



随机过程杂记

作者：邓军

组织：对外经济贸易大学

时间：May 5, 2020

版本：0.1

特别声明和感谢

本文采用Elegant $\LaTeX$ 模板。该模板可以从 [GitHub](#)、[CTAN](#)、[Overleaf](#) 以及 [Gitee](#) 上下载。

本杂记基于 2020 年《应用随机过程》。感谢课程学生对本杂记的阅读和反馈。特别感谢何羽箏同学帮忙录入部分 Tex 文档。

邓军

May 5, 2020

目录

1	概率论基础	1
1.1	为什么需要概率空间	1
1.1.1	理发师悖论 (Barber paradox)	1
1.1.2	贝特朗悖论 (Bertrand's Paradox)	1
1.1.3	非悖论, 生日问题	3
1.2	概率空间	3
1.2.1	可测空间	4
1.2.2	概率空间	8
1.2.3	条件概率	10
1.2.4	全概率公式和 Bayes 公式	11
1.3	随机变量和分布函数	13
1.3.1	数字特征	17
1.3.2	矩函数 (Moment Generating Function)	17
1.3.3	特征函数 (Characteristic function)	18
1.3.4	反演公式及唯一性定理	21
1.3.5	多维随机变量的特征函数	24
1.4	独立性与条件期望	25
1.4.1	独立性	25
1.4.2	条件期望	26
1.4.3	条件分布	27
1.4.4	一般条件期望 $\star$	31
2	随机过程的基本概念与类型	34
2.1	随机过程的背景	34
2.2	基本概念	34
2.3	有限维分布与 Kolmogorov 定理	36
2.3.1	随机过程的数字特征	37
2.4	随机过程的基本类型	40
2.4.1	平稳过程	40
2.4.2	独立增量过程	41
3	Brown 运动 (维纳过程)	44
3.1	基本概念与性质	44
3.2	维纳过程的分布	48

3.3	维纳过程的数字特征	49
3.3.1	二次变差	51
3.4	Brown 运动的鞅性质	55
3.5	Brown 运动的最大值变量及反正弦律	56
3.6	Brown 运动的几种变化	58
3.6.1	Brown 桥	58
3.6.2	几何 Brown 运动	59
4	Poisson 过程	61
4.1	齐次泊松过程	61
4.1.1	Poisson 过程数学模型	61
4.1.2	齐次泊松过程的数字特征	66
4.1.3	时间间隔与等待时间的分布	68
4.1.4	到达时间的条件分布	70
4.1.5	更新计数过程	70
4.2	复合泊松过程	73
4.2.1	复合 Poisson 过程	73
4.3	非齐次泊松过程 (了解内容, 不考察)	75
5	鞅 (Martingale) 过程	77
5.1	基本概念	77
5.2	鞅的停时定理及其应用	83
5.2.1	鞅的停时定理	83
5.3	连续鞅	87

第一章 概率论基础

内容提要

- | | |
|--------------------------------------|---------------------------------|
| ☐ 悖论 | ☐ 随机变量 |
| ☐ 样本空间 | ☐ 分布函数 |
| ☐ σ 代数 | ☐ 特征函数 |
| ☐ 可测空间 | ☐ 独立性 |
| ☐ 概率空间 | ☐ 条件数学期望 |
| ☐ 测度完备化 | |

1.1 为什么需要概率空间

我们需要概率空间的目的有三个：（1）避免语言描述的模糊性（2）模糊性带来的概率悖论，（3）严格的定义随机变量和随机过程。

1.1.1 理发师悖论 (Barber paradox)

小城里的理发师要为所有不给自己刮脸的人刮脸。那么理发师是否应该给自己刮脸呢？显然，如果理发师不给自己刮脸，那么根据他的话，他应该自己刮脸，所以矛盾了。如果理发师给自己刮脸，根据他的描述，他不应该给自己刮脸，这样也矛盾了。

这个例子告诉我们：语言描述模糊性带来了悖论。

1.1.2 贝特朗悖论 (Bertrand's Paradox)

问题描述：在一个半径为 1 的圆内画一个内接等边三角形。如果在圆的周长上随机选择两点，并连接这两个点。求这条弦比内接等边三角形边长要长的概率？

- 第一种计算方式：如下图所示，固定内接等边三角形的一个顶点作为弦的一个点。因此随机选取另外一个点，作为这条弦的另外一个点。显而易见，当这条弦在 60 度-90 度这个范围内比三角形的边长要长。如下图所示的红色线段。因此，概率为 $1/3$ 。
- 第二种计算方式：我们选择和内接等边三角形一条边平行的弦。因此，我们可以过圆心作三角形一条边的垂线。与三条线边平行线中有一半比三条线的边长要长。因此，概率为 $1/2$ 。
- 第三种计算方式：我们再作一个等边三角形的内切圆。只有当弦穿过了内切圆，弦才比三角形的边长要长。因此，概率为

$$\frac{\text{小圆面积}}{\text{大圆面积}} = \frac{1}{4}.$$

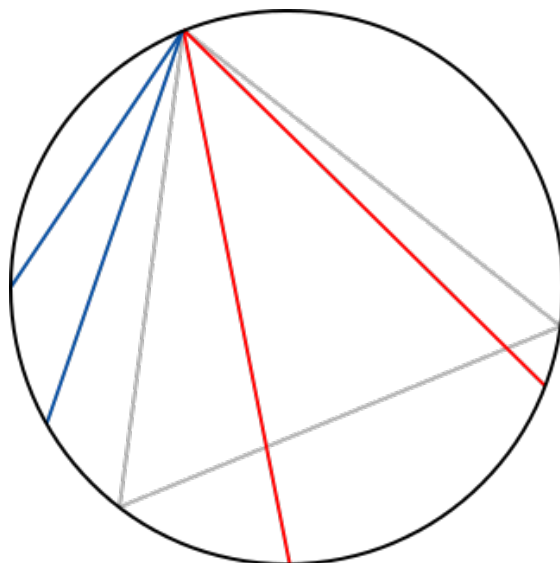


图 1.1: 贝特朗悖论第一种计算方式

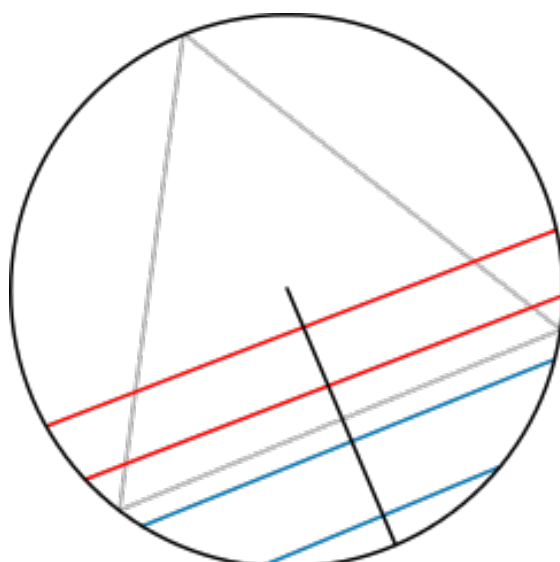


图 1.2: 贝特朗悖论第二种计算方式

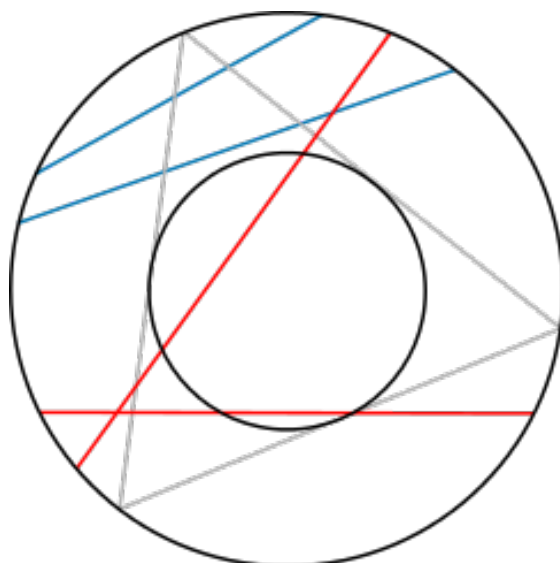


图 1.3: 贝特朗悖论第三种计算方式

这个例子告诉我们：三种计算的概率都是正确的，但是得到的结论呢却不一致。这是因为语言模糊性带来随机选取的不一致，从而带来概率计算的不确定性。

1.1.3 非悖论, 生日问题

生日问题是指，如果在一个房间要多少人，则两个人的生日相同的概率要大于 50%？答案是 23 人。这就意味着在一个典型的标准班级（30 人）中，存在两人生日相同的可能性更高。对于 60 或者更多的人，这种概率要大于 99%。这个问题有时也被称做生日悖论，但从引起逻辑矛盾的角度来说生日悖论并不是一种悖论，它被称作悖论只是因为这个数学事实与一般直觉相抵触而已。大多数人会认为，23 人中有 2 人生日相同的概率应该远远小于 50%。计算与此相关的概率被称为生日问题，在这个问题之后的数学理论已被用于设计著名的密码攻击方法：生日攻击。

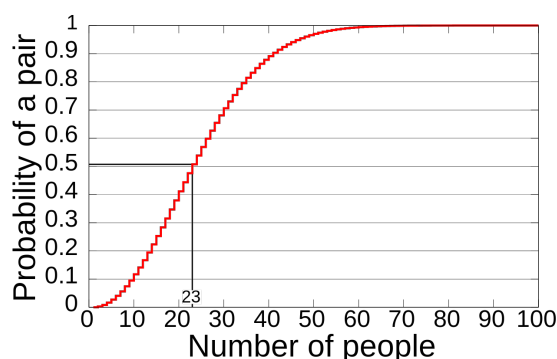


图 1.4: 生日问题

理解 1.1. 悖论

从这些例子，我们看到为了严格的说明什么是随机的，为了避免语言描述带来的模糊性。我们需要严格的定义所讨论的随机所定义的空间！



1.2 概率空间

当我们做一系列的随机试验，能够得到不同的结果，为了研究这些可能的结果。我们把随机试验所有的可能性放在一起，成为样本空间。例如，我们抛掷一个六面的骰子，可能会得到 $\{1, 2, \dots, 6\}$ 中的任何一个数字。因此，我们用 $\Omega = \{1, 2, \dots, 6\}$ 表示抛骰子这个实验的样本空间。

定义 1.1. 样本空间 Ω

我们感兴趣的实验所有结果放在一起构成的集合，称为样本空间 Ω 。其中的元素称为样本空间 Ω 的点。



例 1.1 如果我们想研究抛一次硬币的结果，我们可以用如下的 Ω 来表示样本空间。

$$\Omega = \{\text{正面}, \text{反面}\}. \quad (1.1)$$

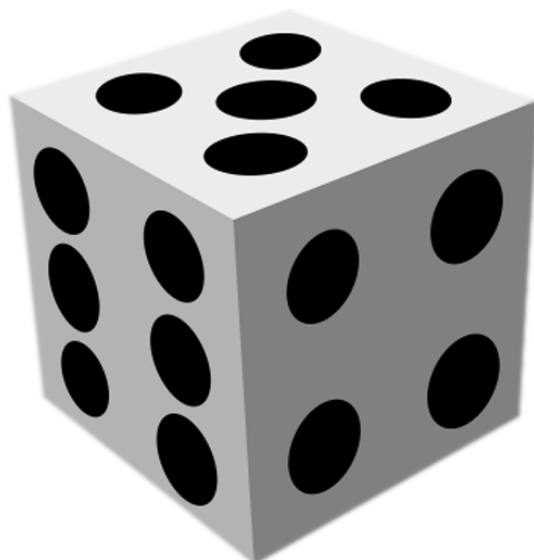


图 1.5: 抛骰子的样本空间

例 1.2 如果我们想研究抛两次硬币的结果, 我们可以用如下的 Ω 来表示样本空间。

$$\Omega = \{\{\text{正面, 正面}\}, \{\text{反面, 正面}\}, \{\text{正面, 反面}\}, \{\text{反面, 反面}\}\}. \quad (1.2)$$

思考抛无穷次硬币的结果的样本空间?

当我们有了样本空间 Ω , 这时我们可以把研究对象限定在该指定的样本空间上。这样就避免了**随机选取**带来的模糊性。例如在贝特朗悖论中, 如果我们事先制定了如何去画弦, 这样就相当于固定了样本空间 Ω , 就能避免前面得到的三种不同的概率。

在样本空间 Ω 基础上, 如果我们还感兴趣**某几个事情一起发生或者都不发生的概率**, 这时我们就希望能够把样本空间 Ω 一些结果当作一个整体放在一起考虑。这就有了子集、事件、 σ 代数的引入。例如我们想研究抛两次硬币第一次出现是正面的概率。在例 1.2 的样本空间 Ω 上, 我们可以选取 $A = \{\{\text{正面, 正面}\}, \{\text{正面, 反面}\}\}$ 作为一个整体事件。下一节我们给出这些概念的详细定义。

1.2.1 可测空间

定义 1.2. 事件 (子集)

样本空间 Ω 中的点构成的集合, 称为 Ω 的子集或者事件。



例 1.3 例如 $\Omega = \{\omega_1, \omega_2, \dots, \omega_{10}\}$ 。我们称 $\omega_i, i = 1, 2, \dots, 10$ 为样本空间 Ω 中的点或者元素。类似 $A = \{\omega_1, \omega_3\}, B = \{\omega_4, \omega_7, \omega_9\}$ 这样由其元素构成的子集就称为事件。注意以后我们在使用符号 ω_i 时, 它可以表示任意的含义, 例如硬币的正反面, 骰子的点数, 一个班级的同学。**它只是一个抽象的符号。**

有了事件的概念, 接下来我们想研究一系列事件发生的概率或者两个事件交集发生的概率, 这时我们就需要 σ 代数的概念。例如在上面例子 1.3 中, 如果我们想知道 A, B 同时发生的概率, 这时我们就需要研究 $A \cap B$; 如果想研究 A, B 有一个发生的概率, 这时

我们就需要研究 $A \cup B$ 。

定义 1.3. σ 代数

Ω 是一个样本空间, $\mathcal{F}$ 是一些由 Ω 子集构成的集合。如果 $\mathcal{F}$ 满足如下三个条件,

1. $\Omega \in \mathcal{F}$;
2. 如果 $A \in \mathcal{F}$, 则我们有 A 的补集 $A^c = \Omega \setminus A \in \mathcal{F}$;
3. 如果一系列集合 $A_n \in \mathcal{F}, n = 1, 2, \dots, \infty$, 则有 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ 。

我们则称 $\mathcal{F}$ 为一个 σ 代数。并称 $(\Omega, \mathcal{F})$ 为可测空间。 $(\Omega, \mathcal{F})$ 中的元素称为可测集或者可测事件。



我们用 $\emptyset$ 表示空集。 $A \in \mathcal{F}$ 表示 A 在 $\mathcal{F}$ 中。对于两个集合 A, B , 我们用 $A \setminus B$ 表示两个的差, 即:

$$A \setminus B = \{\omega | \omega \in A, \omega \notin B\}.$$

理解 1.2. σ 代数

定义 σ 代数是为了去研究 Ω 中单个事件或者事件组合所发生的概率。因此, 我们需要在 $\mathcal{F}$ 包含一系列的事件。但并不是所有的事件放在一起都对我们研究一般问题有效。直观来说,

- $\mathcal{F}$ 必须包含事件的全体, 这样我们才能讨论必然会发生的事情。因此有了 $\Omega \in \mathcal{F}$ 。
- 如果一个事情的正面我们感兴趣, 一般我们也对其反面 (不发生) 感兴趣。因此有了如果 $A \in \mathcal{F}$, 我们要求 A 的补集 $A^c = \Omega \setminus A \in \mathcal{F}$;
- 我们希望如果有无穷多个事情发生的情况或者极限下会怎么样。因此有了如果一系列事件 $(A_n)_{n \geq 1} \in \mathcal{F}$, 我们希望 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ 。
- 只要一个集合在 $\mathcal{F}$ 中, 我们就称为它为关于 $\mathcal{F}$ 的可测集合。所以一个集合是否可测, 依赖于我们定义的 σ 代数 $\mathcal{F}$ 。
- 我们可以把 σ 代数 $\mathcal{F}$ 想象成我们所有知识的集合体。只有在 $\mathcal{F}$ 中的事件, 我们才能够认识!



例 1.4 我们想研究抛一次六面骰子出现 2 或者 3。这时我们可以把这个问题用可测空间来严格描述出来。

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

因为是研究“2 或者 3”, 因此我们把 2, 3 放在一起作为一个整体。因此 $\mathcal{F}$ 包含了事件 $\{2, 3\}$ 。为了让 $\mathcal{F}$ 成为 σ 代数我们还包含了其他的事件。显然在上面的 $\mathcal{F}$ 中, $\{1\}$ 不在里面, 虽然 $\{1\}$ 是 $\{1, 4, 5, 6\}$ 的一个子集。因此, 我们说 $\{1\}$ 不是上面 $\mathcal{F}$ 的可测集, 我们也无法认识她。同样 $\{2\}$ 也不是 $\mathcal{F}$ 的可测集。

例 1.5 如果想研究抛一次六面骰子出现分别出现 2 和出现 3 的概率。这时我们可以把这个问题用可测空间来严格描述出来。

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \mathcal{F}_1 = \{\emptyset, \Omega, \{2\}, \{3\}, \{2, 3\}, \{1, 2, 4, 5, 6\}, \{1, 3, 4, 5, 6\}, \{1, 4, 5, 6\}\}.$$

因为是研究“分别出现 2 和出现 3”，因此我们把 2 和 3 需要单独分离开来研究，而**不能**放在一起作为一个整体。为了让 $\mathcal{F}$ 成为 σ 代数我们还包含了其他的事件。在 σ 代数 $\mathcal{F}_1$ 中， $\{2\}$ 是 $\mathcal{F}_1$ 的可测集，但不是例子 1.4 中 $\mathcal{F}$ 的可测集。

上面例子告诉我们，可测集依赖于我们定义的 σ 代数!!

例 1.6 一些 σ -代数的例子。

- 最小的 σ -代数：平凡 (trivial) σ -代数 $\{\emptyset, \Omega\}$.
- 最大的 σ -代数：全体子集 $2^\Omega := \{A : A \subset \Omega\}$.
- 分割形成的 σ -代数：令 $A_1, A_2, \dots, A_n$ 为 Ω 的不相交的子集，并有 $\bigcup_{i=1}^n A_i = \Omega$.
则由 $A_1, A_2, \dots, A_n$ 进行有限并、补运算产生一个 σ -代数，其中包含 2^n 个集合。
- $\mathcal{F} = \{A, A^c, \emptyset, \Omega\}$.

定义 1.4. σ -代数的生成

对于 Ω 的一个子集族 $\mathcal{S}$ ，如果存在 Ω 上的 σ -代数 $\mathcal{F}$ 使得

- $\mathcal{S} \subset \mathcal{F}$;
- 对于任意包含 $\mathcal{S}$ 的 Ω 上的 σ -代数 $\mathcal{F}'$ ，都有 $\mathcal{F} \subset \mathcal{F}'$,

则称 $\mathcal{F}$ 是由子集系 $\mathcal{S}$ 生成的 (最小的) σ -代数，记作 $\mathcal{F} := \sigma(\mathcal{S})$.



命题 1.1

下面命题成立：

- (1) 对 Ω 的任何子集系 $\mathcal{S}$ ，都存在 $\mathcal{F} := \sigma(\mathcal{S})$.
- (2) 如果 $\mathcal{S}$ 本身是 Ω 上的一个 σ -代数，则 $\sigma(\mathcal{S}) = \mathcal{S}$.



例 1.7

- $\Omega = \{\omega_1, \omega_2, \dots, \omega_4\}$, $\mathcal{S} = \{\{\omega_1\}, \{\omega_2, \omega_3\}\}$ 。则：由 $\mathcal{S}$ 生成的 σ -代数为：

$$\begin{aligned} \sigma(\mathcal{S}) = & \{\emptyset, \Omega, \{\omega_1\}, \{\omega_2, \omega_3, \omega_4\}, \\ & \{\omega_1, \omega_2, \omega_3\}, \{\omega_4\}, \\ & \{\omega_2, \omega_3\}, \{\omega_1, \omega_4\}\} \end{aligned}$$

- 如果 $\mathcal{S} = \{\{\omega_1\}, \{\omega_2\}\}$ 。则：由 $\mathcal{S}$ 生成的 σ -代数为：

$$\begin{aligned} \sigma(\mathcal{S}) = & \{\emptyset, \Omega, \{\omega_1\}, \{\omega_2, \omega_3, \omega_4\}, \\ & \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \\ & \{\omega_2\}, \{\omega_1, \omega_3, \omega_4\}\} \end{aligned}$$

定义 1.5. 集合上下极限

如果 $\{A_n, n = 1, 2, \dots, \infty\}$ 为集合序列。我们定义

$$\limsup_{n \rightarrow +\infty} A_n = \bigcap_{n=1}^{+\infty} \left(\bigcup_{k=n}^{\infty} A_k \right), \quad \liminf_{n \rightarrow +\infty} A_n = \bigcup_{n=1}^{+\infty} \left(\bigcap_{k=n}^{\infty} A_k \right). \quad (1.3)$$

从上面定义可以看到如果定义 $B_n = \bigcup_{k=n}^{\infty} A_k$ ，则 B_n 为一递减集合序列。因此，

B_n 的极限可以定义为无穷的交集, 并且

$$\limsup_{n \rightarrow +\infty} A_n = \lim_{n \rightarrow +\infty} B_n = \bigcap_{n=1}^{+\infty} B_n$$

。如果定义 $C_n = \bigcap_{k=n}^{\infty} A_k$, 则 C_n 为一递增集合序列。因此, C_n 的极限可以定义为无穷的并集, 并且

$$\liminf_{n \rightarrow +\infty} A_n = \lim_{n \rightarrow +\infty} C_n = \bigcup_{n=1}^{+\infty} C_n.$$



理解 1.3. 集合上下极限

类似于数列, 数列的极限不一定存在, 但是我们可以定义数列的上下极限。同理, 对于集合序列一样, 其极限也不一定存在。



例 1.8 如果集合序列 A_n 如下

$$A_n = [0, 1 + \frac{1}{n}].$$

显然, A_n 是一个递减集合序列。可以得到

$$\begin{aligned} \limsup_{n \rightarrow +\infty} A_n &= \bigcap_{n=1}^{+\infty} \left(\bigcup_{k=n}^{\infty} [0, 1 + \frac{1}{k}] \right) = \bigcap_{n=1}^{+\infty} [0, 1 + \frac{1}{n}] = [0, 1], \\ \liminf_{n \rightarrow +\infty} A_n &= \bigcup_{n=1}^{+\infty} \left(\bigcap_{k=n}^{\infty} [0, 1 + \frac{1}{k}] \right) = \bigcup_{n=1}^{+\infty} [0, 1] = [0, 1]. \end{aligned}$$

例 1.9 如果集合序列 A_n 如下

$$A_{2n} = [0, 2 + \frac{1}{n}], \quad A_{2n+1} = [0, 1 + \frac{1}{n}].$$

可以得到

$$\begin{aligned} B_{2n} &= \bigcup_{k=2n}^{\infty} A_k = A_{2n} \cup A_{2n+1} \cup A_{2n+2} \cup \cdots \\ &= (A_{2n} \cup A_{2n+2} \cup \cdots) \cup (A_{2n+1} \cup A_{2n+3} \cup \cdots) \\ &= [0, 2 + \frac{1}{n}] \cup [0, 1 + \frac{1}{n}] = [0, 2 + \frac{1}{n}]. \\ B_{2n+1} &= \bigcup_{k=2n+1}^{\infty} A_k = A_{2n+1} \cup A_{2n+2} \cup \cdots \\ &= (A_{2n+1} \cup A_{2n+3} \cup \cdots) \cup (A_{2n+2} \cup A_{2n+4} \cup \cdots) \\ &= [0, 1 + \frac{1}{n}] \cup [0, 2 + \frac{1}{n+1}] = [0, 2 + \frac{1}{n+1}], \\ \limsup_{n \rightarrow +\infty} A_n &= \bigcap_{n=1}^{+\infty} B_n = \left(\bigcap_{n=1}^{+\infty} B_{2n} \right) \cap \left(\bigcap_{n=0}^{+\infty} B_{2n+1} \right) \\ &= \left(\bigcap_{n=1}^{+\infty} [0, 2 + \frac{1}{n}] \right) \cap \left(\bigcap_{n=0}^{+\infty} [0, 2 + \frac{1}{n+1}] \right) = [0, 2]. \end{aligned}$$



同理可以得到

$$\liminf_{n \rightarrow +\infty} A_n = [0, 1] \neq \limsup_{n \rightarrow +\infty} A_n = [0, 2].$$

1.2.2 概率空间

有了可测空间 $(\Omega, \mathcal{F})$ 我们能够定义可测集，也就是我们可以认识的事件。但现在还没有一把度量的直尺去刻画事件的大小。下面我们就引入这个度量工具：概率。

定义 1.6. 概率空间

$(\Omega, \mathcal{F})$ 为一个可测空间。 P 为一个函数 $P: \mathcal{F} \rightarrow [0, 1]$ 。其定义域为 σ 代数 $\mathcal{F}$ ，值域为 $[0, 1]$ 。如果函数 P 满足：

1. $P(\Omega) = 1$;
2. 对任意的 $\forall A \in \mathcal{F}$, 有 $0 \leq P(A) \leq 1$;
3. 如果 $A_1, A_2, \dots, A_\infty$ 为任意不相交的集合序列，则有

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \quad (1.4)$$

我们把三元组 $(\Omega, \mathcal{F}, P)$ 称为概率空间。



理解 1.4. 概率

概率定义了如何刻画 σ 代数 $\mathcal{F}$ 中事件发生的大小。我们只能度量可以认识事件的概率大小，也即是 σ 代数 $\mathcal{F}$ 中的可测集！



例 1.10 假设 $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $\mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3\}\}$. 定义

$$P(\{\omega_1, \omega_2\}) = 0.5, P(\{\omega_3\}) = 0.5.$$

因此，上面的概率 P 刻画事件 $\{\omega_1, \omega_2\}$ 和 $\{\omega_3\}$ 发生的概率。但是，我们并不知道事件 $\{\omega_1\}$ 发生的概率大小。因为， $\{\omega_1\}$ 关于 $\mathcal{F}$ 不可测，也不在概率 P 的定义域内。

定义 1.7. 零测集

如果一个事件发生的概率为 0，我们称其为零测集。



在随机过程的研究中，零测集通常没有什么作用，对于研究的问题也没有什么影响。但是一个零测集并不代表是空集，依赖于概率的定义。所以我们通常采用下面完备化的技巧是的是这个概率空间性质更好一些。

定义 1.8. 完备的概率测度

令 $\mathcal{N}$ 代表 Ω 的所有 P 零测集的子集的全体，由 $\{\mathcal{F}, \mathcal{N}\}$ 生成的 σ 代数 (即包含 $\mathcal{F}$ 和 $\mathcal{N}$ 的最小 σ 代数) 称为 $\mathcal{F}$ 的完备化，记为 $\hat{\mathcal{F}}$ 。

$\hat{\mathcal{F}}$ 中的每个集合 B 都可以表为 $B = A \cup N$ ，其中 $A \in \mathcal{F}, N \in \mathcal{N}$, 且 $A \cap N = \emptyset$ 。

定义

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

则 P 就被扩张到 $\hat{\mathcal{F}}$ 上。

容易验证, $\hat{P}$ 是 $\hat{\mathcal{F}}$ 上的概率测度, 函数 $\hat{P}$ 称为 P 的完备化。我们总假定 P 是完备的概率测度。



例 1.11 假设 $\Omega = \{\omega_1, \omega_2, \omega_3\}, \mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3\}\}$. 定义

$$P(\{\omega_1, \omega_2\}) = 0, P(\{\omega_3\}) = 1.$$

因此, $\{\omega_1, \omega_2\}$ 为零测集。但是其子集, $\{\omega_1\}, \{\omega_2\}$ 都不是可测集。因此, 如果我们关心的是概率发生的大小, 我们不妨把 $\{\omega_1\}, \{\omega_2\}$ 加入到 $\mathcal{F}$ 中, 从而把她们都变成可测集, 同时也定义她们的概率为 0。因此, 定义

$$\hat{\mathcal{F}} = \{\emptyset, \Omega, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}\}.$$

$$\hat{P}(\{\omega_1\}) = \hat{P}(\{\omega_2\}) = 0, \hat{P}(\{\omega_3\}) = 1.$$

显然, 新得到的 $(\Omega, \hat{\mathcal{F}}, \hat{P})$ 是概率空间 $(\Omega, \mathcal{F}, P)$ 的完备化。

理解 1.5. 概率空间的逻辑

概率空间的逻辑图可以如下表示:

$$\Omega \xrightarrow{\text{感兴趣的集合}} (\Omega, \mathcal{F}) \xrightarrow[\text{也就是可测集}]{\text{可以认识的事件}} (\Omega, \mathcal{F}, P) \xrightarrow[\text{可测事件的概率大小}]{\text{可认识事件的度量大小}} \text{随机变量}$$



命题 1.2. 概率的性质

事件的概率有如下性质:

1. 若 $A, B \in \mathcal{F}$, 则 $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.
2. 若 $A, B \in \mathcal{F}$, 且 $A \subset B$, 则 $P(B - A) = P(B) - P(A)$ (可减性).
3. 若 $A, B \in \mathcal{F}$, 且 $A \subset B$, 则 $P(A) \leq P(B)$ (单调性).
4. 若 $A_n \in \mathcal{F}, n \geq 1$, 则 $P(\bigcup_{n \geq 1} A_n) \leq \sum_{n \geq 1} P(A_n)$.
5. 从下连续: 若 $A_n \in \mathcal{F}$ 且 $A_n \uparrow A \in \mathcal{F}$, 则 $P(A) = \lim_{n \rightarrow \infty} P(A_n)$.
6. 从上连续: 若 $A_n \in \mathcal{F}$ 且 $A_n \downarrow A \in \mathcal{F}$, 则 $P(A) = \lim_{n \rightarrow \infty} P(A_n)$



证明 1、2、3 的证明很显然, 我们省掉了。

4. 引入如下记号:

$$B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus (A_1 \cup A_2), \dots, B_n = A_n \setminus \left(\bigcup_{i=1}^{n-1} A_i \right)$$

通过如上构造, 显然 B_n 为不相交的集合序列。并且对于所有的 $k, \bigcup_{i=1}^k A_i = \bigcup_{i=1}^k B_i, B_n \subset A_n$ 。因此可以得到 $P(B_n) \leq P(A_n)$ 和

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

5. 由于集合序列 A_n 递增, 所以 $A = \bigcup_{n=1}^{\infty} A_n$ 。通过采用 3. 中相同的新集合序列 B_n 。因

此可以得到

$$\begin{aligned} P(A) &= P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n) \\ &= \lim_{k \rightarrow \infty} \sum_{n=1}^k P(B_n) = \lim_{k \rightarrow \infty} P\left(\bigcup_{n=1}^k B_n\right) = \lim_{k \rightarrow \infty} P(A_k). \end{aligned}$$

6. 证明类似 5. 可以考虑 $C_n = \Omega - A_n$ 。这样新的集合序列就是递增的了, 可以直接用 5. 的结论。 ■

例 1.12 设某股票一天的成交笔数为 m , 基本事件为 $\{m\}$, 样本空间 $\mathcal{F}$ 是 Ω 的一切子集组成的集族, 则 $\mathcal{F}$ 是一个 Ω 代数. 定义 $P(\emptyset) = 0$, 并对 $A \in \mathcal{F}$ 令

$$P(A) = \sum_{k \in A} e^{-\lambda} \frac{\lambda^k}{k!}, \quad \lambda > 0.$$

证明: P 为可测空间 $(\Omega, \mathcal{F}, P)$ 上的概率测度.

证明 (1) 我们证明 $P(\Omega) = 1$.

$$P(\Omega) = \sum_{k \in \Omega} e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

(2) 对于 $\mathcal{F}$ 中任意不相交的集合序列 A_n , 可以得到

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{k \in \bigcup_{n=1}^{\infty} A_n} e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{n=1}^{\infty} \sum_{k \in A_n} e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{n=1}^{\infty} P(A_n).$$

因此, P 为可测空间 $(\Omega, \mathcal{F}, P)$ 上的概率测度. ■

1.2.3 条件概率

定义 1.9

概率空间 $(\Omega, \mathcal{F}, P)$ 上面, 如果 $A, B \in \mathcal{F}$ 且 $P(B) > 0$, 则定义

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

称为已知事件 B 发生的条件下, 事件 A 发生的条件概率. ♣

定理 1.1

概率空间 $(\Omega, \mathcal{F}, P)$ 上面, $B \in \mathcal{F}$, 且 $P(B) > 0$, 则对所有的 $A \in \mathcal{F}$, 有 $P(A|B)$ 相对应, 且函数 $P(\cdot|B)$ 满足如下的三条公理:

- $\forall A \in \mathcal{F}, 0 \leq P(A|B) \leq 1$
- $P(\Omega|B) = 1$
- $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 并且 $A_i \cap A_j = \emptyset, i \neq j$, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} P(A_i|B).$$

证明 该定理的证明很显然. 可以从条件概率的定义简单推导得到. ■

定义 1.10

$P_B := P(\cdot|B)$, 则 P_B 是可测空间 $(\Omega, \mathcal{F})$ 上的概率, 称 $(\Omega, \mathcal{F}, P_B)$ 是条件概率空间. 

定理 1.2

A 是概率空间 $(\Omega, \mathcal{F}, P_B)$ 上的正概率事件, $B \in \mathcal{F}$, 且 $P_A(B) > 0$, 则对任意 $C \in \mathcal{F}$ 有

$$P_A(C|B) = P(C|A \cap B).$$


证明 有定义可以得到:

$$\begin{aligned} P_A(C|B) &= \frac{P_A(C \cap B)}{P_A(B)} = \frac{P(C \cap B|A)}{P(B|A)} \\ &= \frac{P(ABC)}{P(A)} / \frac{P(AB)}{P(A)} = \frac{P(ABC)}{P(AB)} = P(C|A \cap B). \end{aligned}$$


理解 1.6. 条件概率

条件概率告诉了我们已经给定了一个事件下, 另外一个事件发生的概率。例如下面图片所示, 可以设想在一个样本空间中事件发生就是在平板上方掉落的球。如果球落在了横板上, 说明该事件就发生了。图片中红色挡板就是事件 A , 蓝色的挡板就是的事件 B 。在给定事件 A 条件下, 事件 B 的条件概率就是下落的小球首先已经触碰到红色挡板 A 的条件下, 并且再次触碰到挡板 B 的概率。该图例的动态演示可以在 <http://setosa.io/conditional/> 查看. 

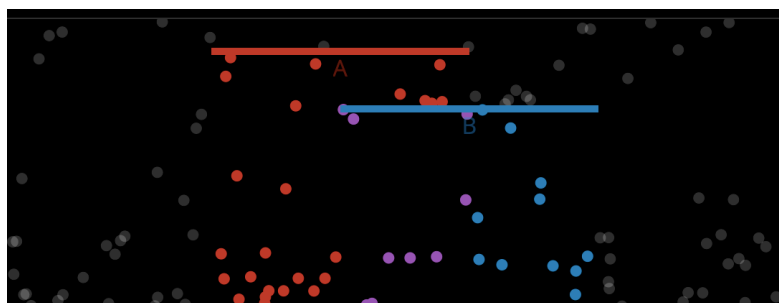


图 1.6: 条件概率

1.2.4 全概率公式和 Bayes 公式**定理 1.3**

概率空间 $(\Omega, \mathcal{F}, P)$ 上, 若

- $A_i \in \mathcal{F}$ 且 $P(A_i) > 0$ 对所有的 $i = 1, 2, \dots$
- $\cup_{i=1}^{\infty} A_i = \Omega$, $A_i \cap A_j = \emptyset$.

则对任意的 $B \in \mathcal{F}$, 有

$$P(B) = \sum_{i=1}^{\infty} P(A_i)P(B|A_i)$$

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{\infty} P(A_i)P(B|A_i)}.$$



理解 1.7. Bayes 公式

我们可以把 B 当作是某一个原因的结果, 而 A_n 是这些所有的可能性。全概率公式告诉我们, 如何通过细分全空间 Ω 得到 B 发生的概率 $P(B)$ 。而 Bayes 公式则反过来, 如果我们看到了结果 B , 去推测是原因 A_j 使得其发生的概率。

Bayes 公式在金融中也有重要的应用。例如在市场做市商要对一个资产提供连续报价时, 做市商需要面对可能的信息交易者和噪音交易者。所以如何报价, 一方面可以为市场提供流动性, 另外一方面又不至于在和信息交易者交易时过多亏损。因此, 做市商可以通过市场提交的订单流来推测信息交易者和噪音交易者的比例, 以此来提高卖价和买价, 以及价差, 来保持自己的做市能力。



例 1.13 高射炮向敌机发射三发炮弹, 每弹击中与否相互独立且每发炮弹击中的概率均为 0.3, 又知敌机若中一弹, 坠毁的概率为 0.2, 若中两弹, 坠毁的概率为 0.6, 若中三弹, 敌机必坠毁。求敌机坠毁的概率。

$$B = \{ \text{敌机坠毁} \}, A_i = \{ \text{敌机中了 } i \text{ 个子弹} \}$$

因此, 我们可以计算

$$\begin{aligned} P(B) &= \sum_{i=0}^3 P(A_i)P(B|A_i) \\ &= (1-0.3)^3 * 0 + C_3^1 * 0.3 * (1-0.3)^2 * 0.2 \\ &\quad + C_3^2 * 0.3^2 * (1-0.3) * 0.6 + C_3^3 * 0.3^3 * 1 \\ &= 0.2286. \end{aligned}$$

计算在敌机坠毁的条件下, 该机中了一发炮弹的概率。

$$\begin{aligned} P(A_1|B) &= \frac{P(A_1)P(B|A_1)}{\sum_{i=0}^3 P(A_i)P(B|A_i)} \\ &= \frac{C_3^1 * 0.3 * (1-0.3)^2 * 0.2}{0.2286} = 0.385826772. \end{aligned}$$

1.3 随机变量和分布函数

内容提要

□ 随机变量

□ 特征函数

□ 分布函数

定义 1.11. Borel σ 代数

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 $\mathcal{B}(\mathbb{R})$, 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 $\mathcal{B}(\mathbb{R}^n)$.



定义 1.12. 随机变量

设 $(\Omega, \mathcal{F}, P)$ 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 $a \in \mathbb{R}$, $\{\omega : X(\omega) < a\} \in \mathcal{F}$, 则称 $X(\omega)$ 是 $\mathcal{F}$ 上的**随机变量**.

$$F(a) = P(\omega : X(\omega) < a), \quad -\infty < a < \infty$$

称为随机变量 X 的**分布函数**.



注: 习惯上将 $\{\omega : X(\omega) \geq a\}$ 记为 $\{X \geq a\}$. 回顾概率空间的逻辑图:

理解 1.8. 概率空间的逻辑

$$\Omega \xrightarrow{\text{感兴趣的集合}} (\Omega, \mathcal{F}) \xrightarrow[\text{也就是可测集}]{\text{可以认识的事件}} (\Omega, \mathcal{F}, P) \xrightarrow[\text{可测事件的概率大小}]{\text{可认识事件的度量大小}} \text{随机变量}$$



理解 1.9. 随机变量

在我们定义了概率空间 $(\Omega, \mathcal{F}, P)$ 后, 我们更感兴趣的是在上面定义一些变量来描述更一般的事物。例如定义一个变量来描述一天气温可能的随机变化; 定义一个变量来描述明天的股票可能的收盘价; 定义一个变量来描述一段时间内股票订单的数量等。但是, 并不是所有这些定义出来的变量都是我们能够认识和研究的。我们需要这些变量关于我们的知识, 即某个 σ 代数 $\mathcal{F}$ 是可测的。

但是对于可测性, 我们在前面只定义了一个集合如果属于 σ 代数 $\mathcal{F}$, 则我们称她关于 $\mathcal{F}$ 可测。因此, 对于定义的变量, 我们希望她取某一个值, 或者在一个区间内的值我们都可以认识。

在定义随机变量 X 时, 对于任意固定的实数 a , $\{\omega : X(\omega) < a\}$ 表示所有在样本空间 Ω 中满足 $\{X(\omega) < a\}$ 的点 ω 。把满足条件所有这些点放在一起就成了样本空间 Ω 的一个子集, 这实际就是 $\{X < a\}$ 在 Ω 的原像。因此, 我们要求 $\{\omega : X(\omega) < a\}$ 是属于 σ 代数 $\mathcal{F}$, 这样我们才能够判断随机变量是否是小于某一

个固定的值 a !!

在定义随机变量 X 时, 虽然只要求对于任意固定的实数 a , $\{\omega : X(\omega) < a\} \in \mathcal{F}$. 但利用如下等式

$$\begin{aligned}\{a_1 \leq X < a_2\} &= \{X < a_2\} \setminus \{X < a_1\} \in \mathcal{F}, \\ \{X = a\} &= \bigcap_{n=1}^{+\infty} \{a - \frac{1}{n} \leq X < a + \frac{1}{n}\} \in \mathcal{F}, \\ \{X \geq a\} &= \Omega \setminus \{X < a\} \in \mathcal{F}.\end{aligned}$$

由此我们可以看到随机变量取一个固定的值 $\{X = a\}$ 或者大于一个值 $\{X \geq a\}$ 都是可测的。



理解 1.10. 随机变量

如果一个变量 X 只取有限个离散的值 $\{y_1, y_2, \dots, y_N\}$. 则只需要对于任意的取值 y_i 都有

$$\{X = y_i\} \in \mathcal{F}$$

则 X 是关于 $\mathcal{F}$ 的随机变量。并称 X 为离散型随机变量。

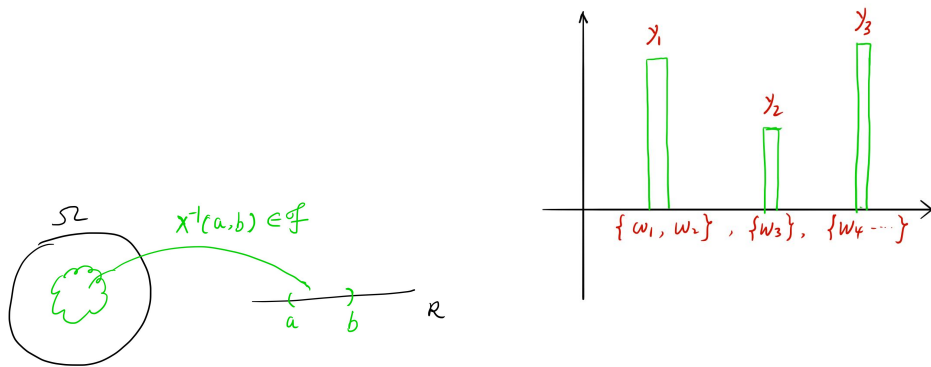


图 1.7: 随机变量



笔记 更直观的可以这样理解随机变量: 左图表示对一个映射, 如果从实数轴上随意选一个区间 (a, b) , 我们去找这个映射在 Ω 的原像都是关于 $\mathcal{F}$ 可测的, 我们就称为随机变量。对于右图所示的离散随机变量, 如果我们看到取值 y_i , 就能用 $\mathcal{F}$ 判断是什么发生了, 那么就是随机变量。

例 1.14 设 $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $\mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3\}\}$. 定义如下的两个变量 X, Y .

$$X = \begin{cases} 1, & \{\omega_1, \omega_2\} \\ 2, & \{\omega_3\} \end{cases}, \quad Y = \begin{cases} 1, & \{\omega_1\} \\ 2, & \{\omega_2, \omega_3\} \end{cases}$$

容易计算,

$$\begin{aligned}\{X = 1\} &= \{\omega_1, \omega_2\} \in \mathcal{F}, \quad \{X = 2\} = \{\omega_3\} \in \mathcal{F}, \\ \{Y = 1\} &= \{\omega_1\} \notin \mathcal{F}, \quad \{Y = 2\} = \{\omega_2, \omega_3\} \notin \mathcal{F}.\end{aligned}$$

由此可见 X 是关于 $\mathcal{F}$ 的随机变量, 而 Y 则不是. 如果定义一个新的 σ 代数 $\hat{\mathcal{F}}$

$$\hat{\mathcal{F}} = \{\emptyset, \Omega, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}\}.$$

容易看到 X 和 Y 都是关于 $\hat{\mathcal{F}}$ 的随机变量. 因此, 在谈论是否是随机变量时需要说明时相对于什么 σ 代数.

命题 1.3

分布函数 $F(x)$ 的性质如下:

1. $F(x)$ 是单调不减函数.
2. $0 \leq F(x) \leq 1, \lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1.$
3. $F(x)$ 是左连续函数, 即 $\forall x \in \mathbb{R}, F(x-) = F(x).$



证明 1.2. 两条性质都是显然的. 注意到:

$$\bigcup_{n=1}^{\infty} \left\{ X < x - \frac{1}{n} \right\} = \{X < x\}$$

由概率测度的从下连续可得:

$$\begin{aligned} F(x) &= P(\{X < x\}) = P\left(\bigcup_{n=1}^{\infty} \{X < x - \frac{1}{n}\}\right) \\ &= \lim_{n \rightarrow +\infty} P(\{X < x - \frac{1}{n}\}) = F(x-). \end{aligned}$$

因此 $F(x)$ 是左连续函数. ■



笔记 如果我们定义随机变量的分布函数为

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

则分布函数 F 是一个右连续函数. 但对我们研究, 这两种定义方式都可以, 只要保持一致就行.

定义 1.13. 连续型和离散型随机变量

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**. ♣

在随机过程中, 我们往往感兴趣的是随机变量或者随机过程的概率性质. 因此对于如果两个变量, 如果他们不相等部分是一个零测集, 我们可以认为她们是一样的.

定义 1.14. 等价随机变量

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**.

注: 对于两个等价的随机变量, 我们视为同一. ♣

例 1.15 在某一概率空间 $(\Omega, \mathcal{F}, P)$ 上,

$$\mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4\}, \dots\}, P(\{\omega_1, \omega_2\}) = 0.2, P(\{\omega_3\}) = 0.8, P(\omega_4) = 0.$$

$$X = \begin{cases} 1, & \{\omega_1, \omega_2\} \\ 2, & \{\omega_3\} \\ 3, & \{\omega_4\} \end{cases} \quad Y = \begin{cases} 1, & \{\omega_1, \omega_2\} \\ 2, & \{\omega_3\} \\ -1, & \{\omega_4\} \end{cases}$$

虽然 X, Y 在 ω_4 上面并不相等, 但是 $P(\omega_4) = 0$ 。因此, 我们仍然称 X, Y 等价随机变量。

在前面随机变量的定义中, 我们定义了小于任意固定常数都是可测的。但在实际应用中, 我们往往还需要考虑随机变量的取值范围或者大于一个值。下面定理告诉我们这些事情都是可测的。

定理 1.4. 随机变量等价定义

下列命题等价:

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



下面定义告诉我们对随机变量进行四则运算和求极限的结果依然还是随机变量。这就保证了, 我们在讨论随机变量求和, 以及以后考虑收敛性是有意义的。

定理 1.5. 随机变量之间的性质

1. 若 X, Y 是随机变量, 则 $\{X < Y\}, \{X \leq Y\}, \{X = Y\}$ 及 $\{X \neq Y\}$ 都属于 $\mathcal{F}$;
2. 若 X, Y 是随机变量, 则 $X + Y, X - Y$ 与 XY 亦然;
3. 若 $\{X_n\}$ 是随机变量序列, 则 $\sup_n X_n, \inf_n X_n, \limsup_{n \rightarrow \infty} X_n$ 和 $\liminf_{n \rightarrow \infty} X_n$ 都是随机变量.



证明 1 和 2. 记 (r_n) 为所有的有理数。则有:

$$\begin{aligned} \{X + Y < a\} &= \{X < a - Y\} \\ &= \bigcup_{n=1}^{\infty} \{X < r_n\} \cap \{a - Y \geq r_n\} \in \mathcal{F}. \\ \{X^2 < a\} &= \{X < \sqrt{a}\} \cap \{X > -\sqrt{a}\} \in \mathcal{F}. \\ XY &= \frac{(X + Y)^2 - (X - Y)^2}{4}. \end{aligned}$$

3. 对于任意的实数 a , 注意到

$$\begin{aligned} \{\sup_n X_n < a\} &= \bigcap_n \{X_n < a\} \in \mathcal{F}, \\ \{\inf_n X_n > a\} &= \bigcap_n \{X_n > a\} \in \mathcal{F}. \end{aligned}$$

因此, $\sup_n X_n, \inf_n X_n$ 都是随机变量。同时

$$\limsup_{n \rightarrow \infty} X_n = \inf_{n \rightarrow \infty} \left(\sup_{1 \leq k \leq n} X_k \right) \in \mathcal{F},$$

$$\liminf_{n \rightarrow \infty} X_n = \sup_{n \rightarrow \infty} \left(\inf_{1 \leq k \leq n} X_k \right) \in \mathcal{F}.$$

因此, $\limsup_{n \rightarrow \infty} X_n$ 和 $\liminf_{n \rightarrow \infty} X_n$ 都是随机变量。 ■

1.3.1 数字特征

有了随机变量的分布后, 我们对一个随机变量的认识还不够, 同时在应用和证明中只知道分布往往还不够直观和便利。因此, 我们刻画随机变量更深刻的一些性质。

定义 1.15. 随机变量一些特征

(1) 取值为 $\{s_k\}$ 的离散型随机变量 X 的**数学期望** (简称为期望) 数学期望 $E[X]$ 定义为

$$E[X] = \sum_k s_k p_k = \sum_k s_k P(X = s_k)$$

如果 $\sum |s_k| p_k < \infty$.

(2) 连续型随机变量 X 的**数学期望 (Expectation)** $E[X]$ 定义为

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

如果 $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, 这里 $F(x)$ 是 X 的分布函数, $f(x)$ 是其密度函数.

利用 Riemann-Stieltjes 积分, 我们可以对离散型随机变量和连续型随机变量的期望给出一个统一的表达式:

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

(3) 设 X 为任一随机变量, 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.

(4) 设 X 为任一随机变量, 对正整数 k , 称 $c_k = E[X - E[X]]^k$ 为 X 的 **k 阶中心矩 (k -th moment function)**. 方差是二阶原点矩.

(5) 设 X, Y 为两个随机变量, 对正整数 k, l , 称 $E[X - E[X]]^k [Y - E[Y]]^l$ 为 X 的 **$k + l$ 阶混合中心矩**. 协方差是二阶混合中心矩. ♣

1.3.2 矩函数 (Moment Generating Function)

定义 1.16. 矩函数

若随机变量 X 的分布函数为 $F_X(x)$, 则称

$$\phi_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} dF_X(x) \quad (1.5)$$

为 X 的矩函数. 如果 F_X 有密度 $f(x)$, 则

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$



命题 1.4

X 的各阶矩与矩函数的关系 (假设对 $\phi(t)$ 求导时, 求导运算与求期望运算可以交换次序)

$$\phi'_X(t) = E(Xe^{tX})$$

$$\phi''_X(t) = E(X^2 e^{tX})$$

$$\vdots$$

$$\phi_X^{(n)}(t) = E(X^n e^{tX}).$$

令 $t = 0$, 得到 $\phi_X^{(n)}(0) = E[X^n]$, $n \geq 1$. 当矩函数存在时, 它唯一地决定分布, 因此我们能够用矩函数刻画随机变量的概率分布.



有时随机变量的矩函数不一定存在, 在这种情况下, 更方便的是特征函数.

1.3.3 特征函数 (Characteristic function)

定义 1.17. 特征函数

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. 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. 特征函数有如下常用性质:

1. 有界性: $|\psi(t)| = |E[e^{itX}]| \leq E[|e^{itX}|] = 1 = \psi(0)$;
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 $t_1, \dots, t_n$ 及复数 $\lambda_1, \dots, \lambda_n$, 有 $\sum_{k=1}^n \sum_{j=1}^n \psi(t_k - t_j) \lambda_k \overline{\lambda_j} \geq 0$.



命题 1.5. 特征函数和矩的关系

设随机变量 X 有 n 阶矩存在, 则它的特征函数可微分 n 次, 且当 $k \leq n$ 时, 有

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

特别地, 特征函数可作如下带皮阿诺型余项的 Taylor 展开:

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



例 1.16 求单点分布 $P(X = c) = 1$ 的特征函数。

$$\psi(t) = E(e^{itc}) = e^{itc}, \quad t \in R.$$

例 1.17 求两点分布 $P(X = 0) = p, P(X = 1) = 1 - p$ 的特征函数。

$$\psi(t) = e^{it \cdot 0}p + e^{it \cdot 1}(1 - p) = p + e^{it}(1 - p), \quad t \in R.$$

例 1.18 求指数分布的特征函数。

解 指数函数的密度函数为:

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

因此可以得到:

$$\begin{aligned} \psi(t) &= \int_0^{\infty} \lambda e^{itx} e^{-\lambda x} dx \\ &= \int_0^{\infty} \lambda e^{-\lambda x} \cos(tx) dx + i\lambda \int_0^{\infty} e^{-\lambda x} \sin(tx) dx \\ &= \lambda \frac{\lambda}{\lambda^2 + t^2} + i\lambda \frac{t}{\lambda^2 + t^2} = \left(1 - \frac{it}{\lambda}\right)^{-1}. \end{aligned}$$

■



笔记 回顾数学知识: 分布积分:

$$\begin{aligned} \int_0^{\infty} e^{-\lambda x} \cos(x) dx &= e^{-\lambda x} \sin(x) \Big|_0^{\infty} + \int_0^{\infty} \lambda e^{-\lambda x} \sin(x) dx \\ &= \int_0^{\infty} \lambda e^{-\lambda x} \sin(x) dx = -\lambda e^{-\lambda x} \cos(x) \Big|_0^{\infty} - \lambda^2 \int_0^{\infty} e^{-\lambda x} \cos(x) dx \\ &= \lambda - \lambda^2 \int_0^{\infty} e^{-\lambda x} \cos(x) dx. \end{aligned}$$

因此可以得到:

$$\int_0^{\infty} e^{-\lambda x} \cos(x) dx = \frac{\lambda}{1 + \lambda^2}.$$

例 1.19 求标准正态分布 $N(0, 1)$ 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

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

解初值问题 $\psi'(t) + t\psi(t) = 0$, $\psi(0) = 1$ 得 $\psi(t) = e^{-\frac{1}{2}t^2}$. ■

理解 1.11. 正态分布 $N(0, 1)$ 的特征函数

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

我们可以把 $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} dx$ 看成均值为 it 的正态分布的累计分布函数. ♣

例 1.20 求正态分布 $N(\mu, \sigma^2)$ 的特征函数.

解 设 $Y \sim N(\mu, \sigma^2)$, $X \sim N(0, 1)$, 则有 $Y = \sigma X + \mu$. 利用特征函数性质 (4) 有 $\psi_Y(t) = e^{it\mu} \psi_X(\sigma t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}$. ■

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 1.6

分布函数由其特征函数唯一决定. 从而说明特征函数与分布函数是相互唯一确定的. ♡

定理 1.7

特征函数与矩的关系, 若随机变量 X 的 n 阶矩存在, 则 X 的特征函数 $\psi(t)$ 的 k 阶导数 $\psi^{(k)}(t)$ 存在, 且

$$E(X^k) = \frac{\psi^{(k)}(0)}{i^k}.$$
♡

例 1.21 随机变量 X 的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{2} \cos(x), & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 0, & \text{others.} \end{cases}$$

求 $E(X)$, $D(X)$.

证明 我们有基本等式:

$$\cos(x) + \cos(y) = 2 \cos\left(\frac{x+y}{2}\right) \cos\left(\frac{x-y}{2}\right).$$

$$\begin{aligned}
\psi(t) &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2} \cos(x) e^{itx} dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2} \cos(x) (\cos(tx) + i \sin(tx)) dx \\
&= 2 \int_0^{\frac{\pi}{2}} \frac{1}{2} \cos(x) \cos(tx) dx \\
&= \frac{1}{2} \int_0^{\frac{\pi}{2}} (\cos((t+1)x) + \cos((t-1)x)) dx \\
&= \frac{1}{2} \left(\frac{1}{t+1} \sin\left((t+1)\frac{\pi}{2}\right) + \frac{1}{t-1} \sin\left((t-1)\frac{\pi}{2}\right) \right).
\end{aligned}$$

上面用到了偶函数在一个对称区间积分为 0. 由于 $\psi^{(1)}(0) = 0$ and $\psi^{(2)}(0) = 2 - \pi^2/4$, 可以得到

$$E(X) = i^{-1}\psi^{(1)}(0) = 0, \quad D(X) = i^{-2}\psi^{(2)}(0) = -2 + \pi^2/4.$$

■

1.3.4 反演公式及唯一性定理

由随机变量 X 的分布函数可惟一确定其特征函数, 能否由 X 的特征函数唯一确定其分布函数??

定理 1.8. 反演公式

设随机变量 X 的分布函数和特征函数分别为 $F(x)$ 和 $\psi(t)$. 则对 $F(x)$ 的任意连续点 x_1, x_2 , 有:

$$F(x_2) - F(x_1) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-itx_1} - e^{-itx_2}}{it} \psi(t) dt.$$

♡



笔记 回顾数学里面的积分公式:

$$\int_0^\infty \frac{\sin(t)}{t} dt = \frac{\pi}{2}, \quad \int_0^\infty \frac{\sin(at)}{t} dt = \operatorname{sgn}(a) \frac{\pi}{2}, \quad a \neq 0.$$

其中, $\operatorname{sgn}(a)$ 表示符号函数, 当 $a > 0$ 时为 1, 当 $a < 0$ 小于 0 时为 -1.

证明 考虑如下的函数 $k(\lambda)$,

$$k(\lambda) := \int_0^\infty e^{-\lambda t} \frac{\sin(t)}{t} dt.$$

对 $k(\lambda)$ 关于 λ 求导, 并利用分布积分可以得到

$$k' := - \int_0^\infty e^{-\lambda t} \sin(t) dt = - \frac{1}{1 + \lambda^2}.$$

由此可以得到

$$k(\lambda) = -\arctan(\lambda) + \text{constant}.$$

根据 $k(\lambda)$ 定义, 由于 $k(+\infty) = 0 = -\arctan(+\infty) + \text{constant} = -\frac{\pi}{2} + \text{constant}$, 因此可知 $\text{constant} = \frac{\pi}{2}$.

$$k(\lambda) = -\arctan(\lambda) + \frac{\pi}{2}.$$

由于 $\int_0^\infty \frac{\sin(t)}{t} dt = k(0) = \frac{\pi}{2}$. 简单的积分变化可以得到 $\int_0^\infty \frac{\sin(at)}{t} dt = \operatorname{sgn}(a) \frac{\pi}{2}$. ■

证明 我们只证明随机变量 X 具有密度函数情况。记 X 的密度函数为 f 。一般的证明类

似。

$$\begin{aligned}
 & \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-itx_1} - e^{-itx_2}}{it} \psi(t) dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-itx_1} - e^{-itx_2}}{it} \int_{-\infty}^{\infty} f(y) e^{ity} dy dt \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \lim_{T \rightarrow \infty} \int_{-T}^T \frac{e^{-it(x_1-y)} - e^{-it(x_2-y)}}{it} dy dt \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \lim_{T \rightarrow \infty} \int_{-T}^T \frac{\cos(t(x_1-y)) - i \sin(t(x_1-y)) - \cos(t(x_2-y)) + i \sin(t(x_2-y))}{it} dy dt \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) 2 \lim_{T \rightarrow \infty} \int_0^T \frac{i \sin(t(x_2-y)) - i \sin(t(x_1-y))}{it} dy dt \\
 &= \frac{1}{\pi} \int_{-\infty}^{\infty} f(y) \int_0^{\infty} \frac{\sin(t(x_2-y)) - \sin(t(x_1-y))}{t} dy dt \\
 &= \frac{1}{\pi} \int_{-\infty}^{\infty} f(y) \frac{\pi}{2} (\operatorname{sgn}(x_2-y) - \operatorname{sgn}(x_1-y)) dy \\
 &= \int_{x_1}^{x_2} f(y) dy = F(x_2) - F(x_1).
 \end{aligned}$$

■

推论 1.1. 唯一性定理

分布函数由其特征函数惟一确定.



推论 1.2

若随机变量 X 的特征函数 $\psi(t)$ 在 R 上绝对可积, 则 X 为连续型随机变量, 其概率密度为

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \psi_X(t) dt.$$



命题 1.6. 离散型反演公式

若随机变量 X 是离散型的, 其分布律为 $p_k = P(X = k)$, $k \in \mathbb{Z}$. 其特征函数为

$$\psi(t) = \sum_{k=-\infty}^{\infty} p_k e^{ikt}, \quad t \in R.$$

则分布 p_k 可以表示为:

$$p_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikt} \psi(t) dt.$$



证明

$$\begin{aligned}
 \int_{-\pi}^{\pi} e^{-imt} \psi(t) dt &= \int_{-\pi}^{\pi} e^{-imt} \sum_{k=-\infty}^{\infty} p_k e^{ikt} dt \\
 &= \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} p_k e^{-imt} e^{ikt} dt \\
 &= p_m \int_{-\pi}^{\pi} dt + \sum_{k=-\infty, k \neq m}^{\infty} \int_{-\pi}^{\pi} p_k e^{-imt} e^{ikt} dt \\
 &= 2\pi p_m.
 \end{aligned}$$

■

利用特征函数和分布函数的唯一性, 我们还可以借用特征函数来求得不容易计算的概率密度函数。

例 1.22 随机变量 X 在 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上服从均匀分布, $Y = \cos(X)$, 利用特征函数求 Y 的概率密度。

证明 X 的密度函数为:

$$f(x) = \begin{cases} \frac{1}{\pi}, & x \in [-\pi/2, \pi/2] \\ 0, & \text{others.} \end{cases}$$

由此 Y 的特征函数为:

$$\begin{aligned}
 \psi_Y(t) &= E(e^{itY}) = \int_{-\pi/2}^{\pi/2} \frac{1}{\pi} e^{it \cos(x)} dx \\
 &= \int_0^1 e^{itu} \frac{2}{\pi} \frac{1}{\sqrt{1-u^2}} du.
 \end{aligned}$$

利用特征函数与分布函数一一对应的惟一性定理, 知随机变量 Y 的概率密度函数为:

$$f_Y(u) = \frac{2}{\pi} \frac{1}{\sqrt{1-u^2}}, \quad 0 < u < 1.$$

■

定理 1.9. 独立随机变量和的特征函数

若随机变量 $X_1, X_2, \dots, X_n$ 相互独立, 令 $Y = \sum_{k=1}^n X_k$, 则

$$\psi_Y(t) = \prod_{k=1}^n \psi_{X_k}(t).$$

♡

例 1.23 随机变量 Y 满足 n 重 Bernoulli 分布, 即 $Y \sim B(n, p)$, 写出其特征函数。

证明 由于

$$Y = \sum_{k=1}^n X_k,$$

其中, $X_k \sim B(1, p)$ 为二项分布. 由此可得:

$$\psi_Y(t) = \prod_{k=1}^n \psi_{X_k}(t) = (q + pe^{it})^n.$$

例 1.24 若 $X_1, X_2, \dots, X_n$ 相互独立, 且 $X_k \sim N(0, 1)$. 证明

$$Y = \frac{1}{\sqrt{n}} \sum_{k=1}^n X_k.$$

也服从 $N(0, 1)$ 分布.

证明 标准正态分布 X_k 的特征函数为 $\psi_{X_k}(t) = e^{-\frac{t^2}{2}}$. 则,

$$\psi_{\sum_{k=1}^n X_k}(t) = \prod_{k=1}^n \psi_{X_k}(t) = e^{-n \frac{t^2}{2}}.$$

由此可以得到:

$$\psi_Y(t) = \psi_{\sum_{k=1}^n X_k}(t/\sqrt{n}) = e^{-\frac{t^2}{2}}.$$

由唯一性定理知, $Y \sim N(0, 1)$.

当然我们也可以直接计算 Y 的均值和方差来证明为标准正态分布.

1.3.5 多维随机变量的特征函数

定义 1.18. 二维特征函数

二维随机变量 (X, Y) 的特征函数定义为

$$\psi(t_1, t_2) = E\left(e^{i(t_1 X + t_2 Y)}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(t_1 x + t_2 y)} dF(x, y).$$

若 (X, Y) 为连续型的随机变量, 则:

$$\psi(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(t_1 x + t_2 y)} f(x, y) dx dy.$$

若 (X, Y) 为离散型的随机变量, 则:

$$\psi(t_1, t_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} e^{i(t_1 m + t_2 n)} P(X = m, Y = n).$$

定义 1.19. n 维特征函数

n 维随机向量 $(X_1, X_2, \dots, X_n)$ 的分布函数为 $F(X_1, X_2, \dots, X_n)$, 则它的特征函数为

$$\psi(t_1, t_2, \dots, t_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i(t_1 x_1 + t_2 x_2 + \dots + t_n x_n)} dF(x_1, x_2, \dots, x_n).$$

引理 1.1

随机变量 $X_1, X_2, \dots, X_n$ 相互独立的充要条件是

$$\psi(t_1, t_2, \dots, t_n) = \prod_{k=1}^n \psi_{X_k}(t_k).$$

引理 1.2

二维随机变量 (X, Y) 的特征函数为 $\psi(t_1, t_2)$. 则, $Z = aX + bY + c$ 的特征函数为

$$\psi_Z(t) = e^{itc} \psi(at, bt).$$

特别的有:

$$\psi_{X+Y}(t) = \psi(t, t).$$



1.4 独立性与条件期望

1.4.1 独立性

定义 1.20. 独立性

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. 更一般地, 设 $A_1, A_2, \dots, A_n$ 为 n 个事件, 如果对任何 $m \leq n$ 及 $1 \leq k_1 < k_2 < \dots < k_m \leq n$, 有

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

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

2. 设 $\{A_i, i \in I\}$ 是一族事件, 若对 I 的任意有限子集 $\{i_1, \dots, i_k\} \neq \emptyset$ 有

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j}) \quad (1.6)$$

则称 $\{A_i, i \in I\}$ 是相互独立的.

3. 设 $\{A_i, i \in I\}$ 是一族事件类, 如果对 I 的任意有限子集 $\{i_1, \dots, i_k\} \neq \emptyset$, 任意 $A_{i_j} \in \mathcal{A}_{i_j}$ 有式成立, 则称 $\{A_i, i \in I\}$ 是独立事件类.
4. 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.



理解 1.12. 随机变量独立

- 两个 σ 代数 $\mathcal{F}, \mathcal{G}$ 相互独立是指我们任意从 $\mathcal{F}, \mathcal{G}$ 找出两个事件 $A \in \mathcal{F}, B \in \mathcal{G}$ 她们都是独立的。
- 随机变量 X, Y 的独立性是指她们生成的 σ 代数 $\sigma(X)$ 和 $\sigma(Y)$ 是相互独立的。
- 由随机变量生成的 σ 代数可以通过下图理解。当我们变动实数轴上一个区间 (a, b) 时, 这样得到映射在 Ω 的原像也会相应发生变化。如果我们把这些原像全部放在一起就是我们所说的有随机变量 X 生成的 σ 代数。



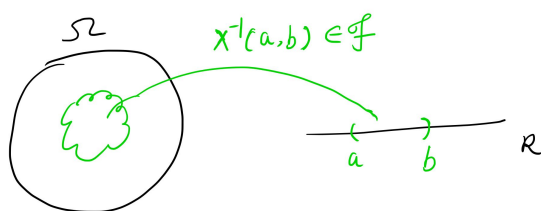


图 1.8: 随机变量



笔记 更直观的可以这样理解随机变量：如果从实数轴上随意选一个区间 (a, b) ，我们去找这个映射在 Ω 的原像都是关于 $\mathcal{F}$ 可测的，我们就称为随机变量。

命题 1.7

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

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



定理 1.10

- (1) 设随机变量 $X_1, \dots, X_n$ 是独立的，则 $E[\prod_{k=1}^n X_k] = \prod_{k=1}^n E[X_k]$.
- (2) 设随机变量 $X_1, \dots, X_n$ 是独立的，则 $Var[\sum_{k=1}^n X_k] = \sum_{k=1}^n Var[X_k]$.



1.4.2 条件期望

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

$$P(A|B) = P(A \cap B) / P(B)$$

定理 1.11. 全概率公式

设 $\{B_n\}$ 是 Ω 的一个分割 (即: $\cup_n B_n = \Omega$, 且 $B_i \cap B_j = \emptyset, i \neq j$), 且使得 $P(B_n) > 0, \forall n$. 如果 $A \in \mathcal{F}$, 则

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

这儿的分割 $\{B_n\}$ 可以是有限的，也还可以是无限的。



定理 1.12. Bayes 公式

设 $\{B_n\}$ 是 Ω 的一个分割，且使得 $P(B_n) > 0, \forall n$ ，如果 $P(A) > 0$ ，则

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



1.4.3 条件分布

定义 1.21. 条件分布

1. 如果 X 与 Y 是离散型随机变量, 对一切使得 $P\{Y = y\} > 0$ 的 y , 给定 $Y = y$ 时, X 的条件概率定义为:

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

X 的条件分布定义为:

$$F(x|y) = P(X \leq x|Y = y)$$

X 的条件期望定义为:

$$E[X|Y = y] = \int x dF(x|y) = \sum_x x P(X = x|Y = y)$$

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

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

X 的条件分布定义为:

$$F(x|y) = P(X \leq x|Y = y) = \int_{-\infty}^x f(x|y) dx$$

X 的条件期望定义为:

$$E[X|Y = y] = \int x dF(x|y) = \int x f(x|y) dx$$



理解 1.13. 条件期望

- 在后面我们会严格的定义条件数学期望 $E[X|Y]$ 。需要注意的是 $E[X|Y]$ 和 $E[X|Y = y]$ 是不同的意思, 虽然我们都叫条件数学期望。 $E[X|Y]$ 表示给定随机变量 Y 时对 X 预测。由于 Y 本身是随机变量, 所以 $E[X|Y]$ 是一个关于 Y 函数, **本身也是一个随机变量**。但这个函数一般没法显示的写出来。在我们作线性回归的时候, 实际上假设了 $E[X|Y] = \alpha + \beta Y$ 这种线性形式。
- 对于任意的实数 y , $E[X|Y = y]$ 表示在给定 $Y = y$ 时, 对 X 的预测值, 所以它是一个数值, 而不是随机变量。但是当我们改变实数 y 时, $E[X|Y = y]$ 也会发生变化, 所以也可以把 $E[X|Y = y]$ 看成是**关于实数 y 的函数**。
- 如果我们知道了 $E[X|Y = y] = f(y)$, 这时可以得到一般条件数学期望 $E[X|Y] = f(Y)$ 。注意一个是普通的实函数, 一个是随机变量。这个技巧在后面计算中会常用到。



例 1.25 设随机变量 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} \frac{1}{2}e^{-y}, & y > |x| \\ 0, & \text{others} \end{cases}$$

求 $E(Y|X = x)$.

证明 由联合密度和边缘密度的关系可得

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\ &= \int_{|x|}^{\infty} \frac{1}{2} e^{-y} dy = \frac{1}{2} e^{-|x|}. \end{aligned}$$

因此,

$$f(y|x) = \frac{f(x, y)}{f_X(x)} = e^{-y+|x|}, \quad y > |x|.$$

最终有:

$$E(Y|X=x) = \int_{|x|}^{\infty} y e^{-y+|x|} dy = 1 + |x|.$$



笔记 从这个例子也可以看到, $E(Y|X=x) = 1 + |x|$, 它是一个关于实数 x 的函数。一般的, 把结果 x 换成随机变量 X , 可以得到 $E(Y|X) = 1 + |X|$ 。

例 1.26 已知 $X \sim N(\mu, \tau^2)$, 给定 $X = x$ 的条件下, Y 的条件分布为 $N(x, \sigma^2)$, 求 Y 的分布及给定 $Y = y$ 的条件下 X 的条件分布。

证明 根据联合分布得到

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(x-\mu)^2}{2\tau^2}} \\ f(y|x) &= \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-x)^2}{2\sigma^2}} \\ f(x, y) &= f_X(x)f(y|x) = \frac{1}{2\pi\tau\sigma} e^{-\frac{(x-\mu)^2}{2\tau^2} - \frac{(y-x)^2}{2\sigma^2}}. \end{aligned}$$

因此有

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y) dx \\ &= \frac{1}{2\pi\tau\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\tau^2} - \frac{(y-x)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi(\tau^2 + \sigma^2)}} e^{-\frac{(y-\mu)^2}{2(\tau^2 + \sigma^2)}}. \\ f(x|y) &= \frac{f(x, y)}{f_Y(y)} = \dots \end{aligned}$$

性质 条件期望的数学性质: 设函数 $g(x)$ 在 R 上连续, 若

$$\int_{-\infty}^{\infty} |g(x)| dF(x|y) < +\infty.$$

则随机变量 $g(X)$ 在 $Y = y$ 条件下的条件数期望为

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

定义 1.22. 条件方差

$$D(X|Y=y) := E(X - E[X|Y=y])^2$$

为 $Y=y$ 的条件下, 随机变量 X 的条件方差.

**定理 1.13. 全期望公式**

对一切随机变量 X 和 Y , 当期望存在时

$$E[X] = E[E[X|Y]] = \int E[X|Y=y] dF_Y(y) \quad (1.7)$$

特别的:

- 当 Y 为一个离散随机变量时, (1.7) 式为

$$E[X] = \sum_y E[X|Y=y] P\{Y=y\}$$

- 当 Y 为一个连续随机变量时, (1.7) 式为

$$E[X] = \int_{-\infty}^{+\infty} E[X|Y=y] f(y) dy$$



例 1.27 随机个随机变量之和 设 $X_1, X_2, \dots$ 是一列与 X 独立同分布的随机变量; 设 N 为一非负整值随机变量, 且与序列 $X_1, X_2, \dots$ 独立. 求 $Y = \sum_{i=1}^N X_i$ 的均值和方差.

解 首先在对 N 取条件的情况下来计算 $Y = \sum_{i=1}^N X_i$ 的矩母函数, 即

$$E \left[\exp \left\{ t \sum_{i=1}^N X_i \right\} | N=n \right] = E \left[\exp \left\{ t \sum_{i=1}^n X_i \right\} \right] = (\phi_X(t))^n$$

其中 $\phi_X(t)$ 是随机变量 X 的矩母函数. 因此, 把上述结果的 n 替换成随机变量 N , 可以得到

$$E \left[\exp \left\{ t \sum_{i=1}^N X_i \right\} | N \right] = (\phi_X(t))^N$$

从而

$$\phi_Y(t) = E \left[\exp \left\{ t \sum_{i=1}^N X_i \right\} \right] = E [(\phi_X(t))^N]$$

现在对 $\phi_Y(t)$ 求导得

$$\phi_Y'(t) = E[N(\phi_X(t))^{N-1} \phi_X'(t)]$$

再求一次导数得

$$\phi_Y''(t) = E[N(N-1)(\phi_X(t))^{N-2} (\phi_X'(t))^2 + N(\phi_X(t))^{N-1} \phi_X''(t)]$$

计算在 $t=0$ 点的值, 得

$$E[Y] = E[NE[X]] = E[N]E[X] \quad (1.8)$$

及

$$\begin{aligned} E[Y^2] &= E[N(N-1)(E[X])^2 + NE[X^2]] \\ &= E[N]Var[X] + E[N^2](E[X])^2. \end{aligned}$$



因此有

$$\begin{aligned} \text{Var}[Y] &= E[Y^2] - (E[Y])^2 \\ &= E[N]\text{Var}[X] + (E[X])^2\text{Var}[N] \end{aligned}$$

■

例 1.28 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 小时的旅程会使他回到矿井；第三个门通向一个隧道，沿此隧道走 3 小时的旅程会使他回到矿井。假定矿工总是等可能地选取任意一个门，用 X 表示矿工到达安全第所需的时间，求 X 的矩母函数。

证明 写出 X 的分布列很困难"故无法直接求其均值！令 Y 表示矿工第一次选择的门"则 $\{Y = i\}$ 表示第一次选择第 i 个门"由题意知

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

由此可得：

$$E[X] = 10.$$

有条件概率公式得到：

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

注意到：

$$\begin{aligned} E[e^{tX}|Y = 1] &= e^{2t} \\ E[e^{tX}|Y = 2] &= E[e^{t(\tilde{X}+3)}] = e^{3t} E[e^{tX}] \\ E[e^{tX}|Y = 3] &= E[e^{t(\tilde{X}+5)}] = e^{5t} E[e^{tX}]. \end{aligned}$$

其中， $\tilde{X}$ 与 X 有相同的分布。由此可得：

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

■

例 1.29 设某段时间内到达商场的顾客人数 N 服从参数为 λ 的泊松分布。每位顾客在该商场的消费额 X 服从 $[a, b]$ 上的均匀分布。各位顾客之间消费是相互独立的且与 N 独立。求顾客在该商场总的消费额。

例 1.30 设随机变量 Y 的 $E(Y)$ 存在， $Y = \sum_{k=1}^N X_k$ 及 N 都是随机变量，并且相互独立。 N 仅取自然数， $E(N)$ 存在。证明：

$$E(Y) = \sum_{k=1}^{\infty} P(N \geq k) E(X_k).$$

例 1.31 已知随机变量 X 服从 $[0, a]$ 上的均匀分布, 随机变量 Y 服从 $[X, a]$ 上的均匀分布, 试求

$$E(Y|X = x), \quad 0 < x < a, \quad \text{and} \quad E(Y).$$

证明 由条件知对 $x > 0$, 有

$$f(y|x) = \begin{cases} \frac{1}{a-x}, & x < y < a \\ 0 & \text{others} \end{cases}$$

对任意的 $0 < x < a$,

$$E(Y|X = x) = \int_x^a \frac{y}{a-x} dy = \frac{a+x}{2}.$$

$$E(Y) = E(E(Y|X)) = E\left(\frac{a+X}{2}\right) = \frac{3a}{4}.$$



笔记 从这个例子也可以看到, $E(Y|X = x) = \frac{a+x}{2}$, 它是一个关于实数 x 的函数。一般的, 把结果 x 换成随机变量 X , 可以得到 $E(Y|X) = \frac{a+X}{2}$ 。

1.4.4 一般条件期望 ★

最后我们将条件期望推广到一般随机变量及 σ 代数情形。

定义 1.23. 一般条件期望

设 X 是随机变量且 $E[|X|] < \infty$. $\mathcal{G}$ 为任意的 σ 代数。如果另外一个随机变量 X^* 满足如下三个条件:

1. $E[|X^*|] < \infty$
2. X^* 是 $\mathcal{G}$ 可测随机变量 (即对任何的 $a \in \mathbb{R}$, 有 $\{X^* \leq a\} \in \mathcal{G}$)
3. 对于所有的 $B \in \mathcal{G}$, 等式 $E[X^* \mathbf{1}_B] = E[X \mathbf{1}_B]$ 成立。也即

$$\int_B E[X|\mathcal{G}] dP = \int_B X dP, \quad \forall B \in \mathcal{G}$$

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



注 上述一般意义下的条件数学期望 X^* 存在, 并且是唯一的 (几乎必然相等的意义下)。这儿 $\mathbf{1}_B(\omega)$ 为一个示性函数 (Indicator function) 如果 ω 在 B 中就取 1, 否则取 0。显然 $\mathbf{1}_B(\omega)$ 也是一个随机变量。

理解 1.14. 条件数学期望

1. 随机变量 X 关于 σ 代数 $\mathcal{G}$ 不一定可测。所以如果想去认识随机变量 X , 就需要利用我们知识 $\mathcal{G}$ 去预测 X 。这个预测就是条件数学期望。
2. 我们要求这个预测的结果 X^* 自然需要关于 $\mathcal{G}$ 可测, 并且满足可积性的条件。
3. 对 X^* 更重要的要求是: 她在平均的意义上和可能不认识的 X 是相等的。这就要求: 我们任意从知识结构里面找一个集合 $B \in \mathcal{G}$, 都有 $E[X^* \mathbf{1}_B] = E[X \mathbf{1}_B]$ 成立。

4. 特别的如果取 $B = \Omega$, 有 $E[X^*] = E[X]$ 。
5. 可以通过如下的图形也理解条件数学期望。前面已经用这个例子解释了条件概率。可以设想在一个样本空间中事件发生就是在平板上方掉落的球。如果球落在了横板上上面, 说明该事件就发生了。图片中红色挡板就是事件 A , 蓝色的挡板就是的事件 B 。在给定事件 A 条件下, 事件 B 的条件概率就是下落的小球首先已经触碰到红色挡板 A 的条件下, 并且再次触碰到挡板 B 的概率。该图例的动态演示可以在 <http://setosa.io/conditional/> 查看。
6. 由于 σ 代数 $\mathcal{G}$ 中就是一系列的事件, 可以把她们想象成一系列的挡板 (A_n) 。当挡板 (A_n) 移动时或和事件 B 形成不同的相交部分, 所以改变挡板 (A_n) 就会得到不同的条件概率和条件数学期望。由于我们并不知道是哪一块挡板 (A_n) 发生了, 所以只能用依赖于 (A_n) 的函数来表示, 就变成了随机变量。

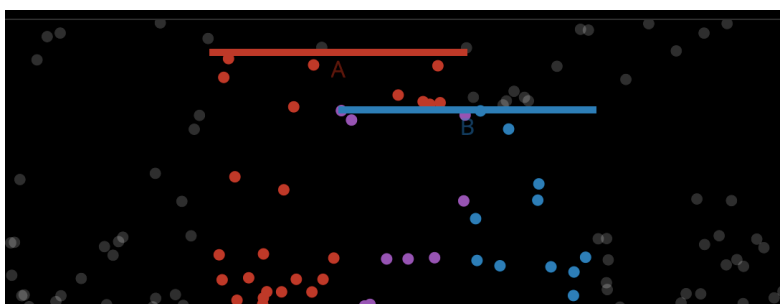


图 1.9: 条件数学期望

例 1.32 假设 B, A 为任意的两个集合。 σ 代数 $\mathcal{G} = \{\emptyset, \Omega, A, A^c\}$ 。可以把 σ 代数 $\mathcal{G}$ 想象成上述图片中的红色挡板 A 和 A^c 。可以通过上面一般条件数学期望的定义证明

$$E[\mathbb{1}_B|\mathcal{G}] = P(B|A)\mathbb{1}_A + P(B|A^c)\mathbb{1}_{A^c}.$$

其中 $\mathbb{1}_B, \mathbb{1}_A, \mathbb{1}_{A^c}$ 都是示性函数。这个验证留给读者。

定理 1.14. 条件期望性质

条件期望有如下基本性质:

1. $E[E[X|\mathcal{G}]] = E[X]$.
2. 若 X 是 $\mathcal{G}$ 可测, 则 $E[X|\mathcal{G}] = X, a.s..$
3. 设 $\mathcal{G} = \{\emptyset, \Omega\}$, 则 $E[X|\mathcal{G}] = E[X], a.s..$
4. 若 $X \leq Y, a.s.$, 则 $E[X|\mathcal{G}] \leq E[Y|\mathcal{G}], a.s..$
5. 若 a, b 为实数, $X, Y, aX + bY$ 的期望存在, 则

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

如果右边和式有意义.

6. $|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}], a.s..$
7. 设 X 及 XY 的期望存在, 且 Y 为 $\mathcal{G}$ 可测, 则

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

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

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

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

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

10. 若 X, Y 是两个独立的随机变量, 函数 $g(x, y)$ 使得 $E[|g(X, Y)|] < +\infty$, 则有

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

这里 $E[g(X, y)]|_{y=Y}$ 的意义是, 先将 y 视为常数, 求得数学期望 $E[g(X, y)]$ 后再将随机变量 Y 代入到 y 的位置.



理解 1.15. 条件数学期望性质

1. 如果随机变量 X 已经是关于我们的知识 $\mathcal{G}$ 可测了, 对她的最佳预测就是自身。则 $E[X|\mathcal{G}] = X$.
2. 如果随机变量 X 关于我们的知识 $\mathcal{G}$ 独立, 对她的最佳预测就是均值。则 $E[X|\mathcal{G}] = E[X]$.



理解 1.16. 条件数学期望

命题 1.8. 最佳预测值

对于任意的 σ 代数 $\mathcal{G}$, 条件数学期望是对随机变量 X 在方差最小化的最佳预测值。

$$\min_{\xi \in \mathcal{G}} E(X - \xi)^2$$

则最优的 $\xi^* = E[X|\mathcal{G}]$.



证明

$$\begin{aligned} E(X - \xi)^2 &= E(X - E[X|\mathcal{G}] + E[X|\mathcal{G}] - \xi)^2 \\ &= E(X - E[X|\mathcal{G}])^2 + 2E[(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - \xi)] + E(E[X|\mathcal{G}] - \xi)^2 \\ &= E(X - E[X|\mathcal{G}])^2 + E(E[X|\mathcal{G}] - \xi)^2 \\ &\geq E(X - E[X|\mathcal{G}])^2. \end{aligned}$$

上述最后一个不等式在 $\xi = E[X|\mathcal{G}]$ 时取等号。同时我们还用到

$$E[(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - \xi)] = E[E[(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - \xi)|\mathcal{G}]] = 0.$$



第二章 随机过程的基本概念与类型

内容提要

- 随机过程的背景和基本概念
- 有限维分布与 Kolmogorov 定理
- 平稳增量过程
- 独立增量过程
- 平稳独立增量过程
- 宽平稳过程
- 正态过程

2.1 随机过程的背景

在许多实际问题中, 不仅需要对随机现象做特定时间点上的一次观察, 且需要做多次的连续不断的观察, 以观察研究对象随时间推移的演变过程. 因此我们需要随机过程来描述。

例 2.1 对某股票的价格进行 T 年的连续观察, 记录得

$$\{X_t, 0 \leq t \leq T\}.$$

例 2.2 对一天气温进行 T 年的连续观察, 记录得

$$\{X_t, 0 \leq t \leq T\}.$$

例 2.3 对上交所市场上每日的成交量连续记录得到

$$\{X_t, 0 \leq t \leq T\}.$$

2.2 基本概念

定义 2.1. 随机过程

随机过程是概率空间 $(\Omega, \mathcal{F}, P)$ 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。当 $T = \{0, 1, 2, \dots\}$ 时称之为随机序列或时间序列。



定义 2.2. 状态空间

用 E 表示随机过程的值域, 称 E 为过程的状态空间。



例 2.4 设 $(\Omega, \mathcal{F}, P)$ 是对应于抛均匀硬币的概率空间: 做 N 次抛硬币独立试验, 引入随机

变量 $(\omega = (\omega_1, \dots, \omega_t, \dots, \omega_N))$

$$X(t, \omega) = \begin{cases} 0, & \omega_t = \text{正面} \\ 1, & \omega_t = \text{反面}. \end{cases}$$

则 $X(t)$ 是一随机过程. 其参数集 $T = \{0, 1, 2, \dots\}$, 状态空间 $E = \{0, 1\}$.

理解 2.1. 随机过程的理解

称 $T \times \Omega = \{(t, \omega) : t \in T, \omega \in \Omega\}$ 为集合 T 与 Ω 的积集 (product set).

随机过程可看成定义在积集 $T \times \Omega$ 上的二元函数.

- 当固定 $t \in T$, $X_t(\omega) = X(t, \omega)$ 是一个随机变量.
- 当固定 ω , 作为 $t \in T$ 的函数, $X(t, \omega)$ 是一个定义在 T 上的普通实值函数. ♣

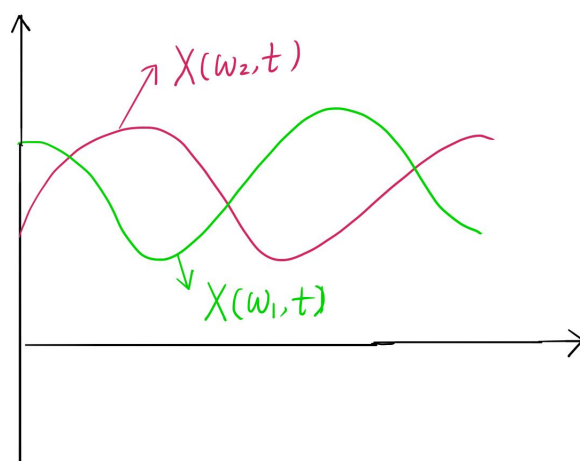


图 2.1: 随机过程轨道

定义 2.3. 轨道

对每一固定 $\omega \in \Omega$, 称 $X_t(\omega)$ 是随机过程 $\{X(t, \omega), t \in T\}$ 的一个样本函数, 或则轨道, 路径, 实现。♣

例 2.5 利用抛硬币的试验定义一个随机过程

$$X(t) = \begin{cases} \cos(\pi t), & \text{出现正面} \\ 2t, & \text{出现反面} \end{cases}$$

设出现正反面的概率相同, 写出 $X(t)$ 的所有样本函数.

证明 记:

$$\omega_1 = \{\text{出现正面}\}, \omega_2 = \{\text{出现反面}\}.$$

则 $X(t)$ 的所有现实为

$$X(\omega_1, t) = \cos(\pi t), \quad X(\omega_2, t) = 2t.$$

例 2.6 (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 2.7 (Brown 运动) 英国植物学家 Brown 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 Brown 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 Brown 运动.

2.3 有限维分布与 Kolmogorov 定理

定义 2.4. 随机过程

随机过程 $X = \{X(t), t \in T\}$, 对 $t \in T$,

$$F(t; x) := P(X(t) < x), \quad x \in R$$

为随机变量 $X(t)$ 的分布函数, 称为过程 X 的一维分布函数.

对任意 $s, t \in R$, 二维随机变量 $(X(s), X(t))$ 联合分布函数

$$F(s, t; x, y) := P(X(s) < x, X(t) < y).$$

称为随机过程 X 的二维分布函数。



定义 2.5. 随机过程 n 维分布

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

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

随机过程的所有的一维分布, 二维分布, $\dots$, n 维分布等等的全体

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

称为随机过程 $\{X(t), t \in T\}$ 的有限分布族。



笔记 知道了随机过程的有限维分布就知道了 $\{X(t), t \in T\}$ 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.

性质 分布族的性质

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

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

2. 相容性: 对于 $m < n$, 有

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

定理 2.1

设分布函数族 $\{F_{t_1, \dots, t_n}(x_1, \dots, x_n), t_1, \dots, t_n \in T, n \geq 1\}$ 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

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

恰好是 $\{X(t), t \in T\}$ 的有限维分布族。





笔记 随机过程的有限维分布函数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.

例 2.8 设随机过程 $\{X(t, \omega), t \in R\}$ 只有两条样本函数

$$X(t, \omega_1) = 2 \cos(t), \quad X(t, \omega_2) = -2 \cos(t), \quad t \in R.$$

且: $P(\omega_1) = 2/3, P(\omega_2) = 1/3$.

求:

- 一维分布函数 $F(0, x)$ 和 $F(\pi/4, x)$;
- 二维分布函数 $F(0, \pi/4; x, y)$.

证明 易知:

$$X(0) = \begin{cases} 2, & \frac{2}{3} \\ -2, & \frac{1}{3} \end{cases} \quad X(\frac{\pi}{4}) = \begin{cases} \sqrt{2}, & \frac{2}{3} \\ -\sqrt{2}, & \frac{1}{3} \end{cases}$$

同时可以得到 $X(0), X(\pi/4)$ 的联合分布

$$(X(0), X(\pi/4)) = \begin{cases} (2, \sqrt{2}), & \frac{2}{3} \\ (-2, -\sqrt{2}), & \frac{1}{3} \end{cases}$$

■

例 2.9 设随机过程

$$X(t) = A \cos(kt + \Theta), \quad t \in R$$

其中 k 是正常数, 随机变量 A 与 Θ 相互独立, $A \sim U(0, 1), \Theta \sim U(-\pi, \pi)$. 试求过程的一维概率密度. 其中 $U(0, 1), U(-\pi, \pi)$ 表示均匀分布.

证明 首先设 $Y(t) = a \cos(kt + \Theta)$, 其中 a 为常数, 易求的 $Y(t)$ 的一维概率密度函数为

$$f_Y(y) = \begin{cases} \frac{1}{\pi \sqrt{a^2 - y^2}}, & |y| < a \\ 0, & \text{others.} \end{cases}$$

注意到: $Y(t) = X(t)|_{A=a}$, 由全概率公式和独立性, 可以得到:

$$\begin{aligned} f_X(t; x) &= \int_{-\infty}^{\infty} f_{X|A}(x|a) dF_A(a) = \int_0^1 f_{X|A}(x|a) da \\ &= \int_x^1 \frac{1}{\pi \sqrt{a^2 - x^2}} da = -\frac{1}{\pi} \log \left(\frac{x}{\sqrt{1 - x^2} + 1} \right). \end{aligned}$$

■

2.3.1 随机过程的数字特征

定义 2.6. n 维特征函数

过程 $\{X(t), t \in T\}$ 的 n 维特征函数定义为

$$\psi(t_1, t_2, \dots, t_n; \theta_1, \theta_2, \dots, \theta_n) = E(\exp(i(\theta_1 X(t_1) + \theta_2 X(t_2) + \dots + \theta_n X(t_n))))$$

称 $\{\psi(t_1, t_2, \dots, t_n; \theta_1, \theta_2, \dots, \theta_n); t_1, t_2, \dots, t_n, n \geq 1\}$ 为 X 的有限维特征函数族. ♣

定义 2.7. 自相关函数

设 $\{X(t), t \in T\}$ 是一随机过程.

- 称 $X(t)$ 的期望 $m_X(t) = E[X(t)]$ 为过程的**均值函数** (如果存在的话).
- 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 $\gamma(t_1, t_2) = E[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))], t_1, t_2 \in T$ 为过程的**协方差函数**;
- 称 $\text{Var}[X(t)] = \gamma(t, t)$ 为过程的**方差函数**;
- 称 $R_X(s, t) = E[X(s)X(t)], s, t \in T$ 为**自相关函数**.



例 2.10 设 p, q 是两个随机变量, 构成随机过程

$$X(t) = p + qt, \quad t \in T$$

均值函数为: $m(t) = E[X(t)] = E(p) + tE(q)$. 自相关函数为

$$R(s, t) = E[(p + qs)(p + qt)] = E(p^2) + E(pq)(s + t) + E(q^2)st.$$

例 2.11 利用抛硬币的试验定义一个随机过程

$$X(t) = \begin{cases} \cos(\pi t), & \omega_1 \\ 2t, & \omega_2. \end{cases}$$

求该过程的均值函数, 方差函数, 相关函数, 协方差函数.

解 因对任意实数 $t \in R$, 有:

$$X(t) = \begin{cases} \cos(\pi t), & \frac{1}{2} \\ 2t, & \frac{1}{2}. \end{cases}$$

$$m_X(t) = E(X(t)) = \frac{1}{2} \cos(\pi t) + t$$

$$E(X^2(t)) = \frac{1}{2} \cos^2(\pi t) + 2t^2$$

$$D_X(t) = \left(\frac{1}{2} \cos(\pi t) - t \right)^2$$

$$\begin{aligned} R(s, t) &= E(X(s)X(t)) = \frac{1}{2} \cos(\pi s) \cos(\pi t) + \frac{1}{2} * 2s * 2t \\ &= \frac{1}{2} \cos(\pi s) \cos(\pi t) + 2st. \end{aligned}$$



例 2.12 设随机过程

$$X(t) = A \cos(\beta t + \Theta)$$

其中 β 是正常数, 随机变量 A 与 Θ 相互独立, $A \sim N(0, 1), \Theta \sim U(0, 2\pi)$. 试求过程的均值函数和相关函数.

解 由定义可以得到

$$\begin{aligned} m_X(t) &= E(X(t)) = E(A \cos(\beta t + \Theta)) \\ &= E(A)E(\cos(\beta t + \Theta)) = 0. \\ R(s, t) &= E(X(s)X(t)) \\ &= E(A^2 \cos(\beta t + \Theta) \cos(\beta s + \Theta)) \\ &= \frac{1}{2} \cos(\beta(t - s)). \end{aligned}$$

■

例 2.13 $X(t) = X_0 + tV, a \leq t \leq b$, 其中 X_0 和 V 是相互独立且服从 $N(0, 1)$ 分布的随机变量.

解 由定义可以得到

$$\begin{aligned} m_X(t) &= E[X(t)] = E(X_0 + tV) = EX_0 + tEV = 0 \\ R(t_1, t_2) &= E[X(t_1)X(t_2)] = E[(X_0 + t_1V)(X_0 + t_2V)] \\ &= E[X_0^2] + t_1t_2E[V^2] = 1 + t_1t_2. \end{aligned}$$

例 2.14 设

$$X(t) = \xi \cos \omega t + \eta \sin \omega t, t \in R$$

$$E(\xi) = E(\eta) = 0, E(\xi^2) = E(\eta^2) = \sigma^2, \omega$$

为常数, ξ 与 η 相互独立同服从正态分布,

- 1) 试求 $X(t)$ 的均值函数和相关函数;
- 2) 写出一维概率密度和二维概率密度.

证明 由定义可以得到

- 1)

$$E\{X(t)\} = E(\xi) \cos \omega t + E(\eta) \sin \omega t = 0$$

因 $E(\xi\eta) = 0$ 故

$$\begin{aligned} R(s, t) &= E\{(\xi \cos \omega t + \eta \sin \omega t)(\xi \cos \omega s + \eta \sin \omega s)\} \\ &= E(\xi^2) \cos \omega t \cos \omega s + E(\eta^2) \sin \omega t \sin \omega s \\ &= \sigma^2 \cos \omega(s - t) = \sigma^2 \cos(\omega\tau), (\tau = s - t). \end{aligned}$$

由此可得:

$$D(X(t)) = R(t, t) = \sigma^2 \cos 0 = \sigma^2.$$

- 2) $X(t)$ 的一维密度为

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

■

2.4 随机过程的基本类型

2.4.1 平稳过程

定义 2.8. 严平稳

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 $t_i + h \in T$) 有,

$(X(t_1 + h), \dots, X(t_n + h))$ 与 $(X(t_1), \dots, X(t_n))$ 具有相同的联合分布, 记为

$$(X(t_1 + h), \dots, X(t_n + h)) \stackrel{d}{=} (X(t_1), \dots, X(t_n)) \quad (2.1)$$

其中 $\stackrel{d}{=}$ 表示同分布。则称 $\{X(t), t \in T\}$ 为**严平稳**的。



例 2.15 投掷一枚均匀的硬币。如果出现正面, $X = 1$; 如果出现反面, $X = -1$ 。则 X 为严平稳过程。

定义 2.9. 宽平稳过程

如果随机过程 $X(t)$ 的所有二阶矩都存在, 并且均值函数 $E[X(t)] = \mu$, 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**。



命题 2.1

- 对于宽平稳过程, 由于 $\gamma(s, t) = \gamma(0, t - s), s, t \in \mathbb{R}$, 可记为 $\gamma(t - s)$;
- $\gamma(t)$ 为偶函数, 且 $\gamma(0) = \text{var}(X(t)), |\gamma(\tau)| \leq \gamma(0)$;
- $\gamma(\tau)$ 具有非负定性, 即对任意时刻 t_k 和实数 $a_k, k = 1, 2, \dots, N$, 有

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



当参数 t 仅取整数值 $0, \pm 1, \pm 2, \dots$ 或 $0, 1, 2, \dots$ 时, 称平稳过程为平稳序列。

例 2.16 (平稳白噪声序列) 设 $\{X_n, n = 0, 1, \dots\}$ 为一列两两互不相关的随机变量序列, 满足 $EX_n = 0, n = 0, 1, 2, \dots$, 且

$$E[X_m X_n] = \begin{cases} 0, & \text{当 } m \neq n \\ \sigma^2, & \text{当 } m = n \end{cases}$$

则白噪声序列 $\{X_n, n = 0, 1, \dots\}$ 为平稳的。这是因为协方差函数 $\text{cov}(X_n, X_m) = E(X_n X_m)$ 只与 $m - n$ 有关。

例 2.17 (滑动平均序列) 设 $\{\varepsilon_n, n = 0, \pm 1, \pm 2, \dots\}$ 为一列互不相关的且有相同均值 m 和方差 σ^2 的随机变量序列。设 $a_1, a_2, \dots, a_k$ 为任意 k 个实数。考虑如下定义序列的平稳性:

$$X_n = a_1 \varepsilon_n + a_2 \varepsilon_{n-1} + \dots + a_k \varepsilon_{n-k+1}, \quad n = 0, \pm 1, \pm 2, \dots$$

解 有定义可知:

$$EX_n = m(a_1 + \dots + a_k).$$



令 $\xi_j = \varepsilon_j - m$, 则由 ε_j 的两两互不相关性知, 协方差函数.

$$\begin{aligned}\gamma(n+\tau, n) &= E[(X_n - m(a_1 + \cdots + a_k))(X_{n+\tau} - m(a_1 + \cdots + a_k))] \\ &= E[(a_1\xi_n + \cdots + a_k\xi_{n-k+1})(a_1\xi_{n+\tau} + \cdots + a_k\xi_{n+\tau-k+1})] \\ &= \begin{cases} \sigma^2(a_{\tau+1}a_1 + \cdots + a_ka_{k-\tau}), & \text{若 } 0 \leq \tau \leq k-1, \\ 0, & \text{若 } \tau \geq k, \end{cases}\end{aligned}$$

即协方差函数仅与时间间隔 τ 有关. 故 $\{X_n, n=0, \pm 1, \cdots\}$ 为平稳序列. ■

2.4.2 独立增量过程

定义 2.10. 独立过程

对任意的正整数 n 及任意的 $t_1, t_2, \cdots, t_n \in T$ 随机变量 $x(t_1), x(t_2), \cdots, x(t_n)$ 相互独立, 称

$$\{x(t), t \in T\}$$

为独立过程。 ♣

注: 独立随机过程的有限维分布由一维分布确定

$$F_n(t_1, \cdots, t_n; x_1, \cdots, x_n) = \prod_{k=1}^n F_k(t_k, x_k)$$

例 2.18 高斯白噪声: 实值时间序列 $\{x(n), n \in N\}$ 的

$$E\{x(n)\} = 0, D[(x(n))] = \sigma^2$$

自相关函数为

$$R(m, n) = \begin{cases} 0, & m \neq n; \\ \sigma^2, & m = n. \end{cases}$$

如果两两序列不相关, 则称 $\{x(n), n \in N\}$ 为离散白噪声 (序列)。又若 $x(n)$ 都服从正态分布, 称 $\{x(n), n \in N\}$ 是高斯白噪声序列。

定义 2.11. 平稳独立增量的过程

如果对任何 $t_1, t_2, \cdots, t_n \in T, t_1 < t_2 < \cdots < t_n$ 随机变量 $X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$ 是相互独立的, 则称 $X(t)$ 为**独立增量过程**。

如果对任何 t_1, t_2 , 有 $X(t_1 + h) - X(t_1) \stackrel{d}{=} X(t_2 + h) - X(t_2)$, 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**平稳独立增量的过程**。 ♣

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

定理 2.2

假设 $\{X(t), t \geq 0\}$ 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

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



证明 必要性显然, 以下证明充分性。由独立增量性有

$$\begin{aligned}\psi_{X(t)}(a)\psi_{X(s)}(a) &= \psi_{X(t+s)}(a) \\ &= E[e^{iaX(t+s)}] \\ &= E\{e^{ia[X(t+s)-X(s)]}\}E[e^{iaX(s)}] \\ &= E\{e^{ia[X(t+s)-X(s)]}\}\psi_{X(s)}(a)\end{aligned}$$

由此可得

$$E\{e^{ia[X(t+s)-X(s)]}\} = \psi_{X(t)}(a) = E[e^{iaX(t)}]$$

即该过程具有平稳的增量。 ■

例 2.19 若 $\{X_n, n \in N\}$ 是独立时间序列, 令

$$Y(n) = \sum_{k=0}^n X(k), \quad X(0) = 0.$$

则 $\{Y(n), n \in N^+\}$ 是独立增量过程。又若 $X(n), n = 1, 2, \dots$ 相互独立同分布, 则 $\{Y(n), n \in N^+\}$ 是平稳独立增量过程。

证明 对于任意的 $n_1, n_2, \dots, n_m$

$$\begin{aligned}Y(n_2) - Y(n_1) &= \sum_{k=0}^{n_2} X(k) - \sum_{k=0}^{n_1} X(k) \\ &= X(n_1 + 1) + \dots + X(n_2). \\ Y(n_3) - Y(n_2) &= \sum_{k=0}^{n_3} X(k) - \sum_{k=0}^{n_2} X(k) \\ &= X(n_2 + 1) + \dots + X(n_3). \\ &\vdots \\ Y(n_m) - Y(n_{m-1}) &= \sum_{k=0}^{n_m} X(k) - \sum_{k=0}^{n_{m-1}} X(k) \\ &= X(n_{m-1} + 1) + \dots + X(n_m).\end{aligned}$$

因为 $X(n), n = 1, 2, \dots$ 相互独立, 所以 $\{Y(n), n \in N^+\}$ 是独立增量过程。 ■

引理 2.1. 独立平稳增量过程的性质

如果 $\{X(t), t \geq 0\}$ 是平稳独立增量过程, $X(0) = 0$, 则

1. 均值函数 $m(t) = kt$, 其中 k 为一常数。
2. 方差函数 $D(t) = \sigma^2 t$.
3. 协方差函数 $C(s, t) = \sigma^2 \min(s, t)$.



证明 1.2. 因均值函数和方差函数满足

$$\begin{aligned} m(s+t) &= E[X(s+t)] = E[X(s+t) - X(s) + X(s)] \\ &= E[X(s+t) - X(s)] + E[X(s)] = m(s) + m(t) \\ D(s+t) &= D(s) + D(t). \end{aligned}$$

把上述的 s 替换成 Δt 。利用 $m(0) = 0$, 可以得到

$$m(t + \Delta t) - m(t) = m(\Delta t) - m(0)$$

由此有:

$$\frac{m(t + \Delta t) - m(t)}{\Delta t} = \frac{m(\Delta t) - m(0)}{\Delta t}$$

让 $\Delta t \rightarrow 0$, 有 $m'(t) = m'(0)$ 。因此 $m(t)$ 是一个线性函数。由此可得到: $m(t) = kt$ 和 $D(t) = \sigma^2 t$ 。

3. 不妨假设 $s < t$, 可以得到:

$$\begin{aligned} C(s, t) &= E[X(t) - m(t)](X(s) - m(s)) \\ &= E[X(t)X(s)] - m(t)m(s) \\ &= E[(X(t) - X(s) + X(s))X(s)] - k^2 ts \\ &= E[X(t) - X(s)] * E[X(s)] + E[X(s)^2] - k^2 ts \\ &= k(t-s) * ks + \sigma^2 s + k^2 s^2 - k^2 ts = \sigma^2 s. \end{aligned}$$

■

引理 2.2

独立增量过程的有限维分布由一维分布和增量的分布确定.



证明 对于独立增量过程 $\{X(t), t \geq 0\}$, 任取 $t_1 < t_2 < \cdots < t_n$, 令

$$Y_1 = X(t_1), Y_2 = X(t_2) - X(t_1), \cdots, Y_n = X(t_n) - X(t_{n-1}).$$

相互独立性, 据此利用特征函数法可证明结论.

第三章 Brown 运动 (维纳过程)

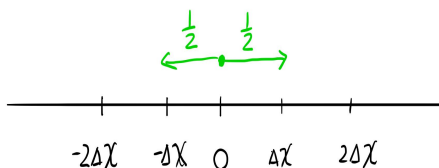
3.1 基本概念与性质

我们从讨论简单的随机游动开始. 设有一个粒子在直线上随机游动, 在每个单位时间内等可能地向左或向右移动一个单位的长度. 现在加速这个过程, 在越来越小的时间间隔中走越来越小的步子. 若能以正确的方式趋于极限, 我们就得到 Brown 运动. 详细地说就是令此过程每隔 Δt 时间等概率地向左或向右移动 Δx 的距离. 如果以 $X(t)$ 记时刻 t 粒子的位置, 则

$$X(t) = \Delta x(\xi_1 + \cdots + \xi_{[t/\Delta t]}) \quad (3.1)$$

其中 $[t/\Delta t]$ 表示 $t/\Delta t$ 的整数部分,

图 3.1: 随机游走



利用如下独立同分布的随机变量序列 ξ_i 来描述随机游走的位置

$$\xi_i = \begin{cases} +1, & \text{如果第 } i \text{ 步向右} \\ -1, & \text{如果第 } i \text{ 步向左} \end{cases}$$

且假设诸 ξ_i 相互独立

$$P\{\xi_i = 1\} = P\{\xi_i = -1\} = \frac{1}{2}.$$

由于 $E[\xi_i] = 0, \text{Var}[\xi_i] = E[\xi_i^2] = 1$, 我们有

$$E[X(t)] = 0, \text{Var}[X(t)] = (\Delta x)^2[t/\Delta t].$$



笔记 令 Δt 趋于零, 要使得上面 X_t 有意义, Δx 不能随意取值. 如果在 Δt 趋于零时, Δx 太大, 说明随机游走变动太剧烈, 以至于会使得 X_t 的方差无穷大. 如果 Δx 太小, 说明随机游走变动太微弱, 以至于会使得 X_t 静止不动. 所以要使得 X_t 成为一个有丰富的性质的随机过程. 我们需要建立 Δx 和 Δt 的合理大小关系.

命题 3.1

- 如果取 $\Delta x = \Delta t$, 令 $\Delta t \rightarrow 0$, 则 $Var[X(t)] \rightarrow 0$, 从而 $X(t) = 0, a.s.$.
- 如果取 $\Delta t = (\Delta x)^3$, 则 $Var[X(t)] \rightarrow \infty$, 这是不合理的. 因为粒子的运动是连续的, 不可能在很短时间内远离出发点.
- 如果 $\Delta x = \sigma\sqrt{\Delta t}$, σ 为某个正常数, 从上面的讨论可见, 当 $\Delta t \rightarrow 0$ 时, $E[X(t)] = 0$ 和 $Var[X(t)] \rightarrow \sigma^2 t$. 该情况下得到的是具有丰富性质的随机过程, 即不变动太剧烈也不是静止不动.



笔记 命题3.1的详细计算过程如下

- 若 $\Delta x = 0$, $X_t = 0$ 此时是静止状态.
- 若 $\Delta x = 1$, 首先来看它的均值:

$$E[X_t] = \Delta x \cdot (E(\xi_1) + \cdots + E(\xi_{[t/\Delta t]}))$$

因为考虑的是对称的随机游走, 且 ξ_i 是独立同分布的, 所以有 $E(\xi_i) = 0$, 且有方差 $Var(\xi_i) = 1$, 则

$$E[X_t] = 0$$

同时有

$$\begin{aligned} Var(X_t) &= Var(\xi_1 + \xi_2 + \cdots + \xi_{[t/\Delta t]}) \\ &= Var(\xi_1) + Var(\xi_2) + \cdots + Var(\xi_{[t/\Delta t]}) \\ &= \underbrace{1 + 1 + \cdots + 1}_{[t/\Delta t]} \\ &= \left[\frac{t}{\Delta t} \right] \xrightarrow{\Delta t \rightarrow 0} \infty \end{aligned}$$

由以上两种情况的结果我们可以看出, 如果 Δx 很小, 过程仍然是一个随机过程, 但是变化太简单了, 以至于静止不动, 没有必要去研究; 若 Δx 取值为 1, 那么变化又太剧烈, 因此以上两种情况都无意义。

那么什么样的 Δx 和 Δt 才是有意义的呢? 我们来看三种情况:

- 若 $\Delta x \sim \Delta t$, 即 Δx 与 Δt 成比例, 二者的大小关系已知, 则根据前面的计算, 均值为:

$$E[X_t] = 0$$

方差为:

$$\begin{aligned} Var(X_t) &= (\Delta x)^2 \cdot (Var(\xi_1 + \cdots + \xi_{[t/\Delta t]})) \\ &= (\Delta t)^2 \cdot \underbrace{(1 + 1 + \cdots + 1)}_{[t/\Delta t]} \\ &= (\Delta t)^2 \cdot \left[\frac{t}{\Delta t} \right] \xrightarrow{\Delta t \rightarrow 0} 0 \end{aligned}$$



也就是说, 这个过程不具备任何的随机性.

- 若 $\Delta x = (\Delta t)^{\frac{1}{3}}$, 均值为:

$$E[X_t] = 0$$

方差为:

$$\text{Var}[X_t] = (\Delta x)^2 \cdot \left[\frac{t}{\Delta t} \right] = (\Delta t)^{\frac{2}{3}} \cdot \left[\frac{t}{\Delta t} \right] \sim \frac{t}{(\Delta t)^{\frac{1}{3}}} \xrightarrow{\Delta t \rightarrow 0} \infty$$

此时这种情况变化也很剧烈.

- 若 $\Delta x = \sigma\sqrt{\Delta t}$, 均值为:

$$E[X_t] = 0$$

方差为:

$$\text{Var}(X_t) = (\Delta x)^2 \cdot \left[\frac{t}{\Delta t} \right] = \sigma^2 \cdot \Delta t \cdot \left[\frac{t}{\Delta t} \right] \rightarrow \sigma^2 \cdot t$$

由于该过程的方差与时间 t 有关, 即使 $\Delta t \rightarrow 0$, 我们依然可以得到 $0 < \sigma^2 \cdot t < +\infty$, 这种过程就是有意义的.

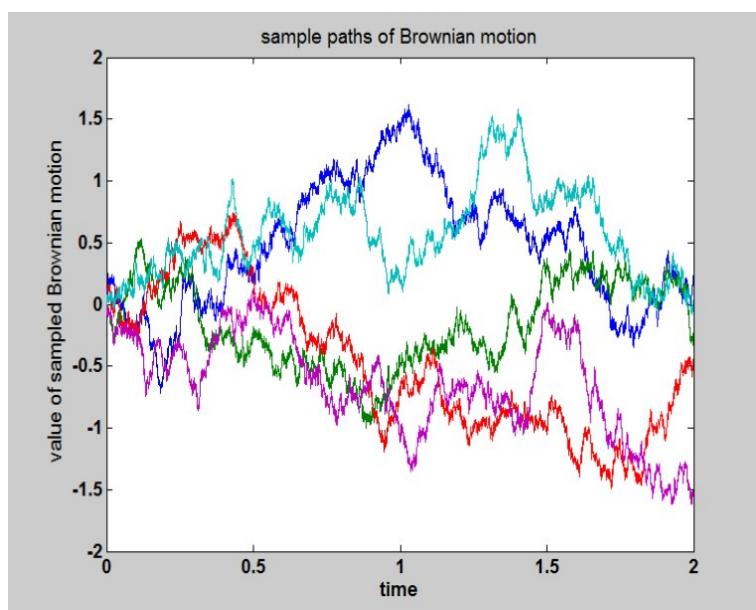
综上所述, 只有满足

$$\begin{cases} E(X_t) &= 0 \\ \text{Var}(X_t) &= \sigma^2 \cdot t \end{cases}$$

这样的过程对我们金融建模才是有意义的.

下面图片是一个随机游走模拟的轨道。我们可以看到每条轨道相比一般物体的运动轨迹要粗糙得多。

图 3.2: 布朗运动轨道



理解 3.1. 轨道的粗糙

从布朗运动轨道可以看到,即使我们放大轨道,观察到的局部轨道依然是很粗糙的。这种粗糙性就是布朗运动最独特的地方,是由前面讨论的 $\Delta x = \sigma\sqrt{\Delta t}$ 带来的。从下面这种图片可以理解为什么布朗运动比一般物体运动要复杂很多。

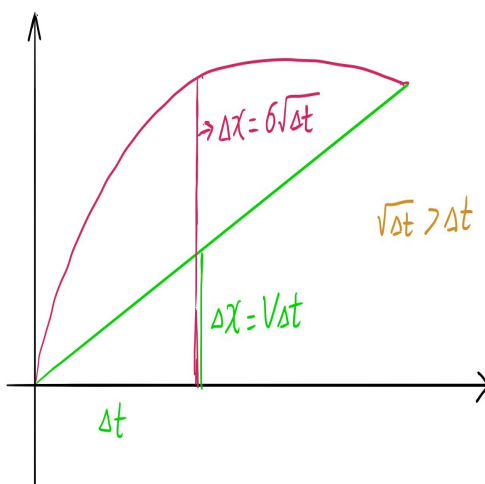
对于一般物体的运动,通过一段时间 Δt ,位置的改变量为 $\Delta x = V\Delta t$ 。这儿 V 为物体运动速度。然而对于布朗运动在 Δt 时间范围内的位置改变量为

$$\Delta x = \sigma\sqrt{\Delta t} = \frac{\sigma}{\sqrt{\Delta t}}\Delta t.$$

由于 $\sqrt{\Delta t} > \Delta t$, 因此布朗运动改变量要比一般物体运动变动量要更大。如果我们把 $\frac{\sigma}{\sqrt{\Delta t}}$ 理解成布朗运动的速度,则当 Δt 越小时,这个速度也越大。这也是布朗运动要更粗糙的原因。



图 3.3: 布朗运动和普通运动



下面我们就给出 Brown 运动的严格定义.

定义 3.1. Brown 运动

随机过程 $\{X(t), t \geq 0\}$ 如果满足:

1. $X(0) = 0$;
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 $N(0, \sigma^2 t)$. 则称 $\{X(t), t \geq 0\}$ 为 **Brown 运动**, 也称为 **Wiener 过程**. 常记为 $\{B(t), t \geq 0\}$ 或 $\{W(t), t \geq 0\}$.



如果 $\sigma = 1$, 我们称之为标准 Brown 运动, 如果 $\sigma \neq 1$, 则可考虑 $\{X(t)/\sigma, t \geq 0\}$, 它是标准 Brown 运动. 故不失一般性, 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.

性质 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) $B(t) - B(s) \sim N(0, t - s)$, 即 $B(t) - B(s)$ 服从均值为 0, 方差为 $t - s$ 的正态分布. 当 $s = 0$ 时, $B(t) - B(0) \sim N(0, t)$.
2. (独立增量) $B(t) - B(s)$ 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.
4. 我们并没有假定 $B(0) = 0$, 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 $\{B^x(t)\}$. 易见,

$$B^x(t) - x = B^0(t) \quad (3.2)$$

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 Brown 运动就可以了, 如不加说明, Brown 运动就是始于 0 的 Brown 运动.

3.2 维纳过程的分布

- 维分布: $B(t) \sim N(0, \sigma^2 t)$;
- 增量分布: $B(t) - B(s) \sim N(0, \sigma^2 |t - s|)$; 设 $t > s$, 因 $B(0) = 0$, 且 $B(t)$ 是平稳独立增量过程, 故 $B(t) - B(s) = B(t - s + s) - B(s)$ 与 $B(t - s) - B(0) = B(t - s)$ 相同分布 $N(0, \sigma^2(t - s))$.

引理 3.1

维纳过程是正态过程.



证明 设维纳过程 $\{B(t), t \geq 0\}$ 的参数是 σ^2 , 任取 n 及 $t_1 < t_2 < \cdots < t_n$,

$$X_k = B(t_k) - B(t_{k-1}), t_0 = 0, k = 1, 2, \cdots, n$$

则 $X_k \sim N(0, \sigma^2(t_k - t_{k-1}))$, 且相互独立, 有

$$\begin{aligned} B(t_k) &= X_1 + X_2 + \cdots + X_k \\ \rightarrow \begin{pmatrix} B(t_1) \\ B(t_2) \\ \vdots \\ B(t_n) \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \end{aligned}$$

■



笔记 我们来回顾一下以上讲到的布朗运动的性质:

- 在任何一个时刻 t , 我们能观测到一个随机变量 B_t , 且满足 $B_t \sim N(0, \sigma^2 t)$. 如果是标准布朗运动, 我们可以假设 $\sigma = 1$.
- 独立增量的性质: 如果我们在任何不相交的时间区间上观测位置的变化量时, 我们可以看到这些增加量, 例如 $B_{t_2} - B_{t_1}$, $B_{t_3} - B_{t_2}$, 它们是相互独立的.

- 平稳增量的性质：从 t_1 到 t_2 的位置增量在分布的意义下只与时间的长度有关，即

$$B_{t_2} - B_{t_1} \sim B_{t_2-t_1} \sim N(0, \sigma^2(t_2 - t_1))$$

$$B_{t_3} - B_{t_2} \sim B_{t_3-t_2} \sim N(0, \sigma^2(t_3 - t_2))$$

3.3 维纳过程的数字特征

命题 3.2. 相关函数

- $E(B(t)) = 0$;
- $R_B(s, t) := E(B(s)B(t)) = \sigma^2 \min(s, t)$



证明 由正态分布可以知道 $E(B(t)) = 0$ 。先不妨假设 $s < t$ ，由独立平稳增量可以得到

$$\begin{aligned} E[B(s)B(t)] &= E[B_s(B_t - B_s + B_s)] = E[B_s(B_t - B_s)] + E[B_s^2] \\ &= E[B_s]E[(B_t - B_s)] + E[B_s^2] = E[B_s^2] = \sigma^2 s. \end{aligned}$$



这儿用到的分开 B_s, B_t 的技巧在随机过程中常用到。

例 3.1 设 $\{B(t), t \geq 0\}$ 是参数为 σ^2 的维纳过程，求下列过程的均值函数和相关函数。

- $X(t) = B^2(t), t \geq 0$;
- $X(t) = tB(1/t), t > 0$

证明 1)

$$m_X(t) = E[X(t)] = E[B^2(t)] = \text{Var}[B(t)] + \{E[B(t)]\}^2 = \sigma^2 t.$$

首先假设 $s < t$ ，可以得到

$$\begin{aligned} R_X(s, t) &= E[X(s)X(t)] = E[B^2(s)B^2(t)] \\ &= E\{B^2(s)[B(t) - B(s) + B(s)]^2\} \\ &= E\{B^2(s)[B(t) - B(s)]^2\} + E[B^4(s)] \\ &\quad + 2E\{B^3(s)[B(t) - B(s)]\} \\ &= E[B^2(s)]E\{[B(t) - B(s)]^2\} + E[B^4(s)] \\ &= \sigma^2 S \sigma^2(t - s) + 3\sigma^4 s^2 = \sigma^4(st + 2s^2) \end{aligned}$$

因此，

$$R_X(s, t) = \sigma^4(st + 2\min^2(s, t))$$

其中因 $B(t) \sim N(0, \sigma^2 t)$;

$$B(t) - B(s) \sim N(0, \sigma^2|t - s|)$$

另因：若 $X \sim N(0, \sigma^2)$ ，有

$$E(X^n) = \begin{cases} 0, & n = 1, 3, 5, \dots; \\ \sigma^n(n-1)(n-3)\cdots 1, & n = 2, 4, 6, \dots \end{cases}$$



2) 根据定义得到

$$\begin{aligned} m_X(t) &= E[X(t)] = tE\left[B\left(\frac{1}{t}\right)\right] = 0, t > 0 \\ R_X(s, t) &= E\left[sB\left(\frac{1}{s}\right) \cdot tB\left(\frac{1}{t}\right)\right] = stE\left[B\left(\frac{1}{s}\right)B\left(\frac{1}{t}\right)\right] \\ &= st\sigma^2 \min\left(\frac{1}{s}, \frac{1}{t}\right) = \sigma^2 \min(s, t). \end{aligned}$$



笔记 我们来看计算过程中可能出现的一种错误情况:

$$\begin{aligned} R_X(s, t) &= E[X(s)X(t)] = E[B^2(s)B^2(t)] \\ &= E[B_s^2(B_t^2 - B_s^2 + B_s^2)] \\ &= E[B_s^2(B_t^2 - B_s^2)] + E(B_s^4) \\ &= E(B_s^2)E(B_t^2 - B_s^2) + E(B_s^4) \end{aligned}$$

最后一步是错误的, 因为 $B_t - B_s$ 和 B_s 是独立的, 但是并不能得出 $B_t^2 - B_s^2$ 和 B_s^2 是独立的, 因为

$$B_t^2 - B_s^2 = (B_t + B_s)(B_t - B_s)$$

并不能得到 $B_t + B_s$ 和 B_s 是独立的.



笔记 这儿需要注意对于布朗运动 B , 如果 $s < t$, 我们知道 B_s 和 $B_t - B_s$ 相互独立. 但是并没有 B_s^2 和 $B_t^2 - B_s^2$ 相互独立!!

下面给出关于 Brown 运动的概率计算的例子.

例 3.2 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 $B(2) \sim N(0, 2)$, 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 $B(2) - B(1)$ 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

$$B(2) = B(1) + (B(2) - B(1)) = \xi + \eta$$

其中为了后面记号方便, 我们用 ξ, η 表示两个独立同分布的标准正态分布. 我们有

$$\begin{aligned} P\{B(1) \leq 0, B(2) \leq 0\} &= P\{\xi \leq 0, \xi + \eta \leq 0\} \\ &= P\{\xi \leq 0, \eta \leq -\xi\} = \int_{-\infty}^0 P\{\eta \leq -x\} \phi(x) dx \\ &= \int_{-\infty}^0 \Phi(-x) \phi(x) dx. \end{aligned}$$

这里 Φ 和 ϕ 分别表示标准正态分布的分布函数和密度函数. 通过如下变量代换可以得到 $u = \Phi(x)$ 和 $du = d\Phi(x) = \phi(x)dx$. 由积分变量替换公式得

$$\begin{aligned} \int_{-\infty}^0 \Phi(-x) \phi(x) dx &= \int_0^{\infty} \Phi(x) \phi(-x) dx = \int_0^{\infty} \Phi(x) \phi(x) dx \\ &= \int_{\frac{1}{2}}^1 u du = \frac{3}{8} \end{aligned}$$

3.3.1 二次变差



笔记 在开始学习二次变差前, 我们先引入一次变差的概念. 为了研究一个函数的粗糙性, 我们可以通过函数值变化的累加值来判断, 即

$$|f(t_2) - f(t_1)| + |f(t_3) - f(t_2)| + \cdots + |f(t_n) - f(t_{n-1})|$$

其中, $t_1, t_2, \cdots, t_n$ 为某一个时间区间 $[0, T]$ 的划分. 于是可以引出一变差的定义.

定义 3.2. 一次变差

在时间 $[0, T]$ 上有函数 $f(t)$, $t_1, t_2, \cdots, t_n$ 为时间区间 $[0, T]$ 的划分. 其一次变差为:

$$I_f[0, T] = \lim_{n \rightarrow \infty} \sum_{i=1}^n |f(t_i) - f(t_{i-1})|, \text{ 其中 } \delta = \max_i (t_i - t_{i-1})$$

**笔记**

- 若把一次变差应用于一般函数 $f(t) = t$,

$$I_f[0, T] = \lim_{n \rightarrow \infty} \sum_{i=1}^n |t_i - t_{i-1}| = \lim_{n \rightarrow \infty} \sum_{i=1}^n (t_i - t_{i-1}) = T < +\infty$$

- 若把一次变差应用于布朗运动 $f(t) = B_t$,

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n |B_{t_i} - B_{t_{i-1}}|$$

我们先来看它的数学期望

$$E(I_B[0, T]) = \lim_{n \rightarrow \infty} \sum_{i=1}^n E|B_{t_i} - B_{t_{i-1}}|$$

我们假设把时间区间等分, $t_i - t_{i-1} = \Delta t = T/n$, 所以有

$$B_{t_i} - B_{t_{i-1}} \sim B_{\Delta t} \sim N(0, \sigma^2 \Delta t)$$

则

$$\begin{aligned} E|B_{t_i} - B_{t_{i-1}}| &= E|B_{\Delta t}| \\ &= \int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi} \cdot \sigma \sqrt{\Delta t}} \cdot e^{-\frac{x^2}{2\sigma^2 \Delta t}} dx \\ &= 2 \int_0^{\infty} x \frac{1}{\sqrt{2\pi} \cdot \sigma \sqrt{\Delta t}} \cdot e^{-\frac{x^2}{2\sigma^2 \Delta t}} dx \\ &\stackrel{u=x^2}{=} \int_0^{\infty} \frac{1}{\sqrt{2\pi} \sigma \sqrt{\Delta t}} \cdot e^{-\frac{u}{2\sigma^2 \Delta t}} du \\ &= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma \sqrt{\Delta t}} \cdot (-2\sigma^2 \Delta t) e^{-\frac{u}{2\sigma^2 \Delta t}} \Big|_0^{\infty} \\ &= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma \sqrt{\Delta t}} \cdot 2\sigma^2 \Delta t \\ &= \frac{1}{\sqrt{2\pi}} \cdot 2\sigma \sqrt{\Delta t} \end{aligned}$$

因此

$$\begin{aligned}
 E(I_B[0, T]) &= \lim_{n \rightarrow \infty} \sum_{i=1}^n E|B_{\Delta t}| \\
 &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{2\sigma}{\sqrt{2\pi}} \cdot \sqrt{\Delta t} \\
 &= \lim_{n \rightarrow \infty} \frac{2\sigma}{\sqrt{2\pi}} \cdot \sqrt{\Delta t} \cdot n \\
 &= \lim_{n \rightarrow \infty} \frac{2\sigma}{\sqrt{2\pi}} \cdot \sqrt{\frac{T}{n}} \cdot n \\
 &= \lim_{n \rightarrow \infty} \frac{2\sigma}{\sqrt{2\pi}} \cdot \sqrt{T} \cdot \sqrt{n} \rightarrow +\infty
 \end{aligned}$$

故有

$$E(I_B[0, T]) = \lim_{n \rightarrow +\infty} \sum_{i=1}^n E|B_{t_i} - B_{t_{i-1}}| \rightarrow +\infty$$

因此, 若用一次变差来衡量布朗运动轨迹的复杂性时, 得到数学期望为 $+\infty$ 的量, 没有研究意义, 因此引入二次变差的概念.

性质 Brown 运动的路径性质 从时刻 0 到时刻 T 对 Brown 运动的一次观察称为 Brown 运动在区间 $[0, T]$ 上的一个路径或一个实现. Brown 运动的几乎所有样本路径 $B(t), 0 \leq t \leq T$ 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) 上都是无限变差的;
5. 对任意 t , 在 $[0, t]$ 上的二次变差等于 $\sigma^2 t$.

为了讨论 Brown 运动的路径性质, 首先给出二次变差的定义.

定义 3.3. 二次变差

Brown 运动的二次变差 $[B, B](t)$ 定义为当 $\{t_i^n\}_{i=0}^n$ 遍取 $[0, t]$ 的分割, 且其模 $\delta_n = \max_{0 \leq i \leq n-1} (t_{i+1}^n - t_i^n) \rightarrow 0$ 时, 依概率收敛意义下的极限

$$[B, B](t) := \lim_{\delta_n \rightarrow 0} \sum_{i=0}^{n-1} |B(t_{i+1}^n) - B(t_i^n)|^2 \quad (3.3)$$



定理 3.1. 二次变差

非标准的布朗运动 ($B_t \sim N(0, \sigma^2 t)$) 的二次变差为 $[B, B](t) = \sigma^2 t$.



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

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



则

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

再由标准正态分布的四阶矩公式, 得

$$\begin{aligned} \text{Var}[S_n] &= \text{Var}\left[\sum_{i=0}^{n-1} (B(t_{i+1}^n) - B(t_i^n))^2\right] \\ &= \sum_{i=0}^{n-1} \text{Var}[(B(t_{i+1}^n) - B(t_i^n))^2] \\ &= \sigma^4 \sum_{i=0}^{n-1} 2(t_{i+1}^n - t_i^n)^2 \leq 2\sigma^4 \max(t_{i+1}^n - t_i^n) t = 2\sigma^4 t \delta_n \\ &\rightarrow 0. \end{aligned}$$

于是有 $S_n - t \rightarrow 0$, a.s., 故 $[B, B](t) = \sigma^2 t$. ■



笔记 考虑二次变差在一般函数的意义. 设有连续函数 $f(t)$, 考虑它在 $[0, T]$ 上的二次变差

$$[f, f](0, T) = \lim_{n \rightarrow +\infty} \sum_{i=1}^n (f(t_{i+1}) - f(t_i))^2$$

将其放缩

$$\begin{aligned} \lim_{n \rightarrow +\infty} \sum_{i=1}^n (f(t_{i+1}) - f(t_i))^2 &\leq \lim_{n \rightarrow +\infty} \sum_{i=1}^n \max(|f(t_{i+1}) - f(t_i)|) \cdot |f(t_{i+1}) - f(t_i)| \\ &= \max(|f(t_{i+1}) - f(t_i)|) \cdot \lim_{n \rightarrow +\infty} \sum_{i=1}^n |f(t_{i+1}) - f(t_i)| \\ &\rightarrow 0 \end{aligned}$$

在上面极限中, 我们用到了 f 为连续函数和一次变差有限. 因此我们可以总结得出

- 对于连续函数实函数 f 有

$$I_f[0, T] < +\infty, [f, f](0, T) = 0$$

- 对于布朗运动 B 有

$$E(I_f[0, T]) \rightarrow +\infty, [B, B](0, T) = \sigma^2 T$$

例 3.3 设 $\{B(t)\}$ 是 Brown 运动, 求 $B(1) + B(2) + B(3) + B(4)$ 的分布.

解 考虑随机向量 $\mathbf{X} = (B(1), B(2), B(3), B(4))'$, 则 $\mathbf{X}$ 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 $\mathbf{A} = (1, 1, 1, 1)$, 则

$$\mathbf{A}\mathbf{X} = X_1 + X_2 + X_3 + X_4 = B(1) + B(2) + B(3) + B(4)$$

具有均值为零, 方差为 $\mathbf{A}\Sigma\mathbf{A}' = 30$ 的正态分布. 于是 $B(1) + B(2) + B(3) + B(4)$ 是服从

均值为 0, 方差为 30 的正态分布.



笔记 下面是另一种解法. 因为正态分布的线性组合也是正态分布, 所以我们有

$$B_1 + B_2 + B_3 + B_4 \sim N(0, ?)$$

因此只需确定方差即可. 又因为

$$\begin{aligned} B_2 &= B_2 - B_1 + B_1 \\ B_3 &= B_3 - B_2 + B_2 - B_1 + B_1 \\ B_4 &= B_4 - B_3 + B_3 - B_2 + B_2 - B_1 + B_1 \end{aligned}$$

令

$$\begin{aligned} \xi_1 &= B_1 \\ \xi_2 &= B_2 - B_1 \\ \xi_3 &= B_3 - B_2 \\ \xi_4 &= B_4 - B_3 \end{aligned}$$

因为 $\xi_1, \xi_2, \xi_3, \xi_4 \sim N(0, 1)$ 所以有

$$\begin{aligned} B_1 + \cdots + B_4 &= 4\xi_1 + 3\xi_2 + 2\xi_3 + \xi_4 \\ \text{Var}(B_1 + \cdots + B_4) &= \text{Var}(4\xi_1 + 3\xi_2 + 2\xi_3 + \xi_4) \\ &= \text{Var}(4\xi_1) + \text{Var}(3\xi_2) + \text{Var}(2\xi_3) + \text{Var}(\xi_4) \\ &= 16 + 9 + 4 + 1 \\ &= 30 \end{aligned}$$

例 3.4 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 $AY = B(\frac{1}{4}) + B(\frac{1}{2}) + B(\frac{3}{4}) + B(1)$ 服从均值为 0, 方差为 $\frac{30}{4}$ 的正态分布.

例 3.5 求概率 $P\{\int_0^1 B(t)dt > \frac{2}{\sqrt{3}}\}$.

解 首先需要指出的是, Brown 运动具有连续路径, 所以对每个路径来说, 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 Riemann 积分的定义, 我们可以从逼近和 $\sum B(t_i)\Delta t_i$ 的极限分布得到. 这里 t_i 是 $[0, 1]$ 的分点, $\Delta t_i = t_{i+1} - t_i$. 例如取 $t_i = \frac{i}{n}$, 当 $n = 4$ 时, 逼近和即为上例中的随机变量. 一般地, 类似地证明, 所有逼近和的分布都是零均值的正态分布, 因此它们的极限分布是正态分布. 于是 $\int_0^1 B(t)dt$ 也是零均值的正态分布. 下面来计算 $\int_0^1 B(t)dt$ 的方差.

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

这样, $\int_0^1 B(t) dt \sim N(0, \frac{1}{3})$. 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.

3.4 Brown 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 3.4. Brown 运动鞅

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

$$E[X(t+s)|\mathcal{F}_t] = X(t), \text{ a.s.} \quad (3.4)$$

这里 $\mathcal{F}_t = \sigma\{X(u), 0 \leq u \leq t\}$ 是由 $\{X(u), 0 \leq u \leq t\}$ 生成的 σ 代数.



定理 3.2

设 $\{B(t)\}$ 是 Brown 运动, 则

1. $\{B(t)\}$ 是鞅;
2. $\{B(t)^2 - t\}$ 是鞅;
3. 对任意实数 u , $\exp \left\{ uB(t) - \frac{u^2}{2}t \right\}$ 是鞅.



证明 首先, 由 $B(t+s) - B(t)$ 与 $\mathcal{F}_t$ 的独立性可知对任意可测函数 $g(x)$, 有

$$E[g(B(t+s) - B(t))|\mathcal{F}_t] = E[g(B(t+s) - B(t))] \quad (3.5)$$

由 Brown 运动的定义, $B(t) \sim N(0, t)$, 所以 $B(t)$ 可积, 且 $E[B(t)] = 0$, 再由其他性质得

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

从而 (1) 得证.

由于 $E[B^2(t)] = t < \infty$, 所以 $B^2(t)$ 可积. 于是得到

$$\begin{aligned} B^2(t+s) &= [B(t) + B(t+s) - B(t)]^2 \\ &= B^2(t) + 2B(t)[B(t+s) - B(t)] + [B(t+s) - B(t)]^2 \end{aligned}$$

$$\begin{aligned} E[B^2(t+s)|\mathcal{F}_t] &= B^2(t) + 2E[B(t)(B(t+s) - B(t))|\mathcal{F}_t] + E[(B(t+s) - B(t))^2|\mathcal{F}_t] \\ &= B^2(t) + s \quad (3.6) \end{aligned}$$

这里我们利用了 $B(t+s) - B(t)$ 与 $\mathcal{F}_t$ 的独立性且具有均值 0, 并对 $g(x) = x^2$ 应用. 在上式两端同时减去 $(t+s)$, 则 (2) 得证.

考虑 $B(t) \sim N(0, t)$ 的矩母函数 $E[e^{uB(t)}] = e^{tu^2/2} < \infty$, 这蕴含着 $e^{uB(t)}$ 是可积的, 并且

$$E[e^{uB(t) - \frac{u^2}{2}t}] = 1$$

取 $g(x) = e^{ux}$, 可得

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

两端同时乘以 $e^{-\frac{u^2}{2}(t+s)}$, 则 (3) 得证. ■



笔记 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 Brown 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 Brown 运动.

3.5 Brown 运动的最大值变量及反正弦律

以 T_x 记 Brown 运动首次击中 x 的时刻, 即

$$T_x = \inf\{t > 0, B(t) = x\}$$

当 $x > 0$ 时, 为计算 $P\{T_x \leq t\}$, 我们考虑 $P\{B(t) \geq x\}$.

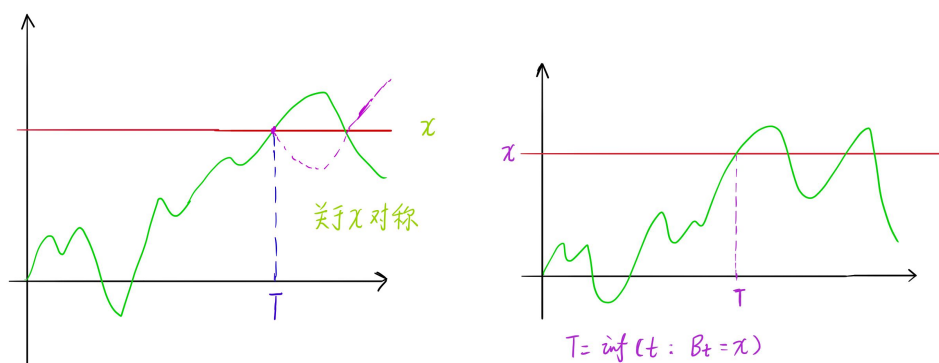
由全概率公式

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

若 $T_x \leq t$, 则 $B(t)$ 在 $[0, t]$ 中的某个时刻击中 x , 由对称性得

$$P\{B(t) \geq x | T_x \leq t\} = \frac{1}{2}$$

图 3.4: 布朗运动反射



再由连续性可知, $B(t)$ 不可能还未击中 x 就大于 x , 所以上面第 2 项为零. 因此

$$\begin{aligned} P\{T_x \leq t\} &= 2P\{B(t) \geq x\} \\ &= \frac{2}{\sqrt{2\pi t}} \int_x^\infty e^{-u^2/2t} du \\ &= \frac{2}{\sqrt{2\pi}} \int_{x/\sqrt{t}}^\infty e^{-y^2/2} dy \end{aligned}$$

由此可见

$$P\{T_x < \infty\} = \lim_{t \rightarrow \infty} P\{T_x \leq t\} = \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-y^2/2} dy = 1$$

对分布函数求导数可得其分布密度

$$f_{T_x}(u) = \begin{cases} \frac{x}{\sqrt{2\pi}} u^{-3/2} e^{-\frac{x^2}{2u}}, & \text{如果 } u > 0 \\ 0, & \text{如果 } u \leq 0 \end{cases}$$

利用式, 可以得到

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

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 Brown 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 Brown 运动的常返性. 由于始于点 a 的 Brown 运动与 $\{a + B(t)\}$ 是相同的, 这里 $\{B(t)\}$ 是始于 0 的 Brown 运

动, 所以

$$P_a\{T_x < \infty\} = P_0\{T_{x-a} < \infty\} = 1$$

即 Brown 运动从任意点出发, 击中 x 的概率都是 1. 当 $x < 0$ 时, 由对称性, T_x 与 T_{-x} 有相同的分布. 于是有

$$P\{T_x \leq t\} = \frac{2}{\sqrt{2\pi}} \int_{|x|/\sqrt{t}}^{\infty} e^{-y^2/2} dy$$

$$f_{T_x}(u) = \begin{cases} \frac{-x}{\sqrt{2\pi}} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}}, & u > 0 \\ 0, & u \leq 0 \end{cases}$$

另一个有趣的随机变量是 Brown 运动在 $[0, t]$ 中达到的最大值 $M(t) = \max_{0 \leq s \leq t} B(s)$ 的分布可由下述等式得到, 对 $x > 0$ 有

$$\begin{aligned} P\{M(t) \geq x\} &= P\{T_x \leq t\} \\ &= 2P\{B(t) \geq x\} \\ &= \frac{2}{\sqrt{2\pi}} \int_{x/\sqrt{t}}^{\infty} e^{-y^2/2} dy \end{aligned}$$

读者不难得到 Brown 运动在 $[0, t]$ 中达到的最小值 $m(t) = \min_{0 \leq s \leq t} \{B(s)\}$ 的分布.

3.6 Brown 运动的几种变化

3.6.1 Brown 桥

由 Brown 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 3.5. Brown 桥

设 $\{B(t), t \geq 0\}$ 是 Brown 运动. 令

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

则称随机过程 $\{B^*(t), 0 \leq t \leq 1\}$ 为 Brown 桥 (Brown Bridge).



因为 Brown 运动是 Gauss 过程, 所以 Brown 桥也是 Gauss 过程, 其 n 维分布由均值函数和方差函数完全确定. 且 $\forall 0 \leq s \leq t \leq 1$, 有

$$\begin{aligned} E[B^*(t)] &= 0 \\ E[B^*(s)B^*(t)] &= E[(B(s) - sB(1))(B(t) - tB(1))] \\ &= E[B(s)B(t) - tB(s)B(1) - sB(t)B(1) + tsB^2(1)] \\ &= s - ts - ts + ts = s(1 - t) \end{aligned}$$

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 Brown 桥名称的由来.

3.6.2 几何 Brown 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 Brown 运动. 由于 Brown 运动的矩母函数为 $E[e^{sB(t)}] = e^{ts^2/2}$, 所以几何 Brown 运动的均值函数与方差函数分别为

$$\begin{aligned} E[X(t)] &= E[e^{B(t)}] = e^{t/2} \\ \text{Var}[X(t)] &= E[X^2(t)] - (E[X(t)])^2 \\ &= E[e^{2B(t)}] - e^t \\ &= e^{2t} - e^t \end{aligned}$$

在金融市场中, 人们经常假定股票的价格按照几何 Brown 运动变化, 在下面的例子中我们如此假定.

例 3.6 (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 Brown 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

$$\begin{aligned} E[\max(X(T) - K, 0)] &= \int_0^\infty P\{X(T) - K > u\} du \\ &= \int_0^\infty P\{ye^{B(T)} - K > u\} du \\ &= \int_0^\infty P\{B(T) > \log \frac{K+u}{y}\} du \\ &= \frac{1}{\sqrt{2\pi T}} \int_0^\infty \int_{\log[(K+u)/y]}^\infty e^{-x^2/2T} dx du \end{aligned}$$



笔记 总结一下本章的内容:

- 对于对称的随机游走, 经过 t 时间位置确定

$$S_t = \Delta x (\xi_1 + \xi_2 + \cdots + \xi_{[t/\Delta t]})$$

其中 $\xi_i = +1$ 或 -1 . 则当 $\Delta t \rightarrow 0$, S_t 即为 Brown 运动.

- (1) 对于 $\forall t$, 有 $B_t \sim N(0, \sigma^2 t)$, 当 $\sigma = 1$ 时为标准的正态分布.
- (2) B_t 是连续的.
- (3) 平稳增量性质: 对于 t_1, t_2, t_3 , 我们可以得到 $B_{t_1}, B_{t_2}, B_{t_3}$, 有 $B_{t_2} - B_{t_1}$ 与 $B_{t_3} - B_{t_2}$ 相互独立, 且

$$B_{t_2} - B_{t_1} \sim N(0, (t_2 - t_1)\sigma^2)$$

$$B_{t_3} - B_{t_2} \sim N(0, (t_3 - t_2)\sigma^2)$$

- 二次变差:

$$[B, B](0, t) = \sigma^2 t$$

$$\Delta[B, B] = \Delta B \cdot \Delta B = \Delta t$$

因此我们可以近似地得到 $\Delta B \sim \sqrt{\Delta t}$

- 一个 Brown 运动从 0 时刻开始得到一条线 T_x , 我们可以得到它分布函数 f_{T_x} , 同时也可以得到 T_x 作为一个随机变量有 $T_x < +\infty$, 且 $E(T_x) = +\infty$
- Brown 运动的鞅性质和指数鞅性质.
- Brown 桥, $B_t^* = B_t - tB(1)$



第四章 Poisson 过程

4.1 齐次泊松过程

在天文，地理，物理，生物，通信，医学，计算机网络，密码学等许多领域，都有关于随机事件流的计数问题，如：

- 盖格计数器上的粒子流；
- 计算机网络上的（图象，声音）流；
- 编码（密码）中的误码流；
- 交通中事故流；
- 细胞中染色体的交换次数，...
- 均构成以时间顺序出现的事件流 $A_1, A_2, \dots$

定义 4.1. 计数过程

随机过程 $\{N(t), t \geq 0\}$ 称为计数过程 (Counting process), 如果 $N(t)$ 表示在 $[0, t]$ 内事件 A 出现的总次数.



计数过程应满足：

1. $N(t) \geq 0$;
2. $N(t)$ 取非负整数值；
3. 如果 $s < t$, 则 $N(s) \leq N(t)$;
4. 对于 $s < t$, $N(t) - N(s)$ 表示时间间隔 $(s, t]$ 内事件出现的次数.

4.1.1 Poisson 过程数学模型

电话呼叫过程: 设 $N(t)$ 为 $[0, t)$ 时间内到达的呼叫次数, 其状态空间为

$$E = \{0, 1, 2, \dots\}$$

此过程有如下特点：

- 1) 零初值性: $N(0) = 0$;
- 2) 独立增量性: 任意两个不相重叠的时间间隔内到达的呼叫次数相互独立;
- 3) 齐次性: 在 (s, t) 时间内到达的呼叫次数仅与时间间隔长度 $t - s$ 有关, 而与起始时间 s 无关;
- 4) 普通性: 在充分小的时间间隔内到达的呼叫次数最多仅有一次, 即对充分小的 Δt , 有

$$\begin{aligned}
P\{N(\Delta t) = 0\} &= P_0(\Delta t) = 1 - \lambda\Delta t + o(\Delta t), \\
P\{N(\Delta t) = 1\} &= P_1(\Delta t) = \lambda\Delta t + o(\Delta t), \\
P\{N(\Delta t) \geq 2\} &= \sum_{k=2}^{\infty} P_k(\Delta t) = o(\Delta t),
\end{aligned}$$

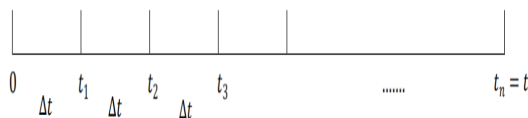
其中 $\lambda > 0$.

为什么实际中有这么多的现象可以用 Poisson 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.

理解 4.1. 齐次泊松过程

实上, 把 $[0, t]$ 划分为 n 个相等的时间区间, 则由条件 (4)' 可知, 当 $n \rightarrow \infty$ 时, 在每个小区间内事件发生两次或两次以上的概率趋于 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. 

图 4.1: n 次独立 Bernoulli 试验



如上图所示, 把 $[0, t]$ 划分为 n 个相等的时间区间 $\Delta t = \frac{t}{n}$. 用 ξ 表示每一次独立同分布的随机变量.

$$\xi = \begin{cases} 1, & \lambda\Delta t := p \\ 0, & 1 - \lambda\Delta t := q \\ 2, & o(\Delta t) \end{cases}$$

$$\begin{aligned}
P(N_t = k) &\approx C_n^k (\lambda\Delta t)^k (1 - \lambda\Delta t)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \left(\lambda \frac{t}{n}\right)^k \left(1 - \lambda \frac{t}{n}\right)^{n-k}.
\end{aligned}$$

让划分的区间趋于无穷时, 可以得到

$$\begin{aligned}
 P(N_t = k) &\approx \lim_{n \rightarrow \infty} \frac{(\lambda t)^k}{k!} \cdot \frac{n!}{(n-k)!} \cdot \frac{1}{n^k} \left(1 - \lambda \frac{t}{n}\right)^{n-k} \\
 &= \frac{(\lambda t)^k}{k!} \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-k+1)}{n \cdot n \cdot \dots \cdot n} \cdot \left(1 - \lambda \frac{t}{n}\right)^{n-k} \\
 &= \frac{(\lambda t)^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \lambda \frac{t}{n}\right)^{n-k} \\
 &= \frac{(\lambda t)^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \lambda \frac{t}{n}\right)^{\frac{-n}{\lambda t} \cdot \frac{\lambda t}{-n} (n-k)} \\
 &= \frac{(\lambda t)^k}{k!} e^{-\lambda t}.
 \end{aligned}$$

可以看到 N_t 刚好是强度为 λt 的泊松分布。由此, 我们可以引出齐次泊松过程的严格定义:

定义 4.2. 齐次泊松过程

计数过程 $\{N(t), t \geq 0\}$ 满足:

- $$\begin{cases}
 (1) N(0) = 0; \\
 (2) \text{是平稳独立增量过程}; \\
 (3) P\{N(\Delta t) = 1\} = \lambda \Delta t + o(\Delta t), \quad \lambda > 0; \\
 (4) P\{N(\Delta t) \geq 2\} = o(\Delta t).
 \end{cases}$$

称 $\{N(t), t \geq 0\}$ 是参数 (或速率, 强度) 为 λ 的齐次泊松过程.



上面关于齐次泊松过程在直观上很显然, 但我们还不知道分布, 所以在实际运用中还不是很方便。下面的定理给出了 N_t 的具体分布, 便于我们讨论齐次泊松过程的数字特征。

定理 4.1. 齐次泊松过程的分布

齐次泊松过程 $\{N(t), t \geq 0\}$ 在时间间隔 $(t_0, t_0 + t)$ 内事件出现 n 次的概率为:

$$P\{[N(t_0 + t) - N(t_0)] = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad (n = 0, 1, 2, \dots)$$



证明 为了记号的方便, 我们引入下面的符号。

$$\begin{aligned}
 P_n(t) &:= P\{N(t) = n\} = P\{[N(t) - N(0)] = n\} \\
 &= P\{N(t_0 + t) - N(t_0) = n\} \quad (1)
 \end{aligned}$$

步骤 1: 我们首先建立 $P_0(t)$ 的方程。由条件 (2)~(4), 得:

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

由此可以得到:



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

令 $h \rightarrow 0$, 得 $\begin{cases} \frac{dP_0(t)}{dt} = -P_0(t)\lambda \\ P_0(0) = 1, \end{cases} \Leftrightarrow N(0) = 0.$ 解得 $P_0(t) = e^{-\lambda t}, t \geq 0.$

步骤 2. 当 $n \geq 1$, 根据全概率公式有

$$\begin{aligned} P_n(t+h) &= P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + o(h) \\ &= (1-\lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h) \end{aligned}$$

由此可以得到

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

令 $h \rightarrow 0$, 得到

$$\frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t)$$

两边同乘以 $e^{\lambda t}$ 后

$$\begin{aligned} d(e^{\lambda t} P_n(t)) &= \lambda e^{\lambda t} P_n(t) dt + e^{\lambda t} dP_n(t) \\ &= \lambda e^{\lambda t} P_n(t) dt + e^{\lambda t} (-\lambda P_n(t) dt + \lambda P_{n-1}(t) dt) \\ &= \lambda e^{\lambda t} P_{n-1}(t) dt \end{aligned}$$

即移项整理得

$$\frac{d[e^{\lambda t} P_n(t)]}{dt} = \lambda e^{\lambda t} P_{n-1}(t), \quad (2)$$

当 $n = 1$, 则

$$\begin{cases} \frac{d[e^{\lambda t} P_1(t)]}{dt} = \lambda e^{\lambda t} P_0(t) = \lambda e^{\lambda t} e^{-\lambda t} = \lambda \\ P_1(0) = 0 \end{cases}$$

解得

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

步骤 3. 假设 $P_{n-1}(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}$ 成立. 带入 (2) 式有

$$\frac{d[e^{\lambda t} P_n(t)]}{dt} = \lambda e^{\lambda t} P_{n-1}(t) = \frac{\lambda(\lambda t)^{n-1}}{(n-1)!} \rightarrow e^{\lambda t} P_n(t) = \frac{(\lambda t)^n}{(n)!} + c$$

利用初始条件 $P_n(0) = 0$, 可证得

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

对一切 $n \geq 0$ 均成立。 ■

基于上面的定理, 得泊松过程的等价定义:



定义 4.3. 泊松过程的等价定义

设计数过程 $\{N(t), t \geq 0\}$ 满足下述条件:

- (1) $N(0) = 0$;
- (2) $N(t)$ 是独立增量过程;
- (3) 对一切 $0 \leq s < t$, $N(t) - N(s) \sim \text{Poisson}(\lambda(t-s))$, 即

$$P\{[N(t) - N(s)] = k\} = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)} \\ (k = 0, 1, 2, \dots)$$

称 $\{N(t), t \geq 0\}$ 是参数 (或速率, 强度) 为 λ 的齐次泊松过程.

特别有

$$P(N(t) = k) = P([N(t) - N(0)] = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, (k = 0, 1, 2, \dots)$$



例 4.1 设 $\{N(t), t \geq 0\}$ 是参数为 λ 的泊松过程, 事件 A 在 $(0, \tau]$ 时间区间内出现 n 次, 试求:

$$P\{N(s) = k | N(\tau) = n\}, 0 < k < n, 0 < s < \tau$$

解

$$\begin{aligned} P\{N(s) = k | N(\tau) = n\} &= \frac{P\{N(s) = k, N(\tau) = n\}}{P\{N(\tau) = n\}} \\ &= P\{N(s) = k, N(\tau) - N(s) = n - k\} n! e^{\lambda \tau} (\lambda \tau)^{-n} \\ &= e^{-\lambda s} \frac{(\lambda s)^k}{k!} e^{-\lambda(\tau-s)} \frac{[\lambda(\tau-s)]^{n-k}}{(n-k)!} n! e^{-\lambda \tau} (\lambda \tau)^{-n} \\ &= \frac{n!}{k!(n-k)!} \left(\frac{s}{\tau}\right)^k \left(1 - \frac{s}{\tau}\right)^{n-k} \\ &= \binom{n}{k} \left(\frac{s}{\tau}\right)^k \left(1 - \frac{s}{\tau}\right)^{n-k}, k = 0, 1, 2, \dots, n \end{aligned}$$



例 4.2 事件 A 的发生形成强度为 λ 的 Poisson 过程 $\{N(t), t \geq 0\}$. 如果每次事件发生时以概率 p 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 $\{M(t), t \geq 0\}$ 是一个强度为 λp 的 Poisson 过程.

解 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, 而且事件发生服从 Poisson 分布. 所以 $M(t)$ 也是具有平稳独立增量的, 故只需验证 $M(t)$ 服从均值为 $\lambda p t$ 的 Poisson 分布. 即对 $t > 0$, 有

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

由于

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

■

例 4.3 若每条蚕的产卵数服从 Poisson 分布, 强度为 λ , 而每个卵变为成虫的概率为 p , 且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 由上例我们立即知道小蚕数服从强度为 λp 的 Poisson 分布, 故所求概率为

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

4.1.2 齐次泊松过程的数字特征

性质 齐次泊松过程的数字特征

1. 因对 $\forall t > 0, N(t) \sim \text{Poisson}(\lambda t)$
2. 均值函数 $m(t) = E(N(t)) = \lambda t$
3. 方差函数 $D(t) = \lambda t$
4. 有 $\lambda = \frac{E\{N(t)\}}{t}$, λ 是单位时间内事件出现的平均次数。
5. 协方差函数 $C(s, t) = \lambda \min(s, t) + \lambda^2 st$.
6. 特征函数 $\Psi_{N(t)}(u) = E(e^{iuN(t)}) = e^{\lambda t(e^{iu}-1)}$.

证明 (2) 因泊松过程 $\{N(t), t \geq 0\}$ 是平稳独立增量过程, 不妨设 $t > s > 0$

均值函数

$$\begin{aligned}
 m(t) &= E(N_t) = \sum_{n=0}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \sum_{n=1}^{\infty} \frac{(\lambda t)(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t} \\
 &= (\lambda t) e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} = \lambda t e^{-\lambda t} e^{\lambda t} = \lambda t.
 \end{aligned}$$

(3) 方差函数

$$\begin{aligned}
D(t) &= \text{Var}(N_t) = E(N_t^2) - (E(N_t))^2 \\
E(N_t^2) &= \sum_{n=0}^{\infty} n^2 \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\
&= e^{-\lambda t} \sum_{n=1}^{\infty} \frac{n(\lambda t)^{n-1}}{(n-1)!} (\lambda t) \\
&= (\lambda t) e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(n-1+1)}{(n-1)!} (\lambda t)^{n-1} \\
&= (\lambda t) e^{-\lambda t} \left(\sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{n=2}^{\infty} \frac{1}{(n-2)!} (\lambda t)^{n-1} \right) \\
&= (\lambda t) e^{-\lambda t} (e^{\lambda t} + \lambda t \cdot e^{\lambda t}) \\
&= \lambda t + (\lambda t)^2 \\
D(t) &= \lambda t + (\lambda t)^2 - (\lambda t)^2 = \lambda t
\end{aligned}$$

(4) 协方差函数

$$\begin{aligned}
R(s, t) &= E\{N(t)N(s)\} = E\{N(s)[N(t) - N(s) + N(s)]\} \\
&= E\{N(s)[N(t) - N(s)]\} + E[N^2(s)] \\
&= E\{N(s)\}E\{N(t) - N(s)\} + E[N^2(s)] \\
&= \lambda s \cdot \lambda(t-s) + (\lambda s + (\lambda s)^2) = \lambda s + \lambda^2 st \\
C(s, t) &= R(s, t) - m(s)m(t) = \lambda s + \lambda^2 st - \lambda s \cdot \lambda t
\end{aligned}$$

一般的有

$$C(s, t) = \lambda \min(s, t), R(s, t) = \lambda \min(s, t) + \lambda^2 st.$$

(5) 回想前面的定义, 因为对任意的随机变量, 有特征函数

$$\phi_{\xi}(u) = E(e^{iu\xi})$$

因此有

$$\begin{aligned}
\phi_{N_t}(u) &= E[e^{iuN_t}] = \sum_{n=0}^{\infty} e^{iun} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t e^{iu})^n}{n!} \\
&= e^{-\lambda t} e^{\lambda t e^{iu}} \\
&= e^{\lambda t(e^{iu}-1)}
\end{aligned}$$

■

例 4.4 设 $N_1(t)$ 和 $N_2(t)$ 分别是强度为 λ_1 和 λ_2 的相互独立的泊松过程,

1. 令 $X(t) = N_1(t) - N_2(t), t > 0$, 求 $X(t)$ 的均值函数和相关函数。
2. 证明 $X(t) = N_1(t) + N_2(t), t > 0$ 是强度为 $\lambda_1 + \lambda_2$ 的泊松过程。
3. 证明 $X(t) = N_1(t) - N_2(t), t > 0$, 不是泊松过程。



解 1) $m_X(t) = E[N_1(t)] - E[N_2(t)] = (\lambda_1 - \lambda_2)t$

$$\begin{aligned}
 R_X(s, t) &= E\{[N_1(1) - N_2(s)][N_1(t) - N_2(t)]\} \\
 &= E[N_1(s)N_1(t)] + E[N_2(s)N_2(t)] - E[N_1(s)N_2(t)] - E[N_2(s)N_1(t)] \\
 &= R_{N_1}(s, t) + R_{N_2}(s, t) - E[N_1(s)]E[N_2(t)] - E[N_2(s)]E[N_1(t)] \\
 &= \lambda_1 \min(s, t) + \lambda_1^2 st + \lambda_2 \min(s, t) + \lambda_2^2 st - 2\lambda_1 \lambda_2 st \\
 &= (\lambda_1 + \lambda_2) \min(s, t) + (\lambda_1^2 + \lambda_2^2)st - 2\lambda_1 \lambda_2 st.
 \end{aligned}$$

2) 由于 N_1, N_2 相互独立, 根据泊松分布的特征函数

$$\begin{aligned}
 \phi_{X_t}(u) &= \Psi_{N_1(t)+N_2(t)}(u) = E[e^{iu(N_1(t)+N_2(t))}] \\
 &= E(e^{iuN_1(t)}) \cdot E(e^{iuN_2(t)}) \\
 &= e^{\lambda_1 t(e^{iu}-1)} \cdot e^{\lambda_2 t(e^{iu}-1)} \\
 &= e^{(\lambda_1+\lambda_2)t(e^{iu}-1)}
 \end{aligned}$$

正好是一个强度为 $\lambda_1 + \lambda_2$ 的泊松分布.

3) $X(t) = N_1(t) - N_2(t)$ 特征函数为

$$\phi_X(u) = \exp\{\lambda_1 t e^{iu} + \lambda_1 t e^{-iu} - (\lambda_1 + \lambda_2)t\}$$

由分布函数与特征函数的一一对应的唯一性定理知 $X(t)$ 不是泊松过程。 ■

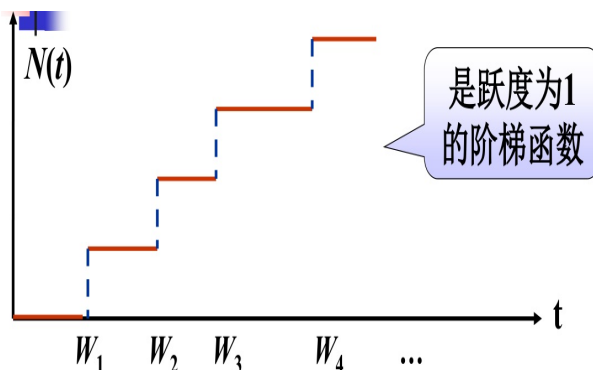
4.1.3 时间间隔与等待时间的分布

用 T_n 表示事件 A 第 $n-1$ 次出现与第 n 次出现的时间间隔。记

$$W_n = \sum_{i=1}^n T_i, \text{ 则 } T_i = W_{i+1} - W_i$$

称 W_n 为事件 A 第 n 次出现的等待时间 (到达时间)。

图 4.2: 时间间隔与等待时间的分布



定理 4.2. 泊松过程等待时间分布

设 $\{T_n, n \geq 1\}$ 是参数为 λ 的泊松过程 $\{N(t), t \geq 0\}$ 的时间间隔序列, 则 $\{T_n, n \geq 1\}$ 相互独立同服从指数分布, 且 $E\{T\} = 1/\lambda$



证明 (1) 因 $\{T_1 > t\} = \{(0, t) \text{ 内事件 } A \text{ 不出现}\}$

$$\rightarrow P\{T_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}$$

$$\rightarrow F_{T_1}(t) = 1 - P\{T > t\} = 1 - e^{-\lambda t}, t \geq 0$$

即 T_1 服从均值为 $1/\lambda$ 的指数分布。

(2) 由泊松过程的平稳独立增量性, 有

$$P\{T_2 > t | T_1 = s\} = P\{\text{在 } (s, t+s) \text{ 内事件 } A \text{ 不出现} | T_1 = s\}$$

$$= P\{N(t+s) - N(s) = 0\}$$

$$= P\{N(t) - N(0) = 0\}$$

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

故 T_2 与 T_1 相互独立, 且 T_2 也服从均值为 $1/\lambda$ 的指数分布。

(3) 对于一般 $n > 1$ 和 $t > 0$, 以及 $r_1, r_2, \dots, r_{n-1} > 0$, 有

$$P\{T_n > t | T_i = r_i, 1 \leq i \leq n-1\}$$

$$= P\{N(t+r_1+\dots+r_{n-1}) - N(r_1+r_2+\dots+r_{n-1}) = 0\}$$

$$= P\{N(t) - N(0) = 0\} = e^{-\lambda t}.$$

$$\text{即 } F_n(t) = P\{T_n \leq t\} = 1 - e^{-\lambda t}, t \geq 0. \quad \blacksquare$$

注 定理4.1.3的结果应该是预料之中的, 由于 Poisson 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.

例 4.5 假定某天文台观测到的流星流是一个 Poisson 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 Poisson 过程, 则有

$$N(4) - N(0) \sim P(3 \times 4)$$

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

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



定理 4.3

参数为 λ 的泊松过程 $\{N(t), t \geq 0\}$, 事件 A 第 n 次出现的等待时间服从 Γ 分布, 其概率密度为:

$$f_{W_n}(t) = \begin{cases} \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, & t \geq 0; \\ 0, & t < 0. \end{cases}$$

在排队论中称 W_n 服从爱尔郎分布。



解证: 因 W_n 是事件 A 第 n 次出现的等待时间, 故

$$\{W_n \leq t\} = \{N(t) \geq n\} = \{(0, t) \text{ 内 } A \text{ 至少出现 } n \text{ 次}\}$$

$$F_{W_n}(t) = P\{W_n \leq t\} = \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t}, t \geq 0$$

$$\begin{aligned} f_{W_n}(t) &= \sum_{k=n}^{\infty} \frac{\lambda(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} - \sum_{k=n}^{\infty} \lambda \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \\ &= \lambda e^{(-\lambda t)} \frac{(\lambda t)^{(n-1)}}{(n-1)!}, t \geq 0. \end{aligned}$$

**4.1.4 到达时间的条件分布****定理 4.4**

设 $\{N(t), t \geq 0\}$ 是 Poisson 过程, 已知在 $(0, t]$ 时间内 A 出现 n 次, 这 n 次到达时间 $W_1, W_2, \dots, W_n$ 的联合条件分布密度为

$$f(t_1, t_2, \dots, t_n | N(t) = n) = \begin{cases} \frac{n!}{t^n}, & 0 < t_1 < \dots < t_n; \\ 0, & \text{其他.} \end{cases}$$



证明 我们证明 $n = 1$ 的情况。

$$\begin{aligned} P(W_1 \leq t_1 | N_t = 1) &= P(W_1 \leq t_1, N_t = 1) / P(N_t = 1) \\ P(W_1 \leq t_1, N_t = 1) &= P(T_1 \leq t_1, T_1 + T_2 > t) \\ &= \int_0^{t_1} \lambda e^{-\lambda x} dx \int_{t-x}^{\infty} \lambda e^{-\lambda y} dy = t_1 \lambda e^{-\lambda t}. \end{aligned}$$

因此, $P(W_1 \leq t_1 | N_t = 1) = t_1/t$.

4.1.5 更新计数过程**定义 4.4. 更新计数过程**

定义: 设 $\{N(t), t \geq 0\}$ 是一个计数过程, 如果它的时间间隔序列 $T_1, T_2, \dots, T_n, \dots$ 相互独立同分布, 称为更新计数过程。



例 4.6 同类型设备的更新, 如一个元件; 一个灯泡; 一个系统... 假定每个更换对象的寿命

具有相同的概率密度, 则相继两次损坏之间的更新时间 $T_1, T_2, \dots$ 相互独立同分布.

定理 4.5

更新计数过程 $\{N(t), t \geq 0\}$ 是泊松过程的充要条件是时间间隔 T 具有指数分布. 

注: 等价于时间间隔序列 $T_1, T_2, \dots, T_n, \dots$ 相互独立同服从相同指数分布.

证明 方法一: 我们用 Poisson 过程的直观定义来证明。

$$\begin{aligned} P(N(\Delta t) = 0) &= P(T_1 > \Delta t) = e^{-\lambda \Delta t} \sim 1 - \lambda \Delta t + o(\Delta t) \\ P(N(\Delta t) = 1) &= P(T_1 < \Delta t, T_1 + T_2 > \Delta t) \\ &= \lambda^2 \int_0^{\Delta t} \int_{\Delta t - t_1}^{+\infty} e^{-\lambda t_1} e^{-\lambda t_2} dt_1 dt_2 \\ &= \lambda \Delta t e^{-\lambda \Delta t} \sim \lambda \Delta t + o(\Delta t). \end{aligned}$$

■

证明 方法 2 (了解): 由定理知必要性, 仅需证充分性, 应有

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

T_i 的特征函数为

$$\varphi_{T_i}(\theta) = \varphi_T(u) = \frac{1}{1 - \frac{i u}{\lambda}}, (i = 1, 2, \dots)$$

等待时间

$$W_k = \sum_{i=1}^k T_i,$$

且 $T_1, T_2, \dots, T_k$ 相互独立, 故

$$\varphi_{W_k}(u) = [\varphi_T(u)]^k = \frac{1}{[1 - \frac{i u}{\lambda}]^k}, (k = 1, 2, \dots)$$

由特征函数的反演公式及唯一性定理知, W_k 的密度函数为

$$f_{w_k}(t) = \begin{cases} \frac{(\lambda)^k}{(k-1)!} t^{k-1} e^{-\lambda t}, & t > 0; \\ 0, & t \leq 0. \end{cases}$$

积分可得

$$F_{w_k}(t) = \begin{cases} 1 - \sum_{i=0}^{k-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t}, & t > 0; \\ 0, & t \leq 0. \end{cases}$$

因

$$\{W_k < t\} = \{N(t) \geq k, t \geq 0\}$$

故

$$\begin{aligned} P\{N(t) = k\} &= P\{N(t) \geq k\} - P\{N(t) \geq k+1\} \\ &= F_{w_k}(t) - F_{w_{k+1}}(t) \\ &= \frac{(\lambda t)^k}{k!} e^{-\lambda t}, (k = 0, 1, 2, \dots, \lambda > 0) \end{aligned}$$

$\{N(t), t \geq 0\}$ 的概率分布为泊松分布, 即 $N(t) \sim P(\lambda t)$. ■

一般地, 设 $\{N(t), t \geq 0\}$ 为更新计数过程, 有:

- 1) 等待时间 $W_k = \sum_{i=1}^k T_i$ 的特征函数为,

$$\varphi_{W_k}(u) = [\varphi_T(u)]^k$$

- 2) 因 $\{W_k < t\} = \{N(t) \geq k\}$, 有

$$F_{W_k}(t) = 1 - F_{N(t)}(k)$$

- 3)

$$\begin{aligned} m(t) = E[N(t)] &= \sum_{k=0}^{\infty} k P\{N(t) = k\} \\ &= \sum_{k=0}^{\infty} k [F_{w_k}(t) - F_{w_{k+1}}(t)] \\ &= \sum_{k=1}^{\infty} F_{w_k}(t) \end{aligned}$$

例 4.7 设到达电影院的观众组成强度为 λ 的 Poisson 流, 如果电影从 t 时刻开演, 计算 $(0, t]$ 内到达电影院的观众等待时间总和的数学期望.

解 设 W_k 是第 k 名观众到达时刻, 在 $(0, t]$ 内到达的观众数为 $N(t)$, 则总等待时间为

$$\sum_{k=1}^{N(t)} (t - W_k)$$

根据全数学期望公式

$$E \left[\sum_{k=1}^{N(t)} (t - W_k) \right] = E \left\{ E \left[\sum_{k=1}^{N(t)} (t - W_k) | N(t) \right] \right\}$$

对 $\forall n \geq 1$,

$$E \left[\sum_{k=1}^{N(t)} (t - W_k) | N(t) = n \right] = nt - E \left[\sum_{k=1}^{N(t)} W_k | N(t) = n \right]$$

由定理知 $W_1, W_2, \dots, W_n$ 与 $[0, t]$ 上均匀分布相互独立随机变量的顺序统计量 $U_{(1)}, U_{(2)}, \dots, U_{(n)}$ 有相同分布.

$$\begin{aligned} &E \left[\sum_{k=1}^{N(t)} W_k | N(t) = n \right] \\ &= E \left(\sum_{k=1}^n U_{(k)} \right) = \sum_{k=1}^n E(U_{(k)}) = \frac{nt}{2} \end{aligned}$$

$$\begin{aligned}
&\Rightarrow E\left[\sum_{k=1}^{N(t)}(t - W_k) | N(t) = n\right] = nt - \frac{nt}{2} = \frac{nt}{2} \\
&\Rightarrow E\left[\sum_{k=1}^{N(t)}(t - W_k) | N(t)\right] = \frac{t}{2}N(t) \\
&\Rightarrow E\left[\sum_{k=1}^{N(t)}(t - W_k)\right] = E\left\{E\left[\sum_{k=1}^{N(t)}(t - W_k) | N(t)\right]\right\} \\
&= E\left[\frac{1}{2}N(t)\right] = \frac{\lambda t^2}{2}
\end{aligned}$$

■

4.2 复合泊松过程

例 4.8 调查城市人员流动情况, 可在关键路口观察公交车的载客情况, 设 $(0, t)$ 内通过的公交车数 $N(t)$ 是一个 poisson 过程, 而每辆车的载客人数为 ξ_n , 则经公交车通过此路口的人数为:

$$X(t) = \sum_{n=1}^{N(t)} \xi_n$$

例 4.9 若将股票交易次数 $N(t)$ 看作一个 poisson 过程, ξ_n 表示第 n 次与第 $n-1$ 次易手前后股票价格差, 则 $X(t)$ 就代表直到 t 时刻股票的价格变化.

4.2.1 复合 Poisson 过程

定义 4.5. 复合 Poisson 过程

定义: 设 $\{N(t), t \geq 0\}$ 是强度为 λ 的齐次 Poisson 过程, $\{\xi_n, n \geq 1\}$ 是相互独立同分布的随机变量序列, 并与 $N(t)$ 相互独立, 称

$$X(t) = \sum_{n=1}^{N(t)} \xi_n$$

为复合 Poisson 过程.



定理 4.6

设 $\{X_t, t \geq 0\}$ 是复合泊松过程

$$X_t = \sum_{n=1}^{N(t)} \xi_n, t \geq 0$$

其中 $\{N(t), t \geq 0\}$ 是强度为 λ 的泊松过程, $\xi_n, n = 1, 2, \dots$ 相互独立与 ξ 同分布, 有:

1. $X(t)$ 的特征函数为

$$\varphi_{X_t}(u) = e^{\lambda t[\varphi_\xi(u)-1]}, t \geq 0$$

2. 均值函数

$$m_X(t) = E[X(t)] = \lambda t E(\xi)$$

3. 方差函数为

$$\text{Var}(X_t) = \lambda t E(\xi^2)$$



证明 (1) 由定义可知:

$$\begin{aligned}\varphi_{X_t}(u) &= E[e^{iuX_t}] = E\left[e^{iu\sum_{n=1}^{N_t}\xi_n}\right] \\ &= E\left[E\left[e^{iu\sum_{n=1}^{N_t}\xi_n} \middle| N_t\right]\right]\end{aligned}$$

$$\begin{aligned}E\left[e^{iu\sum_{n=1}^{N_t}\xi_n} \middle| N_t = k\right] &= E\left[e^{iu\sum_{n=1}^k\xi_n} \middle| N_t = k\right] \\ &= E\left[e^{iu\sum_{n=1}^k\xi_n}\right] = \prod_{n=1}^k E(e^{iu\xi_n}) = (\varphi_\xi(u))^k\end{aligned}$$

$$\begin{aligned}\varphi_{X_t}(u) &= E\left[(\varphi_\xi(u))^{N_t}\right] \\ &= \sum_{n=0}^{\infty} (\varphi_\xi(u))^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\ &= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t \varphi_\xi(u))^n}{n!} \\ &= e^{-\lambda t} e^{\lambda t \varphi_\xi(u)} = e^{(\varphi_\xi(u)-1)\lambda t}\end{aligned}$$

(2) 由特征得到

$$\begin{aligned}E(X_t) &= i^{-1} \cdot \varphi'_{X_t}(u)|_{u=0} \\ &= i^{-1} \cdot e^{(\varphi_\xi(u)-1)\lambda t} \cdot \lambda t \cdot \varphi'_\xi(u)|_{u=0} \\ &= i^{-1} \cdot \lambda t \varphi'_\xi(u)|_{u=0} \\ &= i^{-1} \lambda t \cdot i E(\xi) = \lambda t E(\xi).\end{aligned}$$

(3) 由特征得到

$$\begin{aligned}-E(X_t^2) &= \varphi''_{X_t}(u)|_{u=0} \\ &= e^{(\varphi_\xi(u)-1)\lambda t} \cdot (\lambda t \cdot \varphi'_\xi(u))^2 + e^{(\varphi_\xi(u)-1)\lambda t} \cdot \lambda t \cdot \varphi''_\xi(u)|_{u=0} \\ &= (\lambda t)^2 \cdot (\varphi'_\xi(u))^2|_{u=0} + \lambda t \varphi''_\xi(u)|_{u=0} \\ &= (\lambda t)^2 (i \cdot E(\xi))^2 - \lambda t E(\xi^2) = -(\lambda t)^2 (E(\xi))^2 - \lambda t E(\xi^2)\end{aligned}$$

因此方差为

$$\text{Var}(X_t) = E(X_t^2) - E^2(X_t) = \lambda t E(\xi^2)$$



例 4.10 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 Poisson 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 以 $N_1(t)$ 表示在时间 $(0, t]$ 内进入该商场的顾客数, 则 $\{N_1(t), t \geq 0\}$ 是速率为 $\lambda = 6$ (人/分钟) 的 Poisson 过程. 再以 $N_2(t)$ 表示在时间 $(0, t]$ 内在该商场买东西的顾客数, 并设

$$Y_i = \begin{cases} 1, & \text{如果第 } i \text{ 位顾客在该商场买东西} \\ 0, & \text{如果第 } i \text{ 位顾客在该商场未买东西} \end{cases}$$

则 Y_i 独立同分布于 $B(1, 0.9)$, 与 $\{N_1(t), t \geq 0\}$ 独立, 且

$$N_2(t) = \sum_{i=1}^{N_1(t)} Y_i$$

易知

$$N_2(t) \sim P(5.4t)$$

即 $\{N_2(t), t \geq 0\}$ 是速率为 $\lambda = 5.4$ (人/分钟) 的 Poisson 过程, $t = 12$ 小时 = 720 分钟, 则

$$N_2(720) \sim P(3888)$$

即一天 (12 小时) 在该商场买东西的平均顾客数为 $E[N_2(720)] = 3888$ 人.

注: 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

$$N_3(t) = \sum_{i=1}^{N_1(t)} Z_i$$

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

$$E[N_3] = E[Z_1]E[N_1] = (200 \times 0.5) \times (6 \times 12 \times 60) = 432000 \text{ 元}$$

4.3 非齐次泊松过程 (了解内容, 不考察)

当 Poisson 过程的强度 λ 不再是常数, 而与时间 t 有关时, Poisson 过程被推广为非齐次 Poisson 过程. 一般来说, 非齐次 Poisson 过程是不具备平稳增量的.

在实际中, 非齐次 Poisson 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 Poisson 过程来描述就不合适了, 于是改用非齐次的 Poisson 过程来处理.

定义 4.6. 非齐次 Poisson 过程

如果计数过程满足下列条件:

- 1. $N(0) = 0$;
- 2. $\{N(t), t \geq 0\}$ 是一个独立增量过程;
- 3. $P\{N(t + \Delta t) - N(t) = 1\} = \lambda(t)\Delta t + o(\Delta t)$;
- 4. $P\{N(t + \Delta t) - N(t) \geq 2\} = o(\Delta t)$;

称 $\{N(t), t \geq 0\}$ 是具有速率函数为 $\lambda(t)$ 的非齐次 poisson 过程.



定理 4.7

若 $\{N(t), t \geq 0\}$ 是非齐次泊松过程, 且达到率 $\lambda(t)$ 是连续函数, 则在 $[t_0, t_0 + t]$ 时间内事件 A 出现 k 次的概率为

$$\begin{aligned} P\{[N(t_0 + t) - N(t_0)] = k\} \\ = \frac{[m(t_0 + t) - m(t_0)]^k}{k!} \exp\{-[m(t_0 + t) - m(t_0)]\}, \end{aligned}$$

其中 $k = 0, 1, 2, \dots$, $m(t) = \int_0^t \lambda(s)ds$.



第五章 鞅 (Martingale) 过程

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.

5.1 基本概念

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 $\{Y_n, n = 1, 2, \dots\}$, 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

$$b_n = b_n(Y_1, \dots, Y_{n-1}), \quad n = 2, 3, \dots$$

这儿 $b_n < \infty$ 是第 n 次的赌注, $b_n(Y_1, \dots, Y_{n-1})$ 表示我们的策略只依赖于前面 $n-1$ 实验结果. 若第 n 次赌赢则获利 b_n , 否则输掉 b_n .

设 X_0 是该赌博者的初始赌资, 则他在第 n 次赌博后的赌资可以用如下的 X_n 表示.

$$X_n = X_0 + \sum_{i=1}^n b_i Y_i \quad (5.1)$$

事实上, 我们可以证明

$$E[X_{n+1} | Y_1, \dots, Y_n] = X_n.$$

推理如下: 由式 (5.1) 我们可以得到

$$X_{n+1} = X_n + b_{n+1} Y_{n+1},$$

因此

$$\begin{aligned}
 E[X_{n+1}|Y_1, \dots, Y_n] &= E[X_n|Y_1, \dots, Y_n] + E[b_{n+1}Y_{n+1}|Y_1, \dots, Y_n] \\
 &= X_n + b_{n+1}E[Y_{n+1}|Y_1, \dots, Y_n] \\
 &\quad (\text{因为 } X_n \text{ 与 } b_{n+1} \text{ 由 } Y_1, \dots, Y_n \text{ 确定}) \\
 &= X_n + b_{n+1}E[Y_{n+1}] \\
 &\quad (\text{因为 } \{Y_n\} \text{ 是独立随机变量序列}) \\
 &= X_n \quad (\text{因为 } Y_{n+1} \text{ 为 } \pm 1 \text{ 的二项分布, 所以 } E[Y_{n+1}] = 0, \forall n \geq 0)
 \end{aligned}$$

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.



笔记 因此我们可以得到 $E[X_{n+1}] = E[X_n] = \dots = E[X_0]$

理解 5.1. 鞅的直观本质意义

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. 下面我们定义关于 σ 代数的鞅.



定义 5.1. 域流 Filtration

- 给定样本空间 Ω 和时间指标集 T , 若存在一族 σ -代数 $\{\mathcal{F}_t : t \in T\}$ 使得对任意的 $0 \leq s < t$ 都有 $\mathcal{F}_s \subset \mathcal{F}_t$, 则称 $\{\mathcal{F}_t : t \in T\}$ 为 Ω 上的一个域流 (filtration).
- 对于概率空间 $(\Omega, \mathcal{F}, P)$, 和 Ω 上的一个域流 $\{\mathcal{F}_t\}_{t \in T}$, 则称四元组 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$ 为一个带域流的概率空间, 或者过滤的 (filtered probability space) 概率空间.
- 域流的概念可以动态地刻画随机过程的可测性和信息递增.



笔记 我们考虑 $\Omega = \{\omega_1, \omega_2, \omega_3\}$, 那么在 $t = 0$ 时刻我们只能得到 $\mathcal{F}_0 = \{\phi, \Omega\}$, 而在 $t = 1$ 时刻, 我们假设 $\mathcal{F}_1$ 可能为 $\{\phi, \Omega, \{\omega_1\}, \{\omega_2, \omega_3\}\}$, 在 $t = 2$ 时刻则有 $\mathcal{F}_2 = \{\phi, \Omega, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}\}$, 因此我们有 $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2$, 所以 $\{\mathcal{F}_t : t \in T\}$ 为 Ω 上的一个域流.

定义 5.2. 适应过程 (adapted process)

- 给定一个带域流的概率空间 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$. 如果一个随机过程 $\{X_t\}_{t \in T}$ 满足对任意的 $t \in T$ 都有

$$X_t \in \mathcal{F}_t, \quad \text{亦即 } X_t \text{ 是 } \mathcal{F}_t \text{ 可测的, 或记为}$$

则称 $\{X_t\}$ 为关于域流 $\{\mathcal{F}_t\}_{t \in T}$ 适应的 (adapted) 过程.

- 适应过程在每一时间处都可以由当前已知信息来描述 (而不是给定).
- 本课程中的金融随机过程 (资产价格、投资组合头寸等) 都是适应的.



定义 5.3. 可料过程 (predictable process)

- 给定一个带域流的概率空间 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0,1,\dots,N}, P)$. 如果一个随机过程 $\{X_n\}_{n=0,1,\dots,N}$ 满足对任意的 n 都有

$$X_n \in \mathcal{F}_{n-1}, \quad \text{亦即 } X_n \text{ 是 } \mathcal{F}_{n-1} \text{ 可测的,}$$

则称 $\{X_n\}$ 为关于域流 $\{\mathcal{F}_n\}_{n=0,1,\dots,N}$ **可料** (predictable) 过程。

- 可料过程在每一时间处都可以由前一期信息来描述 (而不是给定)。
- 在离散时间框架下, 我们讨论的策略都是可料的。

**笔记 回忆条件数学期望的性质****定理 5.1**

若 X 为一随机变量, $\mathcal{G}$ 为 σ -代数。则:

- (1) $E[E[X|\mathcal{G}]] = E[X]$.
- (2) 若 X 是 $\mathcal{G}$ 可测, 则 $E[X|\mathcal{G}] = X, a.s.$.
- (3) 设 $\mathcal{G} = \{\emptyset, \Omega\}$, 则 $E[X|\mathcal{G}] = E[X], a.s.$.
- (7) $|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}], a.s.$.
- (9) 设 X 及 XY 的期望存在, 且 Y 为 $\mathcal{G}$ 可测, 则

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

- (10) 若 X 与 $\mathcal{G}$ 相互独立 (即 $\sigma(X)$ 与 $\mathcal{G}$ 相互独立), 则有

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

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

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

**定义 5.4. 离散时间鞅的定义**

设 $\{\mathcal{F}_n, n \geq 0\}$ 是一个 σ 代数流. 随机过程 $\{X_n, n \geq 0\}$ 称为关于 $\{\mathcal{F}_n, n \geq 0\}$ 的鞅, 如果:

- $\{X_n\}$ 是 $\{\mathcal{F}_n\}$ 适应的
- $E[|X_n|] < \infty$
- 并且 $\forall n \geq 0$, 有

$$E[X_{n+1}|\mathcal{F}_n] = X_n \quad (5.2)$$

- 适应列 $\{X_n, \mathcal{F}_n, n \geq 0\}$ 称为**下鞅**, 如果 $\forall n \geq 0$,

$$E[X_n^+] < \infty \quad \text{且} \quad E[X_{n+1}|\mathcal{F}_n] \geq X_n \quad (5.3)$$

- 上鞅**可以类似定义.



在给出例子之前, 先给出由定义直接推出的命题.

命题 5.1

- (1) 适应列 $\{X_n, \mathcal{F}_n, n \geq 0\}$ 是下鞅当且仅当 $\{-X_n, \mathcal{F}_n, n \geq 0\}$ 是上鞅.
- (2) 如果 $\{X_n, \mathcal{F}_n\}, \{Y_n, \mathcal{F}_n\}$ 是两个下鞅, a, b 是两个正常数, 则 $\{aX_n + bY_n, \mathcal{F}_n\}$ 是下鞅.
- (3) 如果 $\{X_n, \mathcal{F}_n\}, \{Y_n, \mathcal{F}_n\}$ 是两个下鞅 (或上鞅), 则 $\{\max(X_n, Y_n), \mathcal{F}_n\}$ ($\{\min(X_n, Y_n), \mathcal{F}_n\}$) 是下鞅 (上鞅).



笔记 如果我们做多步预测, 即求 $E[X_{n+k}|\mathcal{F}_n]$, 其中 $k \geq 1$, 则有

$$\begin{aligned}
 E[X_{n+k}|\mathcal{F}_n] &= E[E[X_{n+k}|\mathcal{F}_{n+k-1}|\mathcal{F}_n]] \\
 &= E[X_{n+k-1}|\mathcal{F}_n] \\
 &= E[E[X_{n+k-1}|\mathcal{F}_{n+k-2}|\mathcal{F}_n]] \\
 &= E[X_{n+k-2}|\mathcal{F}_n] \\
 &= \dots \\
 &= E[X_{n+1}|\mathcal{F}_n] \\
 &= X_n
 \end{aligned}$$

例 5.1 设 $X_1, X_2, \dots$ 是一族独立随机变量序列, 且 $E[|X_i|] < \infty$, 令 $S_0 = 0, S_n = \sum_{k=1}^n X_k$, 则:

- (a) 若对于所有的 $n, E(X_n) = 0$, 则 $\{S_n\}$ 是 (关于 $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$) 的鞅.
- (b) 若 $E(X_k) = \mu \neq 0$, 则 $\{M_n = S_n - n\mu\}$ 是 (关于 $\{\mathcal{F}_n\}$ 的) 鞅.

证明 (a) 当 $E[X_k] = 0, (k = 1, 2, \dots)$ 时, 易见 S_n 是 $\mathcal{F}_n$ 可测的, 而且 $E[|S_n|] \leq \sum_{i=1}^n E[|X_i|] < \infty$, 于是

$$\begin{aligned}
 E[S_{n+1}|\mathcal{F}_n] &= E[X_1 + X_2 + \dots + X_{n+1}|\mathcal{F}_n] \\
 &= E[X_1 + \dots + X_n|\mathcal{F}_n] + E[X_{n+1}|\mathcal{F}_n] \\
 &= S_n
 \end{aligned}$$

从而 $\{S_n\}$ 是一个关于 $\{\mathcal{F}_n\}$ 的鞅. 同理可以证明 (b). ■

例 5.2 在例 5.1 中设 $E[X_k] = \mu \neq 0, E[|X_k|] < \infty, (k = 1, 2, \dots)$, 则有 $E[|S_n|] < \infty$,

$$E[S_{n+1}|\mathcal{F}_n] = E\left[\sum_{i=1}^n X_i + X_{n+1}|\mathcal{F}_n\right] = S_n + \mu$$

显然, 若 $\mu > 0 (\mu < 0)$, 则 $\{S_n\}$ 是一关于 $\{\mathcal{F}_n\}$ 的下鞅 (上鞅).

例 5.3 (倍赌策略) 考虑一个公平博弈问题. 设 $X_1, X_2, \dots$ 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 $X_i (i = 1, 2, \dots)$ 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 $W_0 = 0$, 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 $P\{W_{n+1} = 1|W_n = 1\} = 1$.

下假设前 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 $E[W_{n+1}|\mathcal{F}_n] = W_n$, 这里 $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$, 从而 $\{W_n\}$ 是一个关于 $\{\mathcal{F}_n\}$ 的鞅.

笔记

下面我们给出 $E(W_{n+1}|\mathcal{F}_n) = W_n$ 的证明, 因为当我们对 W_{n+1} 预测时, 有用的信息只有 W_n , 因此实际上是要证明 $E(W_{n+1}|W_n) = W_n$, 即:

1. 如果第 n 次赢了, 则有 $E(W_{n+1}|W_n = 1) = 1 = W_n$
2. 如果前 n 次都输了, 则有 $E(W_{n+1}|W_n = -(2^n - 1)) = 1 \cdot \frac{1}{2} + (-(2^{n+1} - 1)) \cdot \frac{1}{2} = -(2^n - 1) = W_n$

由以上两点我们可以得到 $E(W_{n+1}|W_n) = W_n$.

当游戏终止时, 我们一定会赢得 1 块钱, 也就是说存在 T , 使得 $W_T = 1$, 而在鞅的平均意义上, 应该有 $E(W_{n+1}) = E(W_0) = 0$, 这是否相矛盾? 我们会在后续鞅的停时定理进行解释.

例 5.4 我们可以把例 5.3 再一般化. 设 $X_1, X_2, \cdots$ 仍如例 5.3 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 $\mathcal{F}_{n-1}$ 可测的 (设 B_1 为常数). 仍然令 W_n 同例 5.3 之定义, $W_0 = 0$, 则有

$$W_n = \sum_{j=1}^n B_j X_j$$

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 $\{\mathcal{F}_n\}$ 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 $\mathcal{F}_n$ 可测的, 并且

$$\begin{aligned} E[W_{n+1}|\mathcal{F}_n] &= E\left[\sum_{j=1}^{n+1} B_j X_j \middle| \mathcal{F}_n\right] \\ &= E\left[\sum_{j=1}^n B_j X_j \middle| \mathcal{F}_n\right] + E[B_{n+1} X_{n+1} | \mathcal{F}_n] \\ &= \sum_{j=1}^n B_j X_j + B_{n+1} E[X_{n+1} | \mathcal{F}_n] \\ &= W_n + B_{n+1} E[X_{n+1}] \\ &= W_n. \end{aligned}$$

例 5.5 (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法. 以 X_n 表示第 n 次抽取后坛子中的红球数. 则 $X_0 = 1$,

$$P\{X_{n+1} = k+1 | X_n = k\} = \frac{k}{n+2} \quad (5.4)$$

$$P\{X_{n+1} = k | X_n = k\} = \frac{n+2-k}{n+2} \quad (5.5)$$

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 我们首先计算 $E[X_{n+1} | X_n = k]$. 由(5.4)可以得到

$$\begin{aligned} E[X_{n+1} | X_n = k] &= (k+1)P\{X_{n+1} = k+1 | X_n = k\} + kP\{X_{n+1} = k | X_n = k\} \\ &= (k+1)\frac{k}{n+2} + k\frac{n+2-k}{n+2} = k + \frac{k}{n+2}. \end{aligned}$$

由此得到,

$$E[X_{n+1} | X_n] = X_n + \frac{X_n}{n+2}$$

由于第 $n+1$ 次抽球后的结果, 只和前 n 抽完后坛子里面红、黄球的数量有关. 显然 $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ 中对 X_{n+1} 有影响的信息都包含在 X_n 中 (这个性质我们后面称为 Markov 性), 所以

$$\begin{aligned} E[M_{n+1} | \mathcal{F}_n] &= E[M_{n+1} | X_n] \\ &= E\left[\frac{X_{n+1}}{n+1+2} | X_n\right] \\ &= \frac{1}{n+3} E[X_{n+1} | X_n] \\ &= \frac{1}{n+3} \left(X_n + \frac{X_n}{n+2}\right) \\ &= \frac{X_n}{n+2} = M_n. \end{aligned}$$

本例研究的模型是 Polya 首次引入的, 它适用于描述群体增值和传染病的传播等现象.

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.

定义 5.5. 凸函数

考虑定义在有穷域无穷开区间 I 上的函数 $\varphi(x)$, 称它为凸的, 若 $\forall x, y \in I, 0 < \alpha < 1$, 有

$$\alpha\varphi(x) + (1-\alpha)\varphi(y) \geq \varphi[\alpha x + (1-\alpha)y] \quad (5.6)$$

成立.



引理 5.1. 条件 Jensen 不等式

设 $\varphi(x)$ 为实数集 $\mathbb{R}$ 上的凸函数, 随机变量 M 满足

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

则有

$$E[\varphi(M)|\mathcal{F}_n] \geq \varphi[E[M|\mathcal{F}_n]] \quad (5.7)$$

其中 $\mathcal{F}_n$ 是任意递增的 σ 代数列.



式 (5.7) 是著名的条件 Jensen 不等式, 由它即可得到以下定理.

定理 5.2

设 $\{M_n, n \geq 0\}$ 是关于 $\{\mathcal{F}_n, n \geq 0\}$ 的鞅 (下鞅), $\varphi(x)$ 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 $\{\mathcal{F}_n, n \geq 0\}$ 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.

证明: 根据定义易证, 请读者利用引理 5.1 推导.



5.2 鞅的停时定理及其应用

5.2.1 鞅的停时定理

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 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

$$E[M_n] = E[M_0] \quad (5.8)$$

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

$$E[M_T] = E[M_0] \quad (5.9)$$

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 $\{M_n, n \geq 0\}$ 是一种公平的博弈, M_n 表示局中人第 n 次赌局结束后的赌本. 式 (5.8) 说明他在每次赌局结束时的赌本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 (5.9) 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 5.3 中的赌博者采取的策略, 就可以保证他在赢一元之后结束, 所以我们要为式 (5.9) 的成立附加一些条件.

定义 5.6. (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 $\{\mathcal{F}_n, n \geq 0\}$ 的停时, 如果 T 在 $\{0, 1, 2, \dots, \infty\}$ 中取值, 且对每个 $n \geq 0$, $\{T = n\} \in \mathcal{F}_n$.



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. 仍然回到公平博弈的例子, 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, 如果我下次要输我现在就停止赌博, 这是对停止时刻 T 的第一个要求: 它须是一个停时.

以下看几个停时的例子.

例 5.6 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 5.7 (首次时) $\{X_n, n \geq 0\}$ 是一个随机变量序列, A 是一个事件集, 令

$$T(A) = \inf\{n, X_n \in A\}, \text{ 并约定 } T(\emptyset) = \inf\{n, X_n \in \emptyset\} = \infty,$$

可见 $T(A)$ 是 $\{X_n, n \geq 0\}$ 首次进入 A (即发生了 A 中所含事件) 的时刻, 称 $T(A)$ 是 $\{X_n, n \geq 0\}$ 到集合 A 的首达时. 则: $T(A)$ 是关于 $\{X_n, n \geq 0\}$ 的停时.

证明 事实上

$$\{T(A) = n\} = \{X_0 \notin A, X_1 \notin A, \dots, X_{n-1} \notin A, X_n \in A\}$$

显然 $\{T(A) = n\}$ 完全由 $X_0, X_1, \dots, X_n$ 决定, 从而 $T(A)$ 是关于 $\{X_n, n \geq 0\}$ 的停时.

例 5.8 如果 T 和 S 是两个停时, 则 $T + S, \min(T, S)$ 和 $\max(T, S)$ 也是停时. 这可由下述简单的命题来证明.

命题 5.2

设 T 是取值于 $\{0, 1, 2, \dots, \infty\}$ 的随机变量, 则下述三者等价

- (1) $\{T = n\} \in \sigma(X_0, X_1, \dots, X_n)$;
- (2) $\{T \leq n\} \in \sigma(X_0, X_1, \dots, X_n)$;
- (3) $\{T > n\} \in \sigma(X_0, X_1, \dots, X_n)$.



只要注意到如下等式, 即可证明 (1), (2), (3) 的等价性.

$$\begin{aligned} \{T \leq n\} &= \bigcup_{k=0}^n \{T = k\} \\ \{T > n\} &= \Omega - \{T \leq n\} \\ \{T = n\} &= \{T \leq n\} - \{T \leq n-1\} \end{aligned}$$

这就可以证明 $T + S, \max(T, S)$ 和 $\min(T, S)$ 是停时. 特别地, 由例 5.6 可知, 常数 n 是停时. 设 T 是停时, 令 $T_n = \min\{T, n\}$, 则每个 T_n 都是停时, 并且有 $T_0 \leq T_1 \leq T_2 \leq \dots \leq T_n \leq n, \forall n$.

命题 5.3. 停时定理

设 $\{M_n, n \geq 0\}$ 是一个关于 $\{\mathcal{F}_n\}_{n=0,1,\dots,K}$ 的鞅, T 是一个关于 $\{\mathcal{F}_n\}$ 的停时, 且 $T \leq K$. 则

$$E[M_T | \mathcal{F}_0] = M_0$$

特别地,

$$E[M_T] = E[M_0]$$

停时定理更一般的形式为:

$$E[M_T | \mathcal{F}_n] = M_{n \wedge T} \quad (5.10)$$

这儿, $n \wedge T$ 表示 $\min(n, T)$.

证明 由于 $T \leq K$, 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

$$M_T = \sum_{j=0}^K M_j I_{\{T=j\}} \quad (5.11)$$

对式 (5.11) 关于 $\mathcal{F}_{K-1}$ 取条件期望, 有

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

当 $j \leq K-1$ 时, M_j 和 $I_{\{T=j\}}$ 都是 $\mathcal{F}_{K-1}$ 可测的, 从而

$$E\left[\sum_{j=0}^{K-1} M_j I_{\{T=j\}} | \mathcal{F}_{K-1}\right] = \sum_{j=0}^{K-1} M_j I_{\{T=j\}}$$

又因为 $T \leq K$ 已知, 则 $\{T = K\}$ 与 $\{T > K-1\}$ 是等价的, 由例 5.8 中结果知 $\{T > K-1\} \in \sigma(X_0, \dots, X_{K-1})$, 因此

$$\begin{aligned} E[M_K I_{\{T=K\}} | \mathcal{F}_{K-1}] &= I_{\{T > K-1\}} E[M_K | \mathcal{F}_{K-1}] \\ &= I_{\{T > K-1\}} M_{K-1} \end{aligned}$$

从而

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

重复以上运算, 关于 $\mathcal{F}_{K-2}$ 取条件期望, 我们可以得到

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

继续这样的过程, 最终有

$$E[M_T | \mathcal{F}_0] = I_{\{T \geq 0\}} M_0 = M_0$$

定理 5.3. 鞅停时定理

设 $\{M_n, n \geq 0\}$ 是一个关于 $\{\mathcal{F}_n\}$ 的鞅, T 是停时且满足

- (1) $P\{T < \infty\} = 1$
- (2) $E[|M_T|] < \infty$
- (3) $\lim_{n \rightarrow \infty} E[|M_n| I_{\{T > n\}}] = 0$.

则有

$$E[M_T] = E[M_0]$$



笔记 我们来看鞅的停时定理的一个反例, 主要针对上述需要满足条件的第 (3) 条: 对于 $E[M_n I_{\{T > n\}}]$, 如果赢的概率减小得越快, 则趋近于 0; 如果输钱的数量增长的速度越快, 则不一定趋近于 0. 对于前面提到的倍赌策略, 可以看到:

$$\begin{aligned} E[M_n I_{\{T > n\}}] &= E[(2^n - 1) I_{\{T > n\}}] \\ &= (2^n - 1) P(T > n) \\ &= (2^n - 1) \frac{1}{2^n} \rightarrow 1 \neq 0 \end{aligned}$$

此时不满足条件 (3).

例 5.9 设 $\{X_n\}$ 是在 $\{0, 1, \dots, N\}$ 上的简单随机游动 ($p = \frac{1}{2}$), 并且 0 和 N 为两个吸收壁. 设 $X_0 = a$. 即当 X 触及到 0 或者 N 时, 就停止不动了. 计算: X 被 N 吸收的概率.

证明 容易验证 $\{X_n\}$ 是一个鞅. 令

$$T = \min\{j : X_j = 0 \text{ 或 } N\}$$

, 则 T 是一个停时. 由于 X_n 的取值有界, 故定理 5.2.1 中的条件成立. 从而

$$E[X_T] = E[X_0] = a$$

由于此时 X_T 只取两个值 $N, 0$,

$$E[X_T] = N \cdot P\{X_T = N\} + 0 \cdot P\{X_T = 0\}$$

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 a/N .



笔记 我们也可以利用初等概率论的方法来求解这道题. 我们用 P_a 表示上问中从 a 点出发到达 N 的概率. 根据全概率公式可以得到

$$P_a = \frac{1}{2} \cdot P_{a+1} + \frac{1}{2} \cdot P_{a-1}, \text{ 其中 } P_0 = 0, \quad P_N = 1$$

$$2P_a = P_{a+1} + P_{a-1}$$

$$P_a - P_{a-1} = P_{a+1} - P_a$$

显然, P_a 是一个等差数列, 所以可以得到

$$P_a = \frac{1}{N} \cdot a$$



笔记 思考题: 如果 X 的吸收壁不是 0 和 N 而是 N_1 和 N_2 ($N_1 < a < N_2$). 计算 X 被 N_1 和 N_2 吸收的概率.

5.3 连续鞅

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. 同时利用它们证明 Lundberg-Cramer 破产理论中的 Lundberg 不等式.

首先给出鞅的定义. 设 $(\Omega, \mathcal{F}, P)$ 是一个完备的概率空间; $(\mathcal{F}_t)_{t \geq 0}$ 为一个非降的 $\mathcal{F}$ 的子 σ 代数流. 本书中所涉及的子 σ 代数流都是非降的, 因此以后我们就将“非降的”省略, 简称为子 σ 代数流.

定义 5.7. 连续时间鞅

随机过程 $\{X(t), t \geq 0\}$ (简记为 $\{X_t\}$) 称为 $\{\mathcal{F}_t\}$ 适应的, 如果对每个 $t \geq 0$, X_t 为 $\mathcal{F}_t$ 可测. 一个过程 $\{X_t\}$ 称为关于 $\{\mathcal{F}_t\}$ 的鞅, 如果:

- (1) $\{X_t\}$ 为 $\{\mathcal{F}_t\}$ 适应的过程,
- (2) 每个 X_t 可积 (即 $E[|X_t|] < \infty$)
- (3) 且对一切 $0 \leq s < t$, 有

$$E[X_t | \mathcal{F}_s] = X_s, \quad a.s. \quad (5.12)$$

- 若随机过程 $\{X_t, t \geq 0\}$ 是鞅, 则对 $t > 0$, 有

$$E[X_t] = E[E[X_t | \mathcal{F}_0]] = E[X_0] \quad (5.13)$$

例 5.10 设 ξ 为一可积随机变量 ($E(|\xi|) < +\infty$). 则过程 $X_t := E(\xi | \mathcal{F}_t)$ 为 $(\mathcal{F}_t)$ 的鞅。

证明 对于任意的 $s \leq t$, 由定义可知:

1. 根据条件数学期望的定义显然有 $X_t \in \mathcal{F}_t$
2. $E|X_t| = E|E(\xi | \mathcal{F}_t)| \leq E[E(|\xi| | \mathcal{F}_t)] = E(|\xi|) < +\infty$
3. $E(X_t | \mathcal{F}_s) = E(E(\xi | \mathcal{F}_t) | \mathcal{F}_s) = E(\xi | \mathcal{F}_s) = X_s$

因此过程 $X_t := E(\xi | \mathcal{F}_t)$ 为 $(\mathcal{F}_t)$ 的鞅。 ■

例 5.11 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 $\{X_t, t \geq 0\}$ 是一个关于自身信息流 $(\mathcal{F}_t)_{t \geq 0}$ 鞅. 其中 $\mathcal{F}_t = \sigma(X_r, 0 \leq r \leq t)$ 表示从 $[0, t]$ 上随机过程所产生的信息。

证明 事实上

$$E[|X_t|] = |X_0| E[e^{Y_t}] = |X_0| < +\infty.$$

再对 $0 \leq s < t$, 有

$$\begin{aligned} E[X_t | \mathcal{F}_s] &= E[X_s e^{Y_t - Y_s} | \mathcal{F}_s] \\ &= X_s E[e^{Y_t - Y_s}], \quad \text{用到了 } Y \text{ 的独立增量性质} \\ &= X_s E[e^{Y_{t-s}}] = X_s, \quad \text{用到了 } Y \text{ 的平稳增量性质} \end{aligned}$$

定义 5.8. 连续时间停时

随机变量 τ (即 $\tau: \Omega \rightarrow [0, \infty]$) 称为 $\mathcal{F}_t$ 停时, 如果 $P\{\tau < \infty\} = 1$, 并且对一切 $t \geq 0$, $\{\tau \leq t\}$ 是 $\mathcal{F}_t$ 可测的, 即

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

特别地, 当 $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$ 时, τ 称为关于随机过程 $\{X_t, t \geq 0\}$ 的停时. 若存在常数 $k > 0$ 使得 $P\{\tau \leq k\} = 1$, 则称 τ 为有界停时.



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 (5.13) 中的 t 换成停时 τ 时, 等式仍然成立.

定理 5.4. 连续时间停时定理

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 5.5. 连续时间鞅的收敛定理

设 $\{X_t, t \geq 0\}$ 是一个鞅并且 $X_t \geq 0, \forall t \geq 0$ (简称为非负鞅), 则存在几乎处处收敛的有限极限, 即有

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

