率 1 转到 $i+1$(又一个事件发生)，由 Poisson 过程的独立增量性容易看出它在 i 停留的时间与状态的转移是独立的 (特别是由它的平稳增量性 $\mu_i = \mu_{i+1} = \lambda$, $i = 0, 1, 2, ...$)，从而 Poisson 过程是时齐的连续时间马氏链. 对 $i \in S$, 它的转移概率为

$$\begin{aligned}
p_{i,i}(t) =& P\{N(t+s) = i | N(s) = i\} \\
=& P\{N(t) = 0\} \\
=& e^{-\lambda t} \\
p_{i,i+1}(t) =& P\{N(t+s) = i+1 | N(s) = i\} \\
=& P\{N(t) = 1\} \\
=& \lambda t e^{-\lambda t} \\
p_{i,j}(t) =& \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}, j > i+1 \\
p_{i,j}(t) =& 0, j < i.
\end{aligned}$$

例 5.28 (Yule 过程)**.** 考察生物群体繁殖过程的模型. 设群体中各个生物体的繁殖是相互独立的、强度为 λ 的 Poisson 过程，并且群体中没有死亡，此过程称为 Yule 过程. 我们来说明 Yule 过程是一个连续时间马氏链.

解: 设在时刻 0 群体中有 1 个个体，则群体将有的个体数是 $\{1, 2, 3, ...\}$. 以 $T_i (i \geq 1)$ 表示群体数目从 i 增加到 $i+1$ 所需的时间（注意 T_i 是间隔时间而不是第 i 个事件的发生时间)，由 Yule 过程定义，当群体数目为 i 时, 这 i 个个体是以相互独立的 Poisson 过程来产生后代的. 由 Poisson 过程的可加性知，这相当于一个强度为 λi 的 Poisson 过程；由 Poisson 过程的独立增量性，易知 T_i 与状态的转移是独立的 $(i \geq 1)$，并且 $\{T_i\}$ 是相互独立的参数为 $i\lambda$ 的指数随机变量. 这就说明了 Yule 过程是一个连续时间马氏链，我们来求它的转移概率 $p_{ij}(t)$.

首先

$$
\begin{aligned}
P\{T_1 \le t\} =& 1-e^{-\lambda t} \\
P\{T_1+T_2 \le t\} =& \int_0^t P\{T_1+T_2 \le t | T_1 = x\}\lambda e^{-\lambda x}dx \\
& \text{(利用全期望公式)} \\
=& \int_0^t P\{T_2 \le t-x | T_1 = x\}\lambda e^{-\lambda x}dx \\
=& \int_0^t P\{T_2 \le t-x\}\lambda e^{-\lambda x}dx \\
=& \int_0^t (1-e^{-2\lambda(t-x)})\lambda e^{-\lambda x}dx \\
& \text{(注意}T_2 \sim \mathrm{Exp}(2\lambda)) \\
=& (1-e^{-\lambda t})^2 \\
P\{T_1+T_2+T_3 \le t\} =& \int_0^t P\{T_1+T_2+T_3 \le t | T_1+T_2 = x\} dP\{T_1+T_2 \le x\} \\
=& \int_0^t (1-e^{-3\lambda(t-x)})2\lambda e^{-\lambda x}(1-e^{-\lambda x})dx \\
=& (1-e^{-\lambda t})^3 \\
& \ldots\ldots
\end{aligned}
$$

一般地，用归纳法不难证明

$$P\{T_1+T_2+\cdots+T_j \le t\} = (1-e^{-\lambda t})^j.$$

而

$$\{T_1+T_2+\cdots+T_j \le t\} \iff \{X(t) \ge j+1 | X(0) = 1\},$$

所以

$$
\begin{aligned}
p_{1j}(t) =& P\{X(t) = j | X(0) = 1\} && (5.19) \\
=& P\{X(t) \ge j | X(0) = 1\} - P\{X(t) \ge j+1 | X(0) = 1\} && (5.20) \\
=& (1-e^{-\lambda t})^{j-1} - (1-e^{-\lambda t})^j && (5.21) \\
=& e^{-\lambda t}(1-e^{-\lambda t})^{j-1} \quad (j \ge 1). && (5.22)
\end{aligned}
$$

这是一个几何分布，均值为 $e^{\lambda t}$.

又因为

$$p_{ij}(t) = P\{X(s+t) = j | X(s) = i\},$$

相当于一个总量从 i 个个体开始的 Yule 过程的群体总数在 t 时间内增加到 j 的概率. 而这相当于 i 个独立的服从(5.22)式的几何随机变量的和取值为 j 的概率 (想一想为什么)，是 Pascal 分布，于是

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

例 5.29 (纯生过程)**.** 考虑 Yule 过程的推广。注意 $X(t)$ 仍为计数过程，T_i 为第 i 个事件的到来时间间隔，T_i 服从参数为 $i\lambda$ 的指数分布且相互独立。推广假设这些到来间隔服从参数为 λ_i 的指数分布，则这样的过程称为纯生过程。

对纯生过程，如果间隔时间越来越短，有可能在有限时间内发生无穷个事件。记

$$T = \sum_{i=1}^{\infty} T_i.$$

则事件 $\{T < \infty\}$ 表示在有限时间内发生了无穷次事件。可以证明，

$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i} < \infty \Longrightarrow P(T < \infty) = 1;$$
$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i} = \infty \Longrightarrow P(T < \infty) = 0.$$

(参见 (钱敏平 et al. 2011) P.72 命题 2.2.1。) 据此可知 Yule 过程在有限时间内不会有无穷次事件发生，因此 Yule 过程在任意时间 t 时的个体数 $X(t)$ 都是有穷值（以概率 1 意义下）。

例 5.30 (生灭过程)**.** 仍然考虑一个生物群体的繁殖模型. 每个个体生育后代如例5.28的假定，但是每个个体将以指数速率 μ 死亡，这是一个生灭过程，生灭过程的状态空间为 $\{0, 1, 2, \dots\}$. 在状态 $i (i \geq 1)$，它能转移到 $i+1$(生了一个) 或 $i-1$(死了一个)。以 T_i 记过程从 i 到达 $i+1$ 或 $i-1$ 的时间. 类似例5.28可以得到，T_i 相互独立且与状态会转移到 $i+1$ 或是 $i-1$ 是独立的，T_i 服从参数为 $i\mu + i\lambda$ 的指数分布 (把生灭过程看成两个 Yule 过程之和，一个生，一个灭)，并且下一次转移到 $i+1$ 的概率 $p_{i,i+1}$ 是 $\frac{\lambda}{\mu+\lambda}$(见第 3 章习题 3.3)，到

$i-1$ 的概率 $p_{i,i-1}$ 为 $\frac{\mu}{\mu+\lambda}$. 这样的生灭过程也称为连续时间的分支过程。概率转移 $p_{ij}(t)$ 将在下一小节用 Kolmogorov 微分方程导出。

生灭过程的更一般的定义是当 $X(t)=i$ 时，有参数为 λ_i 的指数分布随机变量 S，和参数为 μ_i 的指数分布随机变量 T，S,T 相互独立，如果 $S<T$ 就在 $t+S$ 后转移到 $i+1$，否则就转移到 $i-1$。这些转移间隔之间独立。

例 5.31 (M/M/S 排队系统)**.** 顾客的来到是参数为 λ 的 Poisson 过程. 服务员数为 S 个，每个顾客接受服务的时间服从速率参数为 μ 的指数分布. 遵循先来先服务，若服务员没有空闲就排队的原则. 以 $X(t)$ 记 t 时刻系统中的总人数，则 $\{X(t),t\geq 0\}$ 是一个生灭过程 (来到看作出生，离去看作死亡)，来到率是恒定参数为 λ 的 Poisson 过程，离去过程的参数会发生变化，以 μ_n 记系统中有 n 个顾客时的离去率，则

$$\mu_n=\begin{cases} n\mu & 1\leq n\leq S,\\ S\mu & n>S.\end{cases}$$

请读者自己证明.

5.5.2 Kolmogorov 微分方程

对于离散时间马氏链，如果已知其转移概率矩阵 $P=(p_{ij})$，则其 n 步转移概率矩阵由其一步转移矩阵的 n 次方可得. 但是对于连续时间马氏链，转移概率 $p_{ij}(t)$ 的求解一般比较复杂. 下面我们先考虑 $p_{ij}(t)$ 的一些性质.

定理 5.21. *时齐连续时间马氏链的转移概率 $p_{ij}(t)$ 满足:*

(1) $p_{ij}(t)\geq 0$;

(2) $\displaystyle\sum_{j\in S}p_{ij}(t)=1$;

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

证明：(1) 和 (2) 由 $p_{ij}(t)$ 的定义易知. 下面证明 (3)。

$$
\begin{aligned}
p_{ij}(t+s) =& P\{X(t+s)=j|X(0)=i\} \\
=& \sum_{k\in S} P\{X(t+s)=j, X(t)=k|X(0)=i\} \\
=& \sum_{k\in S} P\{X(t+s)=j|X(t)=k, X(0)=i\} P\{X(t)=k|X(0)=i\} \\
=& \sum_{k\in S} P\{X(t+s)=j|X(t)=k\} p_{ik}(t) \\
=& \sum_{k\in S} p_{ik}(t) p_{kj}(s)
\end{aligned}
$$

一般称 (3) 为连续时间马氏链的 C-K 方程.

转移概率 $p_{ij}(t)$ 的性质：

命题 5.3. 设 $\{X(t), t\geq 0\}$ 为正则的连续时间马氏链，转移概率函数为 $p_{ij}(t)$。则

(1) $\lim_{h\downarrow 0} X(t+h) = X(t)$, *a.s.*

(2) $p_{ij}(t)$ 在 $t=0$ 连续：$\lim_{t\downarrow 0} p_{ij}(t) = p_{ij}(0)$。

(3) $p_{ij}(t)$ 在 $t\in[0,\infty)$ 一致连续，满足

$$
\sum_{j\in S} |p_{ij}(t+h) - p_{ij}(t)| \leq 2(1-p_{ii}(h)).
$$

(4) $p_{ii(t)} > 0$, $\forall t\geq 0$。

(5) $p_{ij}(t)$ 在 $t=0$ 可导：

$$
\lim_{t\downarrow 0} \frac{p_{ij}(t)-p_{ij}(0)}{t} = \begin{cases} q_{ij}, & i\neq j, \\ -q_{ii}, & i=j. \end{cases}
$$

(6) 若 $q_{ii}=0$，则 $p_{ii}(t)=1$, $\forall t\geq 0$，即 i 是吸收态。

(7) $q_{ii} = \sum_{t>0} \frac{1-p_{ii}(t)}{t}$.

见 (何书元 2008) P.210 定理 3.1 和推论。

定理 5.22.

$$(1)\ \lim_{t\to 0}\frac{1-p_{ii}(t)}{t}=q_{ii}\le\infty;$$

$$(2)\ \lim_{t\to 0}\frac{p_{ij}(t)}{t}=q_{ij}<\infty.$$

称 q_{ij} 为从状态 i 转移到 j 的转移速率. 这个速率的分子是很短时间内转移到 j 的概率，分母是时间区间长度。而 q_{ii} 是从状态 i 离开的速率。

命题 5.4. 连续时间马氏链在状态 i 停留的时间 T_i 服从速率参数为 q_{ii} 的指数分布。

证明：设 T_i 的速率参数为 ν_i，则

$$P(T_i\le t)=1-e^{-\nu_i t}.$$

由正则性，$t\to 0$ 时 $[0,t]$ 时间内最多仅有一次转移，所以

$$\{T_i\le t\}\iff\{\exists u\in(0,t]\ \text{s.t.}\ X(u)\ne i|X(0)=i\}$$

近似等于 $\{X(t)\ne i|X(0)=i\}$（当 $t\to\infty$），于是

$$\begin{aligned}
q_{ii}=&\lim_{t\to 0}\frac{1-p_{ii}(t)}{t}\\
=&\lim_{t\to 0}\frac{1}{t}P(X(t)\ne i|X(0)=i)\\
=&\lim_{t\to 0}\frac{1}{t}P(T_i\le t)=\lim_{t\to 0}\frac{1}{t}(1-e^{-\nu_i t})\\
=&\nu_i.
\end{aligned}$$

这个证明不够严格。

推论 5.13. 对有限状态时齐的连续时间马氏链，有

$$q_{ii}=\sum_{j\ne i}q_{ij}<\infty.$$

证明：由 $\sum_{j\in S}p_{ij}(t)=1$ 有

$$1-p_{ii}(t)=\sum_{j\ne i}p_{ij}(t).$$

故由定理5.22可知

$$\lim_{t\to 0}\frac{1-p_{ii}(t)}{t}=\lim_{t\to 0}\sum_{j\neq i}\frac{p_{ij}(t)}{t}$$
$$=\sum_{j\neq i}\lim_{t\to 0}\frac{p_{ij}(t)}{t}=\sum_{j\neq i}q_{ij}$$
$$<\infty.$$

注：对于无限状态的情况，一般只能得到

$$q_{ii}\geq\sum_{j\neq i}q_{ij}.$$

为了简单起见, 设状态空间为 $S=\{1,2,\dots,n,\dots\}$，此时记

$$Q=\begin{pmatrix} -q_{11} & q_{12} & q_{13} & \cdots & q_{1i} & \cdots \\ q_{21} & -q_{22} & q_{23} & \cdots & q_{2i} & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots \\ q_{i1} & q_{i2} & q_{i3} & \cdots & -q_{ii} & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \ddots \end{pmatrix}$$

称为连续时间马氏链的 Q 矩阵，当矩阵元素 $q_{ii}=\sum_{j\neq i}q_{ij}<\infty$ 时，称该矩阵为**保守的**. 对有限状态时齐的连续时间马氏链，Q 矩阵都是保守的。若所有 q_{ii} 都有限，则 Q 矩阵为保守的。对正则的马氏链，Q 矩阵都是保守的（见 (何书元 2008) 节 5.4.A）。

若 $q_{ii}=\infty$，则马氏链在状态 i 停留时间等于 0，这个状态相当于不起作用；若 $\sum_{j\neq i}q_{ij}<q_{ii}<\infty$，则从状态 i 出发有正概率没有目的地。所以一般仅考虑保守的情况。

下面我们应用这两个定理及推论，来导出一个重要的微分方程.

定理 5.23 (Kolmogorov 微分方程). *设连续时间马氏链 $\{X(t),t\geq 0\}$ 的 Q 矩阵是保守的，即*

$$\sum_{j\neq i}q_{ij}=q_{ii}<\infty,\ \forall i\in S,$$

则对一切 $i,j\in S$ 和 $t\geq 0$，有

(1) 向后方程：

$$p'_{ij}(t) = \sum_{k \neq i} q_{ik} p_{kj}(t) - q_{ii} p_{ij}(t). \tag{5.23}$$

(2) 在适当的正则条件下，有向前方程:

$$p'_{ij}(t) = \sum_{k \neq j} p_{ik}(t) q_{kj} - p_{ij}(t) q_{jj}. \tag{5.24}$$

证明：（1）由连续时间马氏链的 C-K 方程有

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

或等价地

$$p_{ij}(t+h) - p_{ij}(t) p_{ii}(h) = \sum_{k \neq i} p_{ik}(h) p_{kj}(t).$$

变形为

$$p_{ij}(t+h) - p_{ij}(t) = \sum_{k \neq i} p_{ik}(h) p_{kj}(t) - (1 - p_{ii}(h)) p_{ij}(t).$$

于是

$$\lim_{h \to 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} \tag{5.25}$$

$$= \lim_{h \to 0} \sum_{k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) - \lim_{h \to 0} \frac{(1 - p_{ii}(h))}{h} p_{ij}(t). \tag{5.26}$$

若此马氏链状态是有限的，应用定理5.22和推论5.13从上式直接可得 (1)(向后方程).

下面证明对于无限状态情形向后方程成立. 由式(5.26)我们只需证明其中的极限与求和可交换次序即可.

对于固定的 N，有

$$\begin{aligned}
&\liminf_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\\
\geq&\liminf_{h\to 0}\sum_{\substack{k\neq i\\ k<N}}\frac{p_{ik}(h)}{h}p_{kj}(t)\\
=&\sum_{\substack{k\neq i\\ k<N}}\liminf_{h\to 0}\frac{p_{ik}(h)}{h}p_{kj}(t)\\
=&\sum_{\substack{k\neq i\\ k<N}}q_{ik}p_{kj}(t).
\end{aligned}$$

由 N 的任意性，得

$$\liminf_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\geq\sum_{k\neq i}q_{ik}p_{kj}(t). \tag{5.27}$$

又因为 $p_{kj}(t)\leq 1,\ \forall k\in S$，所以

$$\begin{aligned}
&\limsup_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\\
\leq&\limsup_{h\to 0}\left[\sum_{k\neq i,k<N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\sum_{k\geq N}\frac{p_{ik}(h)}{h}\right]\\
=&\limsup_{h\to 0}\left[\sum_{k\neq i,k<N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\left(\sum_{k}\frac{p_{ik}(h)}{h}-\sum_{k<N}\frac{p_{ik}(h)}{h}\right)\right]\\
=&\limsup_{h\to 0}\left[\sum_{k\neq i,k<N}\frac{p_{ik}(h)}{h}p_{kj}(t)+\left(\frac{1-p_{ii}(h)}{h}-\sum_{k\neq i,k<N}\frac{p_{ik}(h)}{h}\right)\right]\\
=&\sum_{k\neq i,k<N}q_{ik}p_{kj}(t)+q_{ii}-\sum_{k\neq i,k<N}q_{ik}.
\end{aligned}$$

同样由 N 的任意性，可知

$$\limsup_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)\leq\sum_{k\neq i}q_{ik}p_{kj}(t). \tag{5.28}$$

$$(\text{因为}q_{ii}=\sum_{k\neq i}q_{ik}<\infty) \tag{5.29}$$

结合(5.27)和(5.28)就证明了

$$\lim_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)=\sum_{k\neq i}q_{ik}p_{kj}(t).$$

于是向后方程得证.

向后方程用矩阵形式写出即是

$$P'(t) = QP(t),$$

其中 $P(t) = (p_{ij}(t))$, $Q = (q_{ij})$.

(2) 在 (1) 中计算 $t+h$ 的状态时是对退后到时刻 h 的状态来取条件的 (所以称为后退方程)，这里我们考虑对时刻 t 的状态取条件，用 C-K 方程

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

同理得到

$$\begin{aligned} &\lim_{h \to 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} \\ =& \lim_{h \to 0} \left[\sum_{k \neq j} p_{ik}(t) \frac{p_{kj}(h)}{h} - \frac{1 - p_{jj}(h)}{h} p_{ij}(t) \right]. \end{aligned}$$

假若上式中极限与求和号可交换，则有定理的结论 (2) 成立，以矩阵形式写出即为

$$P'(t) = P(t)Q.$$

但是这个假定不一定成立，所以在定理中，我们加了“适当正则”这个条件。对于有限状态的马氏链或生灭过程 (特别是只生不灭的纯生过程)，它都是成立的.

例 5.32 (泊松过程). 利用 Kolmogorov 微分方程推导泊松过程的转移概率。

由泊松过程的等价定义，

$$\begin{cases} p_{k,j}(h) = 0, & j < k; \\ p_{k,j}(h) = o(h), & j - k \geq 2; \\ p_{k,k+1}(h) = \lambda h + o(h); \\ p_{k,k}(h) = 1 - \lambda h + o(h) \end{cases}$$

于是

$$\begin{cases} q_{kj} = 0, \quad j < k \text{ 或} j \geq k+2; \\ q_{k,k+1} = \lim_{h \to 0} \frac{p_{k,k+1}(h)}{h} = \lambda; \\ q_{kk} = \lim_{h \to 0} \frac{1 - p_{k,k}(h)}{h} = \lambda. \end{cases}$$

利用 Kolmogorov 向后方程，

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

上式求和中仅 $k = i + 1$ 的项非零，故

$$\begin{aligned} p'_{ij}(t) =& q_{i,i+1}p_{i+1,j}(t) - q_{ii}p_{ij}(t) \quad (5.30) \\ =& \lambda p_{i+1,j}(t) - \lambda p_{ij}(t). \quad (5.31) \end{aligned}$$

因为 $p_{ij}(t) = 0$ 对 $j < i$，所以只需要求解 $j \geq i$ 的情形。并注意

$$p_{ii}(0) = 1, \quad p_{ij}(0) = 0 (j \neq i).$$

对 $j = i$，(5.31)变成

$$p'_{ii}(t) =\lambda p_{i+1,i}(t) - \lambda p_{ii}(t),$$

其中 $p_{i+1,i}(t) = 0$ 故

$$p'_{ii}(t) = -\lambda p_{ii}(t).$$

记 $f(t) = p_{ii}(t)$，则

$$\begin{aligned} f'(t) =& -\lambda f(t), \\ f(0) =& 1. \end{aligned}$$

将方程化为

$$\frac{d}{dt}\log f(t) = -\lambda,$$

则有

$$f(t) = Ce^{-\lambda t},$$

由 $f(0) = 1$ 知 $C = 1$，所以 $f(t) = e^{-\lambda t}$，即

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

对 $j = i + 1$，(5.31)变成

$$\begin{aligned} p'_{i,i+1}(t) =& \lambda p_{i+1,i+1}(t) - \lambda p_{i,i+1}(t) \\ =& \lambda e^{-\lambda t} - \lambda p_{i,i+1}(t). \end{aligned}$$

记 $f(t)=p_{i,i+1}(t)$, 则

$$\begin{cases} f'(t)=\lambda e^{-\lambda t}-\lambda f(t), \\ f(0)=0, \end{cases}$$

变换得

$$\begin{aligned} f'(t)+\lambda f(t)=&\lambda e^{-\lambda t}, \\ f'(t)e^{\lambda t}+f(t)\cdot\lambda e^{\lambda t}=&\lambda, \\ \frac{d}{dt}[f(t)e^{\lambda t}]=&\lambda, \\ f(t)e^{\lambda t}=&c_1+\lambda t, \\ f(t)=&c_1 e^{-\lambda t}+\lambda t e^{-\lambda t}, \end{aligned}$$

利用 $f(0)=0$ 得 $c_1=0$，于是

$$f_{i,i+1}=\lambda t e^{-\lambda t}.$$

对 $j=i+2,i+3,...$ 递推猜测有一般公式

$$p_{i,i+k}(t)=\frac{(\lambda t)^k}{k!}e^{-\lambda t},\ k=0,1,2,\dots, \tag{5.32}$$

归纳地，设(5.32)对 k 成立，只要证明对 $k+1$ 成立。事实上当 $j=k+1$ 时利用(5.31)式有

$$\begin{aligned} p'_{i,i+k+1}(t)=&\lambda p_{i+1,i+k+1}(t)-\lambda p_{i,i+k+1}(t) \\ =&\lambda\frac{(\lambda t)^k}{k!}e^{-\lambda t}-\lambda p_{i,i+k+1}(t), \end{aligned}$$

记 $f(t)=p_{i,i+k+1}(t)$，有

$$f'(t)=\lambda\frac{(\lambda t)^k}{k!}e^{-\lambda t}-\lambda f(t).$$

变换得

$$\begin{aligned} f'(t)+\lambda f(t)=&\lambda\frac{(\lambda t)^k}{k!}e^{-\lambda t}, \\ f'(t)e^{\lambda t}+f(t)\lambda e^{\lambda t}=&\lambda\frac{(\lambda t)^k}{k!}, \\ \frac{d}{dt}[f(t)e^{\lambda t}]=&\frac{\lambda^{k+1}t^k}{k!}, \\ f(t)e^{\lambda t}=&c_1+\frac{(\lambda t)^{k+1}}{(k+1)!}, \\ f(t)=&c_1 e^{-\lambda t}+\frac{(\lambda t)^{k+1}}{(k+1)!}e^{-\lambda t}, \end{aligned}$$

由 $f(0)=0$ 得 $c_1=0$，于是

$$p_{i,i+k+1}(t)=\frac{(\lambda t)^{k+1}}{(k+1)!}e^{-\lambda t}.$$

注意：对泊松过程，

$$p_{i,i+k}(t)=P(N(t+s)-N(s)=k|N(s)=i),\ k=0,1,2,...$$

为参数 λt 的泊松分布的概率分布列。

例 5.33 (Yule 过程)**.** 类似泊松过程，给出 Yule 过程 $\{X(t),t\geq 0\}$ 的转移概率.

解：

Yule 过程是纯生过程，当 $X(t)=i$ 时出生一个个体的速率为 $i\lambda$。这时类似于例5.32的泊松过程，只是速率参数有变化，所以由

$$\begin{cases} p_{k,j}(h)=0, & j<k; \\ p_{k,j}(h)=o(h), & j-k\geq 2; \\ p_{k,k+1}(h)=k\lambda h+o(h); \\ p_{k,k}(h)=1-k\lambda h+o(h) \end{cases}$$

得到

$$\begin{cases} q_{kj}=0, & j<k \text{ 或} j\geq k+2; \\ q_{k,k+1}=\lim_{h\to 0}\frac{p_{k,k+1}(h)}{h}=k\lambda; \\ q_{kk}=\lim_{h\to 0}\frac{1-p_{k,k}(h)}{h}=k\lambda. \end{cases}$$

利用 Kolmogorov 向前方程，

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

上式求和中仅 $k=j-1$ 的项非零，故

$$\begin{aligned} p'_{ij}(t)=&p_{i,j-1}(t)q_{j-1,j}-p_{ij}(t)q_{jj} && (5.33) \\ =&(j-1)\lambda p_{i,j-1}(t)-j\lambda p_{ij}(t). && (5.34) \end{aligned}$$

因为 $p_{ij}(t)=0$ 对 $j<i$，所以只需要求解 $j\geq i$ 的情形。并注意

$$p_{ii}(0)=1,\quad p_{ij}(0)=0(j\neq i).$$

对 $j=i$，由 $p_{i,i-1}(t)=0$，(5.34)变成

$$p'_{ii}(t)=-i\lambda p_{ii}(t),$$

于是由 $p_{ii}(0)=1$ 解得

$$p_{ii}(t)=e^{-i\lambda t}.$$

对 $j=i+1$，代入(5.34)得

$$\begin{aligned}p'_{i,i+1}(t)=&i\lambda p_{ii}(t)-i\lambda p_{i,i+1}(t)\\=&i\lambda e^{-i\lambda t}-(i+1)\lambda p_{i,i+1}(t).\end{aligned}$$

记 $f(t)=p_{i,i+1}(t)$，则

$$f'(t)=i\lambda e^{-i\lambda t}-(i+1)\lambda f(t).$$

变换得

$$f'(t)+(i+1)\lambda f(t)=i\lambda e^{-i\lambda t},$$

两边乘以 $e^{(i+1)\lambda t}$，得

$$\begin{aligned}\frac{d}{dt}[f(t)e^{(i+1)\lambda t}]=&i\lambda e^{\lambda t},\\f(t)e^{(i+1)\lambda t}=&c_1+ie^{\lambda t},\end{aligned}$$

由 $f(0)=0$ 得 $c_1=-i$，于是

$$p_{i,i+1}(t)=f(t)=ie^{-i\lambda t}[1-e^{-\lambda t}].$$

对 $j=i+2,i+3,...$ 递推猜测有一般公式

$$p_{ij}(t)=\binom{j-1}{i-1}e^{-i\lambda t}[1-e^{-\lambda t}]^{j-i},j-i=0,1,2,\dots,\tag{5.35}$$

用归纳法只要证明(5.35)对 j 成立则对 $j+1$ 成立。由(5.34)得

$$\begin{aligned}p'_{i,j+1}(t)=&j\lambda p_{ij}(t)-(j+1)\lambda p_{i,j+1}(t)\\=&j\lambda\binom{j-1}{i-1}e^{-i\lambda t}[1-e^{-\lambda t}]^{j-i}-(j+1)\lambda p_{i,j+1}(t),\end{aligned}$$

记 $f(t)=p_{i,j+1}(t)$，则

$$f'(t)+(j+1)\lambda f(t)=\frac{j!}{(i-1)!(j-i)!}\lambda e^{-i\lambda t}[1-e^{-\lambda t}]^{j-i},$$

两边乘以 $e^{(j+1)\lambda t}$，得

$$\begin{aligned}\frac{d}{dt}[f(t)e^{(j+1)\lambda t}] =&\frac{j!}{(i-1)!(j-i)!}\lambda e^{(j-i+1)\lambda t}[1-e^{-\lambda t}]^{j-i}\\ =&\frac{j!}{(i-1)!(j-i)!}\lambda e^{\lambda t}[e^{\lambda t}(1-e^{-\lambda t})]^{j-i}\\ =&\frac{j!}{(i-1)!(j-i)!}\lambda e^{\lambda t}\\ [e^{\lambda t}-1]^{j-i} =&\frac{j!}{(i-1)!(j-i)!}\frac{1}{j-i+1}\frac{d}{dt}[[e^{\lambda t}-1]^{j-i+1}],\end{aligned}$$

于是

$$\begin{aligned}f(t)e^{(j+1)\lambda t} =&c_1+\frac{j!}{(i-1)!(j-i+1)!}[e^{\lambda t}-1]^{j-i+1}\\ =&c_1+\binom{j}{i-1}e^{(j-i+1)\lambda t}[1-e^{-\lambda t}]^{j-i+1},\\ f(t) =&c_1e^{-(j+1)\lambda t}+\binom{j}{i-1}e^{-i\lambda t}[1-e^{-\lambda t}]^{j-i+1},\end{aligned}$$

由 $f(0)=0$ 得 $c_1=0$，结果得证。

例 5.34 (生灭过程)**.** 导出生灭过程的 Kolmogorov 微分方程。

解:

对生灭过程，

$$p_{ij}(h)=\begin{cases} i\lambda h+o(h), & j=i+1,\\ i\mu h+o(h), & j=i-1,\\ (i\lambda+i\mu)h+o(h), & j=i,\\ o(h), & |j-i|\ge 2,\end{cases}$$

于是

$$\begin{cases} q_{i,i+1}=i\lambda,\\ q_{i,i-1}=i\mu,\\ q_{ii}=i\lambda+i\mu,\\ q_{ij}=0, & j\ne i,i\pm 1.\end{cases}$$

当状态 $i=0$ 时为吸收状态，$p_{00}(t)=1$, $p_{0j}(t)=0(j\neq i)$。

按向后方程，对 $i\geq 1$，$j\geq 0$，有

$$\begin{aligned}
p'_{ij}(t) =& \sum_{k\neq i} q_{ik}p_{kj}(t) - q_{ii}p_{ij}(t) \\
& (\text{仅}k=i\pm 1\text{时非零}) \\
=& q_{i,i-1}p_{i-1,j}(t) + q_{i,i+1}p_{i+1,j}(t) - q_{ii}p_{ij}(t) \\
=& i\mu p_{i-1,j}(t) + i\lambda p_{i+1,j}(t) - (i\lambda + i\mu)p_{ij}(t).
\end{aligned}$$

用向前方程，当 $i\geq 1, j\geq 1$ 时

$$\begin{aligned}
p'_{ij}(t) =& \sum_{k\neq j} p_{ik}(t)q_{kj} - p_{ij}(t)q_{jj} \\
& (\text{仅}k=j\pm 1\text{时非零}) \\
=& q_{j-1,j}p_{i,j-1}(t) + q_{j+1,j}p_{i,j+1}(t) - q_{jj}p_{ij}(t) \\
=& (j-1)\lambda p_{i,j-1}(t) + (j+1)\mu p_{i,j+1}(t) - (j\lambda + j\mu)p_{ij}(t),
\end{aligned}$$

当 $i\geq 1, j=0$ 时，方程中仅 $k=1$ 的求和项非零且 $j\lambda + j\mu = 0$，这时

$$p'_{i0}(t) = \mu p_{i1}(t).$$

5.6 连续时间连续状态的马氏过程

定义 5.17. 设 $\{X(t),\mathscr{F}(t),t\geq 0\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的适应过程，如果对任意有界 Borel 函数 $g(x)$，有

$$E(g(X(t))|\mathscr{F}(s)) = E(g(X(t))|X(s)),\ \forall 0\leq s<t,$$

则称 $\{X(t),\mathscr{F}(t),t\geq 0\}$ 为马氏过程。

对马氏过程 $\{X(t),\mathscr{F}(t),t\geq 0\}$，函数

$$F(s,t;x,y) = P(X(t)\leq y|X(s)=x),\ 0\leq s<t,\ x,y\in\mathbb{R}$$

称为过程的转移概率分布；如果存在非负函数 $f(s,t;x,y)$ 使得

$$F(s,t;x,y) = \int_{-\infty}^{y} f(s,t;x,u)\,du,$$

则称 $f(s,t;x,y)$ 为过程的转移概率密度。如果 $F(s,t;x,y)=F(0,t-s;x,y)$，则称过程是时齐的，记 $F(0,\tau;x,y)$ 为 $F(\tau;x,y)$，对时齐马氏过程，若转移概率密度存在则记为 $f(\tau;x,y)$。

关于时间集合为 T，取值空间为 S 的关于 $\mathscr{F}(t)$ 适应的马氏过程 $\{X(t),\mathscr{F}(t),t\in T\}$，也可以如上定义。可以证明，存在函数族 $\{p(s,t;x,A): s<t,x\in\mathbb{R},A\in\mathcal{B}(\mathbb{R})\}$, 满足：

- (1) 对于任意 $A\in\mathcal{B}(\mathbb{R})$，$p(s,t;x,A)$ 关于 x 为 Borel 函数；
- (2) 对于任意 $x\in\mathbb{R}$, $p(s,t;x,A)$ 关于 A 是 $\mathcal{B}(\mathbb{R})$ 上的概率测度；
- (3) 对于任意的有界 Borel 函数 g，

$$E[g(X(t))|\mathscr{F}(s)]=E[g(X(t))|X(s)]=\int_{\mathbb{R}}g(y)p(s,t;X(s),dy).$$

称函数族 $\{p(s,t;x,A): s<t,x\in\mathbb{R},A\in\mathcal{B}(\mathbb{R})\}$ 为马氏过程 $\{X(t),\mathscr{F}(t),t\in T\}$ 的**转移概率族**。若 $p(s,t;x,A)=p(0,t-s;x,A)$ 恒成立，则称此马氏过程为时齐的，记转换概率函数为 $p(\tau;x,A)$。如果状态空间 S 是 $\mathbb{R}$ 的区间，且对给定 s,t,x，$p(s,t;x,A)$ 关于 A 有密度 $f(s,t;x,y)$，则称 $f(s,t;x,y)$ 为转移概率密度，这时

$$E[g(X(t))|\mathscr{F}(s)]=E[g(X(t))|X(s)]=\int_{\mathbb{R}}g(y)f(s,t;X(s),y)\,dy.$$

例 5.35. 设 $(\Omega,\mathscr{F},\mathscr{F}(t),t\geq 0,P)$ 是带有子 σ 代数流的概率空间，$\{X(t)\}$ 为关于 $\{\mathscr{F}(t)\}$ 的马氏过程，状态空间为 $\mathbb{R}$，转移概率分布族为 $\{p(s,t;x,A)\}$，若转移概率分布族满足空间不变性：

$$p(s,t;x,A)=p(s,t;x+y,A+y),\ \forall s<t,A,y,$$

其中 $A+y=\{x+y: x\in A\}$，则 $\{X(t)\}$ 为独立增量过程。

见 (刘勇 2022) 第二章。

证明：对任意 $A \in \mathcal{B}(\mathbb{R})$，

$$
\begin{aligned}
&E[I_A(X(t)-X(s))|\mathscr{F}(s)]\\
=&E[I_A(X(t)-z)|\mathscr{F}(s)]\Big|_{z=X(s)}\\
=&E[I_{A+z}(X(t))|\mathscr{F}(s)]\Big|_{z=X(s)}\\
=&\int_{\mathbb{R}} I_{A+z}(y)p(s,t;X(s),dy)\Big|_{z=X(s)}\\
=&p(s,t;X(s),A+z)\Big|_{z=X(s)}\\
=&p(s,t;X(s),A+X(s))\\
=&p(s,t;0,A).
\end{aligned}
$$

结果与 $X(s)$ 无关。由此，

$$
\begin{aligned}
&P(X(t)-X(s)\in A)\\
=&E[I_A(X(t)-X(s))]\\
=&E\left\{E[I_A(X(t)-X(s))|\mathscr{F}(s)]\right\}\\
=&p(s,t;0,A).
\end{aligned}
$$

来证明 $X(t)-X(s)$ 与 $\mathscr{F}(s)$ 独立。$\forall A \in \mathcal{B}(\mathbb{R})$，$B \in \mathscr{F}$，

$$
\begin{aligned}
&P(\{(X(t)-X(s)\in A\}\cap B)\\
=&E[I_A(X(t)-X(s))I_B(\omega)]\\
=&E\left\{E[I_A(X(t)-X(s))I_B(\omega)|\mathscr{F}(s)]\right\}\\
=&E\left\{I_B(\omega)\cdot E[I_A(X(t)-X(s))|\mathscr{F}(s)]\right\}\\
=&E\left\{I_B(\omega)p(s,t;0,A)\right\}\\
=&p(s,t;0,A)E[I_B(\omega)]\\
=&P(X(t)-X(s)\in A)P(B),
\end{aligned}
$$

这就证明了 $\{X(t),\mathscr{F}(t)\}$ 是独立增量过程。注意这里的独立增量过程比原来的定义略强，需要 $X(t)-X(s)$ 与 $\mathscr{F}(s)$ 独立，$\mathscr{F}(s)$ 是包含 $\sigma(X(u),u\in[0,s])$ 的。

5.7 补充

5.7.1 级数的极限定理

考虑 $\lim_{n\to\infty}\sum_{i=1}^{\infty} a_i^{(n)}$ 的极限与求和号交换问题。因为级数是 Riemann-Stieljes 积分的特例，也可以看成关于点测度的 Lebesgue 积分，所以积分的收敛定理仍成立。

引理 5.2 (级数的单调收敛). 设 $\{\pi_i \geq 0, i=1,2,\dots\}$，

$$0 \leq a_i^{(n)} \leq a_i^{(n+1)},\ i=1,2,\dots,\ n=1,2,\dots,$$

且

$$\lim_{n\to\infty} a_i^{(n)} = a_i,$$

则

$$\lim_{n\to\infty}\sum_{i=1}^{\infty}\pi_i a_i^{(n)} = \sum_{i=1}^{\infty}\pi_i a_i.$$

引理 5.3 (级数的 Fatou 引理). 设 $\{\pi_i \geq 0, i=1,2,\dots\}$，若

$$\lim_{n\to\infty} a_i^{(n)} = a_i, i=1,2,\dots,$$

则

$$\varliminf_{n\to\infty}\sum_{i=1}^{\infty}\pi_i a_i^{(n)} \geq \sum_{i=1}^{\infty}\pi_i a_i.$$

引理 5.4 (级数的控制收敛定理). 设 $\{a_i^{(n)}\}$ 满足

$$|a_i^{(n)}| \leq b_i,\ \forall i,\ \forall n,$$

其中 $\sum_{i=1}^{\infty} b_i < \infty$，且 $\lim_{n\to\infty} a_i^{(n)} = a_i,\ \forall n$，则

$$\lim_{n\to\infty}\sum_{i=1}^{\infty} a_i^{(n)} = \sum_{i=1}^{\infty} a_i.$$

证明：

$\forall\epsilon>0, \exists N_1$ 使得

$$\sum_{i=N_1+1}^{\infty} b_i < \frac{\epsilon}{2},$$

从而

$$\sum_{i=N_1+1}^{\infty} |a_i^{(n)}| < \frac{\epsilon}{4}.$$

由条件可知 $|a_i| = \lim_n |a_i^{(n)}| \leq b_i$，从而

$$\sum_{i=N_1+1}^{\infty} |a_i| < \frac{\epsilon}{4}.$$

对 $\sum_{i=1}^{N_1} a_i^{(n)}$，存在 $N_2 > N_1$ 使得 $n > N_2$ 时

$$\sum_{i=1}^{N_1} |a_i^{(n)} - a_i| < \frac{\epsilon}{2}.$$

于是 $n > N_2$ 时

$$\begin{aligned}
&\sum_{i=1}^{\infty} |a_i^{(n)} - a_i| \\
=&\sum_{i=1}^{N_1} |a_i^{(n)} - a_i| + \sum_{i=N_1+1}^{\infty} |a_i^{(n)} - a_i| \\
\leq&\frac{\epsilon}{2} + \sum_{i=N_1+1}^{\infty} |a_i^{(n)}| + \sum_{i=N_1+1}^{\infty} |a_i| \\
<& \epsilon.
\end{aligned}$$

即

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} |a_i^{(n)} - a_i| = 0,$$

从而

$$\begin{aligned}
&\lim_{n\to\infty} \left| \sum_{i=1}^{\infty} a_i^{(n)} - \sum_{i=1}^{\infty} a_i \right| \\
\leq& \lim_{n\to\infty} \sum_{i=1}^{\infty} |a_i^{(n)} - a_i| = 0.
\end{aligned}$$

控制收敛定理的另一种叙述是：若 $\{\pi_j \geq 0\}$，$|a_i^{(n)}| \leq c, \forall i, n$，则

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} \pi_i a_i^{(n)} = \sum_{i=1}^{\infty} \pi_i a_i.$$

当 $\sum_j \pi_j = 1$ 时即引理5.4。

5.7.2 定理5.12证明

这里给出定理5.12在无限链情形下的证明。不妨设状态空间 $S=\{1,2,\dots\}$。

(1) 对遍历的马氏链，由定理5.8知

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j} > 0,$$

记为 π_j. 于是极限分布存在（极限分布存在则必唯一）。

先证明 $\{\pi_j, j\in S\}$ 是平稳分布，这要证明

$$\sum_{j\in S} \pi_j = 1, \tag{5.36}$$

和

$$\pi_j = \sum_{i\in S} \pi_i p_{ij},\ \forall j\in S. \tag{5.37}$$

然后再来证明 $\{\pi_j\}$ 是唯一的平稳分布。

对任意 n, M 有

$$1 = \sum_{j\in S} p_{ij}^{(n)} \geq \sum_{j=1}^{M} p_{ij}^{(n)}.$$

固定 M 令 $n\to\infty$ 得

$$\sum_{j=1}^{M} \pi_j \leq 1,$$

从而

$$\sum_{j\in S} \pi_j \leq 1. \tag{5.38}$$

这说明级数 $\sum_{j\in S}\pi_j$ 收敛。

下面证明

$$\pi_j \geq \sum_{i\in S} \pi_i p_{ij}^{(n)},\ \forall j\in S,\ n\geq 1. \tag{5.39}$$

用归纳法。由 C-K 方程，

$$\begin{aligned} p_{ij}^{(n+1)} &= \sum_{l\in S} p_{il}^{(n)} p_{lj} \\ &\geq \sum_{l=1}^{M} p_{il}^{(n)} p_{lj}, \end{aligned}$$

固定 M 令 $n \to \infty$ 则得

$$\pi_j \geq \sum_{l=1}^{M} \pi_l p_{lj},$$

于是

$$\pi_j \geq \sum_{l \in S} \pi_l p_{lj} = \sum_{i \in S} \pi_i p_{ij}^{(1)}, \tag{5.40}$$

即(5.39)对 $n = 1$ 成立。设(5.39)对 n 成立，来证明(5.39)对 $n+1$ 成立。在(5.39)两边乘以 p_{jk} 并关于 j 求和，得

$$\begin{aligned}
&\sum_{j \in S} \pi_j p_{jk} &(5.41)\\
\geq &\sum_{j \in S} \sum_{i \in S} \pi_i p_{ij}^{(n)} p_{jk} &(5.42)\\
= &\sum_{i \in S} \pi_i \sum_{j \in S} p_{ij}^{(n)} p_{jk} &(5.43)\\
= &\sum_{i \in S} \pi_i p_{ik}^{(n+1)}. &(5.44)
\end{aligned}$$

由(5.40)可知(5.44)左边满足

$$\pi_k \geq \sum_{j \in S} \pi_j p_{jk},$$

于是(5.44)变成

$$\pi_k = \sum_{i \in S} \pi_i p_{ik}^{(n+1)}. \tag{5.45}$$

即(5.39)对 $n+1$ 成立。由数学归纳法知(5.39)成立。

下面证明(5.39)只有等号成立。如果不然，设存在 $n \geq 1$ 和 $j_0 \in S$ 使得

$$\pi_j > \sum_{i \in S} \pi_i p_{ij}^{(n)},\ j = j_0,$$

则上式关于 j 求和，得

$$\begin{aligned}
\sum_{j \in S} \pi_j &> \sum_{j \in S} \sum_{i \in S} \pi_i p_{ij}^{(n)}\\
&= \sum_{i \in S} \pi_i \sum_{j \in S} p_{ij}^{(n)}\\
&= \sum_{i \in S} \pi_i,
\end{aligned}$$

矛盾。所以有

$$\pi_j = \sum_{i\in S} \pi_i p_{ij}^{(n)},\ \forall j \in S,\ n \geq 1. \tag{5.46}$$

因 $\pi_i p_{ij}^{(n)} \leq \pi_i$，$\sum_{i\in S} \pi_i < \infty$（见(5.38)），由引理5.4可知极限与求和号可交换，有

$$\pi_j = \sum_{i\in S} \pi_i \pi_j = \pi_j (\sum_{i\in S} \pi_i),$$

由 $0 < \pi_j < \infty$ 可知

$$\sum_{i\in S} \pi_i = 1.$$

即(5.36)成立。在(5.46)中取 $n = 1$ 就得到了(5.37)，所以 $\{\pi_j\}$ 是平稳分布。

下面证明 $\{\pi_j\}$ 是是唯一的平稳分布。若存在离散分布 $\{\tilde{\pi}_j\}$ 也满足

$$\sum_{j\in S} \tilde{\pi}_j = 1, \tag{5.47}$$

$$\tilde{\pi}_j = \sum_{i\in S} \tilde{\pi}_i p_{ij},\ \forall j \in S. \tag{5.48}$$

类似于上面关于(5.39)的证明，可以得到

$$\tilde{\pi}_j = \sum_{i\in S} \tilde{\pi}_i p_{ij}^{(n)},\ \forall j \in S,\ n \geq 1. \tag{5.49}$$

$n = 1$ 时即(5.48)。归纳地，设(5.49)对 n 成立，在(5.49)两边乘以 p_{jk} 并对 j 求和，得

$$\begin{aligned}\tilde{\pi}_k =& \sum_{j\in S} \tilde{\pi}_j p_{jk} = \sum_{j\in S} \sum_{i\in S} \tilde{\pi}_i p_{ij}^{(n)} p_{jk} \\ =& \sum_{i\in S} \tilde{\pi}_i \sum_{j\in S} p_{ij}^{(n)} p_{jk} = \sum_{i\in S} \tilde{\pi}_i p_{ik}^{(n+1)},\end{aligned}$$

由归纳法知(5.49)成立。在(5.49)中令 $n \to \infty$，由引理5.4可知极限与求和号可交换，得

$$\tilde{\pi}_j = \sum_{i\in S} \tilde{\pi}_i \pi_j = \pi_j \sum_{i\in S} \tilde{\pi}_i = \pi_j.$$

所以 $\{\pi_j\}$ 是唯一平稳分布。

(2) 如果所有状态都是非常返或者零常返的，则由定理5.8，

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0,\ \forall i, j \in S.$$

如果存在平稳分布 $\{\pi_j \geq 0\}$，由(5.49)可知

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)},\ \forall j \in S,\ n \geq 1.$$

令 $n \to \infty$，由引理5.4可知极限与求和号可交换，于是

$$\pi_j = \sum_{i \in S} \pi_i \cdot 0 = 0,$$

这与 $\sum_{j \in S} = 1$ 矛盾，所以所有状态都是非常返或者零常返时不存在平稳分布。

5.7.3 强马氏性

设 τ 是取值于 $\{0, 1, 2, \dots\} \cup \{\infty\}$ 的广义随机变量，$\{X_n, n = 0, 1, 2, \dots\}$ 是马氏链，若 $\{\tau = n\}$ 可以被 $X_0, X_1, \dots, X_n$ 的值确定，则称 τ 是关于 $\{X_n\}$ 的一个**停时**。即选择一个时间停止对 $\{X_n\}$ 的观测，如果能够用截止到 n 为止的观测值决定是否停止观测，就是一个停时。

定义 5.18. $\{X_n, n = 0, 1, 2, \dots\}$ 是马氏链，状态空间为 S，设 τ 是取值于 $\{0, 1, 2, \dots\} \cup \{\infty\}$ 的广义随机变量，如果对任意 $n \geq 0$ 和 $i_0, i_1, \dots, i_n \in S$，下面两种情况中必有一种成立：

$$\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subset \{\tau \leq n\};$$
$$\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subset \{\tau >\},$$

则称 τ 是 $\{X_n\}$ 的一个**停时**。

定理 5.24. *$\{X_n, n = 0, 1, 2, \dots\}$ 是马氏链，转移概率矩阵为 P，设 τ 是 $\{X_n\}$ 的一个停时。当 $\tau < \infty$ 时，定义 $Y_n = X_{\tau+n}, n \geq 0$。则在 $\tau < n$，$X_\tau = i$ 的条件下，$\{Y_n, n \geq 0\}$ 是一个从 i 出发，以 P 为转移概率矩阵的马氏链，且它与 $\{X_0, X_1, \dots, X_\tau\}$ 独立。*

参见 (钱敏平 et al. 2011) P.48 定理 1.12.3。显然 $\tau \equiv n$ 是停时，强马氏性推出马氏性；上述定理说明马氏性也推出强马氏性。强马氏性说明从任何一个停时重新出发仍是一个转移概率相同的马氏链且与历史独立。

Chapter 6

鞅

6.1　基本概念

6.1.1　定义

本章将介绍另一类特殊的随机过程——鞅. 近几十年来，鞅理论不仅在随机过程及其他数学分支中占据了重要的地位，而且在实际问题诸如金融、保险和医学上也得到了广泛的应用. 在此我们将阐述鞅的一些基本理论，并以介绍离散时间鞅为主.

鞅的定义是从条件期望出发的，所以对条件期望不熟悉的读者请先学习第 1 章中的相关内容，这对于理解鞅理论是至关重要的.

每个赌博者自然都对能使他在一系列赌博后获得期望收益最大的策略感兴趣. 然而在数学上可以证明，在“公平”的博弈中，是没有这样的赌博策略的.

假设一个赌博者正在进行一系列赌博，每次赌博输赢的概率都是 $\frac{1}{2}$. 令 $\{Y_n, n = 1, 2, \cdots\}$, 是一列独立同分布的随机变量，表示每次赌博的结果

$$P\{Y_n = 1\} = P\{Y_n = -1\} = \frac{1}{2}$$

这里 $\{Y_n = 1\}$ ($\{Y_n = -1\}$) 表示赌博者在第 n 次赌博时的赢 (输).

如果赌博者采用的赌博策略 (即所下赌注) 依赖于前面的赌博结果，那么他的赌

博可以用下面的随机变量序列

$$b_n = b_n(Y_1, \cdots, Y_{n-1}), n = 2, 3, \cdots$$

描述，其中 $b_n < \infty$ 是第 n 次的赌注，若赌赢则获利 b_n，否则输掉 b_n.

设 X_0 是该赌博者的初始赌资，则

$$X_n = X_0 + \sum_{i=1}^{n} b_i Y_i \tag{6.1}$$

是他在第 n 次赌博后的赌资. 可以断言

$$E[X_{n+1}|Y_1, \cdots, Y_n] = X_n.$$

事实上，由式(6.1)我们可以得到

$$X_{n+1} = X_n + b_{n+1} Y_{n+1},$$

因此

$$\begin{aligned} E[X_{n+1}|Y_1, \cdots, Y_n] =& E[X_n|Y_1, \cdots, Y_n] + E[b_{n+1}Y_{n+1}|Y_1, \cdots, Y_n] \\ =& X_n + b_{n+1}E[Y_{n+1}|Y_1, \cdots, Y_n] \\ & \text{(因为}X_n\text{与}b_{n+1}\text{由}Y_1, \cdots, Y_n\text{确定)} \\ =& X_n + b_{n+1}E[Y_{n+1}] \\ & \text{(因为}\{Y_n\}\text{是独立随机变量序列)} \\ =& X_n \quad \text{(因为}E[Y_{n+1}] = 0,\ \forall n \geq 0) \end{aligned}$$

这证明了，如果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，则赌博是“公平的”. 因此任何赌博者都不可能将公平的赌博通过改变赌博策略使得赌博变成有利于自己的赌博.

定义 6.1. 设 $\{X_n, n \geq 0\}$ 和 $\{Y_n, n \geq 0\}$ 是随机过程，对任意 $n \geq 0$，X_n 是 $Y_0, Y_1, \ldots, Y_n$ 的函数，$E|X_n| < \infty$ 且

$$E[X_{n+1}|Y_0, \ldots, Y_n] = X_n,\ \forall n \geq 0, \tag{6.2}$$

则称 $\{X_n\}$ 为关于 $\{Y_n\}$ 的鞅。

定义 6.2. 设 $\{X_n, n \geq 0\}$ 和 $\{Y_n, n \geq 0\}$ 是随机过程，对 $n \geq 0$, X_n 是 $(Y_0, Y_1, \ldots, Y_n)$ 的函数。

如果

$$E[X_n^+] < \infty \text{ 且} E[X_{n+1}|Y_0, Y_1, \dots, Y_n] \geq X_n, \ \forall n \geq 0, \tag{6.3}$$

则称 $\{X_n\}$ 是关于 $\{Y_n\}$ 的**下鞅**。这里 $X_n^+ = \max\{0, X_n\}$.

另一方面，如果

$$E[X_n^-] < \infty \text{ 且} E[X_{n+1}|Y_0, Y_1, \dots, Y_n] \leq X_n, \ \forall n \geq 0, \tag{6.4}$$

则称 $\{X_n\}$ 是关于 $\{Y_n\}$ 的**上鞅**。这里 $X_n^- = \max\{0, -X_n\}$.

注 1: 鞅描述的是“公平”的赌博，下鞅和上鞅分别描述了“有利”赌博与“不利”赌博.

注 2：关于鞅、下鞅、上鞅的期望存在条件要求，因为鞅是条件期望不变，所以只有期望存在有限才有利用价值。下鞅是条件期望上升的，如果允许期望等于 $+\infty$ 就没有利用价值，所以加 $E[X_n^+] < \infty$ 条件。类似地，上鞅是条件期望下降的，如果允许期望等于 $-\infty$ 就没有利用价值，所以加 $E[X_n^-] < \infty$ 条件。

注 3：对鞅 $\{X_n\}$，易见 $E(X_n) = E(X_0)$, $\forall n$。对下鞅，$\{E(X_n), n \geq 1\}$ 单调增；对上鞅，$\{E(X_n), n \geq 1\}$ 单调减。

下面我们定义关于 σ 代数的鞅. 为此，首先介绍有关概念.

定义 6.3. 设 $(\Omega, \mathscr{F}, P)$ 是完备的概率空间，$\{\mathscr{F}_n, n = 0, 1, 2, \dots\}$ 是 $\mathscr{F}$ 内的一列子 σ 代数，并且使得 $\mathscr{F}_n \subset \mathscr{F}_{n+1}, n \geq 0$，称之为 σ 代数流 (filtration). 随机过程 $\{X_n, n = 0, 1, 2, \dots\}$ 称为 $\{\mathscr{F}_n\}$ **适应的**，如果 $\forall n \geq 0$, X_n 是 $\mathscr{F}_n$ 可测的. 此时称 $\{X_n, \mathscr{F}_n, n \geq 0\}$ 为适应列.

σ 代数流反映了随时间变化，逐步增加的对过去和现在的信息。

在定义6.1中定义鞅时，我们假定了 X_n 是 $(Y_0, \dots, Y_n)$ 的函数. 令 $\mathscr{F}_n = \sigma\{Y_0, \dots, Y_n\}, n \geq 0$，则 $\{\mathscr{F}_n, n \geq 0\}$ 是一个 σ 代数流. X_n 是 $Y_0, \cdots, Y_n$ 的函数的确切含义是 $\{X_n\}$ 是 $\{\mathscr{F}_n\}$ 适应的.

例 6.1. 考虑一个无穷尽的掷硬币试验，用 1 表示正面，-1 表示反面，令 Ω 为所有结果序列（由 1 和 -1 组成的序列）的集合，$\mathscr{F}$ 是 Ω 的所有子集构成的 σ 代数。构造 σ 代数流。

解：记 Ω 的元素为 $\omega = (\omega_1, \omega_2, \dots)$，其中 ω_j 取值于 $\{1, -1\}$, 令

$$Y_n(\omega) = \omega_n,$$

即 Y_n 表示第 n 次投掷的结果。令

$$\mathscr{F}_n = \{A \in \mathscr{F} : \forall \omega \in A, (\omega_1, \dots, \omega_n) \in B, B \subset \{1, -1\}^n\},$$

称 $\mathscr{F}_n$ 中的集合为“柱集”，易见 $\mathscr{F}_n$ 构成 σ 代数流，且

$$\mathscr{F}_n = \sigma(Y_1, Y_2, \dots, Y_n),\ n = 1, 2, \dots$$

$\mathscr{F}_n$ 代表了截止到时刻 n 为止可以观测到的试验结果信息。对事件 $A \in \mathscr{F}_n$，A 发生与否仅与前 n 个试验结果有关，与后面的结果无关。

在 $(\{1, -1\}^n, \mathscr{F}_n)$ 中定义概率

$$P(\{(\omega_1, \dots, \omega_n)\}) = p^{\sum_{j=1}^n \omega_j^+} (1-p)^{\sum_{j=1}^n \omega_j^-},$$

其中 $0 < p < 1$，可以将其扩充到 $(\Omega, \mathscr{F})$ 中构成概率空间。

定义 6.4. 设 $\{\mathscr{F}_n, n \geq 0\}$ 是一个 $\mathscr{F}$ 中 σ 代数流. 随机过程 $\{X_n, n \geq 0\}$ 称为关于 $\{\mathscr{F}_n, n \geq 0\}$ 的**鞅**，如果 $\{X_n\}$ 是 $\{\mathscr{F}_n\}$ 适应的，$\forall n \geq 0$，有

$$E[|X_n|] < \infty \text{ 且} E[X_{n+1}|\mathscr{F}_n] = X_n. \tag{6.5}$$

定义 6.5. 适应列 $\{X_n, \mathscr{F}_n, n \geq 0\}$ 称为**下鞅**，如果 $\forall n \geq 0$,

$$E[X_n^+] < \infty \quad \text{且} \quad E[X_{n+1}|\mathscr{F}_n] \geq X_n.$$

适应列 $\{X_n, \mathscr{F}_n, n \geq 0\}$ 称为**上鞅**，如果 $\forall n \geq 0$,

$$E[X_n^-] < \infty \quad \text{且} \quad E[X_{n+1}|\mathscr{F}_n] \leq X_n.$$

若以 X_n 表示一个赌博者在第 n 次赌博后所有的赌资. 式(6.2)表示：平均而言他在下一次赌博结束时的赌资将等于现时的赌资，与他过去赌博的输赢无关. 这也就是说鞅具有一种“无后效性”，同时这体现的正是博弈的公平.

在给出例子之前，先给出由定义直接推出的命题.

命题 6.1. *(1) 适应列 $\{X_n, \mathscr{F}_n, n \geq 0\}$ 是下鞅当且仅当 $\{-X_n, \mathscr{F}_n, n \geq 0\}$ 是上鞅.*

(2) 如果 $\{X_n, \mathscr{F}_n\}$, $\{Y_n, \mathscr{F}_n\}$ 是两个下鞅，a, b 是两个正常数，则 $\{aX_n + bY_n, \mathscr{F}_n\}$ 是下鞅.

(3) 如果 $\{X_n, \mathscr{F}_n\}$, $\{Y_n, \mathscr{F}_n\}$ 是两个下鞅，则 $\{\max(X_n, Y_n), \mathscr{F}_n\}$ 是下鞅; 如果 $\{X_n, \mathscr{F}_n\}$, $\{Y_n, \mathscr{F}_n\}$ 是两个上鞅，则 $\{\min(X_n, Y_n), \mathscr{F}_n\}$ 是上鞅.

在本命题以及其他类似命题中，σ 代数流 $\{\mathscr{F}_n\}$ 可以由 $\{Y_k, k = 1, 2, \cdots, n\}$ 替代，即 $\{X_n\}$ 是关于 $\{Y_n\}$ 的下鞅.

命题6.1（3）证明：由题意，$EX_n^+ < \infty$, $EY_n^+ < \infty$, 所以

$$\begin{aligned}
[\max(X_n, Y_n)]^+ \leq &\max(X_n^+, Y_n^+) \leq X_n^+ + Y_n^+, \\
E[\max(X_n, Y_n)]^+ \leq &E(X_n^+ + Y_n^+) < \infty.
\end{aligned}$$

又

$$\begin{aligned}
E(\max(X_{n+1}, Y_{n+1})|\mathscr{F}_n) \geq E(X_{n+1}|\mathscr{F}_n) \geq X_n, \\
E(\max(X_{n+1}, Y_{n+1})|\mathscr{F}_n) \geq E(Y_{n+1}|\mathscr{F}_n) \geq Y_n,
\end{aligned}$$

从而

$$E(\max(X_{n+1}, Y_{n+1})|\mathscr{F}_n) \geq \max(X_n, Y_n),$$

即 $\{\max(X_n, Y_n), \mathscr{F}_n\}$ 是下鞅。

类似地，如果 $\{X_n, \mathscr{F}_n\}$, $\{Y_n, \mathscr{F}_n\}$ 是两个上鞅，则 $\{\min(X_n, Y_n), \mathscr{F}_n\}$ 是上鞅。

6.1.2 例子

例 6.2. 设 $X_1, X_2, \cdots$ 是一族零均值独立随机变量序列，令 $S_0 = 0$, $S_n = \sum_{k=1}^n X_k$, 则 $\{S_n\}$ 是 (关于 $\mathscr{F}_n = \sigma(X_1, X_2, \cdots, X_n)$ 的) 鞅. 另外，若 $X_k(k = 1, 2, \cdots)$ 均值为 $\mu \neq 0$，则 $\{M_n = S_n - n\mu\}$ 是 (关于 $\{\mathscr{F}_n\}$ 的) 鞅.

证明: 当 $E[X_k] = 0$, $(k = 1, 2, \cdots)$ 时，易见 S_n 是 $\mathscr{F}_n$ 可测的，而且 $E[|S_n|] \leq \sum_{i=1}^n E[|X_i|] < \infty$，于是

$$\begin{aligned}
E[S_{n+1}|F_n] =&E[X_1 + X_2 + \cdots + X_{n+1}|F_n] \\
=&E[X_1 + \cdots + X_n|F_n] + E[X_{n+1}|F_n] \\
=&S_n,
\end{aligned}$$

从而 $\{S_n\}$ 是一个关于 $\{\mathscr{F}_n\}$ 的鞅.

同理可以证明，当 $E[X_k] = \mu \neq 0$, $(k = 1, 2, \cdots)$ 时，$\{M_n\}$ 也是一个关于 $\{\mathscr{F}_n\}$ 的鞅.

注意：设随机变量 X 有均值 $\mu \in (-\infty, \infty)$，则 $E|X| < \infty$。

例 6.3. 在例6.2中设 $E[X_k] = \mu \neq 0$, $(k = 1, 2, \cdots)$，则有 $E[|S_n|] < \infty$，

$$E[S_{n+1}|\mathscr{F}_n] = E[\sum_{i=1}^{n} X_i + X_{n+1}|\mathscr{F}_n] = S_n + \mu,$$

显然，若 $\mu > 0(\mu < 0)$，则 $\{S_n\}$ 是一关于 $\{\mathscr{F}_n\}$ 的下鞅 (上鞅).

例 6.4. 考虑一个公平博弈问题. 设 $X_1, X_2, \cdots$ 独立同分布，分布函数为

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

于是可以将 $X_i(i = 1, 2 \cdots)$ 看做一个投硬币游戏的结果：如果出现正面就赢 1 元，出现反面则输 1 元. 假设我们按以下的规则来赌博，每次投掷硬币之前的赌注都比上一次翻一倍，直到赢了赌博即停. 令 W_n 表示第 n 次赌博后所输 (或赢) 的总钱数, 则 $W_0 = 0$。假设前 n 次投掷的硬币都出现了反面，按照规则，我们已经输了 $1 + 2 + 4 + \cdots + 2^{n-1} = 2^n - 1$ 元，即 $W_n = -(2^n - 1)$. 假如下一次硬币出现的是正面，按规则 $W_{n+1} = 2^n - (2^n - 1) = 1$。由于无论何时只要赢了就停止赌博，所以 W_n 从赢了之后起就不再变化，于是有 $P\{W_{n+1} = 1|W_n = 1\} = 1$.

由公平的前提知道

$$P\{W_{n+1} = 1|W_n = -(2^n - 1)\} = \frac{1}{2}$$
$$P\{W_{n+1} = -2^n - 2^n + 1|W_n = -(2^n - 1)\} = \frac{1}{2},$$

易证 $E[W_{n+1}|\mathscr{F}_n] = W_n$，这里 $\mathscr{F}_n = \sigma(X_1, \cdots, X_n)$，从而 $\{W_n\}$ 是一个关于 $\{\mathscr{F}_n\}$ 的鞅.

例 6.5. 我们可以把例6.4再一般化. 设 $X_1, X_2, \cdots$ 仍如例6.4假定，而每次赌博所下赌注将与前面硬币的投掷结果有关，以 B_n 记第 n 次所下的赌注，则 B_n

是 $X_1,\cdots,X_{n-1}$ 的函数, 换言之 B_n 是 $\mathscr{F}_{n-1}$ 可测的 (设 B_1 为常数). 仍然令 W_n 表示 n 次之后的总输赢，$W_0=0$，则有

$$W_n=\sum_{j=1}^{n}B_jX_j.$$

假设 $E[|B_n|]<\infty$(这保证了每次的赌本都有一定节制)，那么 $\{W_n\}$ 是一个 $\{\mathscr{F}_n\}$ 鞅.

事实上，注意到 $E[|W_n|]<\infty$(这可由 $E[|B_n|]<\infty$ 得到)，而且 W_n 是 $\mathscr{F}_n$ 可测的，并且

$$\begin{aligned}E[W_{n+1}|\mathscr{F}_n]=&E[\sum_{j=1}^{n+1}B_jX_j|\mathscr{F}_n]\\=&E[\sum_{j=1}^{n}B_jX_j|\mathscr{F}_n]+E[B_{n+1}X_{n+1}|\mathscr{F}_n]\\=&\sum_{j=1}^{n}B_jX_j+B_{n+1}E[X_{n+1}|\mathscr{F}_n]\\=&W_n+B_{n+1}E[X_{n+1}]\\=&W_n.\end{aligned}$$

例 6.6. (Polya 坛子抽样模型)

考虑一个装有红、黄两色球的坛子. 假设最初坛子中装有红、黄两色球各一个，每次都按如下规则有放回地随机抽取：如果拿出的是红色的球，则放回的同时再加入一个同色的球；如果拿出是黄色的球也采取同样的做法. 以 X_n 表示第 n 次抽取后坛子中的红球数，则 $X_0=1$，且 $\{X_n\}$ 是一个非时齐的 Markov 链，转移概率为

$$\begin{aligned}P\{X_{n+1}=k+1|X_n=k\}&=\frac{k}{n+2}\\P\{X_{n+1}=k|X_n=k\}&=\frac{n+2-k}{n+2}.\end{aligned}$$

令 M_n 表示第 n 次抽取后红球所占的比例，则 $M_n=\frac{X_n}{n+2}$，并且 $\{M_n\}$ 是一个鞅.

证明： 这是因为

$$
\begin{aligned}
E(X_{n+1}|X_n=k)=&(k+1)\frac{k}{n+2}+k\left(1-\frac{k}{n+2}\right)\\
=&k+\frac{k}{n+2},\\
E(X_{n+1}|X_n)=&X_n+\frac{X_n}{n+2}=\frac{n+3}{n+2}X_n,
\end{aligned}
$$

由于 $\{X_n\}$ 是一个 Markov 链，则 $\mathscr{F}_n=\sigma(X_1,\cdots,X_n)$ 中对 X_{n+1} 有影响的信息都包含在 X_n 中，所以

$$
\begin{aligned}
E[M_{n+1}|\mathscr{F}_n]=&E[M_{n+1}|X_n]\\
=&E\left[\frac{X_{n+1}}{n+1+2}|X_n\right]\\
=&\frac{1}{n+3}E[X_{n+1}|X_n]\\
=&\frac{1}{n+3}\frac{n+3}{n+2}X_n\\
=&\frac{X_n}{n+2}=M_n.
\end{aligned}
$$

本例研究的模型是 Polya 首次引入的，它适用于描述群体增殖和传染病的传播等现象.

例 6.7 (随机利率折现)**.** 考虑按随机利率增长的资产构成的鞅。

设存款本金为 X_0，按日固定利率 r 增值，则

$$X_{n+1}=(1+r)X_n=X_0(1+r)^{n+1},\ n=0,1,2,\ldots.$$

如果利率为连续复利 σ，则不管时间连续还是离散，有

$$X_t=e^{\sigma t}X_0,\quad X_t=e^{\sigma(t-s)}X_s,\ 0\le s<t.$$

设 $\{\mathscr{F}_n,n=0,1,2,\ldots\}$ 为子 σ 代数流，随机连续复利（对数收益率）序列 $\{\sigma_n\}$ 关于 $\{\mathscr{F}_n\}$ 适应，若资金 X_n 满足条件增值关系：

$$E[X_{n+1}|\mathscr{F}_n]=e^{\sigma_n}X_n,$$

其中 X_0 关于 $\mathscr{F}_0$ 可测。将 X_n 折现到 0 时刻得

$$Y_n = \exp(-[\sigma_0 + \sigma_1 + \cdots + \sigma_{n-1}])X_n,$$

则 $\{Y_n, \mathscr{F}_n, n \geq 0\}$ 为鞅。事实上，易见 Y_n 关于 $\mathscr{F}_n$ 可测，且

$$\begin{aligned}
&E(Y_{n+1}|\mathscr{F}_n)\\
=&E\left\{\exp(-[\sigma_0 + \sigma_1 + \cdots + \sigma_n])X_{n+1}\,|\,\mathscr{F}_n\right\}\\
=&\exp(-[\sigma_0 + \sigma_1 + \cdots + \sigma_n]) \cdot E\left\{e^{\sigma_n} X_n\,|\,\mathscr{F}_n\right\}\\
=&\exp(-[\sigma_0 + \sigma_1 + \cdots + \sigma_n]) \cdot e^{\sigma_n} X_n\\
=&\exp(-[\sigma_0 + \sigma_1 + \cdots + \sigma_{n-1}])X_n = Y_n.
\end{aligned}$$

例 6.8 (二叉树模型)**.** 设时间为离散时间 $n = 0, 1, 2, \ldots$，无风险利率为 $R > 0$。某风险资产 $\{S_n, n \geq 0\}$ 的增长可表示为

$$S_{n+1} = (1 + \eta_{n+1})S_n,\ n = 0, 1, 2, \ldots,$$

设其中的简单收益率 η_{n+1} 独立同分布，与 S_0 独立，有 $0 < a < b$ 和 $0 < p < 1$，使得

$$P(1 + \eta_n = a) = p, \quad P(1 + \eta_n = b) = 1 - p.$$

记 $\mathscr{F}_n = \sigma(S_0, \eta_1, \ldots, \eta_n)$。令 M_n 表示 S_n 折现到 0 时刻的值：

$$M_n = (1 + R)^{-n} S_n, n \geq 1.$$

分析 $\{M_n, \mathscr{F}_n, n \geq 1\}$ 成为鞅的条件。

解答： 令 $\mu = E(\eta_n)$, $\sigma^2 = \mathrm{Var}(\eta_n)$。

用 $E\left[\frac{S_{n+1}-S_n}{S_n}|\mathscr{F}_n\right]$ 表示风险资产平均的收益率，则

$$\begin{aligned}
&E\left[\frac{S_{n+1} - S_n}{S_n}|\mathscr{F}_n\right]\\
=&E(\eta_{n+1}|\mathscr{F}_n) = E(\eta_{n+1}) = \mu.
\end{aligned}$$

用 $\frac{1}{S_n}\left[\mathrm{Var}(S_{n+1} - S_n|\mathscr{F}_n)\right]^{1/2}$ 表示资产波动率，这是衡量资产在单位时间每

单位价格波动程度的一个指标。则

$$
\begin{aligned}
&\frac{1}{S_n}\left[\mathrm{Var}(S_{n+1}-S_n|\mathscr{F}_n)\right]^{1/2}\\
=&\frac{1}{S_n}\left[E(S_n^2(\eta_{n+1}-\mu)^2|\mathscr{F}_n)\right]^{1/2}\\
=&[\mathrm{Var}(\eta_{n+1})]^{1/2}=\sigma.
\end{aligned}
$$

为了使得 $\{M_n,\mathscr{F}_n\}$ 成为鞅，需要

$$
\begin{aligned}
&E(M_{n+1}|\mathscr{F}_n)\\
=&(1+R)^{-n-1}E(S_{n+1}|\mathscr{F}_n)\\
=&(1+R)^{-n-1}\left[S_n+S_nE(\eta_{n+1}|\mathscr{F}_n)\right]\\
=&(1+R)^{-n-1}(1+\mu)S_n\\
=&M_n=(1+R)^{-n}S_n,
\end{aligned}
$$

这需要 $\mu=R$。可解得

$$
p=\frac{b-(1+R)}{b-a},
$$

且需要 $0<a<1+R<b$。这个概率称为“风险中性概率”，在此条件下风险资产投资才成为“公平赌博”，也叫做“无套利条件”。在此模型下如果 p 不等于风险中性概率，则可以构造投资策略实现无风险地盈利。

鞅是研究金融市场的重要理论工具。

6.1.3 凸函数变换

下面的引理可以让我们由已知的鞅或下鞅构造出许多新的下鞅.

考虑定义在有穷或无穷开区间 I 上的函数 $\varphi(x)$，称它为凸的，若 $\forall x,y\in I, 0<\alpha<1$，有

$$
\alpha\varphi(x)+(1-\alpha)\varphi(y)\geq\varphi[\alpha x+(1-\alpha)y]
$$

成立.

典型的凸函数：$\varphi(x)=x^2$; $\varphi(x)=|x|$。凸函数的曲线是向下弯曲的，如果函数二次可导且二阶导数大于等于零则为凸函数。

引理 6.1 (条件 Jensen 不等式). 设 $\varphi(x)$ 为实数集 $\mathbb{R}$ 上的凸函数, $\mathscr{G}$ 为 $\mathscr{F}$ 的子 σ 代数, 随机变量 X 满足

$$(1)E[|X|] < \infty;$$
$$(2)E[|\varphi(X)|] < \infty,$$

则有

$$E[\varphi(X)|\mathscr{G}] \geq \varphi(E[X|\mathscr{G}]). \tag{6.6}$$

注：条件 Jensen 不等式的方向有时会记不准。这时，注意 $\varphi(x) = x^2$ 是凸函数，而 $E[X^2|\mathscr{G}] \geq (E[X|\mathscr{G}])^2$ 所以 $E[\varphi(X)|\mathscr{G}] \geq \varphi(E[X|\mathscr{G}])$。

式(6.6)是著名的条件 Jensen 不等式，由它即可得到以下定理.

定理 6.1. *(1)* 设 $\{M_n, n \geq 0\}$ 是关于 $\{\mathscr{F}_n, n \geq 0\}$ 的鞅, $\varphi(x)$ 是 $\mathbb{R}$ 上的凸函数，且满足 $E[\varphi(M_n)^+] < \infty$, $\forall n \geq 0$，则 $\{\varphi(M_n), n \geq 0\}$ 是关于 $\{\mathscr{F}_n, n \geq 0\}$ 的下鞅. 特别地，$\{|M_n|, n \geq 0\}$ 是下鞅；当 $E[M_n^2] < \infty$, $\forall n \geq 0$ 时，$\{M_n^2, n \geq 0\}$ 也是下鞅.

(2) 设 $\{M_n, n \geq 0\}$ 是关于 $\{\mathscr{F}_n, n \geq 0\}$ 的下鞅, $\varphi(x)$ 是 $\mathbb{R}$ 上的非降凸函数，且满足 $E[\varphi(M_n)^+] < \infty$, $\forall n \geq 0$，则 $\{\varphi(M_n), n \geq 0\}$ 是关于 $\{\mathscr{F}_n, n \geq 0\}$ 的下鞅.

定理6.1证明:

(1) 当 $\{M_n\}$ 是鞅且 φ 是凸函数时，由条件 Jensen 不等式，

$$E(\phi(M_{n+1})|\mathscr{F}_n) \geq \phi[E(M_{n+1}|\mathscr{F}_n)] = \phi(M_n),$$

即 $\{\phi(M_n), \mathscr{F}_n, n \geq 0\}$ 是下鞅。

取 $\varphi(M_n) = |M_n|$, 因为 M_n 是鞅所以 $E|M_n| < \infty$, $E(|M_n|^+) = E|M_n| < \infty$, $|\cdot|$ 是函数，所以 $\{|M_n|\}$ 是下鞅。

取 $\varphi(M_n) = M_n^2$，在 $E(M_n^2) < \infty$ 时 $E[\varphi(M_n)^2] = E(M_n^2) < \infty$, x^2 是凸函数，所以 $\{M_n^2\}$ 是下鞅。

(2) 若 $\{M_n, n \geq 0\}$ 是关于 $\{\mathscr{F}_n, n \geq 0\}$ 的下鞅，$\phi(x)$ 是 $\mathbb{R}$ 上的非降凸函数，由条件 Jensen 不等式和 ϕ 的单调性得

$$E(\phi(M_{n+1})|\mathscr{F}_n) \geq \phi[E(M_{n+1}|\mathscr{F}_n)] \geq \phi(M_n),$$

即 $\{\phi(M_n), \mathscr{F}_n, n \geq 0\}$ 是下鞅。

6.1.4 下鞅分解定理

定理 6.2 (Doob 下鞅分解定理)**.** 设 $\{X_n, n = 0, 1, 2, ...\}$ 是关于 $\{\mathscr{F}_n, n \geq 0\}$ 的下鞅且 $E|X_n| < \infty$（$\forall n$），则存在关于 $\{\mathscr{F}_n, n \geq 0\}$ 的鞅 $\{M_n\}$ 和随机变量序列 $\{A_n\}$，满足 $A_0 = 0$, $A_n \leq A_{n+1}$，A_n 关于 $\mathscr{F}_{n-1}$ 可测且 $EA_n < \infty$（$\forall n$），使得

$$X_n = M_n + A_n,\ n = 0, 1, 2, ...$$

证明： 取 $A_0 = 0$，

$$\begin{aligned} A_{n+1} =& A_n + [E(X_{n+1}|\mathscr{F}_n) - X_n] \\ =& \sum_{k=0}^{n} [E(X_{k+1}|\mathscr{F}_k) - X_k], \end{aligned}$$

则易见 A_n 非负、单调增，关于 $\mathscr{F}_{n-1}$ 可测，且 $E|A_n| < \infty$。因为

$$M_{n+1} = X_{n+1} - A_{n+1} = X_{n+1} - E(X_{n+1}|\mathscr{F}_n) + X_n - A_n,$$

有

$$\begin{aligned} E(M_{n+1}|\mathscr{F}_n) =& E(X_{n+1}|\mathscr{F}_n) - E(X_{n+1}|\mathscr{F}_n) + X_n - A_n \\ =& X_n - A_n = M_n. \end{aligned}$$

在分解中，A_n 关于 $\mathscr{F}_{n-1}$ 可测，称 $\{A_n, \mathscr{F}_n\}$ 为**可料** (predictable) 随机序列。

6.2 鞅的停时定理及其应用

6.2.1 停时

本节中我们所讨论的鞅，都是指关于某个随机变量序列的鞅. 所得到的结论对关于子 σ 代数流的鞅也是成立的，为了便于理解和应用，我们没有追求结论的一般性.

对于一个关于 $\{X_n, n \geq 0\}$ 的鞅 $\{M_n, n \geq 0\}$，易知 $\forall n \geq 0$，有

$$E[M_n] = E[M_0], \tag{6.7}$$

我们想知道如果把此处固定的时间 n 换作一个随机变量 T，是否仍然有

$$E[M_T] = E[M_0]. \tag{6.8}$$

一般地，此结论未必成立. 但在一定的条件下可以保证它成立，这就是鞅的停时定理. 鞅的停时定理的意义是："在公平的赌博中，你不可能赢." 设想 $\{M_n, n \geq 0\}$ 是一种公平的博奕, M_n 表示局中人第 n 次赌局结束后的赌本. 式(6.7)说明他在每次赌局结束时的赌本的期望值与他开始时的赌本一样，但是他未必一直赌下去，他可以选择任一时刻停止赌博，这一时刻是随机的. 式(6.8)说明他停止时的赌本和他开始时的赌本相同，然而很容易看出在一般的情况下，这是不正确的. 比如例6.4中的赌博者采取的策略，就可以保证他在赢 1 元之后结束，所以我们要为式(6.8)的成立附加一些条件.

定义 6.6 (停时)**.** 设 $\{X_n, n \geq 0\}$ 是一个随机变量序列，称随机函数 T 是关于 $\{X_n, n \geq 0\}$ 的停时，如果 T 在 $\{0, 1, 2, \cdots, \infty\}$ 中取值，且对每个 $n \geq 0$，$\{T = n\} \in \sigma(X_0, X_1, \cdots, X_n)$.

停时可以看作选择一个停止观测的时间，在时刻 n 是否停止，只能依赖于截止到 n 为止的信息，而不能依赖于 $n+1, n+2, ...$ 时刻的信息。

如果 $P(T < \infty) = 1$，令 $X_T = X_{T(\omega)}(\omega)$，$X_T$ 是一个随机变量（设概率空间 $(\Omega, \mathscr{F}, P)$ 为完备概率空间）。不加 $P(T < \infty) = 1$ 条件时，

$$Y(\omega) = \begin{cases} X_{T(\omega)}(\omega), & \text{当} T(\omega) < \infty, \\ 0, & \text{当} T(\omega) = +\infty \end{cases}$$

是一个随机变量，可以表示为 $Y = X_T I_{\{T<\infty\}}$。

若 $\{\mathscr{F}_n, n = 0, 1, 2, ... \}$ 是 σ 代数流，随机函数 T 在 $\{0, 1, 2, \cdots, \infty\}$ 中取值，且 $\{T = n\} \in \mathscr{F}_n$, $\forall n \geq 0$，称 T 是 $\{\mathscr{F}_n\}$ 的停时。

由定义我们知道事件 $\{T = n\}$ 或 $\{T \neq n\}$ 都应该由 n 时刻及其之前的信息完全确定，而不需要也无法借助将来的情况. 仍然回到公平博奕的例子，赌博者决定何时停止赌博只能以他已经赌过的结果为依据，而不能说，如果我下一次要输我现在就停止赌博，这是对公平赌博的停止时刻 T 的第一个要求：它必须是一个停时.

以下看几个停时的例子.

例 6.9. 确定时刻 $T = n$ 是一个停时，即在赌博开始已确定 n 局之后一定结束，易见这是一个停时.

例 6.10 (首达时)**.** $\{X_n, n \geq 0\}$ 是一个随机变量序列，A 是一个实数集合（严格来说是 Borel 集），令

$$T(A) = \inf\{n : X_n \in A\}, \text{ 并约定 } \ T(\emptyset) = +\infty.$$

可见 $T(A)$ 是 $\{X_n, n \geq 0\}$ 首次进入集合 A 的时刻，称 $T(A)$ 是 $\{X_n, n \geq 0\}$ 到集合 A 的**首达时**，可以证明 $T(A)$ 是关于 $\{X_n, n \geq 0\}$ 的停时.

证明：

$$\{T(A) = n\} = \{X_0 \notin A, X_1 \notin A, \cdots, X_{n-1} \notin A, X_n \in A\},$$

显然 $\{T(A) = n\}$ 完全由 $X_0, X_1, \cdots, X_n$ 决定，从而 $T(A)$ 是关于 $\{X_n, n \geq 0\}$ 的停时.

命题 6.2. 设 T 是取值于 $\{0, 1, 2, \cdots, \infty\}$ 的随机变量，则下述三者等价

$$\begin{aligned}
&(1)\{T = n\} \in \sigma(X_0, X_1, \cdots, X_n); \\
&(2)\{T \leq n\} \in \sigma(X_0, X_1, \cdots, X_n); \\
&(3)\{T > n\} \in \sigma(X_0, X_1, \cdots, X_n).
\end{aligned}$$

证明：只要注意到如下等式，即可证明 (1),(2),(3) 的等价性.

$$\begin{aligned}
&\{T \leq n\} = \bigcup_{k=0}^{n} \{T = k\} \\
&\{T > n\} = \Omega - \{T \leq n\} \\
&\{T = n\} = \{T \leq n\} - \{T \leq n-1\}.
\end{aligned}$$

例 6.11. 如果 T 和 S 是两个停时，则 $T + S, \min(T, S)$ 和 $\max(T, S)$ 也是停时.

证明：这可由命题6.2来证明.

$$
\begin{aligned}
\{T+S=n\} &= \bigcup_{k=0}^{n} (\{T=k\} \cap \{S=n-k\}) \in \mathcal{F}_n; \\
\{\max(T,S) \le n\} &= \{T \le n\} \cap \{S \le n\} \in \mathcal{F}_n; \\
\{\min(T,S) > n\} &= \{T > n\} \cap \{S > n\} \in \mathcal{F}_n.
\end{aligned}
$$

例 6.12. 由例6.9可知常数 n 是停时. 设 T 是停时，令 $T_n = \min\{T, n\}$, 则由例6.11可知每个 T_n 都是停时，并且有

$$T_0 \le T_1 \le T_2 \le \cdots \le T_n \le n, \ \forall n.$$

命题 6.3. 设 $\{M_n, \mathscr{F}_n, n \ge 0\}$ 是鞅，T 是 $\{\mathscr{F}_n\}$ 的停时且 $P(T < \infty) = 1$，令 $T_n = \min(n, T)$, 以及

$$Y_n = M_{T_n}, n = 0, 1, ...$$

则 $\{Y_n, \mathscr{F}_n, n \ge 0\}$ 也是鞅。

证明：

$$
\begin{aligned}
Y_n &= \sum_{i=1}^{n} M_i I_{\{T=i\}} + M_n I_{\{T>n\}} \in \mathscr{F}_n, \\
E(|Y_n|) &\le \sum_{i=1}^{n} E|M_i| + E|M_n| < \infty.
\end{aligned}
$$

而

$$
\begin{aligned}
E(Y_{n+1}|\mathscr{F}_n) &= E\left(\sum_{i=1}^{n+1} M_i I_{\{T=i\}} + M_{n+1} I_{\{T>n+1\}} |\mathscr{F}_n\right) \\
&= \sum_{i=1}^{n} M_i I_{\{T=i\}} + E(M_{n+1} I_{\{T\ge n+1\}}|\mathscr{F}_n),
\end{aligned}
$$

注意到 $\{T \ge n+1\} = \{T > n\} \in \mathscr{F}_n$，所以

$$
\begin{aligned}
E(Y_{n+1}|\mathscr{F}_n) &= \sum_{i=1}^{n} M_i I_{\{T=i\}} + I_{\{T\ge n+1\}} E(M_{n+1}|\mathscr{F}_n) \\
&= \sum_{i=1}^{n} M_i I_{\{T=i\}} + M_n I_{\{T>n\}} = Y_n,
\end{aligned}
$$

证毕。

定理 6.3 (Wald 等式)**.** 设 $\{X_n, n=1,2,...\}$ 独立同分布，$E|X_1|<\infty$, T 是 $\{X_n, n=1,2,...\}$ 的停时，满足 $ET<\infty$，则

$$E\left[\sum_{n=1}^{T} X_n\right] = E(X_1)E(T).$$

证明：

$$\sum_{n=1}^{T} X_n = \sum_{n=1}^{\infty} X_n I_{\{n\le T\}}.$$

注意到

$$\{n\le T\} = \{T\le n-1\}^c$$

由 $X_1,...,X_{n-1}$ 决定，与 X_n 独立。

令

$$Y=\sum_{n=1}^{\infty}|X_n|I_{\{n\le T\}},$$

由单调收敛定理和独立性有

$$\begin{aligned}
E(Y) =& \sum_{n=1}^{\infty} E[|X_n|I_{\{n\le T\}}] \\
=& \sum_{n=1}^{\infty} E[|X_n|]EI_{\{n\le T\}} \\
=& E|X_1|\sum_{n=1}^{\infty} P(T\ge n) \\
=& E|X_1|E(T)<\infty.
\end{aligned}$$

由控制收敛定理可知

$$
\begin{aligned}
&E\sum_{n=1}^{T}X_n = E\sum_{n=1}^{\infty}X_nI_{\{n\le T\}}\\
=&\sum_{n=1}^{\infty}E[X_nI_{\{n\le T\}}] = \sum_{n=1}^{\infty}E[X_n]EI_{\{n\le T\}}\\
=&E(X_1)\sum_{n=1}^{\infty}P(T\ge n)\\
=&E(X_1)E(T).
\end{aligned}
$$

推论 6.1. 设 $\{X_n\}$ 是更新过程 $\{N(t),t\ge 0\}$ 的更新间隔时间，$0<E(X_1)<\infty$。有

$$E[S_{N(t)+1}] = E(X_1)[M(t)+1].$$

这是因为，$N(t)+1$ 关于 $\{X_n\}$ 是停时，因为

$$\{N(t)+1=n\} = \{N(t)=n-1\} = \{S_{n-1}\le t<S_n\}$$

只依赖于 $X_1,\dots,X_n$。所以由 Wald 等式可得结论。

要注意的是，$N(t)$ 关于 $\{X_n\}$ 不是停时，因为

$$\{N(t)=n\} = \{S_n\le t<S_{n+1}\},$$

依赖于 X_{n+1} 的值。

6.2.2 有界停时定理

在给出停时定理之前先给出有界停时的相关结论.

命题 6.4 (有界停时定理)**.** 设 $\{M_n, n\ge 0\}$ 是一个关于 $\{X_n, n\ge 0\}$ 的鞅，T 是一个关于 $\{X_n, n\ge 0\}$ 的停时并且 $T\le K$，$\mathscr{F}_n=\sigma(X_0,X_1,\cdots,X_n)$，则

$$E[M_T|\mathscr{F}_0]=M_0.$$

特别地，

$$E[M_T]=E[M_0].$$

证明：由于 $T \le K$，即 T 只取有限值，且当 $T=j$ 时 $M_T=M_j$，我们可以把 M_T 写作

$$M_T=\sum_{j=0}^{K} M_j I_{\{T=j\}}. \tag{6.9}$$

则

$$E|M_T| \le \sum_{j=0}^{K} E|M_j| < \infty.$$

对式(6.9)关于 $\mathscr{F}_{K-1}$ 取条件期望，有

$$\begin{aligned}E[M_T|\mathscr{F}_{K-1}] =&E[\sum_{j=0}^{K} M_j I_{\{T=j\}}|\mathscr{F}_{K-1}]\\ =&E[\sum_{j=0}^{K-1} M_j I_{\{T=j\}}|\mathscr{F}_{K-1}]+E[M_K I_{\{T=K\}}|\mathscr{F}_{K-1}]\end{aligned}$$

当 $j \le K-1$ 时，M_j 和 $I_{\{T=j\}}$ 都是 $\mathscr{F}_{K-1}$ 可测的，从而

$$E[\sum_{j=0}^{K-1} M_j I_{\{T=j\}}|\mathscr{F}_{K-1}]=\sum_{j=0}^{K-1} M_j I_{\{T=j\}}.$$

又因为 $T \le K$ 已知，则 $\{T=K\}$ 与 $\{T>K-1\}$ 是等价的，由命题6.2知 $\{T>K-1\} \in \sigma(X_0,\cdots,X_{K-1})$，因此

$$\begin{aligned}E[M_K I_{\{T=K\}}|\mathscr{F}_{K-1}] =&I_{\{T>K-1\}}E[M_K|\mathscr{F}_{K-1}]\\ =&I_{\{T>K-1\}}M_{K-1},\end{aligned}$$

从而

$$\begin{aligned}E[M_T|\mathscr{F}_{K-1}] =&I_{\{T>K-1\}}M_{K-1}+\sum_{j=0}^{K-1} M_j I_{\{T=j\}}\\ =&I_{\{T>K-2\}}M_{K-1}+\sum_{j=0}^{K-2} M_j I_{\{T=j\}},\end{aligned}$$

重复以上运算，关于 $\mathscr{F}_{K-2}$ 取条件期望，我们可以得到

$$\begin{aligned}E[M_T|\mathscr{F}_{K-2}] =&E[E[M_T|\mathscr{F}_{K-1}]|\mathscr{F}_{K-2}]\\ =&I_{\{T>K-2\}}M_{K-2}+\sum_{j=0}^{K-2} M_j I_{\{T=j\}}\\ =&I_{\{T>K-3\}}M_{K-2}+\sum_{j=0}^{K-3} M_j I_{\{T=j\}}.\end{aligned}$$

继续这样的过程，有

$$E(M_T|\mathscr{F}_1) = I_{\{T>0\}}M_1 + M_0 I_{\{T=0\}},$$

于是

$$\begin{aligned}
E[M_T|\mathscr{F}_0] =& E[E(M_T|\mathscr{F}_1) \,|\, \mathscr{F}_0] \\
=& E[I_{\{T>0\}}M_1|\mathscr{F}_0] + M_0 I_{\{T=0\}} \\
=& I_{\{T>0\}}E[M_1|\mathscr{F}_0] + M_0 I_{\{T=0\}} \\
=& I_{\{T>0\}}M_0 + M_0 I_{\{T=0\}} \\
=& M_0.
\end{aligned}$$

有界停时定理是鞅停时定理的一种特殊情况，可以看出，它的条件太强了，实际上我们感兴趣的问题中许多都不满足 T 有界这一严格的条件. 假设 T 是一停时并且 $P\{T < \infty\} = 1$，也就是说以概率 1 可以保证会停止 (相对于 $P\{T = \infty\} > 0$). 但与 T 有界不同的是，并没有确定的 K 使 $P\{T \leq K\} = 1$. 例如，上面博奕的例子，赌博者并不能确定在某一时刻之前肯定停止赌博，但可以保证这场赌博不会无限期地延续下去，那么在这种情况下，何种条件下可以得到 $E[M_T] = E[M_0]$ 的结论呢？

6.2.3 停时定理

考虑停时 $T_n = \min\{T, n\}$，注意到

$$M_T = M_{T_n} + M_T I_{\{T>n\}} - M_n I_{\{T>n\}}, \tag{6.10}$$

从而

$$E[M_T] = E[M_{T_n}] + E[M_T I_{\{T>n\}}] - E[M_n I_{\{T>n\}}],$$

可以看出，T_n 是一个有界停时 ($T_n \leq n$)，由上面命题可知 $E[M_{T_n}] = E[M_0]$. 我们希望当 $n \to \infty$ 时，(6.10)后面两项趋于 0。

对于 $E[M_T I_{\{T>n\}}] \to 0$ 来说，这是不困难的，因为 $P\{T < \infty\} = 1$，所以当 $n \to \infty$ 时，$P\{T > n\} \to 0$，$E[M_T I_{\{T>n\}}]$ 相当于对 M_T 限制在一个趋于空集的集合上取期望. 若要求 $E[|M_T|] < \infty$，由控制收敛定理就可保证 $E[M_T I_{\{T>n\}}] \to 0$.

$E[M_n I_{\{T>n\}}] \to 0$ 要麻烦一些，考虑前面的例6.4，在这个例子中，事件 $\{T>n\}$ 相当于事件：前 n 次投掷硬币, 均出现反面. 这个概率是 $(\frac{1}{2})^n$，如果这个事件发生了, 则至少赌博者已经输掉了 2^n-1 元，即 $M_n = 1-2^n$，从而

$$E[M_n I_{\{T>n\}}] = 2^{-n}(1-2^n) = -1+2^{-n} \to -1,\ n\to\infty,$$

不趋于 0, 这也是为什么停时定理的结论在此处不成立的原因. 然而如果 M_n 和 T 满足 $E|M_T|<\infty$ 和

$$\lim_{n\to\infty} E[|M_n| I_{\{T>n\}}] = 0,$$

我们就可以得出结论 $E[M_T] = E[M_0]$. 我们把这个过程写成如下的停时定理.

定理 6.4 (鞅停时定理)**.** 设 $\{M_n, n\ge 0\}$ 是一个关于 $\{\mathscr{F}_n = \sigma(X_0, X_1, \cdots, X_n)\}$ 的鞅，T 是停时，且满足

$$(1) P\{T<\infty\} = 1, \tag{6.11}$$

$$(2) E[|M_T|] < \infty, \tag{6.12}$$

$$(3) \lim_{n\to\infty} E[|M_n| I_{\{T>n\}}] = 0, \tag{6.13}$$

则有

$$E[M_T] = E[M_0].$$

推论 6.2. 设 $\{M_n, n\ge 0\}$ 是一个关于 $\{\mathscr{F}_n = \sigma(X_0, X_1, \cdots, X_n)\}$ 的鞅，T 是停时，且 $P(T<\infty)=1$，设存在随机变量 $Y\ge 0$，$E(Y)<\infty$，使得 $|M_n|\le Y$，$\forall n\ge 0$，则停时定理成立：$E[M_T]=E[M_0]$。上界 Y 的一个特例是 Y 恒等于一个正数。

证明：这时

$$E|M_T| \le E(Y) < \infty,$$
$$\lim_{n\to\infty} E\left[|M_n|\, I_{\{T>n\}}\right] \le \lim_{n\to\infty} E\left[Y\, I_{\{T>n\}}\right] = 0.$$

最后一个极限利用了控制收敛定理。于是，停时定理的三个条件成立。

推论 6.3. 设 $\{M_n, n\ge 0\}$ 是一个关于 $\{\mathscr{F}_n = \sigma(X_0, X_1, \cdots, X_n)\}$ 的鞅，设 T 是停时，且 $P(T<\infty)=1$，令 $T_n = \min(T, n)$，$n\ge 0$，存在随机变量 $Y\ge 0$，$E(Y)<\infty$，使得 $|M_{T_n}|\le Y$，$\forall n\ge 0$，则停时定理成立：$E[M_T]=E[M_0]$。上界 Y 的一个特例是 Y 恒等于一个正数。

证明：因 $P(T<\infty)=1$，可表示

$$M_T=\sum_{n=0}^{\infty} M_n I_{\{T=n\}}=\sum_{n=0}^{\infty} M_{T_n} I_{\{T=n\}}.$$

于是

$$\begin{aligned}|M_T| \leq & \sum_{n=0}^{\infty} |M_{T_n}| I_{\{T=n\}} \leq Y\sum_{n=0}^{\infty} I_{\{T=n\}} \\ =& Y, \\ E|M_T| \leq & E(Y)<\infty.\end{aligned}$$

又

$$\begin{aligned}E|M_n I_{\{T>n\}}|=&E\left[|M_{T_n}|\, I_{\{T>n\}}\right] \\ \leq & E(Y I_{\{T>n\}}) \to 0,\ (n\to\infty)\end{aligned}$$

所以停时定理条件成立。

推论 6.4. 设 $\{M_n, n\geq 0\}$ 是一个关于 $\{\mathscr{F}_n=\sigma(X_0, X_1, \cdots, X_n)\}$ 的鞅，设 T 是停时，且 $P(T<\infty)=1$，令 $T_n=\min(T,n)$，$n\geq 0$，存在常数 $b>0$ 使得使得 $E(M_{T_n}^2)\leq b$, $\forall n\geq 0$，则停时定理成立：$E[M_T]=E[M_0]$。

证明：

$$\begin{aligned}&E\left[M_{T_n}^2 I_{\{T\leq n\}}\right]\leq E(M_{T_n}^2)\leq b, \\ &E\left[M_{T_n}^2 I_{\{T\leq n\}}\right]=E\left[M_{T_n}^2 \sum_{k=0}^{n} I_{\{T=k\}}\right] \\ =&\sum_{k=0}^{n} E\left[M_{T_n}^2 I_{\{T=k\}}\right]=\sum_{k=0}^{n} E\left[M_k^2 I_{\{T=k\}}\right] \\ \to&\sum_{k=0}^{\infty} E\left[M_k^2 I_{\{T=k\}}\right]=E\left[\sum_{k=0}^{\infty} M_k^2 I_{\{T=k\}}\right]=E(M_T^2),\ n\to\infty,\end{aligned}$$

所以 $E(M_T^2)\leq b<\infty$, $E|M_T|\leq\sqrt{E(M_T^2)}<\infty$。

另外，

$$\begin{aligned}&E\left[|M_n|\, I_{\{T>n\}}\right]=E\left[|M_{T_n}|\, I_{\{T>n\}}\right] \\ \leq&\sqrt{E(M_{T_n}^2)}\sqrt{E(I_{\{T>n\}}^2)} \\ \leq&\sqrt{b}\sqrt{P(T>n)}\to 0,\ n\to\infty,\end{aligned}$$

所以停时定理条件成立。

例 6.13. 设 $\{X_n\}$ 是在 $\{0, 1, \cdots, N\}$ 上的简单随机游动 ($p = \frac{1}{2}$)，并且 0 和 N 为两个吸收壁. 设 $X_0 = a$，则 $\{X_n\}$ 是一个马氏链且是鞅，令 $T = \min\{j : X_j = 0或N\}$，则 T 是一个停时，求 $P(X_T = N)$。

解答: $\{X_n\}$ 是有限状态时齐马氏链，分为三组: $\{0\}$, $\{N\}$ 是常返态，$\{1, \dots, N-1\}$ 是瞬态。于是

$$P(\exists n使得m \geq n时X_m \notin \{1, \dots, N-1\}) = 1,$$

因此 $P(T < \infty) = 1$。

来证明 $\{X_n\}$ 是鞅。X_n 有界所以 $E|X_n| < \infty$。由马氏性, $X_{n+1}|\mathcal{F}_n$ 的条件分布等于 $X_{n+1}|X_n$ 的条件分布，对 $i \in \{1, \dots, N-1\}$ 有

$$E(X_{n+1}|X_n = i) = \frac{1}{2}(i-1) + \frac{1}{2}(i+1) = i,$$

对 $i = 0$ 或 $i = N$ 有 $P(X_{n+1} = i|X_n = i) = 1$ 故 $E(X_{n+1} = i|X_n = i) = i$，总之有

$$E(X_{n+1}|X_n = i) = i,\ i = 0, 1, \dots, N,\ n \geq 0,$$

从而

$$E(X_{n+1}|\mathscr{F}_n) = E(X_{n+1}|X_n) = X_n,\ n \geq 0,$$

即 $\{X_n\}$ 是鞅。

由于 X_n 的取值有界，由推论6.2可知

$$E[X_T] = E[X_0] = a.$$

由于此时 X_T 只取两个值 $N, 0$，

$$E[X_T] = N \cdot P\{X_T = N\} + 0 \cdot P\{X_T = 0\} = N \cdot P\{X_T = N\},$$

从而得到

$$P\{X_T = N\} = \frac{E[X_T]}{N} = \frac{a}{N}.$$

即在被吸收时刻它处于 N 点的概率为 $\frac{a}{N}$.

教材例 6.2.5 有错误。$M_n = X_n^2 - n$ 不可能是鞅，否则

$$E(M_0) = E(M_n) = E(X_n^2) - n,$$

但 $0 \le E(X_n^2) \le N^2$ 有界而 $n \to \infty$ 无界。改正的做法见后面的例6.14。

命题 6.5 (马氏链首达时概率（不可约情形）)**.** *对有限状态不可约时齐马氏链 $\{X_n, n \ge 0\}$，设状态空间为 $S = \{1, 2, \dots, K\}$，存在 $C > 0$ 和 $0 < \rho < 1$，使得对任意 $n \ge 0$ 和 $i, j \in S$ 都有*

$$P(X_t \ne j, 1 \le t \le n | X_0 = i) \le C\rho^n.$$

证明参见 §6.6.3.1。

推论 6.5. *在命题6.5条件下，设 A 是 S 的非空子集，则存在 $C' > 0$ 和 $0 < \rho' < 1$ 使得*

$$P(X_t \notin A, 1 \le t \le n | X_0 = i) \le C' \rho'^n.$$

这是节 6.2 和 6.3（停时定理）某些题目需要的引理。证明方法类似于更新过程中证明 $F_n(t) = O(\alpha^n)$ 的证明，见 §6.6.3.2。

命题 6.6 (马氏链首达时概率（可约情形）)**.** *设 $\{X_n\}$ 为有限状态马氏链，状态空间为 $S = \{1, 2, \dots, N\}$。设常返态的集合为 A，瞬态的集合为 $B = S - A$。证明对任意 $i \in S$，存在 $C > 0$ 和 $0 < \rho < 1$ 使得*

$$P(X_t \in B, 1 \le t \le n | X_0 = i) \le C\rho^n.$$

这是节 6.2 和 6.3 的某些题目需要的引理。证明见 §6.6.3.3。

例 6.14. 这是教材例 6.2.5 的改正版本。在例6.13中，$\{X_n\}$ 是双侧吸收壁的对称随机游动。令 $T_n = \min(T, n)$，则 T_n 是停时，令 $M_n = X_{T_n}^2 - T_n$，则 $\{M_n, \mathscr{F}_n, n \ge 0\}$ 是鞅，可以证明 $ET = a(N - a)$。

证明：

$$\begin{aligned} M_n =& X_{T_n}^2 - T_n \\ =& \sum_{k=1}^{n} (X_k^2 - k) I_{\{T=k\}} + (X_n^2 - n) I_{\{T>n\}} \in \mathscr{F}_n, \end{aligned}$$

且 $E|M_n| < \infty$。于是

$$\begin{aligned}E[M_{n+1}|\mathscr{F}_n] =& \sum_{k=1}^{n}(X_k^2 - k)I_{\{T=k\}} \\ &+ E[(X_{n+1}^2 - n - 1)I_{\{T=n+1\}}|\mathscr{F}_n] + E[(X_{n+1}^2 - n - 1)I_{\{T>n+1\}}|\mathscr{F}_n] \\ =& \sum_{k=1}^{n}(X_k^2 - k)I_{\{T=k\}} \\ &+ E[(X_{n+1}^2 - n - 1)I_{\{T>n\}}|\mathscr{F}_n] \\ =& \sum_{k=1}^{n}(X_k^2 - k)I_{\{T=k\}} \\ &+ I_{\{T>n\}}E[(X_{n+1}^2 - n - 1)|\mathscr{F}_n],\end{aligned}$$

由 $\{X_n\}$ 马氏性，

$$E[X_{n+1}^2 - n - 1|\mathscr{F}_n] = E[X_{n+1}^2 - n - 1|X_n],$$

令 Y_{n+1} 表示第 $n+1$ 步的变化，$P(Y_{n+1} = -1) = P(Y_{n+1} = 1) = \frac{1}{2}$ 且 Y_{n+1} 与 X_n 独立，于是对 $i \notin \{0, N\}$,

$$\begin{aligned}E[X_{n+1}^2 - n - 1|X_n = i] =& E[(i + Y_{n+1})^2 - n - 1] \\ =& i^2 + EY_{n+1}^2 + 2iEY_{n+1} - n - 1 \\ =& i^2 + 1 + 0 - n - 1 = i^2 - n,\end{aligned}$$

注意在 $T > n$ 条件下必有 $X_n \notin \{0, N\}$，所以

$$I_{\{T>n\}}E[X_{n+1}^2 - n - 1|\mathscr{F}_n] = I_{\{T>n\}}(X_n^2 - n),$$

$$\begin{aligned}E[M_{n+1}|\mathscr{F}_n] =& \sum_{k=1}^{n}(X_k^2 - k)I_{\{T=k\}} + I_{\{T>n\}}(X_n^2 - n) \\ =& M_n,\end{aligned}$$

即 $\{M_n, \mathscr{F}_n, n \geq 0\}$ 是鞅。

来验证停时定理的条件。例6.13已证明 $P(T < \infty) = 1$，来证明 $E|M_T| < \infty$。注意

$$\begin{aligned}M_T =& X_T^2 - T, \\ |M_T| \leq& N^2 + T,\end{aligned}$$

由命题6.6，

$$\begin{aligned} E(T) =& \sum_{n=1}^{\infty} nP(T = n) \\ \le& \sum_{n=1}^{\infty} nP(X_t \in \{1, \dots, N-1\}, 1 \le t \le n-1 | X_0 = a) \\ \le& \sum_{n=1}^{\infty} nC\rho^{n-1} < \infty, \end{aligned}$$

所以停时定理的第二个条件 $E|M_T| < \infty$ 成立。

再来证明停时定理第三个条件成立。

$$\begin{aligned} E[|M_n| I_{\{T>n\}}] \le& E\left[\sum_{k=1}^{n} |X_k^2 - k| I_{\{T=k\}} I_{\{T>n\}}\right] \\ &+ E\left[|X_n^2 - n| I_{\{T>n\}}^2\right] \\ \le& 0 + (N^2 + n)P(T > n) \\ \le& (N^2 + n)P(X_t \in \{1, \dots, N-1\}, 1 \le t \le n | X_0 = a) \\ \le& (N^2 + n) \cdot C\rho^n \to 0,\ (n \to \infty). \end{aligned}$$

所以，停时定理三个条件都满足，有

$$E(M_T) = E(M_0) = E(X_{T_0}^2 - T_0) = E(X_0^2 - 0) = a^2,$$

但

$$\begin{aligned} E(M_T) =& E(X_T^2) - E(T) = 0 \cdot P(X_T = 0) + N^2 P(X_T = N) - E(T) \\ =& N^2 \frac{a}{N} - E(T) = aN - E(T), \end{aligned}$$

其中用到了例6.13得到的 $P(X_T = N) = \frac{a}{N}$ 结果。于是

$$E(T) = a(N - a).$$

例 6.15. 令 X_n 是一个在 $\{0, \pm 1, \pm 2, \cdots\}$ 上的简单随机游动 $(p = \frac{1}{2})$, $X_0 = 0$，我们已经知道 $\{X_n\}$ 是一个不可约时齐马氏链且是鞅。令 $T = \min\{j : X_j = 1\}$，这是一个首达时，是停时。由例5.18知这个简单随机游动是常返的（零常

返），从而 $P\{T<\infty\}=1$. 又由 $X_T=1$ 知 $E[X_T]=1\neq 0=E[X_0]$，所以停时定理不成立.

实际上，此时停时定理的条件是不满足的，我们不再给出具体证明，只给出如下事实：在这种情况下，

$$P\{T>n\}\sim Cn^{-\frac{1}{2}},$$

其中 C 为常数，从而

$$E[|X_n|I_{\{T>n\}}]\not\to 0.$$

6.2.4 上鞅停时定理

在介绍了鞅的停时定理之后，我们简单地讨论一下有关上鞅停时定理的两个结果. 这在下面的期权值界的例子中是必需的.

定理 6.5. 设 $\{M_n,n\geq 0\}$ 是关于 $\{X_n,n\geq 0\}$ 的上鞅, T 是关于 $\{X_n,n\geq 0\}$ 的停时，$T_n=\min(T,n)$. 设存在一非负随机变量 W，满足 $E[W]<\infty$，且使得

$$M_{T_n}\geq -W,\ \forall n\geq 0,$$

则有

$$E[M_0]\geq E[M_T I_{\{T<\infty\}}].$$

特别地，若 $P\{T<\infty\}=1$，则有

$$E[M_0]\geq E[M_T].$$

证明略.

推论 6.6. 设 $\{M_n,n\geq 0\}$ 是关于 $\{X_n,n\geq 0\}$ 的上鞅, T 是关于 $\{X_n,n\geq 0\}$ 的停时，且 $M_n\geq 0$，则有

$$E[M_0]\geq E[M_T I_{\{T<\infty\}}].$$

我们已经知道对于上鞅，有 $E[M_n]\leq E[M_0], \forall n\geq 0$，此处上鞅停时定理说明当把 n 换为停时 T 时，在附加某些条件的前提下，结论也成立.

6.2.5 停时定理的应用——关于期权值的界

暂略。

6.3 一致可积性

在鞅的停时定理的条件中，式(6.13)即 $\lim_{n\to\infty} E[|M_n|I_{\{T>n\}}]=0$ 一般是很难验证的，为此我们将给出一些容易验证的条件，这些条件蕴含了式(6.13).

首先考虑一个随机变量 X，满足 $E[|X|]<\infty$，$|X|$ 的分布函数为 F, 则

$$\lim_{n\to\infty} E[|X|I_{\{|X|>n\}}]=\lim_{n\to\infty}\int_n^{\infty}|X|dF(x)=0,$$

这也可以用控制收敛定理证明。

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

$$E[|X|I_A]\le E[|X|I_{\{|X|>n\}}],$$

事实上

$$\begin{aligned}
&E[|X|I_A]-E[|X|I_{\{|X|>n\}}]\\
=&E[|X|I_AI_{\{|X|>n\}}]+E[|X|I_AI_{\{|X|\le n\}}]-E[|X|I_AI_{\{|X|>n\}}]-E[|X|I_{A^c}I_{\{|X|>n\}}]\\
=&E[|X|I_AI_{\{|X|\le n\}}]-E[|X|I_{A^c}I_{\{|X|>n\}}]\\
\le&nE[I_AI_{\{|X|\le n\}}]-nE[I_{A^c}I_{\{|X|>n\}}]\\
=&nE[I_AI_{\{|X|\le n\}}]+nE[I_AI_{\{|X|>n\}}]-nE[I_AI_{\{|X|>n\}}]-nE[I_{A^c}I_{\{|X|>n\}}]\\
=&nE(I_A)-nE(I_{\{|X|>n\}}])=0.
\end{aligned}$$

于是，我们可以有以下结论：

命题 6.7. *如果随机变量 X 满足 $E[|X|]<\infty$，则 $\forall\varepsilon>0$, $\exists\delta>0$, 只要 $P(A)<\delta$ 就有*

$$E[|X|\,I_A]<\varepsilon.$$

如果将单个随机变量 X 替换成一列随机变量，就需要推广到一致可积性：

定义 6.7 (一致可积)**.** 一列随机变量 $X_1,X_2,\cdots$ 称为一致可积的，如果 $\forall\varepsilon>0$, $\exists\delta>0$, 使得 $\forall A\in\mathscr{F}$, 只要 $P(A)<\delta$ 就有

$$E[|X_n|I_A]<\varepsilon,\ \forall n\ge 0. \tag{6.14}$$

这个定义的关键在于 δ 不能依赖于 n，并且式(6.14)对任意 n 成立.

先给一个不一致可积的例子.

例 6.16. 考虑例6.4. 令 A_n 是事件 $\{X_1 = X_2 = \cdots = X_n = -1\}$，则 $P(A_n) = \frac{1}{2^n}$，$E[|W_n|I_{A_n}] = 2^{-n}(2^n - 1) \to 1$. 容易看出 $\{W_n\}$ 不满足一致可积的条件.

假设 $\{M_n, n \geq 0\}$ 是一个关于 $\{X_n, n \geq 0\}$ 的一致可积鞅，T 是停时且 $P\{T < \infty\} = 1$ 或等价地 $\lim_{n\to\infty} P\{T > n\} = 0$. 则由一致可积性可得

$$\lim_{n\to\infty} E[|M_n|I_{\{T>n\}}] = 0,$$

即式(6.13)成立。事实上，$\forall \varepsilon > 0$，$\exists \delta > 0$，使得只要 $P(A) < \delta$ 就有

$$E[|M_n|I_A] < \varepsilon.$$

因为 $P(T > n) \to 0$，所以 $\exists N > 0$ 使得 $n \geq N$ 时 $P(T > n) < \delta$，所以

$$E[|M_n|I_{\{T>n\}}] < \varepsilon,$$

这就是 $\lim_{n\to\infty} E[|M_n|I_{\{T>n\}}] = 0$ 的定义。

据此我们给出停时定理的另一种叙述.

定理 6.6 (一致可积条件的停时定理)**.** 设 $\{M_n, n \geq 0\}$ 是一个关于 $\{X_n, n \geq 0\}$ 的一致可积鞅，T 是停时，满足 $P\{T < \infty\} = 1$ 且 $E[|M_T|] < \infty$, 则有 $E[M_T] = E[M_0]$.

一致可积的条件一般比较难验证，下面给出两个一致可积的充分条件.

命题 6.8 (二阶矩有界则一致可积)**.** 假设 $X_1, X_2, \ldots$ 是一列随机变量，并且存在常数 $0 < C < \infty$，使得 $E[X_n^2] \leq C$ 对所有的 n 成立，则此序列是一致可积的.

证明: $\forall \varepsilon > 0$，令 $\delta = \frac{\varepsilon^2}{4C}$，设 $P(A) < \delta$，则

$$\begin{aligned} E[|X_n|I_A] =& E[|X_n|I_{\{A\cap\{|X_n|\geq\frac{2C}{\varepsilon}\}\}}] + E[|X_n|I_{\{A\cap\{|X_n|<\frac{2C}{\varepsilon}\}\}}] \\ \leq& \frac{\varepsilon}{2C} \cdot E[|X_n|^2 I_{\{A\cap\{|X_n|\geq\frac{2C}{\varepsilon}\}}] + \frac{2C}{\varepsilon} \cdot P\{A \cap \{|X_n| < \frac{2C}{\varepsilon}\}\} \\ \leq& \frac{\varepsilon}{2C} E[X_n^2] + \frac{2C}{\varepsilon} P(A) < \varepsilon. \end{aligned}$$

命题 6.9 (有可积上界则一致可积)**.** 设 $\{X_n, n = 0, 1, 2, \ldots\}$ 是随机变量序列，如果存在一个非负随机变量 Y，满足 $E(Y) < \infty$ 且 $|X_n| \leq Y$，$\forall n \geq 0$ 成立，则 $\{X_n\}$ 一致可积.

这只要利用命题6.7。

一致可积的充分条件还有一些，我们不再多列举了.

例 6.17 (分支过程)**.** 令 X_n 表示分支过程第 n 代的个体数. 设每个个体产生后代的分布有均值 μ 和方差 σ^2，则 $\{M_n = \mu^{-n}X_n\}$ 是关于 $\{X_n, n \geq 0\}$ 的鞅 (习题 6.3). 假设 $\mu > 1$，则存在一个常数 C，使得 $\forall n$, $E[M_n^2] \leq C$，从而 $\{M_n\}$ 是一致可积鞅 (习题 6.6(2)).

6.4 鞅收敛定理

鞅论中有两个深刻的结论，一个是上节的停时定理，另一个就是鞅收敛定理. 本节我们将介绍鞅的收敛定理.

鞅收敛定理说明在很一般的条件下，鞅 $\{M_n\}$ 会收敛到一个随机变量，在此记为 M_∞.

我们首先来考虑一个特殊的例子——Polyas 坛子抽样模型 (例6.6). 令 M_n 表示第 n 次摸球后红球所占的比例，当 $n \to \infty$ 时，这个比例会如何变化呢？下面来说明其变化趋势.

设 $0 < a < b < 1$, $M_n < a$，且令

$$T = \min\{j : j \geq n, M_j \geq b\}$$

即 T 表示 n 次摸球之后比例从小于 a 到超越 b 的第一个时刻. 令 $T_m = \min\{T, m\}$，则对于 $m > n$，由停时定理可知

$$E[M_{T_m}] = M_n < a,$$

但是

$$\begin{aligned} E[M_{T_m}] \geq & E[M_{T_m} \cdot I_{\{T \leq m\}}] \\ = & E[M_T I_{\{T \leq m\}}] \\ \geq & b \cdot P\{T \leq m\}, \end{aligned}$$

从而

$$P\{T \leq m\} < \frac{a}{b}.$$

因为上式对一切 $m > n$ 成立，于是有

$$P\{T < \infty\} \leq \frac{a}{b}.$$

这说明至少以概率 $1-\frac{a}{b}$ 红球的比例永远不会超过 b.

现在我们假定这一比例已经超过了 b，那么它能够再一次降回到 a 以下的概率是多少呢？同样的讨论可知，这一概率最大为 $\frac{1-b}{1-a}$.

继续同样的讨论，我们可以知道，从 a 出发超过 b，再小于 a，再大于 b, ……，有 n 个循环的概率应为

$$p_n \leq (\frac{a}{b})(\frac{1-b}{1-a})(\frac{a}{b})\cdots(\frac{a}{b})(\frac{1-b}{1-a}) = (\frac{a}{b})^n(\frac{1-b}{1-a})^n \to 0,\ (n\to\infty).$$

由此可见，这个比例不会在 a,b 之间无限次地跳跃. 由 a,b 的任意性，也表明这一比例不会在任意的两个数之间无限地跳跃. 直观地看，极限 $\lim_{n\to\infty} M_n$ 存在，记为 M_∞. 这一极限是一个随机变量，可以证明 M_∞ 服从 $[0,1]$ 上的均匀分布 (见习题 6.7).

下面我们给出一般的结论.

定理 6.7 (鞅收敛定理). 设 $\{M_n, n\geq 0\}$ 是关于 $\{X_n, n\geq 0\}$ 的鞅，并且存在常数 $0<C<\infty$ 使得 $E[|M_n|]\leq C$ 对任意 n 成立，则当 $n\to\infty$ 时，$\{M_n\}$ *a.s.* 收敛到一个随机变量 M_∞.

证明略去。上面的定理没有要求一致可积性，只要求一阶矩有界。在一致可积性条件下可以给出更强结论：

定理 6.8. 如果 $\{M_n, n\geq 0\}$ 是关于 $\{X_n, n\geq 0\}$ 的一致可积鞅，则 $\lim_{n\to\infty} M_n$ 存在，记为 M_∞，并且

$$E[M_\infty] = E[M_0].$$

证明略.

例 6.18. 令 X_n 表示分支过程中第 n 代的个体数，每个个体生育后代的分布有均值 μ 和方差 σ^2，假定 $X_0=1$，令 $M_n=\mu^{-n}X_n$. 由例6.17已经知道 $\{M_n\}$ 是鞅. 如果 $\mu\leq 1$，由第 5 章的结论已经知道灭绝一定会发生，由此 $M_n\to M_\infty=0$, 从而 $E[M_\infty]\neq E[M_0]$. 在上节我们说明了若 $\mu>1$，则 M_n 是一致可积的，所以在 $\mu>1$ 时，有 $E[M_\infty]=E[M_0]=1$.

例 6.19. 令 $X_1, X_2, \cdots$ 为一独立同分布随机变量序列，$P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}$，令

$$M_n = \sum_{j=1}^{n} \frac{1}{j} X_j.$$

证明收敛定理结论成立。

证明：易见 $\{M_n\}$ 是鞅.

我们来证明 $\{M_n\}$ 是一致可积的，显然 $E[M_n] = 0$，$\mathrm{Var}(X_j) = E(X_j^2) = 1$，则

$$\begin{aligned} E[M_n^2] =& \mathrm{Var}[M_n] \\ =& \sum_{j=1}^{n} \mathrm{Var}[\frac{1}{j} X_j] = \sum_{j=1}^{n} \frac{1}{j^2} \\ \leq& \sum_{j=1}^{\infty} \frac{1}{j^2} < \infty. \end{aligned}$$

从而 $\{M_n\}$ 是一致可积鞅，当 $n \to \infty$ 时，$M_n \to M_\infty$ 并且 $E[M_\infty] = 0$.

例 6.20. 再考虑 Polya 模型 (例6.6). 这里假定最初坛子中有 m 个黄球，k 个红球，所以在第 n 次摸球后坛子中应有 $n + m + k$ 个球. 假定 M_n 为红球的比例. 因为 $0 < M_n < 1$，有上界 $Y = 1$，所以 $\{M_n\}$ 是一致可积鞅，所以当 $n \to \infty$ 时，$M_n \to M_\infty$ 并且 $E[M_\infty] = E[M_0] = \frac{k}{k+m}$. 可以证明，$M_\infty$ 服从 Beta 分布 $B(m, k)$，其分布密度为

$$\frac{\Gamma(k+m)}{\Gamma(k)\Gamma(m)} x^{k-1}(1-x)^{m-1}, \quad 0 < x < 1.$$

证明略.

在 Bayes 统计中，这一结果是很自然的. 假设我们对某一事件发生的概率 p 感兴趣，而对 p 又一无所知，我们就只能假定 p 是 $[0, 1]$ 上的均匀分布 (这就是先验分布). 现在假设我们一共做了 $k + m - 2$ 次试验，事件发生了 $k - 1$ 次，则根据 Bayes 定理, p 的后验分布就应该是参数为 k 和 m 的 Beta 分布.

例 6.21. 令 $\{M_n\}$ 是一关于 $X_0, X_1, ...$ 的鞅，T 是停时，且 $P\{T < \infty\} = 1$，令 $T_n = \min(T, n)$, $Y_n = M_{T_n}$，则 $Y_n \to Y_\infty$ 且 $Y_\infty = M_T$; 若 $\{M_n\}$ 是一致可积的，有 $E[Y_\infty] = E[Y_0]$.

证明: 令 $A = \{\omega : T < \infty\}$, 则 $P(A) = 1$，对 $\omega \in A$，$T(\omega) < \infty$，$\exists N > T(\omega)$，于是 $n \geq N$ 时 $T_n(\omega) = T(\omega)$, $Y_n(\omega) = M_{T(\omega)}(\omega)$, 于是

$$\lim_{n\to\infty} Y_n(\omega) = M_{T(\omega)}(\omega),\ \forall \omega \in A,$$

即 $Y_n \to Y_\infty = M_T$, a.s. 这个结论并不需要鞅收敛定理条件成立。

如果 $\{M_n\}$ 一致可积，则停时定理条件成立，有 $E(M_T) = E(M_0)$，但 $Y_\infty = M_T$, $Y_0 = M_0$，所以也有 $E(Y_\infty) = E(Y_0)$。

例 6.22. 令 $X_1, X_2, \ldots$ 是独立同分布的随机变量序列，$P\{X_i = \frac{3}{2}\} = P\{X_i = \frac{1}{2}\} = \frac{1}{2}$. 令 $M_0 = 1$, 对 $n > 0$, 令 $M_n = X_1 X_2 \cdots X_n$. 则 M_n 有极限但非一致可积。

证明：注意

$$\begin{aligned}
E(X_n) =& \frac{3}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} = 1, \\
E[M_n] =& E[X_1] \cdots E[X_n] = 1, \\
E[M_{n+1}|\mathscr{F}_n] =& E[X_1 \cdots X_{n+1}|\mathscr{F}_n] \\
=& X_1 \cdots X_n \cdot E[X_{n+1}|\mathscr{F}_n] \\
=& X_1 \cdots X_n \cdot E[X_{n+1}] \\
=& X_1 \cdots X_n \cdot 1 = M_n,
\end{aligned}$$

所以 $\{M_n\}$ 是关于 $X_1, X_2, \cdots$ 的鞅. 由于 $E[|M_n|] = E[M_n] = 1$，鞅收敛定理的条件成立，从而

$$M_n \to M_\infty.$$

那么 $\{M_n\}$ 一致可积吗？答案是否定的. 事实上，$M_\infty = 0$(这样 $E[M_\infty] \neq E[M_0]$). 为此考虑

$$\frac{1}{n} \ln M_n = \frac{1}{n} \sum_{j=1}^n \ln X_j,$$

右边是独立同分布随机变量的均值，并且

$$E[\ln X_i] = \frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln \frac{3}{2} = \frac{1}{2} \ln \frac{3}{4} < 0,$$

根据大数定律 $\frac{1}{n} \ln M_n \to \frac{1}{2} \ln \frac{3}{4} < 0$，a.s., 因此 $\ln M_n \to -\infty$, a.s., 从而 $M_n = e^{\ln M_n} \to 0$, a.s., 故 $\{M_n\}$ 不是一致可积的.

6.5 连续鞅

6.5.1 定义

前面我们讨论了鞅的停时定理，也称为可选抽样定理 (optional sampling theorem) 和鞅收敛定理. 请注意这里的鞅都是以离散时间 n 为参数的. 事实上，对于连续参数鞅 (仍称为鞅) 也有类似定理，出于应用的考虑，我们不加证明地给出这些定理. 首先给出连续鞅的定义.

定义 6.8. 设 $(\Omega, \mathscr{F}, P)$ 是一个完备的概率空间，$\{\mathscr{F}(t), t \geq 0\}$ 为一个 σ 代数流. 随机过程 $\{X(t), t \geq 0\}$(简记为 $\{X(t)\}$) 称为 $\{\mathscr{F}(t)\}$ 适应的，如果对每个 $t \geq 0$, $X(t)$ 为 $\mathscr{F}(t)$ 可测. 一个适应过程 $\{X(t)\}$ 称为关于 $\{\mathscr{F}(t)\}$ 的鞅，如果每个 $X(t)$ 一阶矩有限（或称可积)，且对一切 $0 \leq s < t$，有

$$E[X(t)|\mathscr{F}(s)] = X(s), \text{ a.s.} \tag{6.15}$$

特别地，当 $\mathscr{F}(t) = \sigma(X(u), 0 \leq u \leq t)$ 时，式(6.15)变为

$$E[X(t)|X(u), 0 \leq u \leq s] = X(s), \text{ a.s.}$$

此时，简称 $\{X(t)\}$ 为鞅.

可类似定义下鞅和上鞅。

注:$X(t)$ 关于 $\mathscr{F}(t)$ 可测,定义为 $\forall B \in \mathscr{B}$,有 $X^{-1}(t)(B) \in \mathscr{F}(t)$. $\sigma(X(u), 0 \leq u \leq t)$ 表示由 $\{X(u), 0 \leq u \leq t\}$ 生成的 σ 代数，即包含一切形如 $\{X(u) \leq x\}(0 \leq u \leq t,\, x \in \mathbb{R})$ 的事件的最小 σ 代数。这个定义可以推广到 $X(t)$ 取值于 $\mathbb{R}^d$ 的情形。

若随机过程 $\{X(t), t \geq 0\}$ 是鞅，则对任意 $t > 0$, 有

$$E[X(t)] = E[E[X(t)|X(0)]] = E[X(0)]. \tag{6.16}$$

与离散鞅类似，有下述简单例子.

例 6.23. 设 $\{Y_t, t \geq 0\}$ 是零初值具有平稳独立增量的随机过程. 令

$$X_t = X_0 e^{Y_t},$$

其中 X_0 为一常数. 若 $E[e^{Y_t}] = 1$，则 $\{X_t, t \geq 0\}$ 是一个鞅.

证明：事实上

$$E[|X_t|] = |X_0|E[e^{Y_t}] = |X_0| < \infty,$$

再对 $0 \le s < t$，有

$$\begin{aligned} E[X_t|X_r, 0 \le r \le s] =& E[X_s e^{Y_t - Y_s}|X_r, 0 \le r \le s] \\ =& X_s E[e^{Y_t - Y_s}] \\ =& X_s E[e^{Y_{t-s}}] = X_s, \text{ a.s.} \end{aligned}$$

注：$Y(t)$ 的一个取法是取为带漂移的布朗运动

$$Y(t) = -\frac{1}{2}t + B(t).$$

例 6.24. 设 $\{N(t), t \ge 0\}$ 是参数为 λ 的泊松过程，$\mathscr{F}^N(t) = \sigma(\{N(s) : 0 \le s \le t\})$，$M(t) = N(t) - \lambda t$，则 $\{M(t), \mathscr{F}^N(t), t \ge 0\}$ 是鞅，称 $\{M(t)\}$ 为抵偿泊松过程 (compensated Poisson process)。

证明：由泊松过程的独立增量性，对 $t \ge 0$, $u > 0$，

$$\begin{aligned} E(M(t+u)|\mathscr{F}(t)) =& E(N(t+u) - \lambda(t+u)|N(s), 0 \le s \le t) \\ =& E[N(t) + (N(t+u) - N(t))|N(s), 0 \le s \le t] - \lambda(t+u) \\ =& N(t) + E(N(t+u) - N(t)) - \lambda(t+u) \\ =& N(t) + \lambda u - \lambda(t+u) \\ =& N(t) - \lambda t = M(t). \end{aligned}$$

注：将 $N(t)$ 分解为

$$N(t) = M(t) + A(t),$$

其中 $M(t)$ 是鞅，$A(t) = \lambda t$ 为增函数，易见 $N(t)$ 是下鞅，对下鞅有更一般的 Doob-Meyer 下鞅分解定理。

6.5.2 停时定理和极限定理

定义 6.9. 非负广义随机函数 τ 称为 $\{\mathscr{F}(t)\}$ 停时，如果 $P(\tau < \infty) = 1$，并且对一切 $t \ge 0$，$\{\tau \le t\}$ 是 $\mathscr{F}(t)$ 可测的，即

$$\{\tau \le t\} \in \mathscr{F}(t).$$

特别地，当 $\mathscr{F}(t)=\sigma(X(u),0\le u\le t)$ 时，τ 称为关于随机过程 $\{X(t),t\ge 0\}$ 的停时. 若存在常数 $k>0$ 使得 $P\{\tau\le k\}=1$，则称 τ 为有界停时.

非负广义随机函数是指 $\tau:\Omega\longrightarrow[0,\infty]$ 映射。

定义 σ 代数 $\mathscr{F}(\tau)$ 为

$$\mathscr{F}(\tau)=\{A\in\mathscr{F}:A\cap\{\tau\le t\}\in\mathscr{F}(t),\forall t\ge 0\}.$$

如果过程 $\{X(t,\omega),\mathscr{F}(t),t\ge 0,\omega\in\Omega\}$ 为适应过程且 X（二元）可测，τ 是 $\{\mathscr{F}(t)\}$ 停时，则 $X(\tau)=X(\tau(\omega),\omega)$ 是随机变量，关于 $\mathscr{F}(\tau)$ 可测。如果随机过程 X 每条轨道右连续，可测条件满足。泊松过程、布朗运动是轨道右连续的。

下面是鞅论的一个重要结论——停时定理，即在适当条件下，将式(6.16)中的 t 置换成停时 τ 时，等式仍然成立.

定理 6.9. *若 τ 是有界停时，则有*

$$E[X_\tau]=E[X_0].$$

鞅论的另一个重要的结果是收敛定理.

定理 6.10. *设 $\{X_t,t\ge 0\}$ 是一个鞅并且 $X_t\ge 0$, $\forall t\ge 0$ (简称为非负鞅)，则存在几乎处处收敛的有限极限，即有*

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

6.6 补充材料

6.6.1 可测性

因为随机过程 $\{X(t,\omega):t\in[0,\infty),\omega\in\Omega\}$ 是关于时间 t 和样本点 ω 的二元函数，给定 ω 后，作为 t 的函数，$X(\cdot,\omega)$ 代表一条样本路径（轨道、实现），所以两个随机过程在某种意义上相同，有如下三种不同的定义：

1. 如果 $\forall t$, $P(X(t)=Y(t))=1$，称 $\{Y(t)\}$ 是 $\{X(t)\}$ 的修改；
2. $\{X(t)\}$ 和 $\{Y(t)\}$ 具有相同的有限维分布；
3. $P(X(t)=Y(t),t\in[0,\infty))=1$，称 $\{Y(t)\}$ 是 $\{X(t)\}$ 不可区分。

第 3 条是最强的。第 1 条成立时，仅保证了在每个时刻相等，但是样本路径可以可能没有任何一对路径完全重合。

注意第 1 条和第 3 条仅在 $\{X(t)\}$ 和 $\{Y(t)\}$ 属于同一概率空间时有意义，而第 2 条不要求 $\{X(t)\}$ 和 $\{Y(t)\}$ 属于同一概率空间，仅要求其样本空间相同。

如果 $\{Y(t)\}$ 是 $\{X(t)\}$ 的修改且两个过程都是 a.s. 轨道右连续的，则两个过程也不可区分。

随机过程 $\{X(t), t\in[0,\infty)\}$ 的定义是对每个 $t\in[0,\infty)$, $X(t)$ 是 $(\Omega,\mathscr{F})$ 的随机变量，所以不要求 $X(t,\omega)$ 作为 t 和 ω 的二元函数二元可测；但某些问题中会要求二元可测，如随机积分的被积函数。

定义 6.10 (可测过程)**.** 称 d 维的随机过程 X 可测，如果 $\forall B\in\mathscr{B}(\mathbb{R}^d)$, $\{(t,\omega): X(t,\omega)\in B\}$ 属于 σ 域 $\mathscr{B}([0,\infty))\times\mathscr{F}$，即函数 $X(t,\omega)$ 是从 $([0,\infty)\times\Omega, \mathscr{B}([0,\infty))\times\mathscr{F})$ 到 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 的可测变换。

当过程可测时，给定 ω 后的路径 $X(t,\omega)$ 作为 t 的函数是 Borel 可测函数，对给定 t，$X(t,\omega)$ 是随机变量。如果对某个区间 I 有 $\int_I E|X(t)|\,dt<\infty$，由 Fubini 定理则有

$$\begin{aligned}\int_I |X(t,\omega)|\,dt <& \infty,\ \text{a.s.},\\ E\int_I X(t,\omega)\,dt =& \int_I E[X(t)]\,dt.\end{aligned}$$

6.6.2 σ 代数流

随机过程中的 t 一般代表时间，指标集 T 最常见的情形为 $T=\{0,1,2,\dots\}$ 和 $T=[0,\infty)$。随着时间的推进，积累的信息也在增加，可以将时间区分为过去、现在和将来。用 "σ 代数流" 表示积累的信息。

定义 6.11. 对 $T=\{0,1,2,\dots\}$ 或 $T=[0,\infty)$，设有一组 σ 代数 $\mathscr{F}(t)$，使得 $\mathscr{F}(s)\subset\mathscr{F}(t)\subset\mathscr{F}$ 对任意 $0\le s\le t$ 成立，称 $\{\mathscr{F}(t), t\ge 0\}$ 为一个σ **代数流** (filtration)。如果随机过程 $\{X(t), t\in T\}$ 满足 $X(t)$ 关于 $\mathscr{F}(t)$ 可测对任意 $t\in T$ 成立，则称 $\{X(t), t\in T\}$ 是关于 $\{\mathscr{F}(t), t\ge 0\}$ **适应**的。

为了使得 $\{X(t), \mathscr{F}(t), t\ge 0\}$ 是适应过程，最小的 $\mathscr{F}(t)$ 的取法为

$$\mathscr{F}^X(t)=\sigma(\{X(s): 0\le s\le t\}).$$

将 $\mathscr{F}^X(t)$ 视作截止到 t 时刻为止，随机过程 X 所包含的信息。对 $A \in \mathscr{F}^X(t)$，掌握 X 在 $[0,t]$ 的路径信息可以明确判断 A 是否发生。

对 $T=[0,\infty)$，定义

$$\mathscr{F}(t-) = \sigma(\bigcup_{0\le s<t} \mathscr{F}(s)),$$
$$\mathscr{F}(t+) = \bigcap_{\epsilon>0} \mathscr{F}(t+\epsilon).$$

$\mathscr{F}(t-)$ 是时间落在 t 前面的事件的 σ 代数。$\mathscr{F}(t+)$ 是 t 截止到刚过 t 时刻的事件的 σ 代数。如果 $\mathscr{F}(t+)=\mathscr{F}(t)$，$\forall t\ge 0$, 称 $\{\mathscr{F}(t), t\ge 0\}$ 是右连续的；如果 $\mathscr{F}(t-)=\mathscr{F}(t)$，$\forall t\ge 0$, 称 $\{\mathscr{F}(t), t\ge 0\}$ 是左连续的。

对 $\{(X(t), \mathscr{F}^X(t), t\ge 0\}$：

- $\{\mathscr{F}^X(t)\}$ 左连续，意味着观测到 $X(s), 0\le s<t$ 可以获得 $X(t)$ 的信息；
- $\{\mathscr{F}^X(t)\}$ 右连续，意味着观测到 $X(s), 0\le s\le t$ 后，再向前继续观测无穷小时间的信息并没有额外的信息增加。
- 如果 $\{X(t), t\ge 0\}$ 轨道左连续，则 $\{\mathscr{F}^X(t)\}$ 左连续。
- $\{\mathscr{F}^X(t+)\}$ 是右连续的。
- 即使 $\{X(t), t\ge 0\}$ 轨道连续，$\{\mathscr{F}^X(t)\}$ 也不一定右连续，$\{\mathscr{F}^X(t+)\}$ 也不一定左连续。

比可测和适应更强的条件是循序可测 (progressively measurable)。

定义 6.12. 设 $\{X(t), t\ge 0\}$ 是 d 维随机过程，$\{\mathscr{F}(t), t\ge 0\}$ 为 σ 代数流，如果对任意 $t\ge 0$, $X(t,\omega)$ 是 $([0,t]\times\Omega, \mathscr{B}([0,t])\times\mathscr{F}(t))$ 到 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 的可测映射，称 X 关于 $\mathscr{F}(t)$ **循序可测**。

循序可测必为可测，且为适应的。

反过来，如果 $\{X(t), \mathscr{F}(t), t\ge 0\}$ 是可测、适应的，则必存在 X 的修改 Y，使得 $\{Y(t), \mathscr{F}(t), t\ge 0\}$ 是循序可测的。

命题 6.10. *设 $\{X(t), \mathscr{F}(t), t\ge 0\}$ 是适应过程，X 的每条路径都是右连续（或者每条路径都是左连续），则 $\{X(t), \mathscr{F}(t), t\ge 0\}$ 是循序可测过程。*

证明：参见 (Karatzas and Shreve 1998) P.5 命题 1.12。

以右连续情形为例证明。将路径用右连续阶梯函数逼近，对给定 t，任意正整数

n，任意 $s \in (0, t]$，设有整数 k 使得

$$s \in \left(\frac{k}{2^n} t, \frac{k+1}{2^n} t \right],$$

令

$$X_n(s) = X(\frac{k+1}{2^n} t).$$

这是分段阶梯函数，右连续，是 $([0, t] \times \Omega, \mathscr{B}([0, t]) \times \mathscr{F}(t)$ 到 $\mathbb{R}^d$ 的可测映射。根据右连续性，

$$\lim_{n \to \infty} X_n(s) = X(s), \forall (s, \omega) \in [0, t] \times \Omega,$$

因为可测函数的序列极限是可测函数，所以 $X(s, \omega)$ 是 $([0, t] \times \Omega, \mathscr{B}([0, t]) \times \mathscr{F}(t)$ 到 $\mathbb{R}^d$ 的可测映射。即 $\{X(t), \mathscr{F}(t), t \geq 0\}$ 是循序适应过程。

这样，后面讲的布朗运动是循序适应过程。

如果 $\{X(t), t \geq 0\}$ 的每条轨道右连续（或者每条轨道左连续），则 X 必为可测过程。注意这里没有要求适应。

定义 6.13. 称 σ 代数流 $\{\mathscr{F}(t), t \geq 0\}$ 满足“通常条件”（usual conditions），如果其右连续，且 $\mathscr{F}(0)$ 包含 $\mathscr{F}$ 中所有零概率事件的子集。

如果概率空间 $(\Omega, \mathscr{F}, P)$ 完备，则 $\mathscr{F}(0)$ 包含 $\mathscr{F}$ 中所有零概率事件的子集这个条件是满足的。

6.6.3 补充证明

6.6.3.1 命题6.5证明

由于任意状态之间互通，所以对任意 i, j，存在 $N_{ij} \geq 1$ 使得 $p_{ij}^{(N_{ij})} > 0$，令

$$N = \max\{N_{ij} : i, j \in S\},$$
$$\delta = \min\{p_{ij}^{(N_{ij})} : i, j \in S\},$$

则 $N < \infty, \delta > 0$，于是对任意 $i, j \in S$，

$$P(X_t \neq j, 1 \leq t \leq N | X_0 = i) \leq 1 - \delta.$$

来证明对于任意正整数 m，都有

$$P(X_t \neq j, 1 \leq t \leq mN | X_0 = i) \leq (1-\delta)^m, \ \forall i, j \in S. \qquad (*)$$

用归纳法，设上式对 $1, 2, \dots, m-1$ 成立。则

$$\begin{aligned}
& P(X_t \neq j, 1 \leq t \leq mN | X_0 = i) \\
=& \sum_{s \neq j} P(X_t \neq j, 1 \leq t < N, X_N = s; X_t \neq j, N < t \leq mN | X_0 = i) \\
=& \sum_{s \neq j} P(X_t \neq j, 1 \leq t < N, X_N = s | X_0 = i) P(X_t \neq j, N < t \leq mN | X_N = s) \\
\leq& (1-\delta)^{m-1} \sum_{s \neq j} P(X_t \neq j, 1 \leq t < N, X_N = s | X_0 = i) \\
=& (1-\delta)^{m-1} P(X_t \neq j, 1 \leq t \leq N | X_0 = i) \\
\leq& (1-\delta)^m.
\end{aligned}$$

(*) 式得证。

取 $\rho = (1-\delta)^{\frac{1}{N}}$，则 $0 < \rho < 1$，

$$P(X_t \neq j, 1 \leq t \leq n | X_0 = i) \leq \rho^n, \ n = mN, \ \forall i, j \in S.$$

取 $C = \rho^{-N}$，对 $1 \leq n < N$，有

$$\begin{aligned}
& P(X_t \neq j, 1 \leq t \leq n | X_0 = i) \\
\leq& 1 \leq \rho^{n-N} = C\rho^n.
\end{aligned}$$

对 $n = mN + k$, 其中 $m \geq 1$, $0 \leq k < N$，有

$$\begin{aligned}
& P(X_t \neq j, 1 \leq t \leq n | X_0 = i) \\
\leq& P(X_t \neq j, 1 \leq t \leq mN | X_0 = i) \\
\leq& \rho^{mN} = \rho^{n-k} = \rho^{-k}\rho^n \\
\leq& \rho^{-N}\rho^n = C\rho^n, \ \forall n \geq N, \ \forall i, j \in S.
\end{aligned}$$

这就证明了结论对 $n \geq 1$ 成立。

6.6.3.2 推论6.5证明

对 $j \in A$，存在 C_j 和 ρ_j 使得

$$P(X_t \neq j, 1 \leq t \leq n | X_0 = i) \leq C_j \rho_j^n.$$

取 $C' = \max\{C_j, j \in A\}$, $\rho' = \max\{\rho_j, j \in A\} = \rho_m (m \in A)$, 则

$$\begin{aligned}
&P(X_t \notin A, 1 \leq t \leq n | X_0 = i) \\
=&P(X_t \neq j, 1 \leq t \leq n, j \in A | X_0 = i) \\
\leq&P(X_t \neq m, 1 \leq t \leq n | X_0 = i) \\
\leq&C_m \rho_m^n \leq C' \rho'^n, \ n \geq 0.
\end{aligned}$$

6.6.3.3 命题6.6证明

因为从常返态不可能到达瞬态，所以当 i 为常返态时不等式左边为 0，仅需要考虑 $i \in B$ 的情形。

因为 i 是瞬态，所以必存在 $m_i > 0$ 和 $j_i \in A$ 使得 $p_{ij_i}^{(m_i)} > 0$。用反证法证明这一命题，如果不然，则对任意 $n > 0$ 和 $j \in A$ 都有 $p_{ij}^{(n)} = 0$，则由 $\sum_{j \in S} p_{ij}^{(n)} = 1$ 得 $\sum_{j \in B} p_{ij}^{(n)} = 1$，但对 $j \in B$ 有 $p_{ij}^{(n)} \to 0$, $n \to \infty$，$B \subset S$ 是有限集所以 $\sum_{j \in B} p_{ij}^{(n)} \to 0$，矛盾。

于是

$$\begin{aligned}
&P(X_t \in B, 1 \leq t \leq m_i | X_0 = i) \\
\leq&P(X_{m_i} \neq j_i | X_0 = i) \\
=&1 - P(X_{m_i} = j_i | X_0 = i) \\
=&1 - p_{ij_i}^{(m_i)} < 1.
\end{aligned}$$

令 $m = \max\{m_i : i \in B\}$，$\delta = \min\{p_{ij_i}^{(m_i)} : i \in B\} > 0$, 则对所有 $i \in B$ 同时成立

$$\begin{aligned}
&P(X_t \in B, 1 \leq t \leq m | X_0 = i) \\
\leq&P(X_t \in B, 1 \leq t \leq m_i | X_0 = i) \\
=&1 - p_{ij_i}^{(m_i)} \leq 1 - \delta < 1.
\end{aligned}$$

来证明对 $s=1,2,...$ 有

$$P(X_t\in B,1\le t\le sm|X_0=i)\le(1-\delta)^s. \qquad (*)$$

用数学归纳法。当 $s=1$ 时（*）式成立。设（*）式对 s 成立，来证明对 $s+1$ 成立。用 C-K 方程

$$\begin{aligned}
&P(X_t\in B,1\le t\le(s+1)m|X_0=i)\\
\le&\sum_{j\in B}P(X_t\in B,sm+1\le t\le(s+1)m;X_m=j,X_t\in B,1\le t<m|X_0=i)\\
=&\sum_{j\in B}P(X_t\in B,sm+1\le t\le(s+1)m|X_m=j)P(X_m=j,X_t\in B,1\le t<m|X_0=i)\\
\le&(1-\delta)^s\sum_{j\in B}P(X_m=j,X_t\in B,1\le t<m|X_0=i)\\
=&(1-\delta)^sP(X_t\in B,1\le t\le m|X_0=i)\le(1-\delta)^{s+1}.
\end{aligned}$$

令 $\rho=(1-\delta)^{1/m}$，则 $0<\rho<1$。令 $C=\rho^{-m}$, 对 $n=sm$ 有

$$P(X_t\in B,1\le t\le n|X_0=i)\le C\rho^s,$$

对 $1\le n<m$ 有

$$P(X_t\in B,1\le t\le n|X_0=i)\le 1\le\rho^{n-m}=C\rho^n.$$

对 $n=sm+j$, $s\ge1$，$0\le j<m$，有

$$\begin{aligned}
&P(X_t\in B,1\le t\le n|X_0=i)\\
\le&P(X_t\in B,1\le t\le sm|X_0=i)\\
\le&\rho^{sm}\le\rho^{sm+j}\rho^{-m}=C\rho^n,
\end{aligned}$$

证毕。

Chapter 7

布朗运动

7.1 高斯分布

7.1.1 一元正态分布

随机变量 Z 服从标准正态分布 N(0,1)，则密度函数为

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \ -\infty < x < \infty,$$

$\phi(x)$ 是偶函数。分布函数为 $\Phi(x) = \int_{-\infty}^{x} \phi(u)\, du$，满足

$$\Phi(-x) = 1 - \Phi(x).$$

矩母函数为

$$E(e^{tZ}) = e^{\frac{t^2}{2}}, \ -\infty < t < \infty.$$

特征函数为

$$\psi(t) = Ee^{itZ} = e^{-\frac{1}{2}t^2}, \ -\infty < t < \infty.$$

各阶原点矩为

$$E(Z^k) = \begin{cases} 0, & k\text{为奇数}, \\ (k-1)!!, & k\text{为偶数}. \end{cases}$$

特别地，$EZ^4 = 3$。$E|Z| = \sqrt{\frac{2}{\pi}}$。

随机变量 $X \sim \mathrm{N}(\mu, \sigma^2)$，则密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\ -\infty < x < \infty,$$

分布函数为 $F(x) = \Phi(\frac{x-\mu}{\sigma})$。矩母函数为

$$E(e^{uX}) = \exp\{\mu t + \frac{1}{2}\sigma^2 t^2\},\ -\infty < t < \infty.$$

特征函数为

$$\psi_X(t) = Ee^{itX} = \exp\left\{i\mu t - \frac{1}{2}\sigma^2 t^2\right\},\ -\infty < t < \infty.$$

$\frac{X-\mu}{\sigma} \sim \mathrm{N}(0, 1)$。

7.1.2 高斯分布

连续型随机向量 $X = (X_1, \dots, X_n)^T$ 称为（狭义的）多维正态分布随机向量，如果其联合密度为

$$f(x) = (2\pi)^{-\frac{n}{2}} [\det(\Sigma)]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right\},\ x \in \mathbb{R}^n.$$

记此分布为 $\mathrm{N}(\mu, \Sigma)$ 或 $\mathrm{N}_n(\mu, \Sigma)$，其中 $\mu \in \mathbb{R}^n$，Σ 为 n 阶对称正定矩阵。

X 的特征函数为

$$\psi(t) = \exp\left\{i\mu^T t - \frac{1}{2} t^T \Sigma t\right\}, t \in \mathbb{R}^n. \tag{7.1}$$

定义 7.1 (多维高斯分布)**.** 称随机向量 $X = (X_1, X_2, \cdots, X_n)^T$ 服从 n 维高斯分布 (或多元正态分布，多维正态分布), 如果存在 $\mu \in \mathbb{R}^n$, $n \times k$ 常数矩阵 B 和 iid 的标准正态随机变量 $Z_1, Z_2, \dots, Z_k$ 使得

$$X = \mu + BZ.$$

其中 $Z = (Z_1, \dots, Z_k)^T$。这时

$$EX = \mu, \quad \Sigma = \mathrm{Var}(X) = BB^T.$$

记 X 的分布为 $\mathrm{N}_n(\mu, \Sigma)$。

这里我们不特别区分高斯分布和正态分布。多维正态分布要求 Σ 正定，而高斯分布仅要求 Σ 非负定。对一维情形，高斯分布允许方差等于零，这就是退化分布（等于一个常数）。

定理 7.1. n 维高斯分布的特征函数为(7.1); 反之，如果某个随机向量特征函数为(7.1) （其中 Σ 为对称非负定矩阵），则它服从高斯分布 $N_n(\mu, \Sigma)$。

对二元正态分布有

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right).$$

条件分布为

$$Y|X = x \sim \mathrm{N}\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1),\ (1 - \rho^2)\sigma_2^2. \right)$$

若

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right),$$

且 Σ_{11} 正定（满秩），则 $Y|X = x$ 的条件分布为

$$\mathrm{N}\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x - \mu_1),\ \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \right).$$

定理 7.2. $X \sim N(\mu, \Sigma)$ 的充分必要条件是: 对任意 $a \in \mathbb{R}^n$,

$$Y = a^T X \ \sim N(a^T\mu, a^T\Sigma a).$$

见 (何书元 2003) 节 1.4 定理 4.1。这里的一元正态分布是推广的 $\mathrm{N}(\mu, \sigma^2)$，允许 $\sigma^2 = 0$。

定理 7.3. 设 $X \sim N_n(\mu, \Sigma)$, A 为 $m \times n$ 矩阵, $Y = AX$, 则 $Y \sim N_m(A\mu, A\Sigma A^T)$。即多元正态分布的线性组合仍为多元正态分布。

证明：Y 的特征函数为

$$\begin{aligned} Ee^{it^T Y} =& Ee^{it^T AX} = Ee^{i(A^T t)^T X} \\ =& \exp\left\{ i(A^T t)^T\mu - \frac{1}{2}(A^T t)^T \Sigma (A^T t) \right\} \\ =& \exp\left\{ it^T(A\mu) - \frac{1}{2}(t)^T(A\Sigma A^T)t \right\}, \end{aligned}$$

由特征函数与分布的一一对应关系可得定理结论。

定理 7.4. 设 $\xi_n \sim N(\mu_n, \sigma_n^2)$, $n \in \mathbb{N}$, 若 ξ_n 依分布收敛到随机变量 ξ, 则极限

$$\lim \mu_n = \mu,\ \lim \sigma_n^2 = \sigma^2$$

存在，且 $\xi \sim N(\mu, \sigma^2)$.

证明参见王梓坤《随机过程论》P.18。

推论 7.1. 设正态 $\xi_n \sim N(\mu_n, \sigma_n^2)$, $n \in \mathbb{N}$, 若 ξ_n 均方收敛到随机变量 ξ, 则极限

$$\lim \mu_n = \mu,\ \lim \sigma_n^2 = \sigma^2$$

存在，且 $\xi \sim N(\mu, \sigma^2)$.

例 7.1. 边缘分布为一元正态分布不能推出联合分布为多元正态分布。如 Z_1, Z_2 独立同标准正态分布，令

$$X = Z_1, \quad Y = |Z_2| \cdot \operatorname{sgn}(Z_1),$$

则 X, Y 边缘分布都是正态分布但联合分布非多元正态分布，因为 (X, Y) 只在 $\mathbb{R}^2$ 平面的第一、三象限上取值。

7.1.3 高斯过程

定义 7.2 (高斯过程)**.** 对随机过程 $\{X(t), t \geq 0\}$，如果对任意正整数 n 和 $0 \leq t_1 < t_2 < \cdots < t_n$，$(X(t_1), X(t_2), \dots, X(t_n))$ 都服从高斯分布，则称 $\{X(t), t \geq 0\}$ 为高斯过程。

高斯过程的有限维分布完全由其均值函数和协方差函数决定。

7.2 布朗运动概念与性质

7.2.1 对称随机游动

设 $X_1, X_2, \dots$ 是独立的随机变量，

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

令 S_n 表示相应的对应随机游动

$$S_n = X_1 + X_2 + \cdots + X_n.$$

令 $S_0 = 0$。易见

$$
\begin{aligned}
E(X_i) =& 0, \quad \mathrm{Var}(X_i) = 1,\\
E(S_n) =& 0, \quad \mathrm{Var}(S_n) = n.
\end{aligned}
$$

$\{S_n, n = 0, 1, 2, \dots\}$ 是独立平稳增量过程。

令 $\mathscr{F}_n = \sigma(X_1, \dots, X_n)$ 是由 $\{X_i, 1 \le i \le n\}$ 生成的 σ 代数，也可以理解为包含在 $X_1, \dots, X_n$ 中的信息. S_n 是鞅（所有增量均值为 0 的独立平稳增量过程都是鞅）：

$$
E(S_{n+1}|\mathscr{F}_n) = E(S_n + X_{n+1}|\mathscr{F}_n) = S_n + E(X_{n+1}) = S_n, n \ge 0.
$$

递推可得

$$
E(S_n|\mathscr{F}_m) = S_m,\ \forall 0 \le m < n.
$$

考虑上述随机游动 S_n 的每一步实际代表了时间间隔 Δt。给定 $t > 0$ 和正整数 n，当 nt 为正整数时定义可以定义随机过程 $\{X^{(n)}(t), t \ge 0\}$ 使得

$$
X(t) = S_{nt}.
$$

对 $t = 1$，$n = 10$ 时 $X(t)$ 对应 S_{10}，$n = 100$ 时 $X(t)$ 对应 S_{100}，所以 n 越大，相当于在 $[0, t]$ 区间内采样的格子点越多。但是，这会使得 $\mathrm{Var}(S_{nt}) = nt$，所以 $n \to \infty$ 时 $\mathrm{Var}(X^{(n)}(t)) = nt \to \infty$，这样定义的 $\{X^{(n)}(t), t \ge 0\}$ 在极限情况下没有二阶矩。于是，对 $X(t)$ 进行适当伸缩，取

$$
X^{(n)}(t) = \frac{1}{\sqrt{n}} S_{nt},\ t \ge 0, \tag{7.2}
$$

则

$$
\mathrm{Var}(X^{(n)}(t)) = t.
$$

这样给定 n 后对所有使得 nt 为整数的 t 都定义了 $X^{(n)}(t)$ 的值，即在 $\frac{1}{n}$ 间隔的格子点上的值有定义。对于两个格子点中间的 t，其 $X^{(n)}(t)$ 的值用两侧的值进行线性插值，这样 $X^{(n)}(t)$ 是轨道连续的，但仅由时间间隔 $\frac{1}{n}$ 处的值决定。令 $n \to \infty$，$X^{(n)}(t)$ 可以收敛到一个随机过程 $\{X(t), t \ge 0\}$。由 $X^{(n)}(t)$ 的性质，可以猜测 $\{X(t), t \ge 0\}$ 具有如下性质：

(1) $\{X(t), t \ge 0\}$ 轨道连续。

(2) $X(t)$ 服从均值为 0, 方差为 t 的正态分布.

(3) $\{X(t), t \geq 0\}$ 有独立增量.

(4) $\{X(t), t \geq 0\}$ 有平稳增量.

(5) $\{X(t), t \geq 0\}$ 是鞅。

定理 7.5 (Donsker 不变原理)**.** 设 $\{\xi_n, n = 0, 1, ...\}$ 为独立同分布随机变量序列，$E\xi_1 = 0$，$E(\xi_1^2) = 1$。对 $h > 0$，$t \geq 0$，定义

$$\eta_t^h = \sqrt{h} \cdot (\xi_1 + \xi_2 + \cdots + \xi_{[\frac{t}{h}]}).$$

则 $\{\eta_t^h, t \geq 0\}$ 是一个右连左极过程（指轨道右连续、有左极限）。$h \to 0$ 时 $\{\eta_t^h, t \in [0,1]\}$ 收敛到一个标准布朗运动 $\{B(t), t \in [0,1]\}$。

定理中 ξ_n 没有要求服从正态分布，但由中心极限定理可知 $h \to 0$ 时极限分布为正态分布。标准布朗运动定义见下面。

7.2.2 定义

下面我们就给出 Brown 运动的严格定义.

定义 7.3 (布朗运动)**.** 随机过程 $\{B(t), t \geq 0\}$ 如果满足：

(1) $B(0) = 0$;

(2) $\{X(t), t \geq 0\}$ 有平稳独立增量；

(3) 对每个 $t > 0$, $X(t)$ 服从正态分布 $N(0, t)$.

则称 $\{B(t), t \geq 0\}$ 为**标准布朗运动**，也称为维纳过程 (Wiener 过程). 常记为 $\{B(t), t \geq 0\}$ 或 $\{W(t), t \geq 0\}$.

也可以定义 $X(t) = \sigma B(t)$，称 $\{X(t)\}$ 为布朗运动。不失一般性，可以只考虑标准 Brown 运动的情形.

由于这一定义在应用中不是十分方便，我们不加证明地给出下面的性质作为 Brown 运动的等价定义，其证明可以在许多随机过程的著作中找到.

命题 7.1. *Brown 运动是具有下述性质的随机过程 $\{B(t), t \geq 0\}$:*

(1) (正态增量) $B(t) - B(s) \sim N(0, t - s)$. 当 $s = 0$ 时, $B(t) - B(0) \sim N(0, t)$.

(2) (独立增量) 对 $0 < s < t$, $B(t) - B(s)$ 独立于过程的过去状态 $B(u), 0 \leq u \leq s$.

(3) (路径的连续性) $B(t), t \geq 0$ *是* t *的连续函数.*

注命题7.1没有假定 $B(0) = 0$，因此我们称之为始于 x 的 Brown 运动，所以有时为了强调起始点，也记为 $\{B^x(t)\}$. 这样，定义7.3所指的就是始于 0 的 Brown 运动 $\{B^0(t)\}$. 易见，

$$B^x(t) - x = B^0(t). \tag{7.3}$$

(7.3)式按照下面的定义称为布朗运动的**空间齐次性**. 此性质也说明，$B^x(t)$ 和 $x + B^0(t)$ 是相同的，我们只需研究始于 0 的布朗运动就可以了. 如不加说明，布朗运动就是始于 0 的布朗运动.

定义 7.4. 设 $\{X(t), t \geq 0\}$ 是随机过程，如果它的有限维分布是空间平移不变的，即

$$\begin{aligned} &P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n | X(0) = 0\} \\ =&P\{X(t_1) \leq x_1 + x, X(t_2) \leq x_2 + x, \dots, X(t_n) \leq x_n + x | X(0) = x\}, \end{aligned}$$

则称此过程为**空间齐次**的.

如果 $\{X(t), t \geq 0\}$ 的任意有限维分布都是连续型的，则空间齐次当且仅当 $X(0) = 0$ 下 $(X(t_1), \dots, X(t_n))$ 的条件密度 $f(x_1, \dots, x_n)$，在 $X(0) = x$ 条件下变成 $f(x_1 + x, \dots, x_n + x)$，对任意正整数 n 和 $0 \leq t_1 < \cdots < t_n$，$x_1, \dots, x_n \in \mathbb{R}$ 成立。

布朗运动还有如下推广的定义。

定义 7.5. 设 $(\Omega, \mathscr{F}, P)$ 是完备的概率空间，$\{\mathscr{F}(t), t \geq 0\}$ 是 $\mathscr{F}$ 中的一族子 σ 代数，满足 $\mathscr{F}(s) \subset \mathscr{F}(t), \forall 0 \leq s < t$，称 $\{\mathscr{F}(t)\}$ 为 σ 代数流 (filtration)。随机过程 $\{B(t), t \geq 0\}$ 满足 $B(t)$ 关于 $\mathscr{F}(t)$ 可测（$\forall t \geq 0$），称 $\{B(t), \mathscr{F}(t), t \geq 0\}$ 为适应过程。如果 $B(0) = 0$, $B(t)$ 轨道连续，且对任意 $0 \leq s < t$，$B(t) - B(s)$ 与 $\mathscr{F}(s)$ 独立，服从 $\mathrm{N}(0, t-s)$ 分布，则称 $\{B(t), \mathscr{F}(t), t \geq 0\}$ 为标准布朗运动。

当 $\mathscr{F}(t) = \mathscr{F}^B(t) = \sigma(\{B(u) : 0 \leq u \leq t\})$ 时，两种定义等价。

另外，因为 $B(t)$ 轨道连续，所以 $B(t, \omega)$ 是 $([0, \infty) \times \Omega, \mathscr{B}([0, \infty)) \times \mathscr{F})$ 到 $\mathscr{B}(\mathbb{R})$ 的 (二元) 可测变换。

7.2.3 联合分布

先看一个布朗运动有限维分布的计算例子.

例 7.2. 设 $\{B(t), t\geq 0\}$ 是标准 Brown 运动，计算 $P\{B(2)\leq 0\}$ 和 $P\{B(t)\leq 0, t=0,1,2\}$.

解：由于 $B(2)\sim \mathrm{N}(0,2)$，所以

$$P\{B(2)\leq 0\}=\frac{1}{2}.$$

因为 $B(0)=0$，所以

$$P\{B(t)\leq 0, t=0,1,2\}=P\{B(t)\leq 0, t=1,2\}=P\{B(1)\leq 0, B(2)\leq 0\}.$$

虽然 $B(1)$ 和 $B(2)$ 不是独立的，但由命题7.1(2) 和 (3) 可知 $B(2)-B(1)$ 与 $B(1)$ 是相互独立的标准正态分布随机变量，于是利用分解式

$$B(2)=B(1)+(B(2)-B(1))$$

我们有

$$\begin{aligned}
&P\{B(1)\leq 0, B(2)\leq 0\}\\
=&P\{B(1)\leq 0, B(1)+(B(2)-B(1))\leq 0\}\\
=&P\{B(1)\leq 0, B(2)-B(1)\leq -B(1)\}\\
=&\int_{-\infty}^{0} P\{B(2)-B(1)\leq -x\}\phi(x)dx \qquad (*)\\
=&\int_{-\infty}^{0}\Phi(-x)\phi(x)\,dx=\int_{0}^{\infty}\Phi(x)\phi(-x)\,dx\\
=&\int_{0}^{\infty}\Phi(x)\phi(x)\,dx=\int_{0}^{\infty}\Phi(x)\,d\Phi(x)\\
=&\int_{\frac{1}{2}}^{1} y\,dy=\frac{3}{8}.
\end{aligned}$$

其中 (*) 式用了全期望公式:

$$\begin{aligned}
&P\{B(1)\leq 0, B(2)\leq 0\}\\
=&P\{B(1)\leq 0, B(2)-B(1)\leq -B(1)\}\\
=&E\left[I_{\{B(1)\leq 0\}}I_{\{B(2)-B(1)\leq -B(1)\}}\right]\\
=&E\left\{E\left[I_{\{B(1)\leq 0\}}I_{\{B(2)-B(1)\leq -B(1)\}}\middle|\, B(1)\right]\right\}
\end{aligned}$$

而

$$
\begin{aligned}
& E\left[I_{\{B(1)\leq 0\}}I_{\{B(2)-B(1)\leq -B(1)\}}\middle|\, B(1)=x\right] \\
=& E\left[I_{\{x\leq 0\}}I_{\{B(2)-B(1)\leq -x\}}\middle|\, B(1)=x\right] \\
=& I_{\{x\leq 0\}}E\left[I_{\{B(2)-B(1)\leq -x\}}\middle|\, B(1)=x\right] \\
=& I_{\{x\leq 0\}}\Phi(-x),
\end{aligned}
$$

所以

$$
\begin{aligned}
& P\{B(1)\leq 0, B(2)\leq 0\} \\
=& E\left\{I_{\{B(1)\leq 0\}}\Phi(-B(1))\right\} \\
=& \int_{-\infty}^{0}\Phi(-x)\phi(x)\,dx.
\end{aligned}
$$

如果过程从 x 开始，$B(0)=x$，则 $B(t)\sim \mathrm{N}(x,t)$，于是

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

这里，概率 P_x 的下标 x 表示过程始于 x，或 $B(0)=x$ 条件下的条件概率. 积分号中的函数

$$
p_t(x,y)=\frac{1}{\sqrt{2\pi t}}e^{-\frac{(y-x)^2}{2t}} \tag{7.4}
$$

称为布朗运动的**转移概率密度**. 实际上，布朗运动是连续时间、连续状态的马氏过程。利用独立增量性以及转移概率密度，我们可以计算任意布朗运动的有限维分布：

命题 7.2. 对布朗运动 $B^x(t)$, $t\geq 0$，以及 $0\leq t_1<t_2<\cdots<t_n$，有如下联合分布函数

$$
\begin{aligned}
& P_x\left(B(t_1)\leq x_1,\dots,B(t_n)\leq x_n\right) && (7.5)\\
=& \int_{-\infty}^{x_1}p_{t_1}(x,y_1)\,dy_1\int_{-\infty}^{x_2}p_{t_2-t_1}(y_1,y_2)\,dy_2\cdots\int_{-\infty}^{x_n}p_{t_n-t_{n-1}}(y_{n-1},y_n)\,dy_n, && (7.6)
\end{aligned}
$$

即 $[B(t_1),B(t_2),\dots,B(t_n)]$ 的联合密度为

$$
p_{t_1,t_2,\dots,t_n}(y_1,y_2,\dots,y_n)=p_{t_1}(x,y_1)p_{t_2-t_1}(y_1,y_2)\dots p_{t_n-t_{n-1}}(y_{n-1},y_n). \tag{7.7}
$$

证明：在证明中省略 B^x 中的记号 x。只要证明(7.7)。由布朗运动的平稳独立增量性以及正态分布的线性组合服从多元正态分布，可知 $[B(t_1), B(t_2), \dots, B(t_n)]$ 服从联合正态分布。于是其联合密度为

$$\begin{aligned}&p_{t_1,t_2,\dots,t_n}(y_1, y_2, \dots, y_n)\\=&p_{t_1}(y_1)p_{t_2|t_1}(y_2|y_1)\dots p_{t_n|t_1,\dots,t_{n-1}}(y_n|y_1, \dots, y_{n-1}),\end{aligned}$$

其中 $p_{t_n|t_1,\dots,t_{n-1}}(y_n|y_1, \dots, y_{n-1})$ 表示在 $[B(t_1), \dots, B(t_{n-1})] = (y_1, \dots, y_{n-1})$ 条件下 $B(t_n)$ 的条件密度。注意到

$$B(t_n) = [B(t_n) - B(t_{n-1})] + B(t_{n-1}),$$

$[B(t_n) - B(t_{n-1})]$ 与 $B(t_{n-1})$ 独立且 $[B(t_n) - B(t_{n-1})] \sim \mathrm{N}(0, t_n - t_{n-1})$，所以 $B(t_n)|B(t_1) = y_1, \dots, B(t_{n-1}) = y_{n-1}$ 的分布等于 $B(t_n)|B(t_{n-1}) = y_{n-1}$ 的分布，这个条件分布为 $\mathrm{N}(y_{n-1}, t_n - t_{n-1})$，密度为 $p_{t_n - t_{n-1}}(y_{n-1}, y_n)$。命题得证。

7.2.4 二次变差

为了讨论 Brown 运动的路径性质，我们回顾有关有界变差函数的概念，并引入二次变差定义。

闭区间上有界变差函数为两个增函数的差，是几乎处处可微的。

闭区间上连续可微函数必为有界变差，事实上，

$$\begin{aligned}&\sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|\\=&\sum_{i=0}^{n-1} |f'(t_i^*)|\,(t_{i+1} - t_i)\\\le&\max_{u\in[s,t]} |f'(u)|(t - s).\end{aligned}$$

因为增加分点必令 ν_Δ 单调不减，所以对闭区间连续可微函数 f 有

$$\bigvee_s^t(f) = \lim_{n\to\infty} \sum_{i=0}^{n-1} |f'(t_i^*)|\,(t_{i+1} - t_i) = \int_s^t |f'(u)|\,du.$$

其中设 $n \to \infty$ 时分割的最大间距 $\delta_n = \max_{0\le i\le n-1}(t_{i+1} - t_i) \to 0$。

定义 7.6 (二次变差)**.** 对随机过程 $\{X(t), t \geq 0\}$，设 $0 \leq s = t_0 < t_1 < \cdots < t_n = t$，称 $\{t_i\}$ 为区间 $[s,t]$ 的一个分割，记 $\delta_n = \max_i(t_{i+1} - t_i)$,

$$S_n = \sum_{i=0}^{n-1} |X(t_{i+1}) - X(t_i)|^2,$$

如果 $n \to \infty$ 时 $\delta_n \to 0$, 且 S_n 依概率收敛到 ξ，ξ 与分割 $\{t_i\}$ 取法无关，则称 ξ 为 $\{X(t)\}$ 在 $[s,t]$ 的二次变差，记为 $[X,X]([s,t])$。当 $s=0$ 时记为 $[X,X](t)$。

如果 $\{X(t)\}$ 的轨道连续可微，则二次变差等于 0。事实上，这时

$$\begin{aligned}
&\sum_{i=0}^{n-1} |X(t_{i+1}) - X(t_i)|^2 \\
=&\sum_{i=0}^{n-1} |X'(t_i^*)|^2 \, (t_{i+1} - t_i)^2 \\
\leq& \max_{u\in[s,t]} |X'(u)|^2 (t-s)\delta_n \to 0.
\end{aligned}$$

引理 7.1. 设 $\{B(t), t \geq 0\}$ 为标准布朗运动，$\{t_i\}$ 为 $[s,t]$ 的分割，$\delta_n = \max_i(t_{i+1} - t_i)$，则

$$E\left|\sum_{i=0}^{n-1}(B(t_{i+1}) - B(t_i))^2 - (t-s)\right|^2 \leq 2\delta_n(t-s).$$

证明记

$$\Delta t_i = t_{i+1} - t_i, \quad \Delta B(t_i) = B(t_{i+1}) - B(t_i),$$

则

$$\begin{aligned}
&E\left|\sum_{i=0}^{n-1}(B(t_{i+1}) - B(t_i))^2 - (t-s)\right|^2 \\
=&E\left|\sum_{i=0}^{n-1}\left[(\Delta B(t_i))^2 - \Delta t_i\right]\right|^2 \\
=&\sum_{i=0}^{n-1} E\left[(\Delta B(t_i))^2 - \Delta t_i\right]^2 \\
&+ \sum_{i\neq j} E\left\{\left[(\Delta B(t_i))^2 - \Delta t_i\right]\left[(\Delta B(t_j))^2 - \Delta t_j\right]\right\}
\end{aligned}$$

因为 $i \neq j$ 时 $\Delta B(t_i)$ 和 $\Delta B(t_j)$ 相互独立，而 $E[(\Delta B(t_i))^2] = \Delta t_i$，所以上式右侧的第二项为零。利用 $\Delta B(t_i) \sim \mathrm{N}(0, \Delta t_i)$ 以及正态分布的四阶矩，可得

$$\begin{aligned}
&E\left|\sum_{i=0}^{n-1}(B(t_{i+1}) - B(t_i))^2 - (t-s)\right|^2 \\
=&\sum_{i=0}^{n-1}\left[E[(\Delta B(t_i))^4] + (\Delta t_i)^2 - 2(\Delta t_i)E[(\Delta B(t_i))^2]\right] \\
=&\sum_{i=0}^{n-1}\left[3(\Delta t_i)^2 + (\Delta t_i)^2 - 2(\Delta t_i)^2\right] \\
=&2\sum_{i=0}^{n-1}(\Delta t_i)^2 \leq 2\delta_n \sum_{i=0}^{n-1}(\Delta t_i) \\
=&2\delta_n(t-s).
\end{aligned}$$

定理 7.6. 设 $\{B(t), t \geq 0\}$ 为标准布朗运动，$\{t_i\}$ 是 $[s,t]$ 的分割且 $\delta_n = \max_i(t_{i+1} - t_i) \to 0$，则

$$E\left|\sum_{i=0}^{n-1}(B(t_{i+1}) - B(t_i))^2 - (t-s)\right|^2 \to 0,\ n \to \infty,$$

从而 $[B,B]([s,t]) = t - s$, $[B,B](t) = t$。

证明由引理7.1即得极限为 0，从而 $\sum_{i=0}^{n-1}(B(t_{i+1}) - B(t_i))^2$ 均方收敛到 $t-s$，从而依概率收敛到 $t-s$，于是得定理结论。

引理 7.2. 设 $\{B(t), t \geq 0\}$ 为标准布朗运动，$0 \leq s < t$, 取 $s = t_0 < t_1 < \cdots < t_{2^n} = t$ 为 $[s,t]$ 的 2^n 等分点，则

$$\sum_{i=0}^{2^n-1}(\Delta B(t_i))^2 \to t-s,\ a.s.,\ (n \to \infty).$$

证明记

$$S_n = \sum_{i=0}^{2^n-1}(\Delta B(t_i))^2,$$

则

$$ES_n = \sum_{i=0}^{2^n-1}(t_{i+1} - t_i) = t - s,$$

由引理7.1，

$$E|S_n-(t-s)|^2 \leq 2\delta_n(t-s) = 2(t-s)^2 2^{-n},$$

于是

$$\sum_{n=1}^{\infty} E|S_n-(t-s)|^2 < \infty,$$

由单调收敛定理有

$$E\sum_{n=1}^{\infty}|S_n-(t-s)|^2 = \sum_{n=1}^{\infty}E|S_n-(t-s)|^2 < \infty,$$

所以

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

$$|S_n-(t-s)|^2 \to 0,\ \text{a.s.}$$

引理得证。

引理 7.3. 设 $\{B(t), t \geq 0\}$ 为标准布朗运动，$0 \leq s < t$, 取 $s = t_0 < t_1 < \cdots < t_{2^n} = t$ 为 $[s,t]$ 的 2^n 等分点，则

$$\sum_{i=0}^{2^n-1} |\Delta B(t_i)| \to \infty,\ a.s.,\ (n \to \infty).$$

证明记

$$M_n(\omega) = \max_i |B(t_{i+1},\omega) - B(t_i,\omega)|,$$

当 $n \to \infty$ 时 $t_{i+1} - t_i = (t-s)2^{-n}$，由 $B(t,\omega)$ 对固定的 ω 关于 t 连续以及闭区间连续函数一致连续，可知 $M_n(\omega) \to 0$, $n \to \infty$。而

$$\sum_{i=0}^{2^n-1} |\Delta B(t_i)|^2 \leq M_n(\omega) \sum_{i=0}^{2^n-1} |\Delta B(t_i)|,$$

由引理7.2知 $\sum_{i=0}^{2^n-1} |\Delta B(t_i)|^2 \to t-s > 0$, a.s.，于是

$$\sum_{i=0}^{2^n-1} |\Delta B(t_i)| \geq \frac{\sum_{i=0}^{2^n-1} |\Delta B(t_i)|^2}{M_n(\omega)} \to \infty,\ n \to \infty.$$

推论 7.2. 布朗运动的轨道都不是有界变差的。

布朗运动的二次变差会出现在随机积分的 Ito 公式中。一般的 Stiejes 积分并不会出现二次变差。

下面是布朗运动的路径性质. 从时刻 0 到时刻 T 对布朗运动的一次观察称为布朗运动在区间 $[0,T]$ 上的一个**路径**（轨道）或一个**实现**. 布朗运动的几乎所有样本路径 $B(t), 0 \le t \le T$ 都具有下述性质：

(1) 是 t 的连续函数；

(2) 在任意区间 (无论区间多么小) 上都不是单调的;

(3) 在任意点都不是可微的；

(4) 在任意区间 (无论区间多么小) 上都不是有界变差的;

(5) 对任意 $0 \le s < t$，在 $[s,t]$ 上的二次变差等于 $t-s$.

关于性质 1 的直观解释：

$$\lim_{\Delta t \to 0} E|X(t+\Delta t) - X(t)|^2 = \lim_{\Delta t \to 0} \Delta t = 0.$$

关于性质 2 的直观解释：将 $[t, t+\Delta t]$ 等分为 n 分，则每一份上的增量独立同 $\mathrm{N}(0, \frac{\Delta t}{n})$ 分布，这 n 个增量同为正号或者同为负号的概率为

$$2 \times (\frac{1}{2})^n = 2^{-(n-1)} \to 0,$$

所以不可能单调。

性质 3 的直观解释：

$$E\left|\frac{X(t+\Delta t) - X(t)}{\Delta t}\right|^2 = \frac{\Delta t}{\Delta t^2} \to \infty,$$
$$E\left|\frac{X(t+\Delta t) - X(t)}{\Delta t}\right| = \frac{\sqrt{\Delta t} E|Z|}{\Delta t} \to \infty,$$

(其中 Z 表示标准正态分布随机变量) 这提示 $\lim_{\Delta t \to 0} \frac{X(t+\Delta t) - X(t)}{\Delta t}$ 不存在。

性质 4 由推论7.2给出。

性质 5 由定理7.6给出。

7.2.5 高斯过程性质

定理 7.7. 布朗运动是均值函数为 $m(t)=0$，协方差函数为 $\gamma(s,t)=\min(t,s)$ 的高斯过程.

证明: 由定义，对 $0\le t_1<t_2<\cdots<t_n$，可知 $B(t_1), B(t_2)-B(t_1),\dots,B(t_n)-B(t_{n-1})$ 相互独立，都服从正态分布。而 $(B(t_1),B(t_2),\dots,B(t_n))$ 是上述独立的正态分布的线性组合，也服从多维高斯分布，所以布朗运动是高斯过程。

由于布朗运动的均值是 0, 所以其协方差函数为

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

若 $t<s$，则 $B(s)=B(t)+[B(s)-B(t)]$，由独立增量性可得

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

类似地，若 $t>s$，也有 $E[B(t)B(s)]=s$，即 $E[B(t)B(s)]=\min(t,s)$。

对正整数 $n\ge 2$，以及 $0\le t_1<t_2<\cdots<t_n$，$(B(t_1),B(t_2),\dots,B(t_n))$ 服从多元高斯分布，分布由均值和协方差阵决定，期望值等于 0，有

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

从而 $\{B(t),t\ge 0\}$ 是高斯过程。

例 7.3. 设 $\{B(t)\}$ 是布朗运动，求 $B(1)+B(2)+B(3)+B(4)$ 的分布.

解：考虑随机向量 $X=(B(1),B(2),B(3),B(4))^T$，由定理7.7可知 X 服从多元正态分布，且具有零均值和协方差矩阵

$$\Sigma=\begin{pmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{pmatrix}.$$

令 $A=(1,1,1,1)$，则

$$AX=B(1)+B(2)+B(3)+B(4),$$

具有均值为 0，方差为 $A\Sigma A^T=30$ 的正态分布. 于是 $B(1)+B(2)+B(3)+B(4)\sim \mathrm{N}(0,30)$.

例 7.4. 求 $B(\frac{1}{4})+B(\frac{1}{2})+B(\frac{3}{4})+B(1)$ 的分布.

解：考虑随机向量 $Y=(B(\frac{1}{4}),B(\frac{1}{2}),B(\frac{3}{4}),B(1))$. 易见，$Y$ 与上例中的 X 具有相同的情形. 所以它的协方差矩阵为 $\frac{1}{4}\Sigma$. 因此 $AY=B(\frac{1}{4})+B(\frac{1}{2})+B(\frac{3}{4})+B(1)\sim \mathrm{N}(0,\frac{30}{4})$.

例 7.5. 求概率 $P\{\int_0^1 B(t)dt>\frac{2}{\sqrt{3}}\}$.

解：首先需要指出的是，布朗运动具有连续路径，所以对每个路径来说, 黎曼（Riemann）积分 $\int_0^1 B(t)dt$ 存在. 我们只需找出 $\int_0^1 B(t)dt$ 的分布. 由黎曼积分的定义，我们可以从逼近和

$$\sum_{i=1}^n B(t_i)\Delta t_i$$

的极限分布得到. 这里 $0=t_0<t_1<\cdots<t_n=1$ 是 $[0,1]$ 的分点，$\Delta t_i=t_i-t_{i-1}$，例如取 $t_i=\frac{i}{n}$. 当 $n=4$ 时，逼近和即为例7.4中的随机变量. 由于正态分布的线性组合仍为正态分布，所以逼近和的分布都是零均值的正态分布。方差为

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{i=1}^n B(t_i)\Delta t_i\right)=\sum_{i=1}^n\sum_{j=1}^n \mathrm{Cov}(B(t_i),B(t_j))\Delta t_i\Delta t_j\\
=&\sum_{i=1}^n\sum_{j=1}^n \min(t_i,t_j)\Delta t_i\Delta t_j\\
\to&\int_0^1\int_0^1 \min(s,t)\,ds\,dt\\
=&\frac{1}{3},
\end{aligned}$$

由定理7.4可知逼近和的极限 $\int_0^1 B(t)\,dt$ 仍服从正态分布 $\mathrm{N}(0,\frac{1}{3})$。于是所求概

率为

$$P\left\{\int_0^1 B(t)dt > \frac{2}{\sqrt{3}}\right\} = P\left\{\sqrt{3}\int_0^1 B(t)dt > 2\right\} = 1 - \Phi(2).$$

这里 $\Phi(x)$ 是标准正态分布的分布函数.

也可以形式地计算

$$\begin{aligned} E\int_0^1 B(t)\,dt = &\int_0^1 E[B(t)]\,dt = 0,\\ \mathrm{Var}\left(\int_0^1 B(t)\,dt\right) =&E\left(\int_0^1 B(t)\,dt\cdot\int_0^1 B(s)\,ds\right)\\ =&\int_0^1\int_0^1 E[B(t)B(s)]\,ds\,dt\\ =&\int_0^1\int_0^1 \min(t,s)\,ds\,dt\\ =&\frac{1}{3}.\end{aligned}$$

例7.5可以推广为如下的一般结果。

引理 7.4. 设 $\{B(t), t\geq 0\}$ 为标准布朗运动，$g(x)$ 为 $\mathbb{R}$ 上的连续函数，对 $0\leq s<t$，设 $\int_s^t E[|g(B(u))|]\,du<\infty$，则 $\int_s^t g(B(u))\,du$ 存在且

$$E\int_s^t g(B(u))\,du = \int_s^t E[g(B(u))]\,du.$$

证明：因为 $B(t)$ 轨道连续，$g(\cdot)$ 连续故 $g(B(t))$ 轨道连续，$\int_s^t g(B(u))\,du$ 有定义，且 $g(B(t,\omega))$ 是 $([0,\infty)\times\Omega, \mathscr{B}([0,\infty))\times\mathscr{F})$ 到 $\mathscr{B}(\mathbb{R})$ 的可测变换。

因为

$$\int_s^t E[|g(B(u))|]\,du = \int_s^t\int_\Omega |g(B(u,\omega))|\,dP(\omega)\,du < \infty,$$

由富比尼定理即知

$$
\begin{aligned}
& E\int_s^t g(B(u))\,du = \int_\Omega \int_s^t g(B(u,\omega))\,du\,dP(\omega) \\
= & \int_s^t \int_\Omega g(B(u,\omega))dP(\omega)\,du \\
= & \int_s^t E[g(B(u))]\,du.
\end{aligned}
$$

引理 7.5. 设 $\{B(t), t\geq 0\}$ 为标准布朗运动，$g(x,y)$ 为 $\mathbb{R}^2$ 上的连续函数，对 $0\leq s_1<t_1$，$0\leq s_2<t_2$，设

$$\int_{s_1}^{t_1}\int_{s_2}^{t_2} E[|g(B(u),B(v))|]\,dv\,du<\infty,$$

则 $\int_{s_1}^{t_1}\int_{s_2}^{t_2} g(B(u),B(v))\,dv\,du$ 存在且

$$E\int_{s_1}^{t_1}\int_{s_2}^{t_2} g(B(u),B(v))\,dv\,du = \int_{s_1}^{t_1}\int_{s_2}^{t_2} E[g(B(u),B(v))]\,dv\,du.$$

证明仍使用 Fubini 定理。

例 7.6. 对例7.5，用引理7.4和引理7.5求解。

解答： 这时

$$\int_0^1 E|B(u)|\,du = \int_0^1 \sqrt{\frac{2}{\pi}}\sqrt{u}\,du = \sqrt{\frac{4}{3\pi}}<\infty,$$

由引理7.4可知

$$E\int_0^1 B(u)\,du = \int_0^1 E[B(u)]\,du = 0.$$

另外，对 $t<s$，

$$
\begin{aligned}
E|B(t)B(s)| =& E|B(t)[B(t)+(B(s)-B(t))]| \\
=& E[B^2(t)] + E|B(t)|\cdot E|B(s)-B(t)| \\
=& t+\frac{2}{\pi}\sqrt{t(s-t)} \\
\leq& t+s,
\end{aligned}
$$

$t \leq s$ 时不等式仍成立，所以

$$\int_0^1 \int_0^1 E|B(t)B(s)|\,ds\,dt \leq \int_0^1 \int_0^1 (t+s)\,ds\,dt = 1,$$

由引理7.5可知

$$\begin{aligned}
&\mathrm{Var}(\int_0^1 B(u)\,du) \\
=&E\left[\int_0^1 B(u)\,du\right]^2 = E\left[\int_0^1 B(u)\,du \int_0^1 B(v)\,dv\right] \\
=&E\int_0^1 \int_0^1 B(u)B(v)\,dv\,du \\
=&\int_0^1 \int_0^1 E[B(u)B(v)]\,dv\,du \\
=&\int_0^1 \int_0^1 \min(u,v)\,dv\,du \\
=&\frac{1}{3}.
\end{aligned}$$

7.3 布朗运动的鞅性质

本节讨论与 Brown 运动相联系的几个鞅，首先回忆连续鞅的定义.

定义 7.7. 随机过程 $\{X(t), t \geq 0\}$ 称为鞅，如果 $\forall t$, $E[|X(t)|] < \infty$, 且 $\forall s > 0$, 有

$$E[X(t+s)|\mathscr{F}_t] = X(t), \text{ a.s.} \tag{7.8}$$

这里 $\mathscr{F}_t = \sigma\{X(u), 0 \leq u \leq t\}$ 是由 $\{X(u), 0 \leq u \leq t\}$ 生成的 σ 代数，其中的等式(7.8)是几乎必然成立的，在后面有关的证明中，有时也省略 a.s.

更一般的，可以设 $\{X(t), \mathscr{F}(t), t \geq 0\}$ 是适应过程且满足上面的条件。

定理 7.8. 设 $\{B(t)\}$ 是布朗运动，则

(1) $\{B(t)\}$ 是鞅;

(2) $\{B^2(t) - t\}$ 是鞅;

(3) 对任意实数 u,$e^{-\frac{u^2}{2}t}e^{uB(t)}$ 是鞅.

证明：由 $B(t+s)-B(t)$ 与 $\mathscr{F}_t$ 的独立性可知对任意函数 $g(x)$，有

$$E[g(B(t+s)-B(t))|\mathscr{F}_t]=E[g(B(t+s)-B(t))] \tag{7.9}$$

(1) 由布朗运动的定义, $B(t)\sim N(0,t)$, 所以 $B(t)$ 一阶矩有限, 且 $E[B(t)]=0$,

$$\begin{aligned}E[B(t+s)|\mathscr{F}_t]=&E[B(t)+(B(t+s)-B(t))|\mathscr{F}_t]\\=&E[B(t)|\mathscr{F}_t]+E[B(t+s)-B(t)|\mathscr{F}_t]\\=&B(t)+E[B(t+s)-B(t)]=B(t).\end{aligned}$$

(2) 由于 $E[B^2(t)]=t<\infty$，所以 $B^2(t)$ 一阶矩有限，由于

$$\begin{aligned}B^2(t+s)=&[B(t)+(B(t+s)-B(t))]^2\\=&B^2(t)+2B(t)[B(t+s)-B(t)]+[B(t+s)-B(t)]^2,\end{aligned}$$

从而

$$\begin{aligned}&E[B^2(t+s)|\mathscr{F}_t]\\=&B^2(t)+2E[B(t)(B(t+s)-B(t))|\mathscr{F}_t]+E[(B(t+s)-B(t))^2|\mathscr{F}_t]\\=&B^2(t)+2B(t)E[(B(t+s)-B(t))|\mathscr{F}_t]+E[(B(t+s)-B(t))^2]\\=&B^2(t)+0+s,\end{aligned}$$

这里我们利用了 $B(t+s)-B(t)$ 与 $\mathscr{F}_t$ 的独立性且具有均值 0，并对 $g(x)=x^2$ 应用式(7.9). 在上式两端同时减去 $(t+s)$，则 (2) 得证.

(3) 考虑 $B(t)\sim \mathrm{N}(0,t)$ 的矩母函数 $E[e^{uB(t)}]=e^{tu^2/2}<\infty$，这蕴含着 $e^{uB(t)}$ 是一阶矩有限的，并且

$$E\left[e^{-\frac{u^2}{2}t}e^{uB(t)}\right]=1.$$

取 $g(x)=e^{ux}$，利用式(7.9)，可得

$$\begin{aligned}&E\left[e^{uB(t+s)}|\mathscr{F}_t\right]\\=&E\left[\exp\{u\cdot B(t)+u\cdot(B(t+s)-B(t))\}|\mathscr{F}_t\right]\\=&e^{u\cdot B(t)}E\left[\exp\{u\cdot(B(t+s)-B(t))\}|\mathscr{F}_t\right]\\=&e^{u\cdot B(t)}E\left[\exp\{u\cdot(B(t+s)-B(t))\}\right]\\=&e^{uB(t)}e^{\frac{u^2}{2}s},\end{aligned}$$

两端同时乘以 $e^{-\frac{u^2}{2}(t+s)}$, 则 (3) 得证.

注 1: 上述定理所给的这 3 个鞅在理论上也有着十分重要的意义，比如鞅 $\{B^2(t)-t\}$ 就是布朗运动的特征，即，如果连续鞅 $\{X(t)\}$ 使得 $\{X^2(t)-t\}$ 也是鞅，则 $\{X(t)\}$ 是布朗运动.

注 2：结论可以推广到 $\{B(t),\mathscr{F}(t),t\geq 0\}$ 是布朗运动的情形。即 $\mathscr{F}(t)\supset\sigma(\{B(u):0\leq u\leq t\})$ 的情形。

7.4 布朗运动的马氏性

所谓马氏性是指在知道过程的现在与过去的状态的条件下，过程将来的表现（分布）与过去无关. 换言之，过程的将来依赖于现在，但是并不记忆现在的状态是如何得到的，即“无记忆性”. 在第 5 章中我们介绍了马氏链和连续时间离散状态的马氏过程, 而在这里我们所讨论的布朗运动是连续时间连续状态过程，为此我们从连续马氏过程的定义开始.

定义 7.8. 设 $\{X(t),t\geq 0\}$ 是一个连续时间的随机过程, 如果 $\forall t,s>0$, 有

$$P\{X(t+s)\leq y|\mathscr{F}_t\}=P\{X(t+s)\leq y|X(t)\},\ \text{a.s.} \tag{7.10}$$

则称 $\{X(t)\}$ 为马氏过程, 这里 $\mathscr{F}_t=\sigma\{X(u),0\leq u\leq t\}$. 性质(7.10)称为马氏性.

这个定义是利用将来的条件分布，另一种定义是利用将来的状态的函数的条件期望，参见5.17。

定理 7.9. 布朗运动 $\{B(t)\}$ 具有马氏性.

证明：注意到 $B(t+s)-B(t)$ 与 $\mathscr{F}_t$ 独立，用矩母函数方法可以证明 $B(t+s)$ 在给定条件 $\mathscr{F}_t$ 下的分布与在给定条件 $B(t)$ 下的分布是一致的. 事实上，由定理7.8第 (3) 条，$e^{-\frac{u^2}{2}t}e^{uB(t)}$ 是鞅，可得

$$\begin{aligned}
&E\left[e^{uB(t+s)}|\mathscr{F}_t\right]\\
=&e^{\frac{u^2}{2}(t+s)}E\left[e^{-\frac{u^2}{2}(t+s)}e^{uB(t+s)}|\mathscr{F}_t\right]\\
=&e^{\frac{u^2}{2}(t+s)}e^{-\frac{u^2}{2}t}e^{uB(t)}\\
=&e^{\frac{u^2}{2}s}e^{uB(t)}.
\end{aligned}$$

而

$$\begin{aligned}
&E[e^{uB(t+s)}|B(t)]\\
=&E\left[e^{uB(t)}e^{u(B(t+s)-B(t))}|B(t)\right]\\
=&e^{uB(t)}E\left[e^{u(B(t+s)-B(t))}\right]\\
=&e^{uB(t)}e^{\frac{u^2}{2}s}\\
=&E\left[e^{uB(t+s)}|\mathscr{F}_t\right],
\end{aligned}$$

所以 $B(t+s)$ 在给定条件 $\mathscr{F}_t$ 下的分布与在给定条件 $B(t)$ 下的分布是一致的, 即 $\{B(t)\}$ 具有马氏性.

实际上,马氏性也蕴含了:对 $0 \le t < t_1 < t_2 < \cdots < t_n$,$(X(t_1), X(t_2), \dots, X(t_n))$ 在 $\mathscr{F}_t$ 下联合条件分布，等于在 $X(t)$ 条件下的联合条件分布；进一步地，$X(t+s), s \ge 0$ 在 $\mathscr{F}_t$ 下的分布（用函数空间上取值的随机元表示），等于在 $X(t)$ 条件下的分布。参见 (Karatzas and Shreve 1998)§2.5 命题 2.5.15。

连续状态的马氏过程 $\{X(t)\}$ 的转移概率分布定义为在时刻 s 过程处于状态 x 的条件下，过程在时刻 t 的分布函数

$$F(y,t,x,s) = P\{X(t) \le y|X(s) = x\},$$

在布朗运动的情况下，这一分布函数是正态的

$$F(y,t,x,s) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{(u-x)^2}{2(t-s)}\right\} du.$$

布朗运动的转移概率函数满足方程 $F(s,t;x,y) = F(0,t-s,x,y)$，这说明布朗运动是时齐的马氏过程。这等价于

$$P\{B(t) \le y|B(s) = x\} = P\{B(t-s) \le y|B(0) = x\} \tag{7.11}$$

当 $s=0$ 时，$F(0,t;x,y)$ 具有密度函数

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

即布朗运动的转移概率密度。

公式(7.11)给出的性质称为布朗运动的**时齐性**，即条件分布不随时间的平移而变化. 于是由(7.6)可知布朗运动的所有有限维条件分布都是时间平移不变的，但这是条件分布而不是分布，布朗运动并不是严平稳过程.

下面讨论布朗运动的强马氏性，为此给出关于 $B(t)$ 停时的定义.

定义 7.9. 如果非负随机变量 T 可以取无穷值，既 $T:\Omega\to[0,\infty]$，并且 $\forall t$，有 $\{T\le t\}\in\mathscr{F}_t=\sigma\{B(u),0\le u\le t\}$，则称 T 为关于 $\{B(t),t\ge 0\}$ 的**停时**.

所谓强马氏性，实际上是将马氏性中固定的时间 t 用停时 T 来代替. 下面我们不加证明的给出关于布朗运动的强马氏性定理.

定理 7.10. 设 T 是关于布朗运动 $\{B(t)\}$ 的有限停时，

$$\mathscr{F}_T=\{A\in\mathscr{F}:A\cap\{T\le t\}\in\mathscr{F}_t,\forall t\ge 0\},$$

则

$$P\{B(T+t)\le y|\mathscr{F}_T\}=P\{B(T+t)\le y|B(T)\},\ a.s.$$

即布朗运动 $\{B(t)\}$ 具有强马氏性. 进一步地，

$$\hat{B}(t)=B(T+t)-B(T),\ t\ge 0 \tag{7.12}$$

是始于 0 的布朗运动并且独立于 $\mathscr{F}_T$.

参见 (Karatzas and Shreve 1998) 节 2.6 定理 2.6.16.

注：并非所有马氏过程都有强马氏性。

7.5 布朗运动的最大值变量及反正弦律

7.5.1 首达时

考虑从 0 出发的标准布朗运动 $\{B(t),t\ge 0\}$，记 T_x 为到达 x 的首达时，即

$$T_x=\inf\{t>0:B(t)=x\}.$$

因为布朗运动的轨道连续，所以对 $x\ne 0$，$0<T_x\le+\infty$。

当 $x>0$ 时，为计算 $P\{T_x\le t\}$, 我们考虑 $P\{B(t)\ge x\}$. 由全概率公式

$$\begin{aligned}
&P\{B(t)\ge x\}\\
=&P\{B(t)\ge x|T_x\le t\}P\{T_x\le t\}\\
&+P\{B(t)\ge x|T_x>t\}P\{T_x>t\},
\end{aligned}$$

右侧的第二项 $T_x > t$ 表示截止到 t 时刻未达到 $x > 0$ 处，因为 $B(0) = 0$ 以及轨道连续性，这时必有 $B(t) < x$ 成立，即 $P\{B(t) \geq x | T_x > t\} = 0$，所以

$$P\{B(t) \geq x\} = P\{B(t) \geq x | T_x \leq t\} P\{T_x \leq t\}.$$

若 $T_x \leq t$，则 $B(t)$ 在 $[0, t]$ 中的某个时刻击中 x，$B(T_x) = x$, 由布朗运动的强马氏性，可以认为布朗运动从 $B(T_x) = x$ 重新开始，上行与下行概率各 $\frac{1}{2}$，即于是 $P(B(t) \geq x) = P(B(t) \leq x) = \frac{1}{2}$，

$$P\{B(t) \geq x\} = \frac{1}{2} P\{T_x \leq t\}, \tag{7.13}$$

$$P(T_x \leq t) = 2P(B(t) \geq x) \tag{7.14}$$

$$= 2P(\sqrt{t} Z \geq x) \tag{7.15}$$

$$= 2\left[1 - \Phi(\frac{x}{\sqrt{t}})\right] \tag{7.16}$$

$$= 2\Phi(-\frac{x}{\sqrt{t}}) \tag{7.17}$$

$$= 2\int_{\frac{x}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\, dv, x > 0, t > 0. \tag{7.18}$$

对 $x < 0$，同理可得以上结果。于是 T_x 的分布函数为

$$\begin{aligned} P(T_x \leq t) &= 2[1 - \Phi(\frac{|x|}{\sqrt{t}})] \\ &= 2\int_{\frac{|x|}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\, dv,\ t > 0. \end{aligned}$$

由此可见

$$P\{T_x < \infty\} = \lim_{t\to\infty} P\{T_x \leq t\} = 1.$$

由于时刻 0 从点 x_0 出发的布朗运动等同于 $B(t) + x_0$，所以时刻 0 从点 x_0 出发的布朗运动首达 x 的时间等同于 T_{x-x_0}，也是以概率 1 有限的。

求导得 T_x 的密度函数为

$$\begin{aligned} f_{T_x}(t) &= |x| t^{-\frac{3}{2}} \phi(-\frac{|x|}{\sqrt{t}}) \\ &= \frac{|x|}{t\sqrt{2\pi t}} e^{-\frac{x^2}{2t}},\ t > 0. \end{aligned}$$

7.8.2.3给出了 T_x 的如下的 Laplace 变换公式 (类似于矩母函数)：

$$Ee^{-\alpha T_x} = e^{-|x|\sqrt{2\alpha}},\ \forall \alpha > 0.$$

下面证明 $ET_x = +\infty$。事实上，

$$\begin{aligned}
E(T_x) =& \int_0^\infty P(T_x > 0)\, dt \\
=& \int_0^\infty \left\{1 - 2\left[1 - \Phi\left(\frac{|x|}{\sqrt{t}}\right)\right]\right\} dt \\
=& \int_0^\infty \left\{2\Phi\left(\frac{|x|}{\sqrt{t}}\right) - 1\right\} dt \\
=& \int_0^\infty \left\{2\Phi\left(\frac{|x|}{\sqrt{t}}\right) - 2\Phi(0)\right\} dt \\
=& 2\int_0^\infty \int_0^{\frac{|x|}{\sqrt{t}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du\, dt \\
=& 2\int_0^\infty \int_0^{\frac{x^2}{u^2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, dt\, du \\
=& \frac{2x^2}{\sqrt{2\pi}} \int_0^\infty \frac{1}{u^2} e^{-\frac{u^2}{2}}\, du \\
\geq& \frac{2x^2 e^{-\frac{1}{2}}}{\sqrt{2\pi}} \int_0^1 \frac{1}{u^2}\, du = +\infty.
\end{aligned}$$

因此，T_x 虽然几乎必然是有限的，但有无穷的期望. 直观地看，就是布朗运动以概率 1 会击中 x，但它的平均时间是无穷的. 性质 $P\{T_x < \infty\} = 1$ 称为布朗运动的常返性. 由于始于点 a 的布朗运动与 $\{a + B(t)\}$ 是相同的，这里 $\{B(t)\}$ 是始于 0 的布朗运动，所以

$$P_a\{T_x < \infty\} = P_0\{T_{x-a} < \infty\} = 1,$$

即布朗运动从任意点出发，击中 x 的概率都是 1.

7.5.2 最大值和最小值的分布

另一个有趣的随机变量是布朗运动在 $[0, t]$ 中达到的最大值

$$M(t) = \max_{0 \leq u \leq t} B(u).$$

由轨道连续性，$M(t)$ 有定义且取有限值。对 $x > 0$，事件 $M(t) \geq x$ 等价于 $T_x \leq t$。事实上，当 $M(t) \geq x$，由轨道连续性以及 $B(0) = 0$，$[0, t]$ 中一定存在某个 t' 使得 $B(t') = x$，所以 $T_x \leq t$；另一方面，如果 $T_x \leq t$，则 $[0, t]$ 中的某个 $B(t') = x$，所以 $M(t) \geq x$。于是

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

令 $x \to 0$ 得 $P(M(t) \geq 0) = 1$。

求导得 $M(t)$ 的密度函数为

$$f_{M(t)}(x) = \frac{2}{\sqrt{t}}\phi(-\frac{x}{\sqrt{t}}) = \frac{2}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}},\ x > 0.$$

易见 $M(t)$ 与 $|B(t)|$ 同分布。

命题 7.3. $E[M^2(t)] = t.$

证明:

$$\begin{aligned} E[M^2(t)] =& \int_0^\infty \frac{2x^2}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}\,dt \\ =& \int_{-\infty}^\infty \frac{x^2}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}\,dt \\ =& t. \end{aligned}$$

最后的等式是对 $\mathrm{N}(0, t)$ 分布求方差。

考虑最小值的分布，类似可得

$$m(t) = \min_{0 \leq u \leq t} B(u),$$

的分布函数为

$$P(m(t) \leq x) = 2\Phi\left(-\frac{|x|}{\sqrt{t}}\right),\ x < 0.$$

可得 $P(m(t) \leq 0) = 1$，以及 $|m(t)| = -m(t)$ 与 $M(t)$ 同分布。

令

$$H(t) = \max_{0 \leq u \leq t} |B(u)|.$$

关于绝对值的最大值暂未得到简单的分布，但是可以有控制

$$H(t) \leq M(t) + |m(t)|,$$

且 $E[H^2(t)] \leq 4t$。

7.5.3 有零点的概率

如果时刻 τ 使得 $B(\tau)=0$，则称 τ 为布朗运动的零点. 我们有下述定理.

定理 7.11. 设 $\{B^x(t)\}$ 为始于 x 的布朗运动，则 $B^x(t)$ 在 $(0,t)$ 中至少有一个零点的概率为

$$\frac{|x|}{\sqrt{2\pi}}\int_0^t u^{-\frac{3}{2}}e^{-\frac{x^2}{2u}}\,du.$$

证明：如果 $x<0$，则由 $\{B^x(t)\}$ 的连续性得

$$P\{B^x\text{在}(0,t)\text{中至少有一个零点}\} = P\{\max_{0\leq s\leq t} B^x(s)\geq 0\},$$

因为 $B^x(t)=B(t)+x$，我们有

$$\begin{aligned}
&P\{B^x\text{在}(0,t)\text{中至少有一个零点}\}\\
=&P\{\max_{0\leq s\leq t} B^x(s)\geq 0\}\\
=&P\{\max_{0\leq s\leq t} B(s)+x\geq 0\} = P\{\max_{0\leq s\leq t} B(s)\geq -x\}\\
=&P\{T_{-x}\leq t\} = P\{T_x\leq t\}\\
=&\int_0^t f_{T_x}(u)\,du = \frac{|x|}{\sqrt{2\pi}}\int_0^t u^{-\frac{3}{2}}e^{-\frac{x^2}{2u}}\,du.
\end{aligned}$$

对于 $x>0$ 的情况的证明类似，结果相同。

利用定理7.11可以证明如下的定理。

定理 7.12. $B(t)$ 在 $t\in(a,b)$ 中至少有一个零点的概率为

$$\frac{2}{\pi}\arccos\sqrt{\frac{a}{b}}.$$

证明：记

$$h(x) = P\{B\text{在}(a,b)\text{中至少有一个零点} \mid B(a)=x\},$$

由对称性可知 $h(-x)=h(x)$。由马氏性，

$$\begin{aligned}
h(x) =&P\{B^x\text{在}(0,b-a)\text{中至少有一个零点}\}\\
=&\frac{|x|}{\sqrt{2\pi}}\int_0^{b-a} u^{-\frac{3}{2}}e^{-\frac{x^2}{2u}}\,du.
\end{aligned}$$

由全期望公式

$$
\begin{aligned}
&P\{B\text{在}(a,b)\text{中至少有一个零点}\}\\
=&\int_{-\infty}^{\infty} P\{B\text{中至少有一个零点} \mid B(a)=x\} f_{B(a)}(x)\,dx\\
=&\int_{-\infty}^{\infty} h(x)\frac{1}{\sqrt{2\pi a}} e^{-\frac{x^2}{2a}}\,dx\\
=&\frac{2}{\sqrt{2\pi a}}\int_0^{\infty} h(x) e^{-\frac{x^2}{2a}}\,dx\\
=&\frac{2}{\sqrt{2\pi a}}\int_0^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2a}} \int_0^{b-a} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}}\,du\,dx\\
=&\frac{1}{\pi\sqrt{a}}\int_0^{b-a} u^{-\frac{3}{2}} \int_0^{\infty} x e^{-x^2(\frac{1}{2u}+\frac{1}{2a})}\,dx\,du\\
=&\frac{1}{\pi\sqrt{a}}\int_0^{b-a} u^{-\frac{3}{2}} \frac{au}{a+u}\,du = \frac{\sqrt{a}}{\pi}\int_0^{b-a}\frac{u^{-\frac{1}{2}}}{a+u}\,du\\
=&\frac{2\sqrt{a}}{\pi}\int_0^{b-a}\frac{d\sqrt{u}}{a+u} = \frac{2\sqrt{a}}{\pi}\int_0^{\sqrt{b-a}}\frac{dv}{a+v^2}\\
=&\frac{2\sqrt{a}}{\pi}\int_0^{\frac{\sqrt{b-a}}{\sqrt{a}}}\frac{\sqrt{a}dw}{a(1+w^2)} = \frac{2}{\pi}\int_0^{\frac{\sqrt{b-a}}{\sqrt{a}}}\frac{dw}{1+w^2}\\
=&\frac{2}{\pi}\arctan\frac{\sqrt{b-a}}{\sqrt{a}} = \frac{2}{\pi}\arccos\sqrt{\frac{a}{b}}.
\end{aligned}
$$

这里利用了 $\int \frac{1}{1+w^2}dw = \arctan w$, 以及若 $y = \arctan\sqrt{\frac{b-a}{a}}$，则 $1+\tan y^2 = \frac{1}{\cos^2 y} = \frac{b}{a}$ 从而 $y = \arccos\sqrt{\frac{a}{b}}$。

于是，我们得到布朗运动的反正弦律.

定理 7.13. 设 $\{B(t), t\geq 0\}$ 是布朗运动，则

$$P\{B(t)\text{在}(a,b)\text{中没有零点}\} = \frac{2}{\pi}\arcsin\sqrt{\frac{a}{b}}.$$

下面介绍布朗运动在时刻 t 之前最后一个零点以及在 t 之后的第一个零点的分布情况. 令

$$
\begin{aligned}
\zeta_t &= \sup\{s\leq t, B(s)=0\} = t\text{之前最后一个零点},\\
\beta_t &= \inf\{s\geq t, B(s)=0\} = t\text{之后第一个零点},
\end{aligned}
$$

注意 β_t 是一个停时，而 ζ_t 不是停时 (请读者验证). 由反正弦律有

$$P\{\zeta_t \le x\} = P\{B在(x,t)中没有零点\} = \frac{2}{\pi}\arcsin\sqrt{\frac{x}{t}},$$

$$P\{\beta_t \ge y\} = P\{B在(t,y)中没有零点\} = \frac{2}{\pi}\arcsin\sqrt{\frac{t}{y}},$$

$$P\{\zeta_t \le x, \beta_t \ge y\} = P\{B在(x,y)中没有零点\} = \frac{2}{\pi}\arcsin\sqrt{\frac{x}{y}}.$$

由 ξ_t 的分布得

$$P(\zeta_t > 0) = 1 - \frac{2}{\pi}\arcsin\sqrt{\frac{0}{t}}, = 1,\ \forall t > 0.$$

这说明布朗运动 0 时刻从状态 0 出发，在任意的 $(0,t)$ 区间内都存在零点。这说明布朗运动一旦达到零点，会瞬时无穷次达到零点。这从布朗运动的随机游动逼近可以理解。

7.6 布朗运动的几种变化

7.6.1 Brown 桥

由布朗运动，我们可以定义另一类在数理金融中经常用到的过程——Brown 桥过程.

定义 7.10. 设 $\{B(t), t \ge 0\}$ 是布朗运动. 令

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

则称随机过程 $\{B^*(t), 0 \le t \le 1\}$ 为布朗桥.

因为布朗运动是高斯过程，所以布朗桥也是高斯过程，其 n 维分布由均值函数和方差函数完全确定，$\forall 0 \le s \le t \le 1$，有

$$\begin{aligned}
E[B^*(t)] =& 0,\\
E[B^*(s)B^*(t)] =& E[(B(s) - sB(1))(B(t) - tB(1))]\\
=& E[B(s)B(t) - tB(s)B(1) - sB(t)B(1) + tsB^2(1)]\\
=& s - ts - ts + ts = s(1-t)\\
=& \min(s,t)[1 - \max(s,t)].
\end{aligned}$$

此外，由定义可知 $B^*(0) = B^*(1) = 0$，即此过程的起始点是固定的，就象桥一样，这就是布朗桥名称的由来.

布朗桥也可以用来给出条件分布。设 $0 \le a < b < c$，则 $(B(a), B(c), B(b))$ 服从多元正态分布

$$N\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}a & a & a\\ a & c & b\\ a & b & b\end{pmatrix}\right).$$

在 $B(a) = x$, $B(c) = z$ 条件下，由联合正态分布性质可知 $B(b)$ 服从条件正态分布，条件均值为

$$\begin{aligned}&E(B(b)|(B(a), B(c)) = (x, z))\\ =&0 + (a, b)\begin{pmatrix}a & a\\ a & c\end{pmatrix}^{-1}(x, z)^T\\ =&\frac{c-b}{c-a}x + \frac{b-a}{c-a}z.\end{aligned}$$

条件方差为

$$\begin{aligned}&\mathrm{Var}(B(b)|(B(a), B(c)) = (x, z))\\ =&b - (a, b)\begin{pmatrix}a & a\\ a & c\end{pmatrix}^{-1}(a, b)^T\\ =&\frac{b(a+c-b) - ac}{c-a}.\end{aligned}$$

如果 $b = \frac{a+c}{2}$，则条件分布为 $\mathrm{N}(\frac{x+z}{2}, \frac{c-a}{4})$。比如取 $a = 0$, $b = \frac{1}{2}$, $c = 1$，则条件分布为 $\mathrm{N}(\frac{x+z}{2}, \frac{1}{4})$。

利用上述条件分布，可以用来生成 $[0, 1]$ 中的布朗运动的模拟轨道。先抽取 $B(0)$, $B(1)$，利用条件分布抽取 $B(\frac{1}{2})$，然后抽取 $B(\frac{1}{4})$ 和 $B(\frac{3}{4})$，然后抽取 $B(\frac{1}{8})$, $B(\frac{3}{8})$, $B(\frac{5}{8})$, $B(\frac{7}{8})$，如此重复可以得到抽样间隔越来越密的布朗运动轨道的离散化值。

7.6.2 有吸收值的布朗运动

设 T_x 为布朗运动 $\{B(t)\}$ 首次击中 x 的时刻, $x > 0$. 令

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

则 $\{Z(t), t \geq 0\}$ 是击中 x 后，永远停留在那里的布朗运动. $\forall t > 0$，随机变量 $Z(t)$ 的分布有离散和连续两个部分. 离散部分的分布是

$$\begin{aligned} P\{Z(t) = x\} =& P\{T_x \leq t\} \\ =& \frac{2}{\sqrt{2\pi t}} \int_x^{\infty} e^{-\frac{y^2}{2t}}\, dy. \end{aligned}$$

下面求连续部分的分布. $\forall y < x$,

$$P\{Z(t) \leq y\} = P\{B(t) \leq y, \max_{0 \leq s \leq t} B(s) \leq x\} \tag{7.19}$$

$$= P\{B(t) \leq y\} - P\{B(t) \leq y, \max_{0 \leq s \leq t} B(s) > x\} \tag{7.20}$$

现计算(7.20)式中的最后一项，由条件概率公式得

$$P\{B(t) \leq y, \max_{0 \leq s \leq t} B(s) > x\} \tag{7.21}$$

$$= P\{B(t) \leq y | \max_{0 \leq s \leq t} B(s) > x\} P\{\max_{0 \leq s \leq t} B(s) > x\} \tag{7.22}$$

由于事件 $\{\max_{0 \leq s \leq t} B(s) > x\}$ 等价于事件 $\{T_x < t\}$ 及布朗运动的对称性可知：$B(t)$ 在时刻 $T_x(< t)$ 击中 x，为了使在时刻 t 不大于 y，则在 T_x 之后的 $t - T_x$ 这段时间中就必须减少 $x - y$，而减少与增加 $x - y$ 的概率是相等的, 从 T_x 时刻的 $B(T_x) = x$ 增加 $x - y$ 就变成了 $B(t) \geq x + (x - y)$，所以有

$$P\{B(t) \leq y | \max_{0 \leq s \leq t} B(s) > x\} \tag{7.23}$$

$$= P\{B(t) \geq 2x - y | \max_{0 \leq s \leq t} B(s) > x\}. \tag{7.24}$$

从(7.22)和(7.24)式得

$$\begin{aligned} & P\{B(t) \leq y, \max_{0 \leq s \leq t} B(s) > x\} \\ =& P\{B(t) \geq 2x - y, \max_{0 \leq s \leq t} B(s) > x\} \\ =& P\{B(t) \geq 2x - y\} \quad (\text{因为} y < x). \end{aligned}$$

由式(7.20)，有

$$
\begin{aligned}
P\{Z(t)\leq y\} =& P\{B(t)\leq y\}-P\{B(t)\geq 2x-y\}\\
=& P\{B(t)\leq y\}-P\{B(t)\leq y-2x\}\\
=& \frac{1}{\sqrt{2\pi t}}\int_{y-2x}^{y} e^{-\frac{u^2}{2t}}\,du\\
=& \Phi(\frac{y}{\sqrt{t}})-\Phi(\frac{y-2x}{\sqrt{t}}),\ y<x.
\end{aligned}
$$

7.6.3 在原点反射的布朗运动

由 $Y(t)=|B(t)|, t\geq 0$ 定义的过程 $\{Y(t), t\geq 0\}$ 称为在原点反射的布朗运动. 它的概率分布为

$$
\begin{aligned}
P\{Y(t)\leq y\} =& P\{B(t)\leq y\}-P\{B(t)\leq -y\}\\
=& 2P\{B(t)\leq y\}-1\\
=& 2\Phi(\frac{y}{\sqrt{t}})-1\\
=& \frac{2}{\sqrt{2\pi t}}\int_{-\infty}^{y} e^{-\frac{u^2}{2t}}\,du-1,\ y>0.
\end{aligned}
$$

7.6.4 几何布朗运动

由 $X(t)=e^{B(t)}$, $t\geq 0$ 定义的过程 $\{X(t), t\geq 0\}$ 称为几何布朗运动. 由于 $B(t)$ 的矩母函数为 $E[e^{sB(t)}]=e^{ts^2/2}$，所以几何布朗运动的均值函数与方差函数分别为

$$
\begin{aligned}
E[X(t)] =& E[e^{B(t)}]=e^{t/2};\\
\mathrm{Var}[X(t)] =& E[X^2(t)]-(E[X(t)])^2\\
=& E[e^{2B(t)}]-e^{t}\\
=& e^{2t}-e^{t}.
\end{aligned}
$$

几何布朗运动的边缘分布为对数正态分布。令 $X\sim \mathrm{N}(\mu,\sigma^2)$，则 $Y=e^X$ 服从对数正态分布 $\mathrm{LN}(\mu,\sigma^2)$，其期望和方差为

$$
\begin{aligned}
E(Y) =& e^{\mu+\frac{1}{2}\sigma^2},\\
\mathrm{Var}(Y) =& e^{2\mu+\sigma^2}(e^{\sigma^2}-1).
\end{aligned}
$$

在金融市场中，人们经常假定股票的价格按照几何布朗运动变化，在下面的例子中我们如此假定.

例 7.7 (股票期权的价值)**.** 设某人拥有某种股票的交割时刻为 T、交割价格为 K 的欧式看涨期权，即他 (她) 具有在时刻 T 以固定的价格 K 购买一股这种股票的权力. 假设这种股票目前的价格为 y，并按照几何布朗运动变化，我们计算拥有这个期权的平均价值. 设 $X(T)$ 表示时刻 T 的股票价格，若 $X(T)$ 高于 K 时，期权将被实施，因此该期权在时刻 T 的平均价值应为

$$\begin{aligned}
E[\max(X(T)-K,0)] =& \int_0^\infty P\{X(T)-K>u\}\,du \\
=& \int_0^\infty P\{ye^{B(T)}-K>u\}\,du \\
=& \int_0^\infty P\{B(T)>\log\frac{K+u}{y}\}\,du \\
=& \frac{1}{\sqrt{2\pi T}}\int_0^\infty \int_{\log[(K+u)/y]}^\infty e^{-\frac{x^2}{2T}}\,dx\,du.
\end{aligned}$$

7.6.5 漂移布朗运动

设 $\{B(t),t\geq 0\}$ 为标准布朗运动，$\mu\in\mathbb{R}$, $\sigma>0$,

$$X(t) =X(0)+\mu t+\sigma B(t),\ t\geq 0,$$

称 $\{X(t),t\geq 0\}$ 为有漂移的布朗运动, 漂移系数为 μ，方差参数为 σ^2，记为 $\mathrm{BM}(\mu,\sigma^2)$。

等价地，若随机过程 $\{X(t),t\geq 0\}$ 为平稳独立增量过程，$X(t)-X(0)\sim \mathrm{N}(\mu t,\sigma^2 t)$，则称 $\{X(t),t\geq 0\}$ 为漂移布朗运动。

定理 7.14. *设随机过程 $\{X(t),t\geq 0\}$ 满足:*

(1) $\forall 0<t_1<\cdots<t_n$, $X(0),X(t_1)-X(0),\dots,X(t_n)-X(t_{n-1})$ 相互独立;

(2) $\forall s\geq 0$, $t\geq 0$，$X(t)-X(0)$ 与 $X(t+s)-X(s)$ 同分布（没有要求正态分布）;

(3) 对于每一轨道 $\omega\in\Omega$，$X(t,\omega)$ 是 $t\in[0,\infty)$ 的连续函数，

(4) $\forall t\geq 0$，$X(t)-X(0)$ 分布关于 0 对称。

则 $\{X(t)\}$ 是布朗运动（不要求从 0 出发），或者 $X(t)-X(0)$ 恒等于 0。如果仅要求 (1)-(3) 成立，则必存在标准布朗运动 $\{B(t)\}$ 使得

$$X(t)=X(0)+\mu t+\sigma B(t),\ t\geq 0.$$

即这时 $\{X(t)\}$ 为漂移布朗运动。

见 (刘勇 2022) 第四章定理 4.1 和推论 4.1。注意定理中条件 (1)-(4) 仅要求了独立平稳增量、轨道连续和 $X(t)-X(0)$ 分布关于 0 对称，没有涉及到分布正态，而结果的布朗运动和漂移布朗运动都是高斯过程。

7.7 高维布朗运动

定义 7.11. 设 $B^1(t), B^2(t), \cdots, B^d(t)$ 是相互独立的标准布朗运动，则称 $B(t) = (B^1(t), B^2(t), \dots, B^d(t))$ 为 d 维布朗运动。

定理 7.15. 设 $\{X(t), t \geq 0\}$ 为取值于 $\mathbb{R}^d$ 的随机过程，若：

(1) $\forall 0 < t_1 < \cdots < t_n$, $X(0), X(t_1)-X(0), \dots, X(t_n)-X(t_{n-1})$ 相互独立；

(2) $\forall s \geq 0$, $t \geq 0$，$X(t)-X(0)$ 与 $X(t+s)-X(s)$ 同分布（没有要求正态分布）；

(3) 对于每一轨道 $\omega \in \Omega$，$X(t,\omega)$ 是 $t \in [0,\infty)$ 的连续函数。

则存在 $r(r \leq d)$ 维标准布朗运动 $\{B(t)\}$，使得

$$X(t) = X(0) + \mu t + AB(t),\ t \geq 0.$$

即 $\{X(t)\}$ 是多维漂移布朗运动。其中 A 为 $d \times r$ 矩阵，μ 为 $n \times 1$ 常数向量。

见 (刘勇 2022) 第四章推论 4.2。

如同一维布朗运动，d 维布朗运动具有如下转移概率密度：

$$\frac{1}{(2\pi t)^{d/2}} \exp\{-\frac{1}{2t}\sum_{i=1}^{d}(y_i - x_i)^2\},\ x, y \in \mathbb{R}^d.$$

类似于一维布朗运动，对于 d 维布朗运动有如下定理：

定理 7.16. 设 $B(t) = (B^1(t), B^2(t), \dots, B^d(t))$ 为 d 维布朗运动，则有

(1) $\{B(t),\ t \geq 0\}$ 是鞅；

(2) $\{[B^1(t)]^2 + \cdots + [B^d(t)]^2 - td,\ t \geq 0\}$ 是鞅；

(3) 对任意 $u \in \mathbb{R}^d$，$\{\exp(u^T B(t) - \frac{1}{2}tu^T u), t \geq 0\}$ 是鞅，称为指数鞅（d 维布朗运动的拉式变换）.

7.8 补充内容

7.8.1 布朗运动的构造

是否有概率空间 $(\Omega, \mathscr{F}, P)$ 使得其中可以构造标准布朗运动 $\{B(t), t \geq 0\}$? (Karatzas and Shreve 1998) 中证明了三种构造方法。

1. 取 $\Omega = \mathbb{R}^{[0,\infty)}$，其中每个元素是一条从 $[0,\infty)$ 到 $\mathbb{R}$ 的映射的轨道, 表示为 $(\omega(t), t \geq 0)$。取用 $\mathbb{R}^{[0,\infty)}$ 的所有柱集构造的最小 σ 代数为 $\mathscr{F} = \mathscr{B}(\mathbb{R}^{[0,\infty)})$。这里柱集是指
$$C = \{\omega \in \mathbb{R}^{[0,\infty)} : (\omega(t_1), \dots, \omega(t_n)) \in B, B \in \mathscr{B}(\mathbb{R}^n)\}.$$
存在概率测度使得
$$X(t, \omega) = \omega(t)$$
具有标准布朗运动的有限维分布族。只要再满足轨道连续性。存在 $X(t)$ 的修改 $Y(t)$，使得 $Y(t)$ 轨道连续。在闭区间上在有理数处 $X(t)$ 和 $Y(t)$ 是重合的，由 $Y(t)$ 的连续性，$Y(t)$ 的任意有限维分布与 $X(t)$ 的相应分布相同，即 $\{Y(t)\}$ 为布朗运动。

2. 分别构造 $[0,1]$, $[1,2]$，……上的标准布朗运动进行累加。构造 $[0,1]$ 上的布朗运动时，利用布朗桥的思想进行插值。插值到极限时，可以证明插值结果的轨道在 $[0,1]$ 一致地逼近到一个连续轨道。插值步骤是：给定 $B(0)=0$, $B(1) \sim \mathrm{N}(0,1)$ 的值后，$B(\frac{1}{2})$ 条件分布正态，可以产生相应的随机变量，分布均值为 $B(0)$ 和 $B(1)$ 的平均值；给定 $B(0)$ 和 $B(\frac{1}{2})$，$B(\frac{1}{4})$ 条件分布正态，可以产生相应的随机变量；$B(\frac{3}{4})$ 在 $B(\frac{1}{2})$ 和 $B(1)$ 给定后可以产生满足条件分布的随机变量。如此可以不断进行加细一倍的插值，每一层插值看作对目标轨道的一个逼近。

3. 取 $\Omega = C[0,\infty)$，即 $[0,\infty)$ 上所有连续函数构成的线性空间。在其中定义距离
$$\rho(\omega^{(1)}, \omega^{(2)}) = \sum_{n=1}^{\infty} \frac{1}{2^n} \max_{t\in[0,n]} (|\omega^{(1)}(t) - \omega^{(2)}(t))| \wedge 1).$$
$(C[0,\infty), \rho)$ 构成了完备可分距离空间。对 $C[0,\infty)$，定义 $\mathscr{B}(C[0,\infty))$ 为包含 $C[0,\infty)$ 中所有开集的最小 σ 代数，等价于定义为包含 $C[0,\infty)$ 所有柱集的最小 σ 代数。轨道连续的随机过程 $\{X(t), t \geq 0\}$ 可以看成是

取值于 $(C[0,\infty),\mathscr{B}(C[0,\infty)))$ 的随机元。用随机游动进行时间和空间上的伸缩，当时间间隔趋于 0 时，由中心极限定理可得极限的有限维分布与布朗运动相同，这里用了胎紧 (tight) 条件。因为极限是在连续函数空间中，所以极限是轨道连续的。用来逼近的随机游动弱收敛到最终的布朗运动，构造了布朗运动的概率空间 $(C[0,\infty),\mathscr{B}(C[0,\infty)),P_*)$ 称为布朗运动的经典 (canonical) 概率空间，使得 $B(t,\omega)=\omega(t)$ 为布朗运动的概率测度 P_* 称为维纳 (Wiener) 测度。

7.8.2 带漂移的布朗运动的首达时

这是 Shrive(2004) 第二卷习题 3.7。

设 $\{B(t),t\geq 0\}$ 为标准布朗运动，$m>0$, $\mu\in\mathbb{R}$, $\sigma>0$,

$$\begin{aligned}
X(t) =&\mu t+\sigma B(t),\ t\geq 0,\\
\tau_m =&\inf\{t\geq 0: X(t)=m\},\\
Z(t) =&\exp\left\{\sigma X(t)-(\sigma\mu+\frac{1}{2}\sigma^2)t\right\},\ t\geq 0.
\end{aligned}$$

下面逐步讨论 τ_m 的性质。

7.8.2.1 鞅

来证明 $\{Z(t),t\geq 0\}$ 是鞅。

注意

$$Z(t)=\exp\left\{\sigma\mu t+\sigma B(t)-(\sigma\mu+\frac{1}{2}\sigma^2)t\right\}=\exp\left\{\sigma B(t)-\frac{1}{2}\sigma^2 t\right\}.$$

$Z(t)$ 服从对数正态分布所以一阶矩有限。令 $\mathscr{F}(t)=\sigma(\{B(s):0\leq s\leq t\})$。

对 $0\leq s<t$，

$$\begin{aligned}
E[Z(t)|\mathscr{F}(s)] =&E\left[\exp\left\{\sigma B(t)-\frac{1}{2}\sigma^2 t\right\}|\mathscr{F}(s)\right]\\
=&\exp\left\{-\frac{1}{2}\sigma^2 t+\sigma B(s)\right\}E\left[\exp\{\sigma(B(t)-B(s))\}\right]\\
=&\exp\left\{-\frac{1}{2}\sigma^2 t+\sigma B(s)\right\}\exp\{\frac{1}{2}\sigma^2(t-s)\}\\
=&\exp\left\{\sigma B(s)-\frac{1}{2}\sigma^2 s\right\}=X(s).
\end{aligned}$$

7.8.2.2 限时停时

来证明

$$E\left[\exp\left\{\sigma X(t\wedge\tau_m)-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}\right]=1.$$

因为 τ_m 是停时，所以 $t\wedge\tau_m$ 是有界停时，由停时定理可知 $EZ(t\wedge\tau_m)=EZ(0)=1$。

7.8.2.3 限时停时的极限

设 $\mu\ge 0$，来证明对 $\sigma>0$ 有

$$E\left[\exp\left\{\sigma m-(\sigma\mu+\frac{1}{2}\sigma^2)\tau_m\right\}I_{\{\tau_m<\infty\}}\right]=1.$$

由此可证明 $P(\tau_m<\infty)=1$，以及如下的 Laplace 变换公式：

$$Ee^{-\alpha\tau_m}=e^{m\mu-m\sqrt{2\alpha+\mu^2}},\ \forall\alpha>0.$$

当 $\tau_m<\infty$ 时，

$$\begin{aligned}\lim_{t\to+\infty}\exp\{\sigma X(t\wedge\tau_m)\}=&\exp\{\sigma X(\tau_m)\}=e^{\sigma m},\\ \lim_{t\to+\infty}\exp\{-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\}=&\exp\{-(\sigma\mu+\frac{1}{2}\sigma^2)\tau_m\}.\end{aligned}$$

当 $\tau=\infty$ 时，$X(t)<m$,

$$\begin{aligned}\exp\{\sigma X(t\wedge\tau_m)\}\le& e^{\sigma m},\\ \lim_{t\to+\infty}\exp\{-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\}=&\lim_{t\to+\infty}\exp\{-(\sigma\mu+\frac{1}{2}\sigma^2)t\}=0\end{aligned}$$

总有 $e^{\sigma X(t)}\le e^{\sigma m}$ 和 $\exp\{-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\le 1$。

于是

$$\exp\left\{\sigma X(t\wedge\tau_m)-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}\le e^{\sigma m},\ \forall t\ge 0,$$

且

$$\begin{aligned}&\lim_{t\to+\infty}\exp\left\{\sigma X(t\wedge\tau_m)-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}\\ =&\lim_{t\to+\infty}\exp\left\{\sigma X(t\wedge\tau_m)-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}I_{\{\tau_m<\infty\}}\\ &+\lim_{t\to+\infty}\exp\left\{\sigma X(t\wedge\tau_m)-(\sigma\mu+\frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}I_{\{\tau_m=\infty\}}\\ =&\exp\left\{\sigma m-(\sigma\mu+\frac{1}{2}\sigma^2)\tau_m\right\}I_{\{\tau_m<\infty\}}.\end{aligned}$$

由控制收敛定理可知

$$\begin{aligned}1 =& \lim_{t\to+\infty} E\left[\exp\left\{\sigma X(t\wedge\tau_m) - (\sigma\mu + \frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}\right] \\ =&E\left[\exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\} I_{\{\tau_m<\infty\}}\right].\end{aligned} \tag{*}$$

当 $0 < \sigma \le 1$ 时

$$\exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\} I_{\{\tau_m<\infty\}} \le e^{\sigma m} \le e^m,$$

由控制收敛定理，在 (*) 中令 $\sigma \to 0$，得

$$EI_{\{\tau_m<\infty\}} = 1,$$

即 $P(\tau_m < \infty) = 1$。

于是，

$$\begin{aligned}\exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\} I_{\{\tau_m<\infty\}} =& \exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\}, \text{ a.s.}, \\ E\exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\} =&1.\end{aligned}$$

令 $\alpha = \sigma\mu + \frac{1}{2}\sigma^2 > 0$，反解 $\sigma > 0$，得

$$\sigma = \sqrt{2\alpha + \mu^2} - \mu,$$

于是

$$Ee^{-\alpha\tau_m} =e^{-\sigma m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}}, \ \forall \alpha > 0.$$

7.8.2.4 正向漂移时停时的期望

如果 $\mu > 0$，来证明 $E\tau_m < \infty$，给出 $E\tau_m$ 的表达式。

由

$$Ee^{-\alpha\tau_m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}}$$

在两边关于 α 求导得

$$-E[\tau_m e^{-\alpha\tau_m}] = e^{m\mu - m\sqrt{2\alpha+\mu^2}} \frac{-m}{2\alpha + \mu^2},$$

令 $\alpha \to 0$ 得

$$E(\tau_m) = \frac{m}{\mu} \in (0, \infty).$$

7.8.2.5 反向漂移时停时的期望

当 $\mu < 0$ 时，来证明对 $\sigma > -2\mu$，

$$E\left[\exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\} I_{\{\tau_m < \infty\}}\right] = 1.$$

据此证明 $P(\tau_m < \infty) = e^{-2m|\mu|} < 1$，以及 Laplace 变换公式

$$Ee^{-\alpha\tau_m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}},\ \forall\alpha > 0.$$

考虑7.8.2.2的结论，

$$E\left[\exp\left\{\sigma X(t\wedge\tau_m) - (\sigma\mu + \frac{1}{2}\sigma^2)(t\wedge\tau_m)\right\}\right] = 1.$$

若 $\mu < 0$，则当且仅当 $\sigma > -2\mu$ 时 $\sigma\mu + \frac{1}{2}\sigma^2 > 0$。这时，与 (iii) 的证明类似地可以证明

$$E\left[\exp\left\{\sigma m - (\sigma\mu + \frac{1}{2}\sigma^2)\tau_m\right\} I_{\{\tau_m < \infty\}}\right] = 1. \qquad (**)$$

在 (**) 式中令 $\sigma \downarrow -2\mu$ 得

$$P(\tau_m < \infty) = e^{-2m|\mu|} \in (0, \infty).$$

在 (**) 式中令 $\alpha = \sigma\mu + \frac{1}{2}\sigma^2 > 0$，则 $\sigma = -m + \sqrt{2\alpha + \mu^2} > 0$，于是

$$E\left[e^{-\alpha\tau_m} I_{\{\tau_m<\infty\}}\right] = e^{m\mu - m\sqrt{2\alpha+\mu^2}},\ \forall\alpha > 0,$$

而

$$\begin{aligned}
e^{-\alpha\tau_m} =& e^{-\alpha\tau_m} I_{\{\tau_m<\infty\}} + e^{-\alpha\tau_m} I_{\{\tau_m=\infty\}} \\
=& e^{-\alpha\tau_m} I_{\{\tau_m<\infty\}} + 0 \cdot I_{\{\tau_m=\infty\}} \\
=& e^{-\alpha\tau_m} I_{\{\tau_m<\infty\}},
\end{aligned}$$

故

$$Ee^{-\alpha\tau_m} = e^{m\mu - m\sqrt{2\alpha+\mu^2}},\ \forall\alpha > 0.$$

7.8.3 $B(t)/t$ 极限

命题 7.4.

$$\lim_{t\to\infty} \frac{B(t)}{t} = 0,\ a.s.$$

引理 7.6. 设 $\xi_n, n=1,2,...$ 是独立同分布随机变量序列，$E|\xi_1|<\infty$，则

$$\lim_{n\to\infty}\frac{\xi_n}{n}=0,\ a.s.$$

易见 $\frac{\xi_n}{n}=O_p(\frac{1}{n})$ 所以是以概率趋于 0 的。对 $\epsilon>0$，令 $\eta=[|\xi_1/\epsilon|]$，则 $E\eta\le E|\xi_1|/\epsilon$。有

$$\begin{aligned}&\sum_{n=1}^{\infty}P(|\xi_n|\ge n\epsilon)\\=&\sum_{n=1}^{\infty}P(|\xi_1|\ge n\epsilon)\\=&\sum_{n=1}^{\infty}P(|\xi_1|/\epsilon\ge n)\\=&\sum_{n=1}^{\infty}P(\eta\ge n)\\=&E\eta\le E|\xi_1|/\epsilon<0.\end{aligned}$$

由 Borel-Cantelli 引理可知对任意 $\epsilon>0$ 有

$$P(|\xi_n|/n>\epsilon,\ \text{i.o})=0.$$

由 ϵ 任意性可知

$$\lim_{n\to\infty}\frac{\xi_n}{n}=0,\ \text{a.s.}$$

来证明命题7.4。

$$B(n)=B(1)+[B(2)-B(1)]+\cdots+[B(n)-B(n-1)]$$

是独立同分布随机变量和，由强大数律，

$$\lim_{n\to\infty}\frac{B(n)}{n}=E[B(1)]=0.$$

令

$$X_n=\sup_{n\le t\le n+1}(B(t)-B(n)),\quad Y_n=\sup_{n\le t\le n+1}(B(n)-B(t)),$$

则 $X_n, n=0,1,2,...$ 独立同分布，X_0 与 $M(1)=\max_{t\in[0,1]}B(t)$ 同分布。由引理7.6可知

$$\lim_{n\to\infty}\frac{X_n}{n}=0,\ \text{a.s.}$$

Y_n 也具有相同性质。

注意 $n \le t \le n+1$ 时

$$B(n) - Y_n \le B(t) \le X_n - B_n,$$

可得

$$\begin{aligned}
|B(t)| \le & |B(n)| + X_n + Y_n, \\
\frac{|B(t)|}{t} \le & \frac{|B(t)|}{n} \\
\le & \frac{|B(n)|}{n} + \frac{X_n}{n} + \frac{Y_n}{n},
\end{aligned}$$

令 $t \to \infty$，$n = [t]$，有

$$\lim_{t\to\infty} \frac{|B(t)|}{t} = 0, \text{ a.s.}$$

教材习题 7.1 需要这个结论。

Chapter 8

随机积分

本章的目的是引入关于 Brown 运动的积分，讨论其性质并给出在随机分析及金融学中有着重要应用的 Itô 公式. 为了从理论上更严格，这一章的内容参考了北京大学刘勇教授的《应用随机分析》课程讲义 (刘勇 2022)。

8.1 关于随机游动的随机积分

我们从讨论关于简单的随机游动的积分开始. 设 $X_1, X_2, \ldots$ 相互独立，都以各自 $\frac{1}{2}$ 概率分别取 $+1$ 和 -1 值，

$$S_n = \sum_{k=1}^{n} X_k$$

是对称随机游动，$S_0 = 0$，$\mathscr{F}_n = \sigma(X_1, \ldots, X_n)$。

令 B_n 是 $\mathscr{F}_{n-1}$ 可测的随机变量序列，比如它表示第 n 次赌博时所下赌注，则它只能利用第 $n-1$ 次及以前的信息，而不能利用第 n 次赌博的结果. 于是到时刻 n 的收益 Z_n 为（参见例6.4）

$$Z_n = \sum_{i=1}^{n} B_i X_i = \sum_{i=1}^{n} B_i (S_i - S_{i-1}) = \sum_{i=1}^{n} B_i \Delta S_i,$$

这里 $\Delta S_i = S_i - S_{i-1}$, 我们称 Z_n 为 B_n 关于 S_n 的**随机积分**.

容易看出 $\{Z_n\}$ 是关于 $\mathscr{F}_n$ 的鞅。事实上，

$$\begin{aligned}E(Z_n|\mathscr{F}_{n-1}) =& E\left(\sum_{i=1}^{n-1} B_iX_i \,\middle|\, \mathscr{F}_{n-1}\right) + E(B_nX_n|\mathscr{F}_{n-1})\\ =&\sum_{i=1}^{n-1} B_iX_i + B_nE(X_n|\mathscr{F}_{n-1})\\ =&\sum_{i=1}^{n-1} B_iX_i + B_nE(X_n)\\ =&\sum_{i=1}^{n-1} B_iX_i + 0\\ =&Z_{n-1},\end{aligned}$$

易见 $0 \leq m < n$ 有

$$\begin{aligned}&E[Z_n|\mathscr{F}_m]\\ =&E[E(Z_n|\mathscr{F}_{n-1})\,|\,\mathscr{F}_m] = E[Z_{n-1}|\mathscr{F}_m]]\\ =&\cdots = E[Z_{m+1}|\mathscr{F}_m]] = Z_m.\end{aligned}$$

特别地, $E[Z_n] = E[Z_0] = 0.$

如果假定 $E[B_n^2] < \infty$，则

$$\mathrm{Var}[Z_n] = E[Z_n^2] = \sum_{i=1}^{n} E[B_i^2].$$

事实上，

$$Z_n^2 = \sum_{i=1}^{n} B_i^2X_i^2 + 2\sum_{1\leq i<j\leq n} B_iB_jX_iX_j,$$

再注意到 $X_i^2 = 1$，如果 $i < j$，则 B_i, X_i, B_j 都是 $\mathscr{F}_{j-1}$ 可测的，且 X_j 独立于 $\mathscr{F}_{j-1}$，于是

$$\begin{aligned}&E[B_iB_jX_iX_j]\\ =&E\{E[B_iB_jX_iX_j|\mathscr{F}_{j-1}]\}\\ =&E\{B_iB_jX_iE[X_j|\mathscr{F}_{j-1}]\}\\ =&E\{B_iB_jX_iE[X_j]\} = 0.\end{aligned}$$

8.2 Itô 积分

8.2.1 连续可微函数的随机积分

考虑关于布朗运动的积分 $\int_0^T X(t)\,dB(t)$。以 $B(t)$ 为证券价格，$X(t)$ 作为持仓量，则 $\int_0^T X(t)\,dB(t)$ 可以作为 $[0,T]$ 期间的投资收益。但是，这个积分不能看作是黎曼-斯蒂尔杰斯 (Riemann-Stieljes) 积分，因为 $B(t)$ 处处不可微，有非零的二次变差。

如果 $f(t)$ 是连续可微函数，则由分部积分公式可以形式地定义

$$\int_0^t f(t)\,dB(t) = f(t)B(t) - \int_0^t f'(u)B(u)\,du,$$

而 $\int_0^t f'(u)B(u)\,du$ 的被积函数是轨道连续的，可以看成是对每条轨道的黎曼积分。参见8.5.3。

8.2.2 阶梯函数的随机积分

考虑被积函数为简单函数（阶梯函数）的随机积分。

设 $0 = t_0 < t_1 < \cdots < t_n = T$ 是 $[0,T]$ 的一个分割，函数

$$f(t) = \sum_{i=0}^{n-1} c_i I_{(t_i,t_{i+1}]}(t) + c_0' I_{\{0\}}(t),\ t \in [0,T],$$

其中 c_i 和 c_0' 为常量，定义 $\{f(t)\}$ 的伊藤 (Itô) 随机积分为

$$\int_0^T f(t)\,dB(t) = \sum_{i=0}^{n-1} c_i[B(t_{i+1}) - B(t_i)]. \tag{8.1}$$

由布朗运动的独立增量性和高斯过程性质，可知(8.1)服从正态分布，均值为 0，方差为

$$\begin{aligned}
&\mathrm{Var}\left(\int_0^T f(t)\,dB(t)\right)\\
=&\sum_{i=0}^{n-1} c_i^2 (t_{i+1} - t_i)\\
=&\int_0^T f^2(t)\,dt.
\end{aligned}$$

8.2.3 可料阶梯过程的随机积分

考虑被积函数是随机过程的情形。

设 $\{\mathscr{F}(t), t \geq 0\}$ 是 σ 代数流，$\{B(t), \mathscr{F}(t), t \geq 0\}$ 是适应的标准布朗运动，且对任意 $t \geq s \geq 0$，$B(t) - B(s)$ 与 $\mathscr{F}(s)$ 独立。如果取 $\mathscr{F}(t) = \mathscr{F}^B(t) = \sigma(\{B(u) : 0 \leq u \leq t\})$ 则可以满足上述条件。

定义 8.1. 设 $0 = t_0 < t_1 < \cdots < t_n = T$，随机过程 $\{X(t), t \in [0, T]\}$ 为

$$X(t) = \sum_{i=0}^{n-1} \xi_i I_{(t_i, t_{i+1}]}(t) + \xi_0' I_{\{0\}}(t),\ t \in [0, T].$$

其中 ξ_i 是关于 $\mathscr{F}(t_i)$ 可测的随机变量，$E(\xi_i^2) < \infty$，ξ_0' 关于 $\mathscr{F}(0)$ 可测，则称 $\{X(t), \mathscr{F}(t), t \in [0, T]\}$ 为可料阶梯过程（或简单可料适应过程，初等可料适应过程）。

这里“可料”是指过程（轨道）左连续（左连续来源于区间 $(t_i, t_{i+1}]$ 的右闭性质），当 $t \in (t_i, t_{i+1}]$ 时 $X(t) = \xi_i$ 关于 $\mathscr{F}(t_i)$ 可测，而 $\mathscr{F}(t_i) \subset \mathscr{F}(t)$，所以 $\{X(t)\}$ 也是 $\{\mathscr{F}(t)\}$ 适应的。

注意，可料阶梯过程必为可测过程（即 $([0, \infty) \times \Omega, \mathscr{B}([0, \infty)) \times \mathscr{F})$ 到 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 的可测变换）。由左连续性可知其为循序可测的（见命题6.10）。

对可料阶梯过程，定义

$$\int_0^T X(u)\, du = \sum_{i=0}^{n-1} \xi_i (t_{i+1} - t_i). \tag{8.2}$$

可料阶梯过程的每条轨道是阶梯函数，上述积分是对每条轨道的黎曼积分。易见只要 $E|\xi_i| < \infty (i = 0, 1, \dots, n-1)$ 就有

$$E \int_0^T X(u)\, du = \sum_{i=0}^{n-1} E(\xi_i)(t_{i+1} - t_i) = \int_0^T E[X(u)]\, du,$$

即积分与期望可交换次序。

定义 8.2. 设 $\{B(t), \mathscr{F}(t), t \geq 0\}$ 为标准布朗运动，$\{X(t), \mathscr{F}(t), t \geq 0\}$ 为可料阶梯过程，其伊藤随机积分定义为

$$\int_0^T X(t)\, dB(t) = \sum_{i=0}^{n-1} \xi_i [B(t_{i+1}) - B(t_i)]. \tag{8.3}$$

$\int_0^T X(t)\,dB(t)$ 是一个随机变量。

对 $0 < t \le T$，设 $t \in (t_k, t_{k+1}]$，定义

$$\int_0^t X(u)\,dB(u) = \sum_{i=0}^{k-1} \xi_i [B(t_{i+1}) - B(t_i)] + \xi_k [B(t) - B(t_k)]. \tag{8.4}$$

对 $0 \le s < t \le T$，定义

$$\int_s^t X(u)\,dB(u) = \int_0^t X(u)\,dB(u) - \int_0^s X(u)\,dB(u).$$

下面讨论可料阶梯过程的随机积分性质。

8.2.3.1 适应性

$\int_0^t X(u)\,dB(u)$ 关于 $\{\mathscr{F}(t)\}$ 适应。定义(8.4)中 $\xi_i \in \mathscr{F}_{t_i} \subset \mathscr{F}_t$, $\xi_k \in \mathscr{F}_{t_k} \subset \mathscr{F}_t$，$B(t_{i+1}) - B(t_i) \in \mathscr{F}_{t_{i+1}} \subset \mathscr{F}_{t_k} \subset \mathscr{F}_t$，$B(t) - B(t_k) \in \mathscr{F}_t$，所以 $\int_0^t X(u)\,dB(u) \in \mathscr{F}_t$。

8.2.3.2 零均值

这时，

$$\begin{aligned}
&E|\xi_i [B(t_{i+1}) - B(t_i)]| \\
\le&\|\xi_i\| \cdot \|B(t_{i+1}) - B(t_i)\| \quad \text{(Schwarz 不等式)} \\
=&[E(\xi_i^2)]^{1/2} (t_{i+1} - t_i)^{1/2} < \infty,
\end{aligned}$$

且

$$\begin{aligned}
&E\left\{\xi_i [B(t_{i+1}) - B(t_i)]\right\} \\
=&E\left\{E\left[\xi_i (B(t_{i+1}) - B(t_i)) \,|\, \mathscr{F}(t_i)\right]\right\} \\
=&E\left\{\xi_i\, E\left[B(t_{i+1}) - B(t_i) \,|\, \mathscr{F}(t_i)\right]\right\} \\
=&0,
\end{aligned}$$

所以 $E\int_0^T X(t)\,dB(t) = 0$，即随机积分均值为零。

8.2.3.3 常值过程积分

由定义可知

$$\int_a^b \xi\,dB(t) = \int_{[a,b]} \xi\,dB(t) = \int_{(a,b]} \xi\,dB(t) = \xi[B(b) - B(a)].$$

其中 ξ 关于 $\mathscr{F}(a)$ 可测且 $E\xi^2 < \infty$。

8.2.3.4 线性

易见如果 $\{X(t)\}$, $\{Y(t)\}$ 是简单可料过程，α, β 是常数，则

$$\int_0^T [\alpha X(t) + \beta Y(t)]\, dB(t) = \alpha \int_0^T X(t)\, dB(t) + \beta \int_0^T Y(t)\, dB(t).$$

这是随机积分的线性性质。

证明：可以找到 $[0, T]$ 的分割 $0 = t_0 < t_1 < \cdots < t_n = T$ 使得 $\{X(t)\}$ 和 $\{Y(t)\}$ 在每个小区间上都不随 t 变化，且细分后仍保持 $\xi_i \in \mathscr{F}(t_i)$ 这样的要求。于是设

$$X(t) = \sum_{i=0}^{n-1} \xi_i I_{(t_i, t_{i+1}]}(t),$$
$$Y(t) = \sum_{i=0}^{n-1} \eta_i I_{(t_i, t_{i+1}]}(t),$$

则

$$\begin{aligned}
&\alpha X(t) + \beta Y(t) \\
=& \sum_{i=0}^{n-1} (\alpha \xi_i + \beta \eta_i) I_{(t_i, t_{i+1}]}(t), \\
& \int_0^T [\alpha X(t) + \beta Y(t)]\, dB(t) \\
=& \sum_{i=0}^{n-1} (\alpha \xi_i + \beta \eta_i)[B(t_{i+1}) - B(t_i)] \\
=& \alpha \sum_{i=0}^{n-1} \xi_i [B(t_{i+1}) - B(t_i)] + \beta \sum_{i=0}^{n-1} \eta_i [B(t_{i+1}) - B(t_i)] \\
=& \alpha \int_0^T X(t)\, dB(t) + \beta \int_0^T Y(t)\, dB(t).
\end{aligned}$$

8.2.3.5 鞅

$\{\int_0^t X(u)\, dB(u), \mathscr{F}(t), t \in [0, T]\}$ 是鞅。

证明：对任意 $0 \le s \le t \le T$，设 $s \in (t_k, t_{k+1}]$，不妨设 $t = T$($t < T$ 可类似

证明)，则

$$
\begin{aligned}
& E\left(\int_0^T X(u)\,dB(u) \,\middle|\, \mathscr{F}(s)\right) \\
= & \sum_{i=0}^{k-1} E\left(\xi_i[B(t_{i+1}) - B(t_i)] \mid \mathscr{F}(s)\right) \\
& + E\left(\xi_k[B(t_{k+1}) - B(t_k)] \mid \mathscr{F}(s)\right) \\
& + \sum_{i=k+1}^{n-1} E\left(\xi_i[B(t_{i+1}) - B(t_i)] \mid \mathscr{F}(s)\right) \\
= & \sum_{i=0}^{k-1} \xi_i[B(t_{i+1}) - B(t_i)] \\
& + \xi_k E\left([B(t_{k+1}) - B(s)] + [B(s) - B(t_k)] \mid \mathscr{F}(s)\right) \\
& + \sum_{i=k+1}^{n-1} E[E\left(\xi_i[B(t_{i+1}) - B(t_i)] \mid \mathscr{F}(t_i)\right) | \mathscr{F}(s)] \\
= & \sum_{i=0}^{k-1} \xi_i[B(t_{i+1}) - B(t_i)] \\
& + \xi_k[B(s) - B(t_k)] \\
& + \sum_{i=k+1}^{n-1} E\left[\xi_i\, E\left(B(t_{i+1}) - B(t_i) \mid \mathscr{F}(t_i)\right) | \mathscr{F}(s)\right] \\
= & \sum_{i=0}^{k-1} \xi_i[B(t_{i+1}) - B(t_i)] \\
& + \xi_k[B(s) - B(t_k)] \\
= & \int_0^s X(u)\,dB(u).
\end{aligned}
$$

说明：可以将 $\{B(t)\}$ 看成证券价格，t_i 看成是交易日期，$B(t_{i+1}) - B(t_i)$ 是证券价格变化，将 ξ_i 看成是 $(t_i, t_{i+1}]$ 区间的持仓量，ξ_i 可以依赖于 $[0, t_i]$ 的价格信息但不能依赖于 t_i 以后的价格信息，则 $\int_0^t X(u)\,dB(u)$ 是截止到 t 时刻的投资收益。如果是公平市场且不考虑现金的时间价值因素，投资收益应该期望不变，具有鞅性质。

8.2.3.6 等距性

$$
\begin{aligned}
&\operatorname{Var}\left(\int_0^t X(u)\,dB(u)\right) = E\left[\left(\int_0^t X(u)\,dB(u)\right)^2\right] \\
=&E\int_0^t X^2(u)\,du = \int_0^t EX^2(u)\,du.
\end{aligned}
$$

证明：对 $t=T$ 证明，对 $t\in[0,T]$ 证明同理。

$$
\begin{aligned}
&\operatorname{Var}\left(\int_0^T X(u)\,dB(u)\right) \\
=&E\left\{\left(\sum_{i=0}^{n-1}\xi_i[B(t_{i+1})-B(t_i)]\right)^2\right\} \\
=&\sum_{i=0}^{n-1} E\left[\xi_i^2(B(t_{i+1})-B(t_i))^2\right] \\
&+2\sum_{i<j} E\left[\xi_i\xi_j(B(t_{i+1})-B(t_i))(B(t_{j+1})-B(t_j))\right].
\end{aligned}
$$

其中

$$
\begin{aligned}
&E\left[\xi_i^2(B(t_{i+1})-B(t_i))^2\right] \\
=&E\left\{E\left[\xi_i^2(B(t_{i+1})-B(t_i))^2 \mid \mathscr{F}(t_i)\right]\right\} \\
=&E\left\{\xi_i^2\, E\left[(B(t_{i+1})-B(t_i))^2 \mid \mathscr{F}(t_i)\right]\right\} \\
=&E\left\{\xi_i^2\, E\left[(B(t_{i+1})-B(t_i))^2\right]\right\} \\
=&E(\xi_i^2)(t_{i+1}-t_i),
\end{aligned}
$$

而

$$
\begin{aligned}
&E\left[\xi_i\xi_j(B(t_{i+1})-B(t_i))(B(t_{j+1})-B(t_j))\right] \\
=&E\left\{E\left[\xi_i\xi_j(B(t_{i+1})-B(t_i))\,(B(t_{j+1})-B(t_j)) \mid \mathscr{F}(t_j)\right]\right\} \\
=&E\left\{\xi_i\xi_j(B(t_{i+1})-B(t_i))\,E\left[(B(t_{j+1})-B(t_j)) \mid \mathscr{F}(t_j)\right]\right\} \\
=&0,
\end{aligned}
$$

故

$$
\operatorname{Var}\left(\int_0^T X(u)\,dB(u)\right) = \sum_{i=0}^{n-1} E(\xi_i^2)(t_{i+1}-t_i).
$$

注意到

$$X^2(u) = \sum_{i=0}^{n-1} \xi_i^2 I_{(t_i,t_{i+1}]}(u) + \xi_0'^2 I_{\{0\}}(u),$$

所以

$$\begin{aligned}
&\operatorname{Var}\left(\int_0^T X(u)\,dB(u)\right)\\
=&\sum_{i=0}^{n-1} E(\xi_i^2)(t_{i+1}-t_i)\\
=&\int_0^T EX^2(u)\,du\\
=&E\sum_{i=0}^{n-1}\xi_i^2(t_{i+1}-t_i)\\
=&E\int_0^T X^2(u)\,du.
\end{aligned}$$

8.2.3.7 推广的等距性

设 $\{X(t)\}$, $\{Y(t)\}$ 是可料阶梯过程，对 $0 \le s \le t \le T$ 有

$$\begin{aligned}
&\operatorname{Cov}\left(\int_0^s X(u)\,dB(u),\, \int_0^t Y(u)\,dB(u)\right)\\
=&E\left(\int_0^s X(u)\,dB(u)\cdot\int_0^t Y(u)\,dB(u)\right)\\
=&E\int_0^s X(u)Y(u)\,du\\
=&\int_0^s E[X(u)Y(u)]\,du.
\end{aligned}$$

证明：存在 $[0,T]$ 的分割使得 $\{X(t)\}$, $\{Y(t)\}$ 都可以用同一个分割写成可料阶梯过程表示，从而 $X(t)+Y(t)$ 也是简单可料过程。令 $X(u)=0$ 对 $u>s$,

$Y(u)=0$ 对 $u>t$，由恒等式 $xy=\frac{1}{2}[(x+y)^2-x^2-y^2]$ 可得

$$
\begin{aligned}
& E\left(\int_0^s X(u)\,dB(u)\cdot\int_0^t Y(u)\,dB(u)\right) \\
=& E\left(\int_0^T X(u)\,dB(u)\cdot\int_0^T Y(u)\,dB(u)\right) \\
=& \frac{1}{2}E\left\{\left[\int_0^T X(u)\,dB(u)+\int_0^T Y(u)\,dB(u)\right]^2\right\} \\
& -\frac{1}{2}E\left\{\left[\int_0^T X(u)\,dB(u)\right]^2\right\} \\
& -\frac{1}{2}E\left\{\left[\int_0^T Y(u)\,dB(u)\right]^2\right\} \\
=& \frac{1}{2}E\left\{\left[\int_0^T (X(u)+Y(u))\,dB(u)\right]^2\right\} \\
& -\frac{1}{2}E\left\{\left[\int_0^T X(u)\,dB(u)\right]^2\right\} \\
& -\frac{1}{2}E\left\{\left[\int_0^T Y(u)\,dB(u)\right]^2\right\} \\
=& \frac{1}{2}\int_0^T E[(X(u)+Y(u))^2]\,du \\
& -\frac{1}{2}\int_0^T E[X^2(u)]\,du \\
& -\frac{1}{2}\int_0^T E[Y^2(u)]\,du \\
=& \int_0^T E[X(u)Y(u)]\,du.
\end{aligned}
$$

注意到 $X(u)Y(u)$ 的表达式也是可料阶梯过程（不一定二阶矩有限），$E|X(u)Y(u)| \leq \|X(u)\|\cdot\|Y(u)\| < \infty$，所以 $\int_0^T E[X(u)Y(u)]\,du = E\int_0^T X(u)Y(u)\,du$。

8.2.4 被积函数为一般过程的随机积分

设 $(\Omega, \mathscr{F}, P)$ 是完备的概率空间，随机过程的集合 V 定义为

$$V=\Bigg\{\{X(t,\omega):\ t\geq 0,\omega\in\Omega\}:\ \{X(t,\omega)\}\text{关于}\mathscr{B}([0,T])\times\mathscr{F}\text{二元可测},$$

$$\text{关于}\mathscr{F}_t\text{适应},\ \int_0^T E[X^2(t,\omega)]\,dt<\infty\Bigg\}.$$

其中 $\{\mathscr{F}_t, t\geq 0\}$ 为单调增 σ 代数流，$\{B(t), \mathscr{F}_t, t\geq 0\}$ 适应，且 $B(t)-B(s)$ 与 $\mathscr{F}(s)$ 独立对任意 $0\leq s\leq t$。有些教材记 V 为 $\mathscr{L}_T^2$ 或 $\mathscr{L}^2[0,T]$。

在 V 中可以定义内积

$$\langle X, Y\rangle=\int_0^T E[X(t)Y(t)]\,dt,$$

这使得 V 成为 Hilbert 空间（完备内积空间），导出模

$$\|X\|=\left(\int_0^T E\left[X^2(t,\omega)\right]\,dt\right)^{1/2}.$$

如果考虑 $[0,\infty)$ 上的过程，可以将 V 的空间变成 $\int_0^T E[X^2(t,\omega)]\,dt<\infty$, $\forall T>0$，并定义

$$\|X\|=\sum_{n=1}^{\infty}\frac{1}{2^n}\min\left[1,\left(\int_0^n E\left[X^2(t,\omega)\right]\,dt\right)^{1/2}\right].$$

下面我们仅考虑定义在 $[0,T]$ 的过程。

因为要求 $X(t,\omega)$ 二元可测，所以非负可测函数 $X^2(t,\omega)$ 是积分可交换的，有：

$$\begin{aligned}
&\int_0^T E[X^2(t,\omega)]\,dt=\int_0^T\int_\Omega X^2(t,\omega)\,dP(\omega)\,dt\\
=&\int_{[0,T]\times\Omega}X^2(t,\omega)\,dt\,dP(\omega)\\
=&\int_\Omega\int_0^T X^2(t,\omega)\,dt\,dP(\omega)=E\left[\int_0^T X^2(t,\omega)\,dt\right].
\end{aligned}$$

在对 V 中的被积函数可以定义 Itô 随机积分。

8.2.4.1 用可料简单过程逼近

引理 8.1. 对 $\{X(t,\omega)\} \in V$, 存在可料阶梯过程 $\{\phi_n(t,\omega)\}$ 使得 $\|\phi_n - X\| \to 0$, 即

$$\int_0^T E\left[|X(t,\omega) - \phi_n(t,\omega)|^2\right] dt \to 0, \ n \to \infty.$$

证明参考 (张波, 商豪, and 邓军 2020)，(钱敏平 1990) 节 8.2 命题 8.8，(龚光鲁 and 钱敏平 2019) 节 1.2 P.11 命题 1.5。

证明分三个步骤逼近。

步骤一、设 $\{X\} \in V$ 有界，且对每个 $\omega \in \Omega$，$X(t,\omega)$ 关于 t 连续。则存在简单可料过程 $\{\phi_n\}$ 序列：

$$\phi_n(t,\omega) = \sum_{j=0}^{n-1} X(t_j,\omega) I_{(t_j, t_{j+1}]}(t),$$

易见 $\phi_n \in V$，其中 $0 = t_0 < t_1 < \cdots < t_n = T$ 是 $[0,T]$ 的一个分割且 $n \to \infty$ 时 $\delta_n = \max_j(t_{j+1} - t_j) \to 0$。对每个 $\omega \in \Omega$, $n \to \infty$ 时有

$$\int_0^T (X - \phi_n)^2 \, dt \to 0.$$

这是由于 $X(\cdot,\omega)$ 连续则一致连续，所以 $X - \phi_n$ 一致地趋于 0，从而其平方的积分趋于 0。

由于 X 有界故 $(X - \phi_n)^2$ 有界，从而 $\int_0^T (X - \phi_n)^2 \, dt$ 有界，由概率论的有界收敛定理可知

$$\lim_{n\to\infty} E \int_0^T (X - \phi_n)^2 \, dt = E \lim_{n\to\infty} \int_0^T (X - \phi_n)^2 \, dt = 0.$$

步骤二、设 $\{X\} \in V$ 有界（不要求关于 t 连续）。设

$$X(t,\omega) \leq M, \ \forall t, \omega.$$

定义

$$h_n(t,\omega) = \int_0^t K_n(s - t) X(s,\omega) \, ds,$$

其中 $K_n(s)$ 是 $(-\infty,\infty)$ 上的非负连续函数，使得对 $s\notin(-\frac{1}{n},0)$ 有 $K_n(s)=0$，且 $\int_{-\frac{1}{n}}^{0}K_n(s)\,ds=1$，如

$$K_n(s)=\begin{cases}2n+2n^2, & s\in[-\frac{1}{n},0],\\ 0, & \text{其它}.\end{cases}$$

这样定义的 $h_n(\cdot,\omega)$ 是对 $X(\cdot,\omega)$ 的一个左侧局部平均（光滑），易见 $h_n(\cdot,\omega)$ 连续，且仍关于 $\mathscr{F}(t)$ 适应。由勒贝格平方可积空间的逼近性质可知对任意 $\omega\in\Omega$ 都有

$$\int_0^T(h_n(s,\omega)-X(s,\omega))^2\,ds\to 0,\ n\to\infty.$$

由概率论的有界收敛定理可知

$$E\int_0^T(h_n(s,\omega)-X(s,\omega))^2\,ds\to 0,\ n\to\infty.$$

步骤三、设 $\{X\}\in V$。这时，令

$$X_n(t,\omega)=\begin{cases}-n, & \text{若}X(t,\omega)<-n,\\ X(t,\omega), & \text{若}-n\le X(t,\omega)\le n,\\ n, & \text{若}X(t,\omega)>n.\end{cases}$$

则 $X_n(t,\omega)\to X(t,\omega)$, $n\to\infty$。这时 $(X-X_n)^2$ 有上界函数：

$$\begin{aligned}&[X(t,\omega)-X_n(t,\omega)]^2\\ =&[X(t,\omega)-X_n(t,\omega)]^2I_{\{|X(t,\omega)>n\}}\\ \le&[2X^2(t,\omega)+2X_n^2(t,\omega)]I_{\{|X(t,\omega)>n\}}\\ \le&4X^2(t,\omega)I_{\{|X(t,\omega)>n\}}\\ \le&4X^2(t,\omega).\end{aligned}$$

因为 $E\int_0^T X^2(t,\omega)\,dt<\infty$ 所以 $\int_0^T X^2(t,\omega)\,dt<\infty$，a.s.，由 $L^2[0,T]$ 勒贝格平方可积空间的控制收敛定理可知

$$\begin{aligned}&\lim_{n\to\infty}\int_0^T[X(t,\omega)-X_n(t,\omega)]^2\,dt\\ \le&\lim_{n\to\infty}\int_0^T 4X^2(t,\omega)I_{\{|X(t,\omega)>n\}}\,dt\\ =&0,\ \text{a.s.}\end{aligned}$$

因为 $E\left\{\int_0^T 4X^2(t,\omega)\,dt\right\} < \infty$，由概率论的控制收敛定理得

$$\lim_{n\to\infty} E\int_0^T [X(t,\omega) - X_n(t,\omega)]^2\,dt = 0.$$

这样，由步骤三，对 $\{X\} \in V$ 存在有界的 $\{X_n\} \subset V$，使得

$$E\int_0^T [X(t,\omega) - X_n(t,\omega)]^2\,dt < \frac{1}{9n}.$$

由步骤二，对每个 $\{X_n(t,\omega)\}$，存在有界且关于 t 连续的 $\{h_n(t,\omega)\} \in V$ 使得

$$E\int_0^T [X_n(t,\omega) - h_n(t,\omega)]^2\,dt < \frac{1}{9n}.$$

由步骤一，对每个 $\{h_n(t,\omega)\}$ 存在可料阶梯过程 $\{\phi_n(t,\omega)\} \in V$ 使得

$$E\int_0^T [h_n(t,\omega) - \phi_n(t,\omega)]^2\,dt < \frac{1}{9n}.$$

这样，由于

$$[X - \phi_n]^2 \le 3[X - X_n]^2 + 3[X_n - h_n]^2 + 3[h_n - \phi_n]^2,$$

就有

$$E\int_0^T [X(t,\omega) - \phi_n(t,\omega)]^2\,dt \le \frac{1}{n} \to 0,\ n \to \infty.$$

注意到

$$E\int_0^T [X(t,\omega) - \phi_n(t,\omega)]^2\,dt = \int_0^T E[X(t,\omega) - \phi_n(t,\omega)]^2\,dt,$$

引理证毕。

注： 对 V 中的过程 X，存在 “可料修正”$\tilde{X}$，使得 $P(\tilde{X}_t = X_t) = 1, \forall t \in [0,T]$，且 $\tilde{X}(t)$ 左连续。所以可以认为 V 中的过程都是左连续的。

8.2.4.2 伊藤随机积分定义

定义 8.3 (伊藤积分)**.** 对 $X \in V$，设可料简单过程 ϕ_n 使得 $\int_0^T E|\phi_n - X|^2\,dt \to 0$，定义伊藤 (Itô) 积分 $\int_0^T X(t,\omega)\,dB(t)$ 为 $\int_0^T \phi_n(t,\omega)\,dB(t)$ 的均方极限。

定理 8.1. *上述伊藤积分存在且不依赖于 ϕ_n 的选取，不同 ϕ_n 得到的均方极限必 a.s. 相等。*

证明由可料简单函数的伊藤积分的线性性质和等距性，

$$
\begin{aligned}
& E\left|\int_0^T \phi_n(t)\,dB(t) - \int_0^T \phi_m(t)\,dB(t)\right|^2 \\
=& E\left|\int_0^T [\phi_n(t) - \phi_m(t)]\,dB(t)\right|^2 \\
=& \int_0^T E|\phi_n(t) - \phi_m(t)|^2\,dt \\
\le& 2\int_0^T E|\phi_n(t) - X(t)|^2\,dt + 2\int_0^T E|\phi_m(t) - X(t)|^2\,dt \\
\to& 0,\ n, m \to \infty,
\end{aligned}
$$

所以 $\int_0^T \phi_n(t)\,dB(t)$ 是 $L^2(\Omega, \mathscr{F}, P)$（即 $(\Omega, \mathscr{F}, P)$ 中所有二阶矩有限随机变量组成的 Hilbert 空间）中的基本列，所以有极限且极限有二阶矩。这里利用了非负可测函数的积分与期望可交换次序。

设 ϕ_n, ψ_n 都可以逼近 X，则类似地有

$$
\begin{aligned}
& E\left|\int_0^T \phi_n(t)\,dB(t) - \int_0^T \psi_n(t)\,dB(t)\right|^2 \\
=& E\left|\int_0^T [\phi_n(t) - \psi_n(t)]\,dB(t)\right|^2 \\
=& \int_0^T E|\phi_n(t) - \psi_n(t)|^2\,dt \\
\le& 2\int_0^T E|\phi_n(t) - X(t)|^2\,dt + 2\int_0^T E|\psi_n(t) - X(t)|^2\,dt \\
\to& 0,\ n \to \infty,
\end{aligned}
$$

由 $L^2(\Omega,\mathscr{F},P)$ 中内积的连续性和上式可知，

$$
\begin{aligned}
&\lim_{n\to\infty} E\left|\int_0^T \phi_n(t)\,dB(t)-\int_0^T \psi_n(t)\,dB(t)\right|^2\\
=&E\left|\lim_{n\to\infty}\int_0^T \phi_n(t)\,dB(t)-\lim_{n\to\infty}\int_0^T \psi_n(t)\,dB(t)\right|^2\\
=&0,
\end{aligned}
$$

即 $\int_0^T \phi_n(t)\,dB(t)$ 和 $\int_0^T \psi_n(t)\,dB(t)$ 的 L^2 极限 a.s. 相等。

8.2.4.3 伊藤随机积分定义推广

可以放松 V 中对 $\int_0^T E[X^2(t,\omega)]\,dt<\infty$ 的要求，仅要求 $\int_0^T X^2(t,\omega)\,dt<\infty$, a.s.，这样的适应过程的空间记为 V^*，可以在 V^* 上定义伊藤随机积分。有些教材记 V^* 为 $\mathscr{L}^{2,\mathrm{loc}}[0,T]$。进一步将对固定 T 的限制去掉，令

$$
\begin{aligned}
V^*=\Bigg\{&\{X(t,\omega):\ t\geq 0,\omega\in\Omega\}:\ \{X(t,\omega)\}\text{关于}\mathscr{B}([0,\infty))\times\mathscr{F}\text{二元可测},\\
&\text{关于}\mathscr{F}_t\text{适应},\ \int_0^T X^2(t,\omega)\,dt<\infty,\ \text{a.s.},\forall T>0\Bigg\}.
\end{aligned}
$$

伊藤随机积分的另一定义为：设 $\{X(t)\}\in V^*$，对 $[0,T]$ 的分割 $0=t_0<t_1<\cdots<t_n=T$，满足 $\lim_{n\to\infty}\max_j(t_{j+1}-t_j)=0$，如果

$$
\sum_{j=0}^{n-1}X(t_j,\omega)[B(t_{j+1}-B(t_j))]
$$

均方极限存在且不依赖于分割的选取,则定义极限为伊藤积分 $\int_0^T X(t,\omega)\,dB(t)$。

命题 8.1. *若* $Z\sim N(0,1)$，*则*

$$
Ee^{uZ^2}=\begin{cases}(1-2u)^{-\frac{1}{2}}, & 0<u<\frac{1}{2},\\ +\infty, & u\geq\frac{1}{2}.\end{cases}
$$

证明:

$$
\begin{aligned}
E(e^{uZ^2}) &= \int_{-\infty}^{\infty} e^{uz^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,dz\\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(1-2u)z^2}\,dz,
\end{aligned}
$$

当 $u \geq \frac{1}{2}$ 时

$$E(e^{uZ^2}) \geq \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\cdot 1\,dz = +\infty,$$

当 $0 < u < 1$ 时，

$$
\begin{aligned}
E(e^{uZ^2}) =&(1-2u)^{-1/2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}(1-2u)^{-1/2}}\exp\{-\frac{1}{2}\frac{z^2}{(1-2u)^{-1}}\}\,dz\\
=&(1-2u)^{-1/2}.
\end{aligned}
$$

例 8.1. 设 f 是连续函数，考虑 $\int_0^1 f(B(t))\,dB(t)$. 因为 $B(t)$ 有连续的路径，所以 $f(B(t))$ 也在 $[0,1]$ 上连续, 从而 $\int_0^1 f(B(t))dB(t)$ 有定义. 然而根据 f 的不同，这个积分可以有 (或没有) 有限的矩. 例如：

(1) 取 $f(t) = t$, 则由于

$$\int_0^1 E[B^2(t)]\,dt = \int_0^1 t\,dt = \frac{1}{2} < \infty,$$

所以被积函数 $B(t) \in V$，于是 $E[\int_0^1 B(t)dB(t)] = 0$，并且由 Itô 积分的等距性有

$$E\left[\int_0^1 B(t)\,dB(t)\right]^2 = \int_0^1 E[B^2(t)]\,dt = \frac{1}{2}.$$

实际上，可以利用定义计算出 $\int_0^t B(u)\,dB(u)$。将 $[0,t]$ 等分为 n 段，取 $B(u)$ 的可料简单过程逼近，在每一段 $(t_i, t_{i+1}]$ 用左端点的值 $B(t_i)$ 代表，用下面的随机积分逼近 $\int_0^t B(u)\,dB(u)$:

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

其中 $t_i = \frac{i}{n}$，简记 $B_i = B(t_i)$。注意 $B_0 = B(0) = 0$, $B_n = B(t)$，且

$$\begin{aligned}
&\frac{1}{2}\sum_{i=0}^{n-1}[B_{i+1} - B_i]^2 \\
=&\frac{1}{2}\sum_{i=0}^{n-1}B_{i+1}^2 + \frac{1}{2}\sum_{i=0}^{n-1}B_i^2 - \frac{1}{2}\sum_{i=0}^{n-1}B_iB_{i+1} \\
=&\frac{1}{2}B_n^2 + \sum_{i=0}^{n-1}B_i^2 - \frac{1}{2}\sum_{i=0}^{n-1}B_iB_{i+1} \\
=&\frac{1}{2}B^2(t) + \sum_{i=0}^{n-1}B_i[B_i - B_{i+1}],
\end{aligned}$$

从而

$$\begin{aligned}
&\sum_{i=0}^{n-1}B_i[B_{i+1} - B_i] \\
=&\frac{1}{2}B^2(t) - \frac{1}{2}\sum_{i=0}^{n-1}[B_{i+1} - B_i]^2 \\
\to&\frac{1}{2}B^2(t) - \frac{1}{2}[B, B](t) = \frac{1}{2}B^2(t) - \frac{1}{2}t.
\end{aligned}$$

即

$$\int_0^t B(u)\,dB(u) = \frac{1}{2}B^2(t) - \frac{1}{2}t.$$

在第8.4中我们将用 Itô 公式直接算出这个结果。

(2) 取 $f(t) = e^{t^2}$，此时考虑 $\int_0^1 e^{B^2(t)}dB(t)$. 令 $X(t) = e^{B^2(t)}$，则

$$\int_0^1 X^2(t)\,dt = \int_0^1 e^{2B^2(t)}\,dt < \infty, \text{ a.s.},$$

这是因为被积函数轨道连续。于是，$\{X(t)\} \in V^*$，随机积分 $\int_0^1 e^{B^2(t)}dB(t)$ 存在。但是，

$$E[e^{2B^2(t)}] = E[e^{2tZ^2}] = \begin{cases} (1-4t), & 0 < t < \frac{1}{4}, \\ +\infty, & t \geq \frac{1}{4}, \end{cases}$$

所以 $\{X(t)\} \notin V$，$E\left(\int_0^1 e^{B^2(t)}dB(t)\right)^2 = \infty$。

例 8.2. 求积分 $J = \int_0^1 t\,dB(t)$ 的均值与方差.

解：因为 t^2 是 $[0,1]$ 上的连续函数所以 $\int_0^1 t^2\,dt<\infty$，$X(t)=t$ 非随机，必为 $\{\mathscr{F}_t\}$ 适应的（取常数值的随机变量关于任一 σ 代数可测），即有 $\{X(t)\}\in V$，从而由 Itô 积分性质有 $E[J]=0$,

$$E[J^2]=\int_0^1 t^2dt=\frac{1}{3}.$$

例 8.3. 求使得积分 $\int_0^1(1-t)^{-\alpha}\,dB(t)$ 满足 V 中被积函数条件的 α 的值.

解：取 α 令

$$\int_0^1(1-t)^{-2\alpha}\,dt<\infty,$$

这个积分的瑕点是 $t=1$，只要 $\alpha<\frac{1}{2}$，这时被积函数属于 V。

8.2.4.4 Itô 积分不是黎曼-斯蒂尔杰斯积分

例 8.4. 当 $f(t)$ 为非随机连续可微函数时，$\int_a^b f(t)\,dB(t)$ 与轨道的黎曼-斯蒂尔杰斯积分相同，尽管 $B(t)$ 的轨道不是有界变差的。如果 $f(t)$ 是随机过程，则因为 $\{B(t),t\geq 0\}$ 的处处连续但处处不可微，所以轨道的黎曼-斯蒂尔杰斯积分可能不存在。给出反例。

证明：设 $\{\xi(t),t\geq 0\}$ 为随机过程，则沿轨道定义的黎曼-斯蒂尔杰斯积分 $\int_a^b\xi(t)\,dB(t)$ 不一定存在。例如，考虑

$$\int_0^t B(s)\,dB(s).$$

按照黎曼-斯蒂尔杰斯积分，应该取 $[0,t]$ 的分割 $0=t_0<t_1<\cdots<t_n=t$，使得最大区间间距 $\delta_n\to 0$，可任取 $t_i^*\in[t_i,t_{i+1}]$ 构造达布和

$$S_n=\sum_{i=0}^{n}B(t_i^*)(B(t_{i+1})-B(t_i)),$$

达布和收敛到积分，且不依赖于 t_i^* 的具体选取。令

$$S_n=\sum_{i=0}^{n-1}B(t_i)(B(t_{i+1})-B(t_i)),$$

$$S_n'=\sum_{i=0}^{n-1}B(t_{i+1})(B(t_{i+1})-B(t_i)),$$

则 $n\to\infty$ 时

$$\|S_n-S_n'\|^2=\left\|\sum_{i=0}^{n-1}(B(t_{i+1})-B(t_i))^2\right\|^2\to t\neq 0,$$

说明黎曼-斯蒂尔杰斯积分不存在。所以在伊藤积分中 t_i^* 固定地取为区间的左端点 t_i，这可以保证伊藤积分过程的鞅性。

8.3 Itô 积分定义的鞅

设 $\{B(t),\mathscr{F}(t),t\geq 0\}$ 为适应的标准布朗运动，$B(t)-B(s)$ 与 $\mathscr{F}(s)$ 独立 $(\forall t\geq s\geq 0)$。设 $\{X(t,\omega)\}\in V$，

$$Y(t,\omega)=\int_0^t X(s,\omega)\,dB(s),\ t\in[0,T],$$

$\{Y(t,\omega),t\in[0,T]\}$ 为随机过程，此随机过程有如下性质：

性质：

- (1) 适应性：$\{Y(t),\mathscr{F}(t),t\in[0,T]\}$ 是适应过程。
- (2) 零均值：对 $X\in V$ 有 $EY(t)=0$。
- (3) 线性性质：对 $\alpha_1,\alpha_2\in\mathbb{R}$, $\xi,\eta\in V$，有 $\int_0^T(\alpha_1\xi+\alpha_2\eta)\,dB(t)=\alpha_1\int_0^T\xi\,dB(t)+\alpha_2\int_0^T\eta\,dB(t)$。
- (4) 等距性：$E\left[\int_0^T X\,dB(t)\right]^2=\int_0^T EX^2\,dt$。
- (5) 推广的等距性：对 $\xi,\eta\in V$ 有 $E\left[\int_0^T\xi(t)\,dB(t)\ \int_0^T\eta(t)\,dB(t)\right]=\int_0^T E(\xi(t)\eta(t))\,dt$。
- (6) 鞅性：$\{Y(t),\mathscr{F}(t),t\in[0,T]\}$ 是鞅。
- (7) 轨道连续性：存在轨道连续的 $\{Z(t),t\in[0,T]\}$ 使得 $P(Y(t)=Z(t))=1$，$\forall t\in[0,T]$(即 Y 的修改)。可以认为 $\{Y(t)\}$ 轨道连续。

8.3.1 鞅性

关于 Itô 积分过程的鞅性质：

定理 8.2. 设 $\{X(t)\} \in V$，对 $0 \le s < t$，$B(t)-B(s)$ 与 $\mathscr{F}(s)$ 独立，则

$$Y(t)=\int_0^t X(s)dB(s),\ 0 \le t \le T,$$

是零均值的连续的二阶矩有界的鞅.

证明：零均值用可料阶梯过程逼近可得。连续性证明略去。

关于二阶矩有界，由等距性，

$$E\left(\int_0^t X(u)\,dB(u)\right)^2=\int_0^t E[X^2(u)]\,du \le \int_0^T E[X^2(u)]\,du < \infty,$$

故二阶矩有界。

来证明鞅性。对 $\{X(t)\}$ 存在可料阶梯过程 $\phi_n(t)$ 使得

$$\begin{aligned}
&\int_0^T E|X(t)-\phi_n(t)|^2\,dt \to 0\ (n\to\infty),\\
&E\left|\int_0^t X(t)\,dB(t)-\int_0^t \phi_n(t)\,dB(t)\right|^2 \to 0,\ (n\to\infty), \forall 0<t\le T.
\end{aligned}$$

对 $0 \le s < t \le T$，$\forall A \in \mathscr{F}_s$，

$$\begin{aligned}
&\left|E\left\{I_A\int_0^t X(u)\,dB(u)\right\}-E\left\{I_A\int_0^t \phi_n(u)\,dB(u)\right\}\right|\\
=&\left|E\left\{I_A\int_0^t [X(u)-\phi_n(u)]\,dB(u)\right\}\right|\\
\le&\left(EI_A^2\right)^{1/2}\left(\int_0^t E[X(u)-\phi_n(u)]^2\,du\right)^{1/2}\\
\le&\left(P(A)\right)^{1/2}\left(\int_0^T E[X(u)-\phi_n(u)]^2\,du\right)^{1/2}\\
\to&0.
\end{aligned}$$

同理上面 t 替换成 s 时也成立。由可料阶梯过程的 Itô 积分的鞅性（见8.2.3.5），有

$$E\left(\int_0^t \phi_n(u)\,dB(u)|\mathscr{F}_s\right)=\int_0^s \phi_n(u)\,dB(u),$$

于是

$$
\begin{aligned}
& E\left(I_A \int_0^t \phi_n(u)\, dB(u)\right) \\
=& E\left\{E\left(I_A \int_0^t \phi_n(u)\, dB(u) \middle| \mathscr{F}_s\right)\right\} \\
=& E\left\{I_A E\left(\int_0^t \phi_n(u)\, dB(u) \middle| \mathscr{F}_s\right)\right\} \\
=& E\left\{I_A \int_0^s \phi_n(u)\, dB(u)\right\},
\end{aligned}
$$

上式左侧趋于 $E\left(I_A \int_0^t X(u)\, dB(u)\right)$, 右侧趋于 $E\left(I_A \int_0^s X(u)\, dB(u)\right)$, 所以有

$$
E\left(I_A \int_0^t X(u)\, dB(u)\right) = E\left(I_A \int_0^s X(u)\, dB(u)\right), \ \forall A \in \mathscr{F}_s,\ 0 \le s < t.
$$

即

$$
E\left(\int_0^t X(u)\, dB(u) \middle| \mathscr{F}_s\right) = \int_0^s X(u)\, dB(u),\ 0 \le s < t.
$$

得证。

推论 8.1. 对任意有界的 *Borel* 可测函数 f, $\int_0^t f(B(s))\, dB(s)$ 是零均值、轨道连续、二阶矩有界的鞅.

证明 $X(t) = f(B(t))$ 是可测适应的, 并且由 f 的有界性可知存在常数 $K > 0$ 使得 $|f(x)| < K$, 于是 $\int_0^T E[f^2(B(s))]\, ds \le K^2 T < \infty$. 由定理8.2可得结论。

8.3.2 非随机被积函数时的高斯过程

上述定理提供了构造鞅的方法. 在节8.2中我们已经证明，非随机的阶梯函数的 Itô 积分是正态分布的随机变量. 更一般地，我们有下述定理.

定理 8.3. 如果 $g(t)$ 非随机的函数，且 $\int_0^T g^2(s)\, ds < \infty$，则

$$
Y(t) = \int_0^t g(s)\, dB(s),\ 0 \le t \le T
$$

是高斯过程，其均值函数为零，协方差函数为

$$Cov(Y(t),\, Y(t+u)) = \int_0^t g^2(s)\, ds,\ 0 \le t \le t+u \le T.$$

$\{Y(t)\}$ 也是二阶矩有界的轨道连续的鞅.

证明：关于 $\{Y(t)\}$ 是高斯过程，可以用阶梯函数逼近 $g(t)$，用阶梯函数的随机积分逼近 $g(t)$ 的随机积分，利用正态分布的极限分布为正态分布，以及有限维分布服从多元正态分布，当且仅当其任意线性组合服从一元正态分布。具体证明略。

因为被积函数是非随机的，所以

$$\int_0^t E[g^2(s)]ds = \int_0^t g^2(s)ds < \infty.$$

由定理8.2知 Y 为零均值、轨道连续、二阶矩有界的鞅。

为计算协方差函数，利用鞅性得

$$\begin{aligned}
&\mathrm{Cov}(Y(t), Y(t+u)) \\
=&E\left[\int_0^t g(s)\, dB(s) \cdot \int_0^{t+u} g(s)\, dB(s)\right] \\
=&E\left[E\left(\int_0^t g(s)\, dB(s) \cdot \int_0^{t+u} g(s)\, dB(s) \middle| \mathscr{F}_t\right)\right] \\
=&E\left[\int_0^t g(s)\, dB(s) \cdot E\left(\int_0^{t+u} g(s)\, dB(s) \middle| \mathscr{F}_t\right)\right] \\
=&E\left[\int_0^t g(s)\, dB(s) \int_0^t g(s)\, dB(s)\right] \\
=&E\left[\int_0^t g(s)\, dB(s)\right]^2 \\
=&\int_0^t E[g^2(s)]\, ds = \int_0^t g^2(s)\, ds.
\end{aligned}$$

证毕。

证明也可参见 (Shreve 2004) 定理 4.4.9 和例 4.7.3。

例 8.5. 根据定理8.3可得 $J = \int_0^t s dB(s) \sim N(0, \frac{t^3}{3})$.

定理 8.4. 设 $\{X(t)\} \in V^*$，对 $0 \le s < t$，$B(t) - B(s)$ 与 $\mathscr{F}(s)$ 独立，

$$Y(t) = Y(0) + \int_0^t X(s)dB(s),\ 0 \le t \le T,$$

如果 $Y(t) \ge 0,\ \forall t \in [0,T]$，则 $\{Y(t), \mathscr{F}(t), t \in [0,T]\}$ 是上鞅；如果进一步地 $E[Y(t)] = c$ 不依赖于 t，则 $\{Y(t), \mathscr{F}(t), t \in [0,T]\}$ 是鞅。

证明略，参见 (Glasserman 2004) 定理 B.3.1。

8.3.3 二次变差

下面讨论 Itô 积分的二次变差.

定义 8.4. 设 $Y(t) = \int_0^t X(s)\, dB(s),\ 0 \le t \le T$ 是 Itô 积分，如果在依概率收敛的意义下，极限

$$\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} |Y(t_{i+1}^n) - Y(t_i^n)|^2$$

当 $\{t_i^n\}_{i=0}^n$ 遍取 $[0,t]$ 的分割，且其模 $\delta_n = \max_{0 \le i \le n-1}(t_{i+1}^n - t_i^n) \to 0$ 时存在，则称此极限为 Y 的**二次变差**，记为 $[Y,Y](t)$.

定理 8.5. 设 $Y(t) = \int_0^t X(s)\, dB(s),\ 0 \le t \le T$ 是 *Itô* 积分过程，则 Y 的二次变差为

$$[Y,Y](t) = \int_0^t X^2(s)\, ds.$$

证明： 这里仅考虑 $\{X(s)\}$ 为可料阶梯过程的情形，对一般情形，我们可以用可料阶梯过程逼近的方法得到。

不妨假定 $X(s)$ 在 $[0,T]$ 上只取两个不同的值, 取任意有限多个值的情形可类似证之. 为简单起见，设 $T = 1$，$X(t)$ 在 $[0, 1/2]$ 上取 ξ_0，在 $(1/2, 1]$ 上取 ξ_1，即

$$X(t) = \xi_0 I_{[0,\frac{1}{2}]}(t) + \xi_1 I_{(\frac{1}{2},1]}(t).$$

于是

$$\begin{aligned} Y(t) &= \int_0^t X(s)\, dB(s) \\ &= \begin{cases} \xi_0 B(t), & \text{若} 0 \le t \le \frac{1}{2} \\ \xi_0 B(\frac{1}{2}) + \xi_1 (B(t) - B(\frac{1}{2})), & \text{若} \frac{1}{2} < t \le 1. \end{cases} \end{aligned}$$

因此，对 $[0,t]$ 的任何分割，有

$$Y(t_{i+1}^n)-Y(t_i^n)=\begin{cases}\xi_0(B(t_{i+1}^n)-B(t_i^n)) & \text{当} t_i^n<t_{i+1}^n\le\frac{1}{2}\\ \xi_1(B(t_{i+1}^n)-B(t_i^n)) & \text{当} \frac{1}{2}\le t_i^n<t_{i+1}^n\le 1.\end{cases}$$

当 $t\le\frac{1}{2}$ 时，

$$\begin{aligned}[Y,Y](t)=&\lim_{n\to\infty}\sum_{i=0}^{n-1}(Y(t_{i+1}^n)-Y(t_i^n))^2\\ =&\xi_0^2\lim_{n\to\infty}\sum_{i=0}^{n-1}(B(t_{i+1}^n)-B(t_i^n))^2\\ =&\xi_0^2[B,B](t)=\xi_0^2t=\int_0^t X^2(s)ds.\end{aligned}$$

当 $t>\frac{1}{2}$ 时，

$$\begin{aligned}[Y,Y](t)=&\lim_{n\to\infty}\sum_{i=0}^{n-1}(Y(t_{i+1}^n)-Y(t_i^n))^2\\ =&\xi_0^2\lim_{n\to\infty}\sum_{t_i<\frac{1}{2}}(B(t_{i+1}^n)-B(t_i^n))^2+\xi_1^2\lim_{n\to\infty}\sum_{t_i>\frac{1}{2}}(B(t_{i+1}^n)-B(t_i^n))^2\\ =&\xi_0^2[B,B](\frac{1}{2})+\xi_1^2[B,B]((\frac{1}{2},t])=\int_0^t X^2(s)\,ds.\end{aligned}$$

这里的极限都是当 $\delta_n\to 0$ 时在依概率收敛意义下的极限.

对同一个布朗运动 $\{B(t)\}$ 关于两个不同被积函数的 Itô 积分

$$Y_1(t)=\int_0^t X_1(s)\,dB(s),\quad Y_2(t)=\int_0^t X_2(s)\,dB(s),$$

由于

$$Y_1(t)+Y_2(t)=\int_0^t(X_1(s)+X_2(s))\,dB(s),$$

我们可以定义 Y_1 和 Y_2 的**二次协变差**:

$$[Y_1,Y_2](t)=\frac{1}{2}\left([Y_1+Y_2,Y_1+Y_2](t)-[Y_1,Y_1](t)-[Y_2,Y_2](t)\right).$$

由定理8.5，有

$$[Y_1,Y_2](t)=\int_0^t X_1(s)X_2(s)\,ds.$$

二次协变差也可以定义为

$$[Y_1, Y_2](t) = \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} [Y_1(t_{i+1}^n) - Y_1(t_i^n)][Y_2(t_{i+1}^n) - Y_2(t_i^n)],$$

其中极限为依概率极限。

事实上，

$$\begin{aligned}
&\lim_{\delta_n \to \infty} \sum_{i=0}^{n-1} [Y_1(t_{i+1}^n) - Y_1(t_i^n)][Y_2(t_{i+1}^n) - Y_2(t_i^n)] \\
=&\frac{1}{2} \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} \{[Y_1(t_{i+1}^n) - Y_1(t_i^n) + Y_2(t_{i+1}^n) - Y_2(t_i^n)]^2 \\
&-[Y_1(t_{i+1}^n) - Y_1(t_i^n)]^2 - [Y_2(t_{i+1}^n) - Y_2(t_i^n)]^2\} \\
=&\frac{1}{2} \left\{ \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} [(Y_1(t_{i+1}^n) + Y_2(t_{i+1}^n)) - (Y_1(t_i^n) + Y_2(t_i^n))]^2 \right. \\
&- \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} [Y_1(t_{i+1}^n) - Y_1(t_i^n)]^2 \\
&\left. - \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} [Y_2(t_{i+1}^n) - Y_2(t_i^n)]^2 \right\} \\
=&\frac{1}{2} \{[Y_1 + Y_2, Y_1 + Y_2](t) - [Y_1, Y_1](t) - [Y_2, Y_2](t)\}.
\end{aligned}$$

8.4 Itô 公式

Itô 公式，随机分析中的变量替换公式或链锁法则，是随机分析中的一个主要工具，许多重要的公式，例如 Dynkin 公式，Feynman-Kac 公式以及分部积分公式，都是由 Itô 公式导出的.

问题：设 $f(\cdot)$ 为可微函数，是否有

$$f(B(T)) - f(0) = \int_0^T f'(B(u))\, dB(u)?$$

可以证明这个公式不成立。Itô 公式则是解决了类似的问题。

8.4.1 布朗运动微分的运算

因为布朗运动在 $[0,t]$ 上的二次变差为 t，即在依概率收敛的意义下

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

这里 $\{t_i^n\}$ 是 $[0,t]$ 的分割，$\delta_n=\max_{0\le i\le n-1}(t_{i+1}^n-t_i^n)$, 上式从形式上可表示为

$$\int_0^t [dB(s)]^2=\int_0^t ds=t,$$

或

$$[dB(t)]^2=dt.$$

上面结果的严格意义为以下定理.

定理 8.6. 设 $g(x)$ 是 $\mathbb{R}$ 上的连续函数，满足 $\int_0^T E[g^2(B(u))]\,du<\infty$，$\{t_i^n\}$ 是 $[0,t]$ 的分割，则对 $B(t_i^n)$ 和 $B(t_{i+1}^n)$ 之间的任意值 θ_i^n，依概率收敛意义下有

$$\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}g(\theta_i^n)[B(t_{i+1}^n)-B(t_i^n)]^2=\int_0^t g(B(s))\,ds. \tag{8.5}$$

可以将结论记作

$$[dB(t)]^2=dt.$$

证明: 首先取 $\theta_i^n=B(t_i^n)$，由 $g(x)$ 的连续性和 $B(s)$ 的轨道连续性可知 $g(B(s))$ 的轨道 a.s. 地可积，按黎曼积分定义有

$$\sum_{i=0}^{n-1}g(B(t_i^n))(t_{i+1}^n-t_i^n)\to\int_0^t g(B(s))\,ds,\ \text{a.s.} \tag{8.6}$$

存在有限。这个极限是 a.s. 收敛，也依概率收敛。

我们来证明

$$\left\|\sum_{i=0}^{n-1}g(B(t_i^n))(B(t_{i+1}^n)-B(t_i^n))^2-\sum_{i=0}^{n-1}g(B(t_i^n))(t_{i+1}^n-t_i^n)\right\|^2\to 0\ (n\to\infty), \tag{8.7}$$

这样，$\sum_{i=0}^{n-1}g(B(t_i^n))(B(t_{i+1}^n)-B(t_i^n))^2$ 与 $\sum_{i=0}^{n-1}g(B(t_i^n))(t_{i+1}^n-t_i^n)$ 有相同的依概率极限，由(8.6)和(8.7)就可知

$$\sum_{i=0}^{n-1}g(B(t_i^n))(B(t_{i+1}^n)-B(t_i^n))^2\to\int_0^t g(B(s))\,ds, \tag{8.8}$$

其中的极限是依概率收敛。

记 $\Delta B_i = B(t_{i+1}^n) - B(t_i^n)$, $\Delta t_i = t_{i+1}^n - t_i^n$，则由布朗运动的独立增量性可知 ΔB_j 与 $\mathscr{F}_{t_j}$ 独立，有

$$
\begin{aligned}
& E\{g(B(t_i^n))g(B(t_j^n))\ [(\Delta B_i)^2 - \Delta t_i]\ [(\Delta B_j)^2 - \Delta t_j]\} \\
=& E\left\{E\left(g(B(t_i^n))g(B(t_j^n))[(\Delta B_i)^2 - \Delta t_i][(\Delta B_j)^2 - \Delta t_j] \mid \mathscr{F}_{t_j}\right)\right\} \\
=& E\left\{g(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]g(B(t_j^n))\ E\left([(\Delta B_j)^2 - \Delta t_j] \mid \mathscr{F}_{t_j}\right)\right\} \\
=& E\left\{g(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]g(B(t_j^n))\ E[(\Delta B_j)^2 - \Delta t_j]\right\} \\
& = 0,\ \forall i < j,
\end{aligned}
$$

于是

$$
\begin{aligned}
& E\left[\sum_{i=0}^{n-1} g(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]\right]^2 \\
=& E\left(\sum_{i=0}^{n-1} g^2(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]^2\right) \\
& + 2\sum_{i<j} E\left(g(B(t_i^n))g(B(t_j^n))[(\Delta B_i)^2 - \Delta t_i][(\Delta B_j)^2 - \Delta t_j]\right) \\
=& E\left(\sum_{i=0}^{n-1} g^2(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]^2\right) + 0 \\
=& \sum_{i=0}^{n-1} E\left(g^2(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]^2\right) \\
=& \sum_{i=0}^{n-1} E\left\{E\left(g^2(B(t_i^n))[(\Delta B_i)^2 - \Delta t_i]^2 \mid \mathscr{F}_{t_i}\right)\right\} \\
=& \sum_{i=0}^{n-1} E\left\{g^2(B(t_i^n))E\left([(\Delta B_i)^2 - \Delta t_i]^2 \mid \mathscr{F}_{t_i}\right)\right\} \\
=& \sum_{i=0}^{n-1} E\left\{g^2(B(t_i^n))E\left([(\Delta B_i)^2 - \Delta t_i]^2\right)\right\},
\end{aligned}
$$

其中 $\Delta B_i \sim \mathrm{N}(0, \Delta t_i)$，令 $\Delta B_i = \sqrt{\Delta t_i} Z$，$Z \sim \mathrm{N}(0,1)$，可得

$$
\begin{aligned}
E\left([(\Delta B_i)^2 - \Delta t_i]^2\right) =& (\Delta t_i)^2 E(Z^2 - 1)^2 \\
=& (\Delta t_i)^2 (EZ^4 - 2EZ^2 + 1) = 2(\Delta t_i)^2.
\end{aligned}
$$

从而

$$
\begin{aligned}
&E\left[\sum_{i=0}^{n-1} g(B(t_i^n))[(\Delta B_i)^2-\Delta t_i]\right]^2\\
=&2\sum_{i=0}^{n-1} E[g^2(B(t_i^n))]\Delta t_i^2\\
\leq&2\delta_n\sum_{i=0}^{n-1} E[g^2(B(t_i^n))]\Delta t_i\\
\to&0\times\int_0^t E[g^2(B(t))]\,dt=0.
\end{aligned}
$$

因此，在均方收敛的意义下

$$
\sum_{i=0}^{n-1} g(B(t_i^n))((\Delta B_i)^2-\Delta t_i)\to 0,
$$

其依概率极限也趋于 0，从而 $\sum_{i=0}^{n-1} g(B(t_i^n))(B(t_{i+1}^n)-B(t_i^n))^2$ 与 $\sum_{i=0}^{n-1} g(B(t_i^n))(t_{i+1}^n-t_i^n)$ 有相同的依概率极限。

再来考虑一般的 θ_i^n 的选取。对 $B(t_i^n)$ 和 $B(t_{i+1}^n)$ 之间的任意值 θ_i^n，当 $\delta_n\to 0$ 时，

$$
\begin{aligned}
&\sum_{i=0}^{n-1}[g(\theta_i^n)-g(B(t_i^n))](\Delta B_i)^2\\
\leq&\max_i[(g(\theta_i^n)-g(B(t_i^n))]\sum_{i=0}^{n-1}[B(t_{i+1}^n)-B(t_i^n)]^2,
\end{aligned}
$$

由 $g(x)$ 连续以及 $B(t)$ 轨道连续，可知 $g(B(t))$ 每条轨道为闭区间 $[0,t]$ 上的连续函数，每条轨道一致连续从而 $\max_i[g(\theta_i^n)-g(B(t_i^n)]\to 0$, a.s. 由布朗运动二次变差性质得 $\sum_{i=0}^{n-1}[B(t_{i+1}^n)-B(t_i^n)]^2\to t$(依概率收敛)。于是当 $\delta_n\to 0$ 时，

$$
\sum_{i=0}^{n-1}[g(\theta_i^n)-g(B(t_i^n))](\Delta B_i)^2\to 0,
$$

极限为依概率收敛。因此 $\sum_{i=0}^{n-1} g(\theta_i^n)(\Delta B_i)^2$ 与 $\sum_{i=0}^{n-1} g(B(t_i^n))(\Delta B_i)^2$ 具有相同的依概率收敛意义的极限 $\int_0^t g(B(s))ds$, 定理结论成立。

注：若 $\|\xi_n-\eta_n\|^2=E(\xi_n-\eta_n)^2\to 0$，则 $\xi_n-\eta_n$ 依概率趋于 0，若 $\xi_n\to\xi$

为依概率收敛，则 $\eta_n \to \xi$ 依概率收敛。事实上

$$\eta_n - \xi = (\xi_n - \xi) + (\eta_n - \xi_n) = o_p(1) + o_p(1) = o_p(1),$$

其中 $o_p(1)$ 表示依概率趋于 0 的随机变量序列，两个这样的序列的和仍为依概率趋于 0。

如果 $\xi_n \to 0$(依概率收敛)，$\eta_n \to \eta$(依概率收敛)，则 $\xi_n \eta_n \to 0$(依概率收敛)。这个性质记为 $o_p(1) \cdot O_p(1) = o_p(1)$。$O_p(1)$ 表示依概率有界的随机变量序列，依概率收敛或者依分布收敛都推出依概率有界。$\{X_n\}$ 称为依概率有界，若 $\forall \epsilon > 0, \exists M > 0$ 使得

$$\sup_{n \geq 1} P(|X_n| > M) < \epsilon.$$

与 $[dB(t)]^2 = dt$ 类似的结论有 $(dt)^2 = 0$, $dB(t)\, dt = 0$，其严格含义为如下定理。

定理 8.7. 设 $g(x)$ 是 $\mathbb{R}$ 上的连续函数，$\{t_i^n\}$ 是 $[0,t]$ 的分割，则对 $B(t_i^n)$ 和 $B(t_{i+1}^n)$ 之间的任意值 θ_i^n，有

$$\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} g(\theta_i^n)(t_{i+1}^n - t_i^n)^2 = 0, \ \ a.s., \tag{8.9}$$

$$\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} g(\theta_i^n)[B(t_{i+1}^n) - B(t_i^n)](t_{i+1}^n - t_i^n) = 0, \ \ a.s. \tag{8.10}$$

将这两个结论记作

$$dt\, dt = 0, \quad dB(t)\, dt = 0.$$

证明：

$$\begin{aligned}
&\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} |g(\theta_i^n)|(t_{i+1}^n - t_i^n)^2 \\
\leq &\lim_{\delta_n \to 0} \delta_n \sum_{i=0}^{n-1} |g(\theta_i^n)|(t_{i+1}^n - t_i^n) \\
= &\lim_{\delta_n \to 0} \delta_n \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} |g(\theta_i^n)|(t_{i+1}^n - t_i^n) \\
= &0 \times \int_0^t |g(B(s))|\, ds = 0, \text{ a.s.}
\end{aligned}$$

这时因为 $g(x)$ 连续而 $B(s)$ 轨道连续则 $|g(B(s))|$ 轨道连续，于是可积。

另一方面，

$$
\begin{aligned}
&\sum_{i=0}^{n-1}|g(\theta_i^n)|\cdot|B(t_{i+1}^n)-B(t_i^n)|\cdot(t_{i+1}^n-t_i^n)\\
\le&\max_i|B(t_{i+1}^n)-B(t_i^n)|\cdot\sum_{i=0}^{n-1}|g(\theta_i^n)|(t_{i+1}^n-t_i^n),
\end{aligned}
$$

因为 $B(s)$ 轨道连续所以在 $[0,t]$ 轨道一致连续，有

$$\lim_{\delta_n\to 0}\max_i|B(t_{i+1}^n)-B(t_i^n)|=0,\ \text{a.s.}$$

而

$$\sum_{i=0}^{n-1}|g(\theta_i^n)|(t_{i+1}^n-t_i^n)\to\int_0^t|g(B(s))|\,ds,\ \text{a.s.},$$

所以有

$$\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}|g(\theta_i^n)|\cdot|B(t_{i+1}^n)-B(t_i^n)|\cdot(t_{i+1}^n-t_i^n)=0.$$

8.4.2 Itô 公式

对于普通的微积分，我们有微分链式法则：

$$df(g(t))=f'(g(t))\,dg(t)=f'(g(t))g'(t),$$

和积分变量替换法则：

$$
\begin{aligned}
&\int_0^t f'(g(s))\,dg(s)\\
=&\int_{g(0)}^{g(t)}f'(u)\,du\\
=&f(g(t))-f(g(0)).
\end{aligned}
$$

对上式求导，则得到

$$f'(g(t))dg(t)=d[f(g(t))].$$

如果是 Itô 积分，是否有：

$$df(B(t))=f'(B(t))dB(t),$$

或者

$$\int_0^t f'(B(s))dB(s) = f(B(t)) - f(B(0))?$$

答案是否定的，因为 $B(t)$ 不可微。

我们给出布朗运动的变换的 Itô 公式，这相当于是对上述问题的修正结果。公式比普通微积分多出了一项 $\frac{1}{2}\int_0^t f''(B(s))\,ds$，这是由布朗运动的二次变差非零引起的。注意连续可微函数的二次变差等于零。

定理 8.8. 如果 $f(x)$ 是二次连续可微函数，则对任何 $t \geq 0$，有

$$f(B(t)) = f(0) + \int_0^t f'(B(s))\,dB(s) + \frac{1}{2}\int_0^t f''(B(s))\,ds. \tag{8.11}$$

证明：因为 $f'(B(s))$ 轨道连续，适应，$\int_0^t [f'(B(s))]^2\,ds < \infty$, a.s.，所以 $\int_0^t f'(B(s))\,dB(s)$ 的被积函数属于 V^*，是 Itô 积分。因为 $f''(B(s))$ 轨道连续所以 $\int_0^t f''(B(s))\,ds$ 在 a.s. 的黎曼积分意义下有定义。

取 $[0,t]$ 的分割 $\{t_i^n\}$，$0 = t_0 < t_1 < \cdots < t_n = t$, 有

$$f(B(t)) = f(0) + \sum_{i=0}^{n-1}[f(B(t_{i+1}^n)) - f(B(t_i^n))].$$

对 $f(B(t_{i+1}^n)) - f(B(t_i^n))$ 每条轨道应用 Taylor 公式得

$$f(B(t_{i+1}^n))-f(B(t_i^n)) = f'(B(t_i^n))[B(t_{i+1}^n)-B(t_i^n)]+\frac{1}{2}f''(\theta_i^n)[B(t_{i+1}^n)-B(t_i^n)]^2,$$

其中 θ_i^n 取值于 $B(t_i^n)$ 和 $B(t_{i+1}^n)$ 之间. 于是

$$\begin{aligned} f(B(t)) =& f(0) + \sum_{i=0}^{n-1} f'(B(t_i^n))[B(t_{i+1}^n) - B(t_i^n)] \\ &+ \frac{1}{2}\sum_{i=0}^{n-1} f''(\theta_i^n)[B(t_{i+1}^n) - B(t_i^n)]^2. \end{aligned}$$

令 $\delta_n \to 0$ 取极限，则上式的第一个和收敛于 Itô 积分 $\int_0^t f'(B(s))\,dB(s)$, 利用定理8.6可知第二个和收敛于 $\int_0^t f''(B(s))\,ds$, 都是依概率极限。

注 1：式(8.11)称为布朗运动的伊藤公式（Itô 公式）。也可以形式地写成

$$df(B(t)) = f'(B(t))\,dB(t) + \frac{1}{2}f''(B(t))\,dt. \tag{8.12}$$

为了帮助记忆(8.12)，可以利用 $f(B(t))$ 的泰勒展开形式地认识到

$$\begin{aligned}&df(B(t))\\=&f'(B(t))\,dB(t)+\frac{1}{2}f''(B(t))[dB(t)]^2+o([dB(t)]^2)\\=&f'(B(t))\,dB(t)+\frac{1}{2}f''(B(t))\,dt+o(dt),\end{aligned}$$

其中 $[dB(t)]^2=dt$ 的严格含义是定理8.6。

注 2：在定理8.8的证明中，为什么不直接使用一阶泰勒展开，如：

$$f(B(t_{i+1}^n))-f(B(t_i^n))=f'(\theta_i^n)[B(t_{i+1}^n)-B(t_i^n)],$$

其中 θ_i^n 取值于 $B(t_i^n)$ 和 $B(t_{i+1}^n)$ 之间, 从而

$$\begin{aligned}f(B(t))=&f(0)+\lim_{\delta_n\to 0}\sum_{i=0}^{n-1}f'(\theta_i^n)[B(t_{i+1}^n)-B(t_i^n)]\\=&f(0)+\int_0^t f'(B(s))\,dB(s)?\end{aligned}$$

这是因为 θ_i^n 的选取会影响到极限结果，参见例8.4。于是进行了二阶泰勒展开，使得 $f'(\cdot)$ 的自变量固定取为小区间左端点处的值 $B(t_i^n)$，这恰好符合 Itô 积分定义。

注 3：为了使得定理更严格，可以增加条件

$$\int_0^T E\{[f'(B(s))]^2\}\,ds<\infty,\quad \int_0^T |f''(B(s))|\,ds<\infty,\ \text{a.s.}$$

但 Itô 公式在更一般的条件下适用。

例 8.6. 计算 $\int_0^t B(u)\,dB(u)$。

解：在例8.1中我们已经计算了这个积分的均值和方差，并用定义给出了积分结果。下面用 Itô 公式计算。这里 $f'(x)=x$，从而 $f(x)=\frac{1}{2}x^2$，$f''(x)=1$，则

由伊藤公式

$$\begin{aligned}\frac{1}{2}B^2(t) =&f(B(t))\\ =&f(0)+\int_0^t f'(B(s))\,dB(s)+\frac{1}{2}\int_0^t f''(B(s))\,ds\\ =&0+\int_0^t B(s)\,dB(s)+\frac{1}{2}\int_0^t 1\,ds\\ =&\int_0^t B(s)\,dB(s)+\frac{t}{2}.\end{aligned}$$

即有

$$\int_0^t B(s)\,dB(s)=\frac{1}{2}B^2(t)-\frac{t}{2}.$$

如果 $B(t)$ 换成某个可微函数 $h(t)$，$h(0)=0$，则结果应为

$$\int_0^t f'(h(u))dh(u)=\int_0^t df(h(u))=\frac{1}{2}h^2(t),$$

没有 $-\frac{t}{2}$ 这一项。

例 8.7. 计算 $d(e^{B(t)})$。

解： 令 $f(B(t))=e^{B(t)}$，注意 $f'(x)=f''(x)=e^x$, 所以

$$\begin{aligned}d(e^{B(t)}) =&df(B(t))\\ =&f'(B(t))\,dB(t)+\frac{1}{2}f''(B(t))\,dt\\ =&e^{B(t)}\,dB(t)+\frac{1}{2}e^{B(t)}\,dt.\end{aligned}$$

8.4.3 Itô 过程

由式(8.11)看出，布朗运动的函数可以表示为一个 Itô 积分加上一个轨道绝对连续的过程. 我们称这类过程为 Itô 过程，严格地，我们有下面定义.

定义 8.5. 如果过程 $\{Y(t),0\le t\le T\}$ 可以表示为

$$Y(t)=Y(0)+\int_0^t \mu(s)\,ds+\int_0^t \sigma(s)dB(s),\ 0\le t\le T, \tag{8.13}$$

其中过程 $\{\mu(t)\}$ 和 $\{\sigma(t)\}$ 满足

(1) $\mu(t)$ 是适应的并且 $\int_0^T |\mu(t)|\,dt < \infty$, a.s.,

(2) $\{\sigma(t)\} \in V^*$.

则称 $\{Y(t)\}$ 为 **Itô 过程**.

有时我们也将 Itô 过程(8.13)记为微分的形式

$$dY(t) = \mu(t)\,dt + \sigma(t)\,dB(t),\ 0 \le t \le T. \tag{8.14}$$

其中函数 $\mu(t)$ 称为漂移系数, $\sigma(t)$ 称为扩散系数, 它们可以依赖于 $Y(t)$ 或 $B(t)$，甚至过去的路径 $\{B(s), 0 \le s \le t\}$，例如 $\mu(t) = \cos(M(t)+t)$，这里 $M(t) = \max_{0\le s\le t} B(s)$.

一类非常重要的情形是 $\mu(t)$ 与 $\sigma(t)$ 仅仅通过 $Y(t)$ 依赖于 t，在这种情况下，式(8.14)应改写为

$$dY(t) = \mu(Y(t))\,dt + \sigma(Y(t))\,dB(t),\ 0 \le t \le T. \tag{8.15}$$

例 8.8. 股票投资的随机微分方程。

解：在金融应用中，股票的价格 $S(t)$ 是用随机微分方程

$$dS(t) = \mu S(t)\,dt + \sigma S(t)\,dB(t)$$

描述的. 上述方程也可以写成

$$\frac{dS(t)}{S(t)} = \mu\,dt + \sigma\,dB(t),$$

即股价增长率包括一个恒定速率的部分和一个随机部分，随机部分的方差为 $\sigma^2\,dt$。

如果 $a(t)$ 表示在时刻 t 投资者的股票持仓量，那么在整个时间区间 $[0,T]$ 内的收益为

$$\int_0^T a(t)\,dS(t) = \mu\int_0^T a(t)S(t)\,dt + \sigma\int_0^T a(t)S(t)\,dB(t).$$

关于 Itô 过程可以给出其链式法则的 Itô 公式：

定理 8.9. 设 $\{X(t)\}$ 是由

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dB(t)$$

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

$$\{Y(t) = g(t, X(t)), t \geq 0\}$$

仍为 *Itô* 过程，并且

$$dY(t) = \frac{\partial g}{\partial t}(t, X(t))\,dt + \frac{\partial g}{\partial x}(t, X(t))\,dX(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X(t)) \cdot (dX(t))^2, \tag{8.16}$$

其中 $(dX(t))^2 = (dX(t)) \cdot (dX(t))$ 按照下面规则计算：

$$dt \cdot dt = dt \cdot dB(t) = dB(t) \cdot dt = 0, \quad dB(t) \cdot dB(t) = dt.$$

即

$$dY(t) = \left(\frac{\partial g}{\partial t}(t, X(t)) + \frac{\partial g}{\partial x}(t, X(t))\mu(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X(t))\sigma^2(t)\right)\,dt \tag{8.17}$$

$$+ \frac{\partial g}{\partial x}(t, X(t))\sigma(t)dB(t). \tag{8.18}$$

特别地，如果 $g(t,x) = g(x)$ 只是 x 的函数，(8.18)简化为

$$dY(t) = \left[g'(X(t))\mu(t) + \frac{1}{2}g''(X(t))\sigma^2(t)\right]\,dt \tag{8.19}$$

$$+ g'(X(t))\sigma(t)\,dB(t). \tag{8.20}$$

证明：我们形式地对公式进行证明。将 $g(t,x)$ 进行二阶 Taylor 展开并取 $x = X(t)$，有

$$\begin{aligned}
dg(t, X(t)) =& \frac{\partial}{\partial t}g(t, X(t))\,dt + \frac{\partial}{\partial x}g(t, X(t))\,dX(t) \\
&+ \frac{1}{2}\frac{\partial^2}{\partial t^2}g(t, X(t))\,dt\,dt \\
&+ \frac{1}{2}\frac{\partial^2}{\partial x^2}g(t, X(t))\,dX(t)\,dX(t) \\
&+ \frac{\partial^2}{\partial t\partial x}g(t, X(t))\,dt\,dX(t) \\
=& \frac{\partial}{\partial t}g(t, X(t))\,dt + \frac{\partial}{\partial x}g(t, X(t))\,dX(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}g(t, X(t))\,dX(t)\,dX(t).
\end{aligned}$$

将 $dX(t) = \mu(t)\,dt + \sigma(t)\,dB(t)$ 代入，并注意

$$\begin{aligned}
&dX(t)\,dX(t)\\
=&\mu^2(t)\,dt\,dt + \sigma^2(t)\,dB(t)\,dB(t) + 2\mu(t)\sigma(t)\,dt\,dB(t)\\
=&\sigma^2(t)\,dt,
\end{aligned}$$

简记 $\frac{\partial}{\partial t}g(t, X(t))$ 为 $\frac{\partial g}{\partial t}$ 并类似简化，则有

$$\begin{aligned}
&dg(t, X(t))\\
=&\frac{\partial g}{\partial t}\,dt + \frac{\partial g}{\partial x}[\mu(t)\,dt + \sigma(t)\,dB(t)]\\
&+ \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\sigma^2(t)\,dt\\
=&\left[\frac{\partial g}{\partial t} + \frac{\partial g}{\partial x}\mu(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\sigma^2(t)\right]\,dt\\
&+ \frac{\partial g}{\partial x}\sigma(t)\,dB(t).
\end{aligned}$$

若 $g(t, x) = g(x)$，则上式中 $\frac{\partial g}{\partial t} = 0$, $\frac{\partial g}{\partial x} = g'(X(t))$, $\frac{\partial^2 g}{\partial x^2} = g''(X(t))$，公式变成

$$\begin{aligned}
&dg(X(t))\\
=&\left[g'(X(t))\mu(t) + \frac{1}{2}g''(X(t))\sigma^2(t)\right]\,dt\\
&+ g'(X(t))\sigma(t)\,dB(t).
\end{aligned}$$

例 8.9. 讨论几何布朗运动的模型。

解： 设 $\{X(t)\}$ 满足

$$dX(t) = (r - \frac{1}{2}\sigma^2)\,dt + \sigma dB(t).$$

设 $S(0)$ 为常数，

$$S(t) = S(0)e^{X(t)},$$

令 $g(x) = S(0)e^x$，则 $g'(X(t)) = g''(X(t)) = S(t)$, $dX(t)\,dX(t) = \sigma^2\,dt$，

$$
\begin{aligned}
&dS(t)\\
=&\left[g'(X(t))(r-\frac{1}{2}\sigma^2)+\frac{1}{2}g''(X(t))\sigma^2\right]\\
&+g'(X(t))\sigma\,dB(t)\\
=&\left[S(t)(r-\frac{1}{2}\sigma^2)+\frac{1}{2}S(t)\sigma^2\right]\,dt\\
&+S(t)\sigma\,dB(t)\\
=&S(t)r\,dt+S(t)\sigma\,dB(t).
\end{aligned}
$$

所以价格 $S(t)$ 的模型也是 Itô 过程，漂移系数为 $rS(t)$，扩散系数为 $\sigma S(t)$。

将方程改写为

$$\frac{dS(t)}{S(t)} = r\,dt + \sigma\,dB(t),$$

左侧表示股票的瞬时收益率，或对数收益率 $d\log(S(t))$，它服从 $\mathrm{N}(r\,dt, \sigma^2\,dt)$。由

$$d\log(S(t)) = r\,dt + \sigma\,dB(t)$$

可得

$$
\begin{aligned}
\int_0^t d\log(S(u)) =&\log(S(t)) - \log(S(0))\\
=&\int_0^t r\,dt + \int_0^t \sigma\,dB(u)\\
=&rt + \sigma B(t) \sim \mathrm{N}(rt, \sigma^2 t),
\end{aligned}
$$

即 $S(t)/S(0)$ 服从对数正态分布 $\mathrm{LN}(rt, \sigma^2 t)$。

反过来，已知 $S(t)$ 服从如下的随机微分方程：

$$\frac{dS(t)}{S(t)} = r\,dt + \sigma\,dB(t),$$

则 $S(t)$ 的漂移系数为 $rS(t)$，扩散系数为 $\sigma S(t)$，取 $g(S(t)) = \log(S(t)) -$

$\log(S(0))$，有 $g'(S(t)) = [S(t)]^{-1}$, $g''(S(t)) = -[S(t)]^{-2}$, 用 Itô 公式得

$$\begin{aligned}
dX(t) =& dg(S(t)) \\
=& \left[g'(S(t))rS(t) + \frac{1}{2}g''(S(t))\sigma^2 S^2(t) \right] dt \\
& + g'(S(t))\sigma S(t)\, dB(t) \\
=& \left[[S(t)]^{-1}rS(t) + \frac{1}{2}(-1)[S(t)]^{-2}\sigma^2 S^2(t) \right] dt \\
& + [S(t)]^{-1}\sigma S(t)\, dB(t) \\
=& (r - \frac{1}{2}\sigma^2)\, dt + \sigma\, dB(t).
\end{aligned}$$

8.4.4 多维 Itô 公式

设 $B(t) = (B_1(t), \dots, B_d(t))^T$, $t \geq 0$ 是多维标准布朗运动，即每个分量为一维布朗运动且各分量随机过程之间相互独立，设 $\{\mathscr{F}_t, t \geq 0\}$ 为 σ 代数流，$\{B(t), \mathscr{F}_t, t \geq 0\}$ 为适应过程，且对任意 $0 \leq s < t$，$B(t) - B(s)$ 与 $\mathscr{F}_s$ 独立。

定义如下的多维 Itô 过程 $X(t) = (X_1(t), \dots, X_n(t))^T$:

$$dX_i(t) = \mu_i(t)\, dt + \sum_{j=1}^{d} \sigma_{ij}(t)\, dB_j(t),\ i = 1, 2, \dots, n.$$

用矩阵形式写成

$$dX(t) = \mu(t) + \Sigma(t)\, dB(t),$$

其中 $\mu(t) = (\mu_1(t), \dots, \mu_n(t))^T$，$\Sigma(t)$ 的 (i, j) 元素为 $\sigma_{ij}(t)$，是 $n \times d$ 矩阵。

设 $g(t, x) = (g_1(t, x), \dots, g_m(t, x))^T$ 是 $[0, \infty) \times \mathbb{R}^n$ 到 $\mathbb{R}^m$ 的二次连续可微函数，令

$$Y(t) = g(t, X(t)),\ 0 \leq t \leq T,$$

则 $\{Y(t)\}$ 仍为多维 Itô 过程，有如下的表达式。

定理 8.10 (多维 Itô 公式).

$$
\begin{aligned}
dY_k(t) =&\frac{\partial g_k(t,X(t))}{\partial t}\,dt \\
&+\sum_{j=1}^{n}\frac{\partial g_k(t,X(t))}{\partial x_j}dX_j(t) \\
&+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 g_k(t,X(t))}{\partial x_i\partial x_j}\,dX_i(t)\,dX_j(t) \\
=&\frac{\partial g_k(t,X(t))}{\partial t}\,dt \\
&+\frac{\partial g_k(t,X(t))}{\partial x^T}dX(t) \\
&+\frac{1}{2}dX^T(t)\frac{\partial^2 g_k(t,X(t))}{\partial x^T\partial x}dX(t), \\
&k=1,2,\dots,m.
\end{aligned}
$$

其中

$$
\begin{aligned}
dX(t) =&(dX_1(t),\dots,dX_n(t))^T, \\
\frac{\partial g_k(t,X(t))}{\partial x^T} =&\left(\frac{\partial g_k(t,X(t))}{\partial x_1},\dots,\frac{\partial g_k(t,X(t))}{\partial x_n}\right), \\
\frac{\partial^2 g_k(t,X(t))}{\partial x^T\partial x} =&\left(\frac{\partial^2 g_k(t,X(t))}{\partial x_i\partial x_j}\right)_{1\leq i\leq n;1\leq j\leq n}.
\end{aligned}
$$

在展开 $dX_i(t)$ 时，利用

$$
dB_i(t)\,dB_j(t)=\delta_{ij}\,dt,\quad dt\,dt=dt\,dB_i(t)=dB_i(t)\,dt=0.
$$

推论 8.2 (分部积分公式). 设 $X(t)$,$Y(t)$ 是基于标准布朗运动 $\{B(t),\mathscr{F}_t,t\geq 0\}$ 的 *Itô* 积分,

$$
\begin{aligned}
dX(t) =&\mu_X(t)\,dt+\sigma_X(t)\,dB(t), \\
dY(t) =&\mu_Y(t)\,dt+\sigma_Y(t)\,dB(t),
\end{aligned}
$$

则

$$
\begin{aligned}
d(X(t)Y(t)) =&X(t)\,dY(t)+Y(t)\,dX(t)+dX(t)\,dY(t) \\
=&X(t)\,dY(t)+Y(t)\,dX(t)+\sigma_X(t)\sigma_Y(t)\,dt,
\end{aligned}
$$

从而

$$
\int_0^t Y(u)\,dX(u)=X(t)Y(t)-\int_0^t X(u)\,dY(u)-\int_0^t \sigma_X(u)\sigma_Y(u)\,du.
$$

如果 $\sigma_X(t) \equiv 0$ 或者 $\sigma_Y(t) \equiv 0$，则分部积分公式与通常的微积分的分部积分形式相同：

$$\int_0^t Y(u)\,dX(u) = X(t)Y(t) - \int_0^t X(u)\,dY(u).$$

证明：用多维 Itô 公式，令 $g(t,x,y) = xy$，则

$$\frac{\partial g}{\partial t} = 0,\ \frac{\partial g}{\partial x} = y,\ \frac{\partial g}{\partial y} = x,$$
$$\frac{\partial^2 g}{\partial x^2} = \frac{\partial^2 g}{\partial y^2} = 0,\ \frac{\partial^2 g}{\partial x \partial y} = 1,$$

于是

$$\begin{aligned}
&d(X(t)Y(t)) = dg(t, X(t), Y(t)) \\
=&\frac{\partial g}{\partial t}\,dt + \frac{\partial g}{\partial x}\,dX(t) + \frac{\partial g}{\partial y}\,dY(t) \\
&+ \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX(t))^2 + \frac{1}{2}\frac{\partial^2 g}{\partial y^2}(dY(t))^2 \\
&+ \frac{\partial^2 g}{\partial x \partial y} dX(t)dY(t) \\
=&Y(t)\,dX(t) + X(t)\,dY(t) + dX(t)dY(t),
\end{aligned}$$

其中

$$\begin{aligned}
&dX(t)dY(t) \\
=&\mu_X(t)\mu_Y(t)\,dt\,dt + \mu_X(t)\sigma_Y(t)\,dt\,dB(t) + \sigma_X(t)\mu_Y(t)\,dB(t)\,dt \\
&+ \sigma_X(t)\sigma_Y(t)\,dB(t)\,dB(t) \\
=&\sigma_X(t)\sigma_Y(t)\,dt.
\end{aligned}$$

例 8.10. 用分部积分公式计算 $\int_0^t u\,dB(u)$。

解答：令 $dX(t) = 0\,dt +\ dB(t)$，$dY(t) = dt + 0\,dB(t)$, 由分部积分公式

$$\begin{aligned}
&\int_0^t u\,dB(u) = \int_0^t Y(u)\,dX(u) \\
=&X(t)Y(t) - \int_0^t X(u)\,dY(u) - \int_0^t 0 \cdot 1\,du \\
=&tB(t) - \int_0^t B(u)\,du.
\end{aligned}$$

8.5 补充：随机积分引入

先考虑以布朗运动的变换为被积函数的积分。

8.5.1 布朗运动连续变换的积分

8.5.1.1 存在性

定理 8.11. 设 $\{B(t), t \geq 0\}$ 为标准布朗运动，$g(t, x)$ 为连续函数，则

$$X = \int_a^b g(t, B(t))\, dt$$

为随机变量，其中 $0 \leq a < b$。

证明： 因为 $B(t)$ 轨道连续，所以 $g(t, B(t))$ 为 t 的连续函数（每条轨道）。取 $[a, b]$ 的分割 $\{t_i, i = 0, 1, \dots, n\}$, 使得 $\lim_{n\to\infty} \delta_n = \max\{t_{i+1} - t_i\} = 0$，令

$$S_n = \sum_{i=0}^{n-1} g(t_i^*, B(t_i^*))(t_{i+1} - t_i)$$

为达布和，其中 $t_i^* \in [t_i, t_{i+1}]$，则 S_n 为随机变量且 $n \to \infty$ 时 $S_n \to \int_a^b g(t, B(t))\, dt$, $\forall \omega \in \Omega$。随机变量序列有点点极限则极限必为随机变量。

注：设 $(\Omega, \mathscr{F}, P)$ 为概率空间，$\{X_n\}$ 为随机变量序列，

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \in (-\infty, \infty),\ \forall \omega \in \Omega,$$

则 X 为随机变量。事实上

$$\{X > x\} = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{X_m > x\},$$

由 $\{X_m > x\} \in \mathscr{F}$ 知 $\{X > x\} \in \mathscr{F}$，即 X 为随机变量。

8.5.1.2 期望

定理 8.12. 设 $g(t, x)$ 是 $[0, \infty) \times (-\infty, \infty)$ 上的可测函数，$0 \leq a < b < \infty$，$\int_a^b E|g(u, B(u))|\, du < \infty$，则

$$E \int_a^b g(u, B(u))\, du = \int_a^b E[g(u, B(u))]\, du.$$

证明：由 Fubini 定理，

$$
\begin{aligned}
& E\int_a^b g(u,B(u))\,du \\
=& \int_\Omega \int_a^b g(u,B(u))\,du\,dP(\omega) \\
=& \int_a^b \int_\Omega g(u,B(u))\,dP(\omega)\,du \\
=& \int_a^b E[g(u,B(u))]\,du.
\end{aligned}
$$

推论 8.3. 设 $g(t,x)$ 是 $[0,\infty)\times(-\infty,\infty)$ 上的有界连续函数，$0\le a<b<\infty$，则

$$E\int_a^b g(u,B(u))\,du = \int_a^b E[g(u,B(u))]\,du.$$

证明：设 $|g(t,x)|\le C$, 则

$$\int_a^b E|g(u,B(u))|\,du \le C(b-a)<\infty.$$

定理 8.13. 设 $g(x)$ 是 $(-\infty,\infty)$ 上可微函数，且 $|g'(x)|\le C$, $\forall x$。则

$$E\int_a^b g(B(u))\,du = \int_a^b E[g(B(u))]\,du.$$

证明: 取 $[a,b]$ 的分割 $\{t_i, i=0,1,\dots,n-1\}$, 使得 $\lim_{n\to\infty}\max_i(t_{i+1}-t_i)=0$，记

$$S_n=\sum_{i=0}^{n-1} g(B(t_i^*))(t_{i+1}-t_i),$$

其中 $t_i^*\in[t_i,t_{i+1}]$，则

$$\int_a^b g(B(u))\,du=\lim_{n\to\infty}S_n.$$

来证明 S_n 均方收敛，因为均方收敛和 a.s. 收敛必收敛到相同的随机变量，所以 S_n 均方收敛到 $\int_a^b g(B(s))\,ds$。

只要证明 $\{S_n\}$ 是 L^2 空间（有二阶矩的随机变量空间）的基本列。对 n,m，存在分割 $\{t_i, i=0,1,\dots,k\}$ 使得

$$S_n = \sum_{i=0}^{k-1} g(B(t_i'))(t_{i+1}-t_i),$$
$$S_m = \sum_{i=0}^{k-1} g(B(t_i''))(t_{i+1}-t_i),$$

其中 $t_i', t_i'' \in [t_i, t_{i+1}]$。记 $\delta_k = \max_i (t_{i+1}-t_i)$，则

$$\begin{aligned}
\|S_n - S_m\| \le & \sum_{i=0}^{k-1} \|g(B(t_i')) - g(B(t_i''))\| \, (t_{i+1}-t_i) \\
= & \sum_{i=0}^{k-1} \|g'(x_i^*)[g(B(t_i')) - g(B(t_i''))]\| \, (t_{i+1}-t_i) \\
\le & C \sum_{i=0}^{k-1} \|[g(B(t_i')) - g(B(t_i''))]\| (t_{i+1}-t_i) \\
= & C \sum_{i=0}^{k-1} |t_i' - t_i''|^{1/2} (t_{i+1}-t_i) \\
\le & C \delta_k^{1/2} (b-a) \to 0, \ n,m \to \infty,
\end{aligned}$$

其中 x_i^* 取值于 $B(t_i')$ 和 $B(t_i'')$ 组成的闭区间内。于是，S_n 均方收敛，可知其均方极限有二阶矩从而有一阶矩，由 L^2 空间中内积的连续性可得

$$\begin{aligned}
& E\int_a^b g(B(u))\,du = E \lim_{n\to\infty} S_n = \lim_{n\to\infty} ES_n \\
= & \lim_{n\to\infty} \sum_{i=0}^{n} E[g(B(t_i^*))](t_{i+1}-t_i) = \int_a^b E[g(B(s))]\,ds.
\end{aligned}$$

8.5.2 布朗运动乘以连续函数作为被积函数

定理 8.14. 设 $\{B(t), t\ge 0\}$ 是标准布朗运动，$0 \le a < b < \infty$，函数 $f(t)$ 在 $[a,b]$ 连续，$g(t)$ 在 $[c,d]$ 连续，则

$$Y = \int_a^b f(t)B(t)\,dt$$

为随机变量，服从正态分布，期望为 0，方差为

$$2\int_a^b \int_a^t f(t)f(s)s\,ds\,dt$$

又 $\left(\int_a^b f(t)B(t)\,dt, \int_c^d g(s)B(s)\,ds\right)$ 服从联合正态分布，且

$$Cov\left(\int_a^b f(t)B(t)\,dt, \int_c^d g(s)B(s)\,ds\right) = \int_a^b \int_c^d f(t)g(s)(t\wedge s)\,ds\,dt.$$

证明：

(i) 令 $A = \max(\max_{t\in[a,b]} f(t), \max_{s\in[c,d]} g(s))$。令 $\xi = \max_{s\in[0,b\vee d]} B(s) + \left|\min_{s\in[0,b\vee d]} B(s)\right|$。由节7.5.2可知 $E\xi^2 < \infty$。

设 $\{t_i\}$，$\{s_i\}$ 分别是 $[a,b]$ 和 $[c,d]$ 的分割，且最大间距 $\delta_n \to 0$。令

$$S_n = \sum_{i=0}^{n-1} f(t_i^*)B(t_i^*)(t_{i+1}-t_i)$$

为积分 $\int_a^b f(t)B(t)\,dt$ 的达布和，其中 $t_i^* \in [t_i, t_{i+1}]$，则 S_n 是多元正态分布随机变量的线性组合，也服从正态分布。因为 $B(t)$ 连续，所以对每一条轨道，$\int_a^b f(t)B(t)\,dt$ 积分存在有限，S_n 点点收敛到积分 $\int_a^b f(t)B(t)\,dt$，正态分布的极限 $\int_a^b f(t)B(t)\,dt$ 仍为正态分布。

(ii) 来证明 S_n 均方收敛到 $\int_a^b f(t)B(t)\,dt$。$\forall\epsilon>0$，对 S_n, S_m, 取达布和 S_k 使得 S_k 的分点包含 S_n 和 S_m 的所有分点，存在 N 使得 $n,m\ge N$ 时

$$\max(\delta_n,\delta_m,\delta_k) < \left(\frac{\epsilon}{4A(b-a)}\right)^2,$$

由 $f(x)$ 在闭区间的一致连续性可知当 k 充分大时对任意 $|x-y|<\delta_k$ 都有

$$|f(x)-f(y)| < \delta = \frac{\epsilon}{4b^{1/2}(b-a)}.$$

这时设

$$S_n = \sum_{i=0}^{k-1} f(t_i')B(t_i')(t_{i+1}-t_i),$$
$$S_k = \sum_{i=0}^{k-1} f(t_i'')B(t_i'')(t_{i+1}-t_i)$$

则

$$
\begin{aligned}
&\|S_n - S_k\| \\
\leq &\left\|\sum \left|f(t_i')B(t_i') - f(t_i'')B(t_i'')\right| (t_{i+1} - t_i)\right\| \\
\leq &\sum \|f(t_i')B(t_i') - f(t_i'')B(t_i'')\| (t_{i+1} - t_i) \\
\leq &\sum |f(t_i')| \cdot \|B(t_i') - B(t_i'')\| (t_{i+1} - t_i) \\
&+ \sum |f(t_i') - f(t_i'')| \cdot \|B(t_i'')\| (t_{i+1} - t_i) \\
\leq &A \sum |t_i' - t_i''|^{1/2} (t_{i+1} - t_i) + \sum \delta b^{1/2} (t_{i+1} - t_i) \\
\leq &A\delta_k^{1/2}(b-a) + \delta b^{1/2}(b-a) \\
\leq &\frac{\epsilon}{2},
\end{aligned}
$$

同理 $\|S_m - S_k\| \leq \frac{\epsilon}{2}$，所以

$$
\begin{aligned}
&\|S_n - S_m\| \\
\leq &\|S_n - S_k\| + \|S_m - S_k\| \leq \epsilon,
\end{aligned}
$$

即 $\{S_n\}$ 是 L^2 空间中的基本列，收敛到一个二阶矩随机变量，记为 $\int_a^b f(t)B(t)\,dt$，另外 S_n 也 a.s. 收敛，这两个极限是 a.s. 相等的。

(iii) 由于 S_n 服从正态分布，由 S_na.s. 收敛或者均方收敛都推出依分布收敛，正态分布的依分布极限仍是正态分布，极限分布期望为 $EY = \lim_n ES_n$, 方差 $\mathrm{Var}(Y) = \lim_n \mathrm{Var}(S_n)$。

因为 $ES_n = 0$，故

$$
E \int_a^b f(t)B(t)\,dt = 0.
$$

证明 $EY = \lim_n ES_n$ 和 $\mathrm{Var}(Y) = \lim_n \mathrm{Var}(S_n)$，还可以用内积的连续性。因为 $\|S_n - Y\| \to 0$，所以

$$
E(S_n \cdot 1) \to E(Y \cdot 1) = EY, \quad E(S_n \cdot S_n) \to E(Y \cdot Y) = \mathrm{Var}(Y).
$$

下面给出 $\lim_n E(S_n^2)$ 的值。

$$\begin{aligned}
E(S_n^2) =& \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} f(t_i)f(t_j)E[B(t_i^*)B(t_j^*)](t_{i+1}-t_i)(t_{j+1}-t_j) \\
=& \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} f(t_i)f(t_j)(t_i^*\wedge t_j^*)(t_{i+1}-t_i)(t_{j+1}-t_j) \\
\to& \int_a^b\int_a^b f(t)f(s)(t\wedge s)\,ds\,dt \\
=& 2\int_a^b\int_a^t f(t)f(s)s\,ds\,dt.
\end{aligned}$$

设 S_n 和 S_n' 分别为 $\int_a^b f(t)B(t)\,dt$ 和 $\int_c^d g(s)B(s)\,ds$ 的达布和，由 S_n 和 S_n' 的均方收敛性可知

$$\begin{aligned}
&\left\|\langle S_n, S_n'\rangle - \left\langle \int_a^b f(t)B(t)\,dt,\ \int_c^d g(s)B(s)\,ds\right\rangle\right\| \\
\le& \left\|\left\langle S_n - \int_a^b f(t)B(t)\,dt,\ S_n'\right\rangle\right\| + \left\|\left\langle \int_a^b f(t)B(t)\,dt,\ S_n' - \int_c^d g(s)B(s)\,ds\right\rangle\right\| \\
\le& \left\|S_n - \int_a^b f(t)B(t)\,dt\right\|\cdot\|S_n'\| + \left\|\int_a^b f(t)B(t)\,dt\right\|\cdot\left\|S_n' - \int_c^d g(s)B(s)\,ds\right\| \\
&\to 0,\ (n\to\infty),
\end{aligned}$$

而 (S_n, S_n') 为二元正态分布，其极限分布也服从二元正态分布。

由内积的连续性可得

$$\begin{aligned}
&\operatorname{Cov}\left(\int_a^b f(t)B(t)\,dt,\ \int_c^d g(s)B(s)\,ds\right) \\
=& E\left(\int_a^b f(t)B(t)\,dt\cdot\int_c^d g(s)B(s)\,ds\right) \\
=& \lim_{n\to\infty} E(S_nS_n') \\
=& \lim_{n\to\infty}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} f(t_i)g(s_j)(t_i^*\wedge s_j^*)(t_{i+1}-t_i)(s_{j+1}-s_j) \\
=& \int_a^b\int_c^d f(t)g(s)(t\wedge s)\,ds\,dt.
\end{aligned}$$

下面考虑 $\int_a^b f(t)\,dB(t)$ 这样的随机积分。

8.5.3 关于连续可微函数的随机积分

定理 8.15. 设函数 $f(t)$, $g(t)$ 在 $[0,\infty)$ 连续可微，则

$$X(t)=\int_0^t f(s)\,dB(s)=f(t)B(t)-\int_0^t f'(s)B(s)\,ds,$$

为随机变量，$X(t)$ 服从正态分布，且

$$\begin{aligned}
E\int_0^t f(s)\,dB(s) =&0,\\
Var\left(\int_0^t f(s)\,dB(s)\right) =&\int_0^t f^2(s)\,ds,\\
Cov\left(\int_0^t f(u)\,dB(u),\ \int_0^s g(u)\,dB(u)\right) =&\int_0^{t\wedge s} f(u)g(u)\,du.
\end{aligned}$$

证明：利用分部积分公式即可得

$$X(t)=\int_0^t f(s)\,dB(s)=f(t)B(t)-\int_0^t f'(s)B(s)\,ds,$$

由定理8.14可知 $X(t)$ 服从正态分布，均值为 0，且可计算

$$\mathrm{Var}(X(t)) \tag{8.21}$$

$$=tf^2(t)+\int_0^t\int_0^t f'(s)f'(u)(s\wedge u)\,du\,ds-2f(t)\int_0^t f'(s)s\,ds. \tag{8.22}$$

(8.22)的第二项为

$$
\begin{aligned}
&\int_0^t \int_0^t f'(s)f'(u)(s\wedge u)\,du\,ds\\
=&2\int_0^t \int_s^t f'(s)f'(u)s\,du\,ds\\
=&2\int_0^t sf'(s)\left[\int_s^t f'(u)\,du\right]\,ds\\
=&2\int_0^t sf'(s)[f(t)-f(s)]\,ds\\
=&2f(t)\int_0^t sf'(s)\,ds-2\int_0^t sf(s)f'(s)\,ds
\end{aligned}
$$

所以

$$
\mathrm{Var}(X(t)) = tf^2(t)-2\int_0^t sf(s)f'(s)\,ds,
$$

其中

$$
\begin{aligned}
\int_0^t sf(s)f'(s)\,ds =& \int_0^t sf(s)\,df(s)\\
=&tf^2(t)-\int_0^t f^2(s)\,ds-\int_0^t sf(s)f'(s)\,ds,
\end{aligned}
$$

所以

$$
2\int_0^t sf(s)f'(s)\,ds = tf^2(t)-\int_0^t f^2(s)\,ds,
$$

于是

$$
\mathrm{Var}(X(t)) = \int_0^t f^2(s)\,ds.
$$

设 $0<s<t$，

$$
\begin{aligned}
&\mathrm{Cov}\left(\int_0^t f(u)\,dB(u),\ \int_0^s g(v)\,dB(v)\right) && (8.23)\\
=&E\left\{\left[f(t)B(t)-\int_0^t f'(u)B(u)\,du\right]\left[g(s)B(s)-\int_0^s g'(v)B(v)\,dv\right]\right\} && (8.24)\\
=&f(t)g(s)s && (8.25)\\
&+\int_0^s\int_0^t f'(u)g'(v)(u\wedge v)\,du\,dv && (8.26)\\
&-f(t)\int_0^s g'(v)v\,dv && (8.27)\\
&-g(s)\int_0^t f'(u)(u\wedge s)\,du. && (8.28)
\end{aligned}
$$

(8.28)式第四项为

$$
\begin{aligned}
&-g(s)\int_0^t f'(u)(u\wedge s)\,du\\
=&-g(s)\int_0^s f'(u)u\,du-g(s)\int_s^t f'(u)s\,du\\
=&-g(s)\int_0^s f'(u)u\,du-f(t)g(s)s+f(s)g(s)s,
\end{aligned}
$$

化简得

$$
\begin{aligned}
&\mathrm{Cov}\left(\int_0^t f(u)\,dB(u),\ \int_0^s g(v)\,dB(v)\right) && (8.29)\\
=&\int_0^s\int_0^t f'(u)g'(v)(u\wedge v)\,du\,dv && (8.30)\\
&-f(t)\int_0^s g'(v)v\,dv && (8.31)\\
&-g(s)\int_0^s f'(u)u\,du+f(s)g(s)s. && (8.32)
\end{aligned}
$$

(8.32)式中

$$
\begin{aligned}
&\int_0^s \int_0^t f'(u)g'(v)(u \wedge v)\,du\,dv \\
=&\int_0^s \int_0^v f'(u)g'(v)u\,du\,dv + \int_0^s \int_v^t f'(u)g'(v)v\,du\,dv \\
=&\int_0^s \int_u^s f'(u)g'(v)u\,dv\,du + \int_0^s \int_v^t f'(u)g'(v)v\,du\,dv \\
=&\int_0^s uf'(u)\left[\int_u^s g'(v)\,dv\right]du + \int_0^s vg'(v)\left[\int_v^t f'(u)\,du\right]dv \\
=&\int_0^s uf'(u)[g(s)-g(u)]\,du + \int_0^s vg'(v)[f(t)-f(v)]\,dv \\
=&g(s)\int_0^s uf'(u)\,du - \int_0^s uf'(u)g(u)\,du + f(t)\int_0^s vg'(v)\,dv - \int_0^s vf(v)g'(v)\,dv
\end{aligned}
$$

将上式代入(8.32)中化简得

$$
\begin{aligned}
&\mathrm{Cov}\left(\int_0^t f(u)\,dB(u),\ \int_0^s g(v)\,dB(v)\right) \\
=&-\int_0^s uf'(u)g(u)\,du - \int_0^s vf(v)g'(v)\,dv + sf(s)g(s) \\
=&-\int_0^s ud[f(u)g(u)] + sf(s)g(s) \\
=&-sf(s)g(s) + \int_0^s f(u)g(u)\,du + sf(s)g(s) \\
=&\int_0^s f(u)g(u)\,du.
\end{aligned}
$$

得证。

证明中的一阶矩和二阶矩可以直接推论：因为 $\int_0^t f^2(s)\,ds < \infty$，所以 $f \in V$，所以 $\int_0^t f(s)\,dB(s)$ 期望为零，根据等距性可得方差和协方差的结论，正态性则利用了定理8.14的结论。

推论 8.4. 设 $f(t)$ 在 $[0,\infty)$ 连续可微，对 $0 \leq a < b$，$t > 0$，定义

$$
\begin{aligned}
X(t) =& \int_0^t f(t)\,dB(t), \\
\int_a^b f(u)\,dB(u) =& X(b) - X(a) = \int_0^b f(u)\,dB(u) - \int_0^a f(u)\,dB(u),
\end{aligned}
$$

则 $\{X(t), t \geq 0\}$ 是高斯过程，有独立增量，增量 $\int_a^b f(u)\,dB(u)$ 服从正态分布，均值为 *0*，方差为

$$
Var\left(\int_a^b f(u)\,dB(u)\right) = \int_a^b f^2(u)\,du.
$$

证明由节8.5.2可知 $(X(a), X(b))$ 服从二元正态分布，同理可知 $\{X(t), t \geq 0\}$ 的任意有限维分布服从多元正态分布。于是 $X(b) - X(a)$ 均值为 0，方差为

$$
\begin{aligned}
\text{Var}\left(\int_a^b f(u)\,dB(u)\right) =& \text{Var}(X(b) - X(a)) \\
=& \int_0^b f^2(u)\,du + \int_0^a f^2(u)\,du - 2\int_0^a f^2(u)\,du \\
=& \int_a^b f^2(u)\,du.
\end{aligned}
$$

对任意 $0 \leq t_1 < t_2 \leq t_3 < t_4$，有

$$
\begin{aligned}
& \text{Cov}(X(t_2) - X(t_1),\, X(t_4) - X(t_3)) \\
=& \text{Cov}(X(t_2), X(t_4)) - \text{Cov}(X(t_2), X(t_3)) - \text{Cov}(X(t_1), X(t_4)) + \text{Cov}(X(t_1), X(t_3)) \\
=& \int_0^{t_2} f^2(u)\,du - \int_0^{t_2} f^2(u)\,du - \int_0^{t_1} f^2(u)\,du + \int_0^{t_1} f^2(u)\,du \\
=& 0,
\end{aligned}
$$

因为 $\{X(t), t \geq 0\}$ 是高斯过程，不相关即独立，所以 $\{X(t), t \geq 0\}$ 是独立增量过程。

补充：推论8.4高斯过程结论的另外证明方法，来自 (Shreve 2004) 定理 4.4.9 和例 4.7.3。

先证明 $X(t)$ 服从正态分布。显然 $\int_0^T E[f^2(s)]\,ds = \int_0^T f^2(s)\,ds < \infty$，所以 $\int_0^t f(s)\,dB(s)$ 满足空间 V 的条件，其均值和方差由 Itô 积分的鞅性质可得。来证明 $X(t)$ 的矩母函数等于

$$E[e^{uX(t)}] = \exp\left\{\frac{1}{2}u^2\int_0^t f^2(s)\,ds\right\},\ \forall u\in\mathbb{R}. \tag{8.33}$$

因为 f 非随机，上式等价于

$$E\exp\left\{uX(t) - \frac{1}{2}u^2\int_0^t f^2(s)\,ds\right\} = 1,\ \forall u\in\mathbb{R}.$$

证明上式需要利用更多的理论，参见 (Glasserman 2004) 定理 B.2.3。(Shreve 2004) 中的讨论不够严格，没有注意到被积函数是否在 V 空间的条件。

推论 8.5. 若 $f(t)$ 是非随机的连续可微函数，则

$$\int_0^T f(t)\,dB(t) = \lim_{n\to\infty}\sum_{i=0}^{n-1} f(t_i^*)[B(t_{i+1}) - B(t_i)],$$

其中 $0 = t_0 < t_1 < \cdots < t_n = T$ 是 $[0,T]$ 的任意分割，满足 $\delta_n = \max_i(t_{i+1} - t_i) \to 0$，$t_i^* \in [t_i, t_{i+1}]$ 任意选取，极限为 L^2 极限且不依赖于分割和 t_i^* 的选取。

证明：易见 $f(t)$ 看作一个随机过程满足空间 V 的条件，定理8.1已经证明了当 $t_i^* = t_i$ 的情形。这里的结论是对非随机连续可微被积函数的一个增强结果，即代表点不限于区间左端点。记

$$S_n = \sum_{i=0}^{n-1} f(t_i^*)[B(t_{i+1}) - B(t_i)],$$

定理8.15已证明 S_n 构成 L^2 基本列，有均方极限在 L^2 空间中，定义其均方极限为 $\int_0^T f(t)\,dB(t)$。下面证明极限不依赖于分割和 t_i^* 的选择。

设 S_n 和 S_n' 都是上述积分逼近序列但分割和代表点选择不同，可以合并其分割点，使得 S_n 和 S_n' 可以写成

$$\begin{aligned}
S_n &= \sum_{i=0}^{N_n} f(t_i')[B(t_{i+1}) - B(t_i)],\\
S_n' &= \sum_{i=0}^{N_n} f(t_i'')[B(t_{i+1}) - B(t_i)],
\end{aligned}$$

由独立增量性，可得

$$
\begin{aligned}
\|S_n - S_n'\|^2 &= \left\| \sum_{i=0}^{N_n} [f(t_i') - f(t_i'')] \, [B(t_{i+1}) - B(t_i)] \right\|^2 \\
&= \sum_{i=0}^{N_n} [f(t_i') - f(t_i'')]^2 \, \|B(t_{i+1}) - B(t_i)\|^2 \\
&= \sum_{i=0}^{N_n} [f(t_i') - f(t_i'')]^2 \, (t_{i+1} - t_i) \\
&= \sum_{i=0}^{N_n} [f'(\tilde{t}_i)]^2 (t_i' - t_i'')^2 \, (t_{i+1} - t_i) \\
&\le \max_{t\in[0,T]} [f'(t)]^2 \sum_{i=0}^{N_n} (t_{i+1} - t_i)^2 \, (t_{i+1} - t_i) \\
&\le \max_{t\in[0,T]} [f'(t)]^2 \delta_n^2 t \to 0, \ n \to \infty.
\end{aligned}
$$

因此极限不依赖于分割和代表点的选择。

8.5.3.1 空间 V 的完备性

显然 V 是线性空间。定义模

$$
\|X\| = \left(\int_0^T E[X^2(t,\omega)] \, dt \right)^{1/2} \tag{8.34}
$$

后，按照此模导出的距离，V 是可分 Banach 空间，即定义定义了模 (范数) 和距离的线性空间，完备（基本列都有极限且关于极限封闭），可分（有可数的稠密子集）。参见 (钱敏平 1990) 节 8.2 P.294，(龚光鲁 and 钱敏平 2019) 节 1.2 P.8 命题 1.4。

由 Fubini 定理，(8.34)式也可以写成

$$
\|X\| = \left(E \left[\int_0^T X^2(t,\omega) \, dt \right] \right)^{1/2}.
$$

来证明(8.34)式是 V 的模。易见 $\|X\| \ge 0$，且 $\|X\| = 0$ 当且仅当 $X(t,\omega) = 0$, a.s.。实际上，

$$E\left[\int_0^T X^2(t,\omega)\,dt\right]=0,$$

则依概率 1 地 $\int_0^T X^2(t,\omega)\,dt=0$，从而对给定的 ω，$X(t,\omega)$ 作为 t 的函数是几乎处处为 0 的。这作为 V 中 0 元素的理解。对实数 c 有 $\|cX\|=|c|\cdot\|X\|$。要证明三角不等式 $\|X+Y\|\le\|X\|+\|Y\|$，只要证明 $\|X+Y\|^2\le\|X\|^2+\|Y\|^2+2\|X\|\,\|Y\|$。事实上

$$\begin{aligned}
\|X+Y\|^2 =&E\left[\int_0^T (X+Y)^2\,dt\right]\\
=&E\left[\int_0^T X^2\,dt\right]+E\left[\int_0^T Y^2\,dt\right]+E\left[\int_0^T 2XY\,dt\right]\\
\le&\|X\|^2+\|Y\|^2+2E\left[\left(\int_0^T X^2\,dt\right)^{1/2}\left(\int_0^T Y^2\,dt\right)^{1/2}\right]\\
\le&\|X\|^2+\|Y\|^2+2\left\{E\left[\left(\int_0^T X^2\,dt\right)^{1/2}\right]^2\right\}^{1/2}\left\{E\left[\left(\int_0^T Y^2\,dt\right)^{1/2}\right]^2\right\}^{1/2}\\
=&\|X\|^2+\|Y\|^2+2\left(E\int_0^T X^2\,dt\right)^{1/2}\left(E\int_0^T Y^2\,dt\right)^{1/2}\\
=&\|X\|^2+\|Y\|^2+2\|X\|\,\|Y\|.
\end{aligned}$$

这里利用了 $[0,T]$ 上平方可积函数空间 $L^2[0,T]$ 的 Schwarz 不等式，和随机变量的 Schwarz 不等式。这就证明了 $\|\cdot\|$ 是 V 的模（范数），V 构成赋范线性空间。

将 V 中的 X 看成是 $([0,T]\times\Omega,\mathscr{B}([0,T])\times\mathscr{F},dt\times dP)$ 测度空间上的平方可积函数组成的空间，则按照测度论中 L_p 空间的理论可知 $\|\cdot\|$ 是模，完备。可分性从引理8.1可以看出。

由泛函分析的一般理论，在 V 中定义内积

$$\langle\phi,\psi\rangle=E\int_0^T \phi(u)\psi(u)\,du,$$

在 V 构成完备的内积空间（即 Hilbert 空间），且可分（有可数稠密子集）。

8.5.3.2 等距性证明

被积函数是可料阶梯过程时，$E[(\int XdB)^2] = \int EX^2dt$。性质可以推广到 $X \in V$ 的情形。

这里给出协方差的公式，这也包含了方差的公式。

定理 8.16. 设 $\{X(u), u \in [0,s]\}$ 和 $\{Y(u), u \in [0,t]\}$ 是 V 中的随机过程，其中 $0 \le s \le t$，则

$$
\begin{aligned}
& Cov\left(\int_0^s X(u)\,dB(u),\ \int_0^t Y(u)\,dB(u)\right) \\
=& \int_0^s E[X(u)Y(u)]\,du.
\end{aligned}
$$

证明： 随机积分期望为零，协方差等于乘积的期望，由随机积分定义知存在简单过程 $\{\phi_n(u)\}, \{\psi_n(u)\} \in V$，使得

$$
\begin{aligned}
E\int_0^s [X(u)-\phi_n(u)]^2du \to& 0, \\
\left\| \int_0^s X(u)\,dB(u) - \int_0^s \phi_n(u)\,dB(u) \right\| \to& 0, \\
E\int_0^t [Y(u)-\psi_n(u)]^2du \to& 0, \\
\left\| \int_0^s Y(u)\,dB(u) - \int_0^t \psi_n(u)\,dB(u) \right\| \to& 0.
\end{aligned}
$$

由 $L^2(\Omega, \mathscr{F}, P)$ 空间中内积的连续性和可料简单函数的推广等距性可知

$$
\begin{aligned}
& \mathrm{Cov}\left(\int_0^s X(u)\,dB(u),\ \int_0^t Y(u)\,dB(u)\right) \\
=& E\left(\int_0^s X(u)\,dB(u) \cdot \int_0^t Y(u)\,dB(u)\right) \\
=& \lim_{n\to\infty} E\left(\int_0^s \phi_n(u)\,dB(u) \cdot \int_0^t \psi_n(u)\,dB(u)\right) \\
=& \lim_{n\to\infty} \int_0^s E[\phi_n(u)\psi_n(u)]\,du
\end{aligned}
$$

计算

$$\begin{aligned}
&\left|\int_0^s E[\phi_n(u)\psi_n(u) - X(u)Y(u)]\,du\right| \\
\le &\int_0^s E[|\phi_n(u) - X(u)|\cdot|Y(u)|]\,du \\
&+ \int_0^s E[|\phi_n(u)|\cdot|\psi_n(u) - Y(u)|]\,du \\
\le &\int_0^s \|\phi_n(u) - X(u)\|\cdot\|Y(u)\|\,du \\
&+ \int_0^s \|\phi_n(u)\|\cdot\|\psi_n(u) - Y(u)\|\,du \\
\le &\left(\int_0^s \|\phi_n(u) - X(u)\|^2\,ds\right)^{1/2}\left(\int_0^s \|Y(u)\|^2\,du\right)^{1/2} \\
+ &\left(\int_0^s \|\phi_n(u)\|^2\,du\right)^{1/2}\left(\int_0^s \|\psi(u) - Y(u)\|^2\,du\right)^{1/2} \\
= &\left(\int_0^s E[\phi_n(u) - X(u)]^2\,ds\right)^{1/2}\left(\int_0^s E[Y^2(u)]\,du\right)^{1/2} \\
+ &\left(\int_0^s E[\phi_n(u)]^2\,du\right)^{1/2}\left(\int_0^s E[\psi_n(u) - Y(u)]^2\,du\right)^{1/2},
\end{aligned}$$

其中

$$\begin{aligned}
&\int_0^s E[\phi_n(u) - X(u)]^2\,ds \to 0, \\
&\int_0^s E[\psi_n(u) - Y(u)]^2\,du \to 0, \\
&\int_0^s E[Y^2(u)]\,du < \infty, \\
&\int_0^s E[\phi_n(u)]^2\,du \le 2\int_0^s E[\phi_n(u) - X(u)]^2\,du + 2\int_0^s E[X^2(u)]\,du = O(1),
\end{aligned}$$

于是有

$$\left|\int_0^s E[\phi_n(u)\psi_n(u) - X(u)Y(u)]\,du\right| \to 0,$$

$$\begin{aligned}
&\operatorname{Cov}\left(\int_0^s X(u)\,dB(u),\ \int_0^t Y(u)\,dB(u)\right) \\
= &\int_0^s E[X(u)Y(u)]\,du.
\end{aligned}$$

证明中利用了

$$\left|\int_0^s f(u)g(u)du\right| \leq \left(\int_0^s f^2(u)\,du\right)^{1/2}\left(\int_0^s g^2(u)\,du\right)^{1/2}.$$

Chapter 9

随机过程在金融中的应用

9.1 金融市场的术语与基本假定

本节我们给出金融市场的一些术语与基本假定。

9.1.1 无套利原则

金融市场有一个最基本的假设，即市场不允许没有初始投资的无风险利润。若违背了这个市场原则，则可以得到套利机会。

所谓套利，即指在开始时无资本, 经过资本的市场运作后, 变成有非负的随机资金, 而且有正资金的概率为正。

套利机会在实际操作中很少存在，违背无套利原则的情况一般是短暂且难以把握的，因为在出现套利机会时, 大量的投机者就会涌向市场进行套利, 他们追逐套利利润的积极性将有效地消除套利机会。经过一个相对较短的混乱时期后, 市场就会重返正常, 即回复到无套利状态。在金融衍生证券的定价理论中, 并不讨论这段混乱时期, 因此, 在研究中普遍地排除套利, 即假定正常运行的市场没有套利机会，这样的市场也称为可行市场。

9.1.2 看涨期权

看涨期权是指一种不附带义务的未来购买的权利。在时刻 0 时买方与卖方有一个合约, 按此合约规定买方有一个权利, 能在时刻 T （到期日）以价格 K（执行价）从卖方买进股票。如果时刻 T 时股票的市场价格 S_T 低于执行价格 K，买方可以拒绝支付执行价；而如果时刻 T 时股票的市场价格 S_T 高于执行价格 K, 买方就一定会选择支付执行价同时获得高价格的股票，称之为期权被执行了。综合起来, 买方在时刻 T 净得随机收益（现金流）为

$$(S_T - K)^+ = \max\{0, S_T - K\}.$$

因为买方希望 S_T 尽量大, 以便有更多的获利, 也就是有选择权的买方盼望股票上涨, 这种合约称为看涨期权。

前面讨论的这种期权中，买方只有在到期时才能执行期权，这种看涨期权称为欧式看涨期权。若解除执行期权的时间限制，允许买方在到期日钱的任何时间行使期权，则称为美式看涨期权。

由于这个合约能给买方带来随机收益, 就需要买方在 $t = 0$ 时刻用钱从卖方购买。这个合约在 $t = 0$ 时刻的价格, 称为它的贴水或保证金。如何确定这个合约在时刻 $t < T$ 的价格 (包括贴水) 即是我们要重点讨论的问题。欧式期权由于有执行时间的限制，未来的现金流收入较之美式期权要低些，但是欧式期权现金流收入有明确的表达式，所以估计这种期权的价格会更容易些。

9.1.3 看跌期权

看跌期权是指一种不附带义务的未来出售的权利。在时刻 0 时买方与卖方有一个合约, 按此合约规定买方有一个权利, 能在时刻 T 以价格 K（执行价）卖给合约卖方一股股票或者等量的一股股票的市场价格。如果时刻 T 时股票的市场价格 S_T 高于执行价格 K，买方可以拒绝执行期权；而如果时刻 T 时股票的市场价格 S_T 低于执行价格 K, 买方就一定会选择执行期权而获利。综合起来, 买方在时刻 T 净得随机收益（现金流）为

$$(K - S_T)^+ = \max\{0, K - S_T\}.$$

因为买方希望 S_T 尽量小, 以便有更多的获利, 也就是有选择权的买方盼望股票下跌, 这种合约称为看跌期权。

如果买方只能在最终时刻 T 行使期权, 则称为欧式看跌期权; 若买方能在任意时刻 T 行使期权, 则称为美式看跌期权。同样由于这个合约能给买方带来随机收益, 就需要买方在 $t=0$ 时刻用钱从卖方购买。这个合约在 t=0 时刻的价格, 称为它的贴水或保证金。

比看涨期权与看跌期权更为一般的欧式期权是：甲方卖给乙方一个由证券组合组成的合约, 此合约能在时刻 T 给乙方带来随机收益 $f(S_T)$ 称为欧式未定权益。

9.1.4 无套利原则下的看涨与看跌平权关系

未定权益为 S_T 的欧式权益，称为在时刻 T 到期的远期合约，远期合约在任意时刻 $t(<T)$ 的价格为证券的即时价格 S_t.

当 $t=T$ 时有

$$(S_T-K)^+-(K-S_T)^+=S_T-K,$$

即在交割时刻,如果买进一张看涨期权,卖出一张看跌期权,收益为:若 $S_T>K$,$(S_T-K)^+-(K-S_T)^+=S_T-K+0$,若 $S_T\le K$,则 $(S_T-K)^+-(K-S_T)^+=0-(K-S_T)$，均等价于买入标的资产同时卖出一张马上到期的数额为 K 的存款。

于是

$$e^{-r(T-t)}(S_T-K)^+-e^{-r(T-t)}(K-S_T)^+=e^{-r(T-t)}S_T-Ke^{-r(T-t)},$$

而

$$\begin{aligned}C_t=&E\left[e^{-r(T-t)}(S_T-K)^+\right],\\P_t=&E\left[e^{-r(T-t)}(K-S_T)^+\right],\end{aligned}$$

且在 T 时刻的权益值为 S_T 的远期合约在 t 时刻的价值为

$$S_t=E\left[e^{-r(T-t)}S_T\right],$$

所以有

$$C_t-P_t=S_t-Ke^{-r(T-t)}.$$

其中，K 表示执行价格，r 表示银行利率。

有了这个平权关系，欧式看涨期权与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。

9.2 Black-Scholes 模型

期权的价格是是期权合约中唯一随市场不断变化而改变的量，它受到期权合约中的期限，执行价格，标的资产的价格，无风险利率等众多因素的影响，期权价格的高低直接影响到购买者和出售者的盈亏状况，如何对期权进行定价是期权交易的核心问题。对期权进行定价的过程中，首当其冲的关键性问题就是如何构建合适的模型来描述标的资产的价格变化过程?

法国数学家 Barchelier 是研究期权定价问题的先驱人物，在其博士论文《The Theory of Speculation》中首次利用随机游走的思想给出了股票价格运行的随机模型。1942 年日本数学家伊藤清对布朗运动引进随机积分，开创了随机微分方程理论，是一个重要的里程碑式的工作. 而在 1965 年著名经济学家 Samuelson 将提出的随机分析学作为工具引入到金融学中，他对 Barchelier 的股票模型进行了修正，首次提出了用几何布朗运动来描述股票价格过程，这种思想得到了广泛而长期的应用。几何布朗运动模型克服了 Barchelier 的模型中可能使得股票价格出现负值的这种与现实问题不符合的情况，基于这个模型，Samuelson 研究了看涨期权的定价问题。

Samuelson 所得出的定价公式有一个遗憾的地方，就是依赖于投资者的个人风险偏好，这就限制了一些问题。Black 和 Scholes 在 1973 年找到了弥补这一遗憾的新方法，他们建立期权定价模型的关键突破点在于，构造一个由标的股票和无风险债券的适当组合（投资组合），直观意义即是不要把鸡蛋放在同一个篮子里面。这个投资组合具有如下特点：投资组合的损益特征与期权在到期日的损益特征是相同的，不管标的股票的价格在未来的时间里怎样变化，两种资产构成的投资组合产生无风险的回报。根据无套利原理（即零投资只能得到零回报），他们用动态复制的方法推导出了欧式期权（看跌或看涨）的定价公式，并且是精确的显式解，他们考虑的情况是基于无红利支付的情况。Black-Scholes 的开创性工作即在于在他们的模型中，所有的投资者的投资回报率都是同一个无风险利率，而不依赖于投资者的个人风险偏好。Black-Scholes 定价模型具有划时代的意义，自从它问世以来，有关欧式期权，美式期权，障碍期权，股票期权、股指期权、利率期权、外汇期权等各式各样的期权定价问题得到了广泛而深入的讨论。

Black-Scholes 期权定价模型给出了如下假设：

- （1）市场不存在无风险套利机会。

- （2）期权是欧式的，即只能在到期日执行期权。
- （3）市场无摩擦，即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，无风险利率和波动率在合约期限内均为常数，投资者可以用无风险利率自由借贷。
- （5）股票不分发股利也不作其他任何形式的利润分配。
- （6）对卖空无任何限制。
- （7）交易时间及价格变动是连续的。

9.2.1 期权定价问题的描述

以下考虑欧式看涨期权，我们给出 Black-Sholes 期权定价公式。

设某种风险资产（如股票）在 t 时刻的价格为 S_t, 并设它满足如下模型:

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad t \in [0, T].$$

其中常数　表示风险资产的平均回报率, (>0) 表示风险资产的波动率，(B_t) 表示标准 Brown 运动（在概率测度 P 下），T 表示期权的到期日。

方程(9.2.1)的解是几何 Brown 运动：

$$S_t = S_0 \exp\{(\mu - \frac{\sigma^2}{2})t + \sigma B_t\}.$$

给出一个无风险资产，由如下微分方程来刻画：

$$d\beta_t = r\beta_t dt, \quad t \in [0, T].$$

假定当前的银行利率为常数 r, 而且不随时间变化（实际情况中，银行利率会随时间发生变化，此处作简化处理）。

用 a_t 表示 t 时刻投资于风险资产的资金数量，b_t 表示 t 时刻投资于无风险资产的资金数量，(a_t, b_t) 称为一个投资组合。在 t 时刻，由数量为 a_t 的风险资产和数量为 b_t 的无风险资产构成的投资组合的财富值为

$$V_t = a_t S_t + b_t \beta_t, \quad t \in [0, T].$$

假定投资组合是自融资的，即财富的增量仅由 S_t 和 β_t 的变动引起，即

$$dV_t = a_t dS_t + b_t d\beta_t.$$

现在我们寻找一个自融资的策略 (a_t, b_t) 和一个相应的财富过程 V_t，使得

$$V_t = a_t S_t + b_t \beta_t = u(T-t, S_t), \quad t \in [0, T].$$

其中 $u(t,x)$ 为待求的光滑函数。

因为在到期日 T 时刻，投资组合的值 V_T 应为 T 时刻的现金流 $(S_T - K)^+$，故可得到一个终端条件。

$$V_T = U(0, S_T) = (S_T - K)^+.$$

令 $f(t,x) = u(T-t,x)$，则 $V_t = f(t, S_t)$，已知 S_t 满足如下积分方程

$$S_t = S_0 + \mu \int_0^t S_s ds + \sigma \int_0^t S_s dB_s.$$

由 Itô 公式可得：

$$\begin{aligned}
V_t - V_0 =& f(t, S_t) - f(0, S_0) \\
=& \int_0^t \left[\frac{\partial f(s, S_s)}{\partial t} - \mu S_s \frac{\partial f(s, S_s)}{\partial x} + \frac{1}{2}\sigma^2 S_s^2 \frac{\partial f(s, S_s)}{\partial x^2} \right] ds \\
&+ \int_0^t \sigma S_s \frac{\partial f(s, S_s)}{\partial x} dB_s \\
=& \int_0^t \left[-\frac{\partial u(T-s, S_s)}{\partial t} + \mu S_s \frac{\partial u(T-s, S_s)}{\partial x} + \frac{1}{2}\sigma^2 S_s^2 \frac{\partial u(T-s, S_s)}{\partial x^2} \right] ds \\
&+ \int_0^t \sigma S_s \frac{\partial T-s, S_s)}{\partial x} dB_s
\end{aligned}$$

另一方面，(a_t, b_t) 是自融资的，故

$$V_t - V_0 = \int_0^t a_s dS_s + \int_0^t b_s d\beta_s.$$

由 $d\beta_t = r\beta_t dt$ 以及 $V_t = a_t S_t + b_t \beta_t$ 可知

$$b_t = \frac{V_t - a_t S_t}{\beta_t}.$$

$$
\begin{aligned}
V_t - V_0 &= \int_0^t a_s dS_s + \int_0^t \frac{V_s - a_s S_s}{\beta_s} r\beta_s ds \\
&= \int_0^t a_s dS_s + \int_0^t r(V_s - a_s S_s) ds \\
&= \int_0^t [(\mu - r) a_s S_s + rV_s] ds \\
&\quad + \int_0^t \sigma a_s S_s dB_s.
\end{aligned}
$$

比较(9.2.1)与(9.2.1)得

$$a_t = \frac{\partial u(T-t, S_t)}{\partial x}.$$

即得到如下偏微分方程：

$$
\begin{aligned}
&-\frac{\partial u(T-t,S_t)}{\partial t} + rS_t \frac{\partial u(T-t,S_t)}{\partial x} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 u(T-t,S_t)}{\partial x^2} \\
&-ru(T-t,S_t) = 0, \qquad t \in [0,T].
\end{aligned}
$$

即得到 $u(t,x)$ 满足的偏微分方程为

$$
\begin{aligned}
&\frac{\partial u(t,x)}{\partial t} + rx\frac{\partial u(t,x)}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u(t,x)}{\partial x^2} - ru(t,x) = 0 \\
&x > 0,\ t \in [0,T].
\end{aligned}
$$

由 $V_T = u(0, S_T) = (S_T - K)^+$ 可得

$$u(0,x) = (x-K)^+, \quad x > 0.$$

9.2.2 Black-Scholes 公式

一般来说，求解一个偏微分方程的显示解是比较困难的事情，故很多时候都只能寻求数值解。Black 和 Scholes 基于无红利支付的情况，根据无套利原理, 用动态复制的方法推导出了欧式期权（看跌或看涨）的定价公式，并且是精确的显式解。

偏微分方程(9.2.1)的解为：

$$u(t,x) = x\Phi(d_1(t,x)) - Ke^{-rt}\Phi(d_2(t,x)),$$

其中 $\Phi(x)$ 是标准正态分布函数，

$$d_1(t,x) = \frac{\ln(x/K) + (r + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}, \qquad d_2 = d_1 - \sigma\sqrt{t}.$$

即有

$$V_0 = u(T, S_0) = S_0\Phi(d_1(T, S_0)) - Ke^{-rT}\Phi(d_2(T, S_0))$$

其中

$$d_1(T, S_0) = \frac{ln(S_0/K) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, \quad d_2 = d_1(T, S_0) - \sigma\sqrt{T}.$$

随机过程 $V_1 = u(T-t, S_t)$ 表示的是自融资的投资组合在时刻 $t \in [0, T]$ 的价值，自融资投资策略 (a_t, b_t) 为

$$a_t = \frac{\partial u(T-t, S_t)}{\partial x}, \quad b_t = \frac{u(T-t, S_t) - a_t S_t}{\beta_t}$$

我们称式(9.2.2)为 Black-Scholes 期权定价公式，可知期权价格与平均回报率 μ 是无关的，但与波动率 σ 有关。

如何理解在无套利原则下，期权的初始价格为 $q = u(T, S_0)$?

假设初始期权价格 $p \neq q$. 若 $p > q$，可采用如下投资策略：在时刻 $t = 0$，以价格 p 将期权卖出，同时根据公式(9.2.2)所提供的投资策略投资 q. 故可得到一个初始纯利润 $p - q > 0$. 在到期 T 时刻，投资组合的价值为 $a_T S_T + b_T \beta_T = (S_T - K)^+$,并且有义务支付 $(S_T - K)^+$ 给期权的购买者。这意味着,若 $S_T > K$，必须以价格 S_T 购买股票，而以执行价格 K 将股票卖给期权的买方，损失为 $S_T - K$，正好与投资组合当前的价值 $(S_T - K)^+ = S_T - K$ 持平，净利润为 $p - q$；若 $S_T \leq K$，期权不会被执行，故净利润为 $p - q + S_T - K$，总是有正利润的。若 $p < q$，会导致类似情形的发生。

故在无套利原则下，期权的初始价格一定为 $q = u(T, S_0)$.

9.2.3 Black-Scholes 公式的推导

Harrison 和 Kreos 在 1979 年提出了一种鞅定价方法来解决期权的定价问题。为 Black-Sholes 公式的推导作准备，我们首先介绍等价概率测度与重要的 Girsanov 定理。

定义 9.1. 令 P 和 Q 是定义在 σ 代数 $\mathscr{F}$ 上的两个概率测度，若存在一个非负函数 f_1，使得

$$Q(A) = \int_A f_1(w) dP(w), \quad A \in \mathscr{F}.$$

则称 f_1 为概率测度 Q 关于概率测度 P 的密度，且称概率测度 Q 关于概率测度 P 绝对连续，记为 $Q \ll P$.

定义 9.2. 若 P 关于 Q 绝对连续，且 Q 关于 P 绝对连续，即 $Q \ll P$ 与 $P \ll Q$ 同时成立，则称 P 和 Q 是等价的概率测度。

一般地，我们用 $B = (B_t, t \in [0,T])$ 表示在概率测度 P 下的 Brown 运动。若考虑如下形式的过程

$$\tilde{B}_t = B_t + qt, \quad t \in [0,T],$$

其中 q 为某个常数。当 $q \neq 0$ 时，$\tilde{B}_t$ 不是标准 Brown 运动。如果将概率测度 P 转换成一个新的合适的概率测度 Q 使得 B_t 在新的概率测度 Q 下是一个标准 Brown 运动，这即是 Girsanov 定理解决的问题。一般地，令

$$\mathscr{F}_t = \sigma(B_t, s \leq t), \quad t \in [0,T]$$

表示由 Brown 运动生成的自然 σ 代数流。

定理 9.1. *令 $Y(t) \in \mathbb{R}^n$ 为一个 Itô 过程:*

$$dY_t = \rho_t dt + \theta_t dB_t,$$

其中 $B_t \in \mathbb{R}^m$, $\rho_t \in \mathbb{R}^n$, $\theta_t \in \mathbb{R}^{m\times n}$,假设存在过程 u_t 和 α_t,使得 $\theta_t u_t = \rho_t - \alpha_t$，令

$$M_t = \exp\left\{-\int_0^t u_s dB_s - \frac{1}{2}\int_0^t u_s^2 ds\right\}, \quad dQ(w) = M_T(w)dP(w),$$

其中 M_t 是关于概率测度 P 的鞅，则 Q 也是概率测度；过程 $\tilde{B}_t = \int_0^t u_s ds + B_t$, $t \leq T$ 是关于概率测度 Q 的一个标准 Brown 运动；且过程 Y_t 满足如下方程：

$$dY_t = \alpha_t dt + \theta_t d\tilde{B}_t.$$

测度变换的目的是消去随机微分方程中的漂移项。

例 9.1. 假设 $Y(t) = (Y_1(t), Y_2(t))^T \in \mathbb{R}^2$，且有

$$\begin{aligned}
dY_1(t) =& 2dt + dB_1(t) + dB_2(t),\\
dY_2(t) =& 4dt + dB_1(t) - dB_2(t),\\
dY(t) =& \begin{pmatrix} 2 \\ 4 \end{pmatrix} dt + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} dB(t); \quad B(t) = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix}.
\end{aligned}$$

选择 $\alpha_t = 0$，则有

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

有唯一解 $u_1 = 3,\ u_2 = -1$。故令 $dQ(w) = \exp\{-3B_1(T) + B_2(T) - 5T\}dP(w)$，

$$d\tilde{B}(t) = \begin{pmatrix} 3 \\ -1 \end{pmatrix} dt + dB(t),$$

在概率测度 Q 下，$B(t)$ 是一个标准 Brown 运动，且

$$dY(t) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} d\tilde{B}(t).$$

以下运用 Girsanov 定理来推导 Black-Scholes 公式：

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 Q 下，$\tilde{S}_t$ 是一个鞅。令 $f(t,x) = e^{-rt}x$，由 Itô 公式有

$$d\tilde{S}_t = \sigma\tilde{S}_t d\tilde{B}_t,\ \ \tilde{B}_t = B_t + \frac{\mu - r}{\sigma}t,\ t \in [0,T].$$

由 Girsanov 定理知，存在一个等价鞅测度 Q，使得 $\tilde{B}_t$ 在此测度下是一个标准 Brown 运动，上式解为：

$$\tilde{S}_t = \tilde{S}_0 \exp\left\{-\frac{1}{2}\sigma^2 t + \sigma\tilde{B}_t\right\},$$

在概率测度 Q 下为一个鞅。

定理 9.2. 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

$$V_t = a_t S_t + b_t \beta_t,\quad t \in [0,T],$$

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

$$V_T = h(S_T),$$

则有投资组合在到期日 t 时刻的价值为

$$V_t = E_Q\left[e^{-r(T-t)}h(S_T)\,|\,\mathscr{F}_t\right],\quad t \in [0,T],$$

其中 $E_Q(A|\mathscr{F}_t)$ 表示在新的概率测度 Q 下随机变量 A 关于 $\mathscr{F}_t = \sigma(B_s, s \le t)$ 的条件期望。

证明：考虑贴现的财富过程：

$$V_t = e^{-rt}V_t = e^{-rt}(a_tS_t + b_t\beta_t).$$

由 Itô 公式有

$$d\tilde{V}_t = -r\tilde{V}_t dt + e^{-rt}dV_t,$$

且有

$$\begin{aligned} d\tilde{V}_t =& - re^{-rt}(a_tS_t + b_t\beta_t)dt + e^{-rt}(a_t dS_t + b_t d\beta_t) \\ =& a_t(-re^{-rt}S_t dt + e^{-rt}dS_t) \\ =& a_t d\tilde{S}_t. \end{aligned}$$

$\tilde{V}_0 = V_0$，故有

$$\tilde{V}_0 = V_0 + \int_0^t a_s d\tilde{S}_s = V_0 + \sigma\int_0^t a_s \tilde{S}_s d\tilde{B}_s,$$

在等价鞅测度 Q 下，$\tilde{B}$ 为标准 Brown 运动，$\{a_t\tilde{S}_t, t\in[0,T]\}$ 适应的。故 $\{\tilde{V}_t\}$ 为关于 $(\mathscr{F}_t)$ 的鞅。且由鞅的性质有

$$\tilde{V}_t = E_Q[\tilde{V}_T \,|\, \mathscr{F}_t] \;\; t\in[0,T],$$

但

$$\tilde{V}_T = e^{-rt}V_T = e^{-rT}h(S_T),$$

故

$$e^{-rt}V_t = E_Q[e^{-rT}h(S_T) \,|\, \mathscr{F}_t],$$

即

$$V_t = E_Q[e^{-r(T-t)}h(S_T) \,|\, \mathscr{F}_t],$$

定理得证。

由定理9.2，我们计算欧式看涨期权的定价公式：令 $\theta = T - t$, $t\in[0,T]$，故投资组合在时刻 t 的价值为：

$$\begin{aligned} V_t =& E_Q[e^{-r\theta h(S_T)}|\mathscr{F}_t] \\ =& E_Q[\exp\{-r\theta h(S_t\exp\{(r-\frac{1}{2}\sigma^2)\theta + \sigma(\tilde{B}_T - \tilde{B}_t)\})\} \,|\, \mathscr{F}_t]. \end{aligned}$$

在时刻 t，S_t 是 B_t 的函数，故有 $\sigma(S_t) \subset \mathscr{F}_t$；另外，在概率测度 Q 下，$\tilde{B}_T - \tilde{B}_t$ 关于 $\mathscr{F}_t$ 独立并且具有正态分布。由条件期望的性质有：$V_t = f(t, S_t)$，其中

$$f(t,x) = e^{-r\theta} \int_{-\infty}^{+\infty} h\left[x \exp\left\{(r - \frac{1}{2}\sigma^2)\theta + \sigma y \theta^{\frac{1}{2}}\right\}\right] \varphi(y) dy,$$

$\varphi(y)$ 表示标准正态分布的密度函数。

对于欧式看涨期权来说，

$$h(x) = (x-K)^+ = \max\{0, x-K\}.$$

故

$$\begin{aligned} f(t,x) = & \int_{-z_2}^{+\infty} \left(x \exp\{-\frac{1}{2}\sigma^2\theta + \sigma y \theta^{\frac{1}{2}}\} - Ke^{-r\theta}\right) \varphi(y) dy \\ = & x\Phi(z_1) - Ke^{-r\theta}\Phi(z_2), \end{aligned}$$

其中 $\Phi(x)$ 表示标准正态分布的分布函数，且

$$z_1 = \frac{\ln \frac{x}{K} + (r + \frac{1}{2}\sigma^2)\theta}{\sigma\sqrt{\theta}}, \qquad z_2 = z_1 - \sigma\sqrt{\theta}.$$

对于欧式看跌期权，$h(x) = (K-x)^+$，采用类似的方式得出：

$$f(t,x) = Ke^{-r\theta}\Phi(-z_2) - x\Phi(-z_1).$$

本节介绍的是经典的 Black-Scholes 公式。在对股票价格波动率的实际研究中，波动率呈现出一些统计特征：如波动率微笑，肥尾分布，群聚效应，均值回复，杠杆作用等。针对这些现象，许多学者对 Black-Scholes 模型中标的资产服从几何 Brown 运动、波动率为常数等假设作了改进。对标的资产的分布的修正中，提出用跳扩散过程等来描述股价过程。对波动率为常数的修正中，根据波动率函数的特点分为两大类：确定性波动率模型和随机波动率模型。

Chapter 10

随机过程在保险中的应用

10.1 基本概念

10.2 经典破产理论介绍

Chapter 11

MCMC 方法

参考文献

Glasserman, Paul. 2004. *Monte Carlo Methods in Financial Engineering*. Springer.

Karatzas, I., and S. E. Shreve. 1998. *Brownian Motion and Stochastic Calculus*. 2nd ed. Springer, New York.

Ross, Sheldon M. 2019. *Introduction to Probability Models*. 12th ed. Academic Press. https://doi.org/https://doi.org/10.1016/C2017-0-01324-1.

Shreve, Steven E. 2004. *Stochastic Calculus for Finance II Continuous Time Models*. Springer.

严加安. 2023. 金融数学引论. 2nd ed. 科学出版社.

何书元. 2003. 应用时间序列分析. 北京大学出版社.

———. 2008. 随机过程. 北京大学出版社.

何声武. 1999. 随机过程导论. 高等教育出版社.

刘勇. 2022. 应用随机分析. https://www.math.pku.edu.cn/teachers/liuyong/asa/lectnote22.pdf.

张波, 商豪, and 邓军. 2020. 应用随机过程. 人民大学出版社.

林元烈. 2002. 应用随机过程. 清华大学出版社.

王寿仁. 1997. 概率论基础和随机过程. 科学出版社.

谢衷洁. 1990. 时间序列分析. 北京大学出版社.

钱敏平. 1990. 随机过程引论. 北京大学出版社.

钱敏平, 龚光鲁, 陈大岳, and 章复熹. 2011. 应用随机过程. 高等教育出版社.

龚光鲁, and 钱敏平. 2019. 随机微分方程引论. 第 3 版 ed. 电子工业出版社.