基础随机过程

Basic Stochastic Processes

预览版

原著作者 Zdzisław Brzeźniak, Tomasz Zastawniak（英国）

中文翻译 临江仙

更新发布 https://afdian.com/a/solitairemiya

更新日期 2026 年 4 月 5 日

惊鸿舟数海，一笔临江仙。

目录

第一章 概率论回顾

在本章中回顾概率论中的一些基本概念和结果. 以下是需要回顾的简要清单:

(1) 概率空间、σ-域和测度;

(2) 随机变量及其分布;

(3) 期望和方差;

(4) 由随机变量生成的 σ-域;

(5) 独立性、条件概率.

建议读者查阅概率论教材以获取更多细节.

第 1 节　事件与概率

定义 1.1. 设 Ω 是一个非空集合. $\mathcal{F}$ 是 Ω 上的一个 σ-域，如果 $\mathcal{F}$ 是 Ω 的子集族且满足

(1) 空集 $\emptyset$ 属于 $\mathcal{F}$;

(2) 若 A 属于 $\mathcal{F}$，则其补集 $\Omega \setminus A$ 也属于 $\mathcal{F}$;

(3) 若 $A_1, A_2, \ldots$ 是 $\mathcal{F}$ 中的一列集合，则它们的并集 $A_1 \cup A_2 \cup \cdots$ 也属于 $\mathcal{F}$.

例 1.1. 在本课程中，$\mathbb{R}$ 表示实数集. *Borel* 集族 $\mathcal{F} = \mathcal{B}(\mathbb{R})$ 是 $\mathbb{R}$ 上的一个 σ-域. 回顾 $\mathcal{B}(\mathbb{R})$ 是包含 $\mathbb{R}$ 中所有区间的最小 σ-域.

定义 1.2. 设 $\mathcal{F}$ 是 Ω 上的一个 σ-域. 概率测度 P 是一个函数

$$P : \mathcal{F} \to [0, 1]$$

满足

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

(2) 若 $A_1, A_2, \ldots$ 是两两不交的集合（即当 $i \neq j$ 时，$A_i \cap A_j = \varnothing$）且属于 $\mathcal{F}$，则

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots.$$

三元组 $(\Omega, \mathcal{F}, P)$ 称为概率空间. 属于 $\mathcal{F}$ 的集合称为事件. 当 $P(A) = 1$ 时，称事件 A 几乎必然（*a.s.*）发生.

例 1.2. 取单位区间 $\Omega = [0,1]$，其上的 σ-域 $\mathcal{F} = \mathcal{B}([0,1])$ 由 $[0,1]$ 中的 *Borel* 集 $B \subset [0,1]$ 组成，概率测度为 $P = \mathrm{Leb}$（*Lebesgue* 测度）. 则 $(\Omega, \mathcal{F}, P)$ 是一个概率空间. 回顾 *Lebesgue* 测度是定义在 *Borel* 集上的唯一测度，满足对任意区间 $[a,b]$ 有

$$\mathrm{Leb}[a,b] = b - a$$

（事实上 *Lebesgue* 测度可以延拓到更大的 σ-域，但这里只需 *Borel* 集）.

练习 1.1. 证明：若 $A_1, A_2, \ldots$ 是一个递增序列的事件，即

$$A_1 \subset A_2 \subset \cdots,$$

则

$$P(A_1 \cup A_2 \cup \cdots) = \lim_{n\to\infty} P(A_n).$$

类似地，若 $A_1, A_2, \ldots$ 是一个递减序列的事件，即

$$A_1 \supset A_2 \supset \cdots,$$

则

$$P(A_1 \cap A_2 \cap \cdots) = \lim_{n\to\infty} P(A_n).$$

提示 将 $A_1 \cup A_2 \cup \cdots$ 写成一列不交事件的并：从 A_1 开始，然后添加一个不交的集合得到 $A_1 \cup A_2$，再添加一个不交的集合得到 $A_1 \cup A_2 \cup A_3$，依此类推. 这样就得到一列不交集合，可以使用概率测度的定义. 对于交集 $A_1 \cap A_2 \cap \cdots$，利用 *De Morgan* 法则将其写成某些事件的并.

引理 1.1（Borel-Cantelli）. 设 $A_1, A_2, \ldots$ 是一列事件，满足 $P(A_1) + P(A_2) + \cdots < \infty$，且设 $B_n = A_n \cup A_{n+1} \cup \cdots$. 则 $P(B_1 \cap B_2 \cap \cdots) = 0$.

练习 1.2. 证明上述 *Borel-Cantelli* 引理.

提示 $B_1, B_2, \ldots$ 是一个递减序列的事件.

第 2 节 随机变量

定义 1.3. 若 $\mathcal{F}$ 是 Ω 上的一个 σ-域，则函数 $\xi : \Omega \to \mathbb{R}$ 称为 $\mathcal{F}$-可测的，如果对每个 *Borel* 集 $B \in \mathcal{B}(\mathbb{R})$ 都有

$$\{\xi \in B\} \in \mathcal{F}.$$

若 $(\Omega, \mathcal{F}, P)$ 是概率空间，则这样的函数 ξ 称为随机变量.

注记 1.1. 对于事件 $\{\xi \in B\}$ 这类记号，下面使用简写形式以避免混乱. 精确地说，应该用

$$\{\omega \in \Omega : \xi(\omega) \in B\}$$

代替 $\{\xi \in B\}$. 顺便指出，$\{\xi \in B\}$ 只是逆像 $\xi^{-1}(B)$ 的一种方便的写法.

定义 1.4. 由随机变量 $\xi : \Omega \to \mathbb{R}$ 生成的 σ-域 $\sigma(\xi)$ 由所有形如 $\{\xi \in B\}$ 的集合组成，其中 B 是 $\mathbb{R}$ 中的 *Borel* 集.

定义 1.5. 由一族随机变量 $\{\xi_i : i \in I\}$ 生成的 σ-域 $\sigma\{\xi_i : i \in I\}$ 定义为包含所有形如 $\{\xi_i \in B\}$ 的事件的最小 σ-域，其中 B 是 $\mathbb{R}$ 中的 *Borel* 集且 $i \in I$.

练习 1.3. 称 $f : \mathbb{R} \to \mathbb{R}$ 为 *Borel* 函数，如果对任意 *Borel* 集 $B \subset \mathbb{R}$，其逆像 $f^{-1}(B)$ 也是 *Borel* 集. 证明：若 f 是 *Borel* 函数且 ξ 是随机变量，则复合函数 $f(\xi)$ 是 $\sigma(\xi)$-可测的.

提示 考虑事件 $\{f(\xi) \in B\}$，其中 B 是任意 *Borel* 集. 这个事件能否写成 $\{\xi \in A\}$ 的形式，其中 A 是某个 *Borel* 集？

引理 1.2（Doob-Dynkin）. 设 ξ 是随机变量. 则每个 $\sigma(\xi)$-可测的随机变量 η 都可以写成

$$\eta = f(\xi)$$

的形式，其中 $f : \mathbb{R} \to \mathbb{R}$ 是某个 *Borel* 函数.

这个高度非平凡的结果的证明在此省略.

定义 1.6. 每个随机变量 $\xi : \Omega \to \mathbb{R}$ 都产生一个概率测度

$$P_\xi(B) = P\{\xi \in B\}$$

定义在 $\mathbb{R}$ 上，定义域为 *Borel* 集 $B \in \mathcal{B}(\mathbb{R})$ 的 σ-域. 称 P_ξ 为 ξ 的分布. 函数 $F_\xi : \mathbb{R} \to [0, 1]$ 定义为

$$F_\xi(x) = P\{\xi \leqslant x\}$$

称为 ξ 的分布函数.

练习 1.4. 证明分布函数 F_ξ 是单调递增的、右连续的，且

$$\lim_{x\to-\infty} F_\xi(x)=0,\quad \lim_{x\to+\infty} F_\xi(x)=1.$$

提示 例如，为验证右连续性，证明对任意递减序列 x_n 满足 $x_n\to x$，有 $F_\xi(x_n)\to F_\xi(x)$. 练习 *1.1* 的结果可能很有用.

定义 1.7. 若存在 *Borel* 函数 $f_\xi:\mathbb{R}\to\mathbb{R}$ 使得对任意 *Borel* 集 $B\subset\mathbb{R}$ 有

$$P\{\xi\in B\}=\int_B f_\xi(x)\,\mathrm{d}x,$$

则称 ξ 是具有绝对连续分布的随机变量，f_ξ 称为 ξ 的密度函数. 若存在一列（有限或无限）两两不同的实数 $x_1,x_2,\ldots$ 使得对任意 *Borel* 集 $B\subset\mathbb{R}$ 有

$$P\{\xi\in B\}=\sum_{x_i\in B}P\{\xi=x_i\},$$

则称 ξ 具有离散分布，其取值为 $x_1,x_2,\ldots$，在 x_i 处的质量为 $P\{\xi=x_i\}$.

练习 1.5. 设 ξ 具有密度为 f_ξ 的连续分布. 证明若 f_ξ 在 x 处连续，则

$$\frac{\mathrm{d}}{\mathrm{d}x}F_\xi(x)=f_\xi(x).$$

提示 将 $F_\xi(x)$ 表示为 f_ξ 的积分.

练习 1.6. 证明：若 ξ 具有离散分布，取值为 $x_1,x_2,\ldots$，则 F_ξ 在每个不含任何 x_i 的区间 $(s,t]$ 上为常数，且在每个 x_i 处有大小为 $P\{\xi=x_i\}$ 的跳跃.

提示 增量 $F_\xi(t)-F_\xi(s)$ 等于落在区间 $[s,t)$ 中的 x_i 的总质量.

定义 1.8. 若干随机变量 $\xi_1,\ldots,\xi_n$ 的联合分布是 $\mathbb{R}^n$ 上的概率测度 $P_{\xi_1,\ldots,\xi_n}$，满足

$$P_{\xi_1,\ldots,\xi_n}(B)=P\left\{(\xi_1,\ldots,\xi_n)\in B\right\}$$

对任意 $\mathbb{R}^n$ 中的 *Borel* 集 B. 若存在 *Borel* 函数 $f_{\xi_1,\ldots,\xi_n}:\mathbb{R}^n\to\mathbb{R}$ 使得

$$P\left\{(\xi_1,\ldots,\xi_n)\in B\right\}=\int_B f_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n$$

对任意 $\mathbb{R}^n$ 中的 *Borel* 集 B，则称 $f_{\xi_1,\ldots,\xi_n}$ 为 $\xi_1,\ldots,\xi_n$ 的联合密度.

定义 1.9. 随机变量 $\xi:\Omega\to\mathbb{R}$ 称为可积的，如果

$$\int_\Omega |\xi|\,\mathrm{d}P<\infty.$$

此时

$$E(\xi)=\int_\Omega \xi\,\mathrm{d}P$$

存在, 称为 ξ 的期望. 可积随机变量 $\xi:\Omega\to\mathbb{R}$ 的全体记为 L^1, 或在可能有歧义时记为 $L^1(\Omega,\mathcal{F},P)$.

例 1.3. 集合 A 的指示函数 $\mathbf{1}_A$ 在 A 上等于 1，在 A 的补集 $\Omega\setminus A$ 上等于 0. 对任意事件 A

$$E\left(\mathbf{1}_A\right)=\int_\Omega \mathbf{1}_A\,\mathrm{d}P=P(A).$$

称 $\eta:\Omega\to\mathbb{R}$ 为阶梯函数，如果

$$\eta=\sum_{i=1}^n \eta_i\mathbf{1}_{A_i},$$

其中 $\eta_1,\dots,\eta_n$ 是实数且 $A_1,\dots,A_n$ 是两两不交的事件. 则

$$E(\eta)=\int_\Omega \eta\,\mathrm{d}P=\sum_{i=1}^n\eta_i\int_\Omega \mathbf{1}_{A_i}\,\mathrm{d}P=\sum_{i=1}^n\eta_iP(A_i).$$

练习 1.7. 证明对任意使得 $h(\xi)$ 可积的 *Borel* 函数 $h:\mathbb{R}\to\mathbb{R}$ 有

$$E(h(\xi))=\int_{\mathbb{R}} h(x)\,\mathrm{d}P_\xi(x).$$

提示 首先对阶梯函数 $h:\mathbb{R}\to\mathbb{R}$ 验证等式成立，然后对非负函数通过阶梯函数逼近，利用积分的单调收敛性，最后对任意 *Borel* 函数通过分解为正部和负部 $h=h^+-h^-$ 得到结果.

特别地，练习 1.7 表明若 ξ 具有密度为 f_ξ 的绝对连续分布，则

$$E(h(\xi))=\int_{-\infty}^{+\infty}h(x)f_\xi(x)\,\mathrm{d}x.$$

若 ξ 具有离散分布，其（有限或无限个）两两不同的取值为 $x_1,x_2,\dots$，则

$$E\left(h(\xi)\right)=\sum_i h\left(x_i\right)P\left\{\xi=x_i\right\}.$$

定义 1.10. 随机变量 $\xi:\Omega\to\mathbb{R}$ 称为平方可积的，如果

$$\int_\Omega |\xi|^2\,\mathrm{d}P<\infty.$$

此时 ξ 的方差可以定义为

$$\mathrm{var}(\xi)=\int_\Omega(\xi-E(\xi))^2\,\mathrm{d}P.$$

平方可积随机变量 $\xi:\Omega\to\mathbb{R}$ 的全体记为 $L^2(\Omega,\mathcal{F},P)$，或在无歧义时简记为 L^2.

注记 1.2. 下面的练习 *1.8* 的结果表明可以在方差的定义中写入 $E(\xi)$.

练习 1.8. 证明：若 ξ 是平方可积随机变量，则它是可积的.

提示 使用 *Schwarz* 不等式（见式 *(1.1)*）并适当选择 η.

$$[E(\xi\eta)]^2\leqslant E(\xi^2)E(\eta^2).$$

练习 1.9. 证明：若 $\eta:\Omega\to[0,\infty)$ 是非负平方可积随机变量，则

$$E(\eta^2)=2\int_0^\infty tP(\eta>t)\,\mathrm{d}t.$$

提示 用 η 的分布函数 $F_\eta(t)$ 表示 $E(\eta^2)$，然后分部积分.

第 3 节　条件概率与独立性

定义 1.11. 对任意满足 $P(B)\neq 0$ 的事件 $A,B\in\mathcal{F}$，给定 B 时 A 的条件概率定义为

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

练习 1.10. 证明全概率公式

$$P(A)=P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+\cdots$$

对任意事件 $A\in\mathcal{F}$ 和任意满足 $B_1\cup B_2\cup\cdots=\Omega$ 且对所有 n 有 $P(B_n)\neq 0$ 的两两不交事件列 $B_1,B_2,\ldots\in\mathcal{F}$ 成立.

提示 $A=(A\cap B_1)\cup(A\cap B_2)\cup\cdots$.

定义 1.12. 两个事件 $A,B\in\mathcal{F}$ 称为独立的，如果

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

一般地，称 n 个事件 $A_1, \ldots, A_n \in \mathcal{F}$ 是独立的，如果对任意指标 $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ 有

$$P\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right) = P\left(A_{i_1}\right) P\left(A_{i_2}\right) \cdots P\left(A_{i_k}\right).$$

练习 1.11. 设 $P(B) \neq 0$. 证明 A 和 B 是独立事件当且仅当 $P(A|B) = P(A)$.

提示 若 $P(B) \neq 0$，则可以除以它.

定义 1.13. 两个随机变量 ξ 和 η 称为独立的，如果对任意 *Borel* 集 $A, B \in \mathcal{B}(\mathbb{R})$，两个事件

$$\{\xi \in A\} \quad 和 \quad \{\eta \in B\}$$

是独立的. 称 n 个随机变量 $\xi_1, \ldots, \xi_n$ 是独立的，如果对任意 *Borel* 集 $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$，事件

$$\{\xi_1 \in B_1\}, \ldots, \{\xi_n \in B_n\}$$

是独立的. 一般地，一族（有限或无限）随机变量称为独立的，如果这族中任意有限个随机变量都是独立的.

命题 1.1. 若两个可积随机变量 $\xi, \eta : \Omega \to \mathbb{R}$ 是独立的，则它们是不相关的，即

$$E(\xi\eta) = E(\xi)E(\eta),$$

前提是乘积 $\xi\eta$ 也是可积的. 若 $\xi_1, \ldots, \xi_n : \Omega \to \mathbb{R}$ 是独立的可积随机变量，则

$$E\left(\xi_1 \xi_2 \cdots \xi_n\right) = E\left(\xi_1\right) E\left(\xi_2\right) \cdots E\left(\xi_n\right),$$

前提是乘积 $\xi_1 \xi_2 \cdots \xi_n$ 也是可积的.

定义 1.14. 两个包含于 $\mathcal{F}$ 的 σ-域 $\mathcal{G}$ 和 $\mathcal{H}$ 称为独立的，如果任意两个事件

$$A \in \mathcal{G} \quad 和 \quad B \in \mathcal{H}$$

是独立的. 类似地，包含于 $\mathcal{F}$ 的有限个 σ-域 $\mathcal{G}_1, \ldots, \mathcal{G}_n$ 称为独立的，如果任意 n 个事件

$$A_1 \in \mathcal{G}_1, \ldots, A_n \in \mathcal{G}_n$$

是独立的. 一般地，一族（有限或无限）σ-域称为独立的，如果其中任意有限个都是独立的.

练习 1.12. 证明两个随机变量 ξ 和 η 是独立的当且仅当由它们生成的 σ-域 $\sigma(\xi)$ 和 $\sigma(\eta)$ 是独立的.

提示 $\sigma(\xi)$ 和 $\sigma(\eta)$ 中的事件分别形如 $\{\xi \in A\}$ 和 $\{\eta \in B\}$，其中 A 和 B 是 *Borel* 集.

有时方便起见，会讨论随机变量和 σ-域的独立性组合.

定义 1.15. 称随机变量 ξ 独立于 σ-域 $\mathcal{G}$，如果 σ-域

$$\sigma(\xi) \quad \text{和} \quad \mathcal{G}$$

是独立的. 这可以推广到任意（有限或无限）由随机变量或 σ-域或两者组合而成的族. 具体地说，这样一个族称为独立的，如果对族中任意有限个随机变量 $\xi_1, \ldots, \xi_m$ 和 σ-域 $\mathcal{G}_1, \ldots, \mathcal{G}_n$，$\sigma$-域

$$\sigma(\xi_1), \ldots, \sigma(\xi_m), \mathcal{G}_1, \ldots, \mathcal{G}_n$$

是独立的.

第 4 节　解答

解答 1.1. 若 $A_1 \subset A_2 \subset \cdots$，则

$$A_1 \cup A_2 \cup \cdots = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \cdots,$$

其中集合 $A_1, A_2 \setminus A_1, A_3 \setminus A_2, \ldots$ 是两两不交的. 因此，由概率测度的定义

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots) &= P(A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \cdots) \\ &= P(A_1) + P(A_2 \setminus A_1) + P(A_3 \setminus A_2) + \cdots \\ &= \lim_{n \to \infty} P(A_n). \end{aligned}$$

最后一个等式成立是因为上述级数的部分和为

$$\begin{aligned} P(A_1) + P(A_2 \setminus A_1) + \cdots + P(A_n \setminus A_{n-1}) &= P(A_1 \cup \cdots \cup A_n) \\ &= P(A_n). \end{aligned}$$

若 $A_1 \supset A_2 \supset \cdots$，则等式

$$P(A_1 \cap A_2 \cap \cdots) = \lim_{n \to \infty} P(A_n)$$

通过对 A_n 取补集并应用 *De Morgan* 法则得到

$$\Omega \setminus (A_1 \cap A_2 \cap \cdots) = (\Omega \setminus A_1) \cup (\Omega \setminus A_2) \cup \cdots.$$

解答 1.2. 由于 B_n 是递减序列的事件，由练习 *1.1* 的结果可知

$$\begin{aligned} P\left(B_1 \cap B_2 \cap \cdots\right) &= \lim_{n\to\infty} P\left(B_n\right) \\ &= \lim_{n\to\infty} P(A_n \cup A_{n+1} \cup \cdots) \\ &\leqslant \lim_{n\to\infty} \left(P(A_n) + P(A_{n+1}) + \cdots\right) \\ &= 0. \end{aligned}$$

最后一个等式成立是因为级数 $\sum_{n=1}^{\infty} P(A_n)$ 是收敛的. 上述不等式成立是因为次可加性

$$P\left(A_n \cup A_{n+1} \cup \cdots\right) \leqslant P\left(A_n\right) + P\left(A_{n+1}\right) + \cdots .$$

因此 $P\left(B_1 \cap B_2 \cap \cdots\right) = 0$.

解答 1.3. 若 B 是 $\mathbb{R}$ 中的 *Borel* 集且 $f : \mathbb{R} \to \mathbb{R}$ 是 *Borel* 函数，则 $f^{-1}(B)$ 也是 *Borel* 集. 因此

$$\{f(\xi) \in B\} = \{\xi \in f^{-1}(B)\}$$

属于由 ξ 生成的 σ-域 $\sigma(\xi)$. 由此可知复合函数 $f(\xi)$ 是 $\sigma(\xi)$-可测的.

解答 1.4. 若 $x \leqslant y$，则 $\{\xi \leqslant x\} \subset \{\xi \leqslant y\}$，所以

$$F_\xi(x) = P\{\xi \leqslant x\} \leqslant P\{\xi \leqslant y\} = F_\xi(y).$$

这意味着 F_ξ 是单调递增的.

接下来，任取序列 $x_1 \geqslant x_2 \geqslant \cdots$ 并设 $\lim_{n\to\infty} x_n = x$. 则事件

$$\{\xi \leqslant x_1\} \supset \{\xi \leqslant x_2\} \supset \cdots$$

构成一个递减序列，其交集为

$$\{\xi \leqslant x\} = \{\xi \leqslant x_1\} \cap \{\xi \leqslant x_2\} \cap \cdots .$$

由练习 *1.1* 可知

$$F_\xi(x) = P\{\xi \leqslant x\} = \lim_{n\to\infty} P\{\xi \leqslant x_n\} = \lim_{n\to\infty} F_\xi(x_n).$$

这证明了 F_ξ 是右连续的.

由于事件

$$\{\xi \leqslant -1\} \supset \{\xi \leqslant -2\} \supset \cdots$$

构成一个交集为 $\varnothing$ 的递减序列，且

$$\{\xi \leqslant 1\} \subset \{\xi \leqslant 2\} \subset \cdots$$

构成一个并集为 Ω 的递增序列，由练习 *1.1* 可知

$$\begin{aligned}
\lim_{x\to-\infty} F_\xi(x) &= \lim_{n\to\infty} F_\xi(-n) = \lim_{n\to\infty} P\{\xi \leqslant -n\} = P(\varnothing) = 0,\\
\lim_{x\to+\infty} F_\xi(x) &= \lim_{n\to\infty} F_\xi(n) = \lim_{n\to\infty} P\{\xi \leqslant n\} = P(\Omega) = 1,
\end{aligned}$$

因为 F_ξ 是单调递增的.

解答 1.5. 若 ξ 具有密度 f_ξ，则分布函数 F_ξ 可以写成

$$F_\xi(x) = P\{\xi \leqslant x\} = \int_{-\infty}^{x} f_\xi(y)\,\mathrm{d}y.$$

因此，若 f_ξ 在 x 处连续，则 F_ξ 在 x 处可导且

$$\frac{\mathrm{d}}{\mathrm{d}x}F_\xi(x) = f_\xi(x).$$

解答 1.6. 若 $s < t$ 是实数使得对任意 i 都有 $x_i \notin (s,t]$，则

$$F_\xi(t) - F_\xi(s) = P\{\xi \leqslant t\} - P\{\xi \leqslant s\} = P\{\xi \in (s,t]\} = 0,$$

即 $F_\xi(s) = F_\xi(t)$. 由于 F_ξ 是单调递增的，这意味着 F_ξ 在 $(s,t]$ 上为常数. 为证明 F_ξ 在每个 x_i 处有大小为 $P\{\xi = x_i\}$ 的跳跃，计算

$$\begin{aligned}
\lim_{t\searrow x_i} F_\xi(t) - \lim_{s\nearrow x_i} F_\xi(s) &= \lim_{t\searrow x_i} P\{\xi \leqslant t\} - \lim_{s\nearrow x_i} P\{\xi \leqslant s\}\\
&= P\left\{\xi \leqslant x_i\right\} - P\left\{\xi < x_i\right\} = P\left\{\xi = x_i\right\}.
\end{aligned}$$

解答 1.7. 若 h 是阶梯函数，

$$h = \sum_{i=1}^{n} h_i \mathbf{1}_{A_i},$$

其中 $h_1, \ldots, h_n$ 是实数且 $A_1, \ldots, A_n$ 是两两不交的覆盖 $\mathbb{R}$ 的 *Borel* 集，则

$$\begin{aligned} E(h(\xi)) &= \sum_{i=1}^{n} h_i E(\mathbf{1}_{A_i}(\xi)) = \sum_{i=1}^{n} h_i P\{\xi \in A_i\} \\ &= \sum_{i=1}^{n} h_i P_\xi(A_i) = \sum_{i=1}^{n} \int_{A_i} h(x)\,\mathrm{d}P_\xi(x) = \int_{\mathbb{R}} h(x)\,\mathrm{d}P_\xi(x). \end{aligned}$$

接下来，任意非负 *Borel* 函数 h 都可以用一列单调递增的阶梯函数逼近. 对这样的 h，结果由积分的单调收敛性得到. 最后，这蕴含了对所有 *Borel* 函数 h 的所求等式，因为每个 *Borel* 函数都可以分解为正部和负部 $h = h^+ - h^-$，其中 $h^+, h^- \geqslant 0$.

解答 1.8. 由 *Schwarz* 不等式（*1.1*）取 $\eta = 1$，若 ξ 是平方可积的，则

$$[E(|\xi|)]^2 = [E(1 \cdot |\xi|)]^2 \leqslant E(1^2)E(\xi^2) = E(\xi^2) < \infty,$$

即 ξ 是可积的.

为证明解答 1.9，设 $F(t) = P\{\eta \leqslant t\}$ 是 η 的分布函数. 则

$$E(\eta^2) = \int_0^\infty t^2\,\mathrm{d}F(t).$$

由于 $P(\eta > t) = 1 - F(t)$，需要证明下式成立:

$$\int_0^\infty t^2\,\mathrm{d}F(t) = 2\int_0^\infty t(1 - F(t))\,\mathrm{d}t. \tag{1.1}$$

解答 1.9. 首先，建立（*1.2*）的一个版本，其中 ∞ 替换为有限数 a. 分部积分，得到（*1.3*）.

可以看到（*1.2*）可由（*1.3*）推出，只要（*1.4*）成立.

但是

$$0 \leqslant a^2(1 - F(a)) = a^2 P(\eta > a) \leqslant (n+1)^2 P(\eta > n) \leqslant 4n^2 P(\eta \geqslant n),$$

其中 n 是 a 的整数部分，且

$$E(\eta^2) = \sum_{k=0}^{\infty} \int_{\{k \leqslant \eta \leqslant k+1\}} \eta^2\,\mathrm{d}P < \infty.$$

因此，当 $n \to \infty$ 时

$$n^2 P(\eta \geqslant n) \leqslant \int_{\{\eta \geqslant n\}} \eta^2\,\mathrm{d}P = \sum_{k=n}^{\infty} \int_{\{k \leqslant \eta < k+1\}} \eta^2\,\mathrm{d}P \to 0,$$

这证明了（*1.4*）（见式 *(1.5)*）.

$$\begin{aligned}\int_0^a t^2\,\mathrm{d}F(t) &= \int_0^a t^2\,\mathrm{d}(F(t)-1)\\ &= t^2(F(t)-1)\Big|_0^a - 2\int_0^a t(F(t)-1)\,\mathrm{d}t \\ &= -a^2(1-F(a)) + 2\int_0^a t(1-F(t))\,\mathrm{d}t.\end{aligned} \tag{1.2}$$

$$a^2(1-F(a)) \to 0, \quad \text{当} a\to\infty. \tag{1.3}$$

$$n^2P(\eta\geqslant n) \leqslant \int_{\{\eta\geqslant n\}} \eta^2\,\mathrm{d}P = \sum_{k=n}^{\infty}\int_{\{k\leqslant\eta<k+1\}} \eta^2\,\mathrm{d}P \to 0. \tag{1.4}$$

解答 1.10. 由于 $B_1\cup B_2\cup\cdots=\Omega$，

$$A = A\cap(B_1\cup B_2\cup\cdots) = (A\cap B_1)\cup(A\cap B_2)\cup\cdots,$$

其中

$$(A\cap B_i)\cap(A\cap B_j) = A\cap(B_i\cap B_j) = A\cap\varnothing = \varnothing.$$

由可数可加性

$$\begin{aligned}P(A) &= P\left((A\cap B_1)\cup(A\cap B_2)\cup\cdots\right)\\ &= P\left(A\cap B_1\right) + P\left(A\cap B_2\right) + \cdots\\ &= P\left(A\mid B_1\right)P\left(B_1\right) + P\left(A\mid B_2\right)P\left(B_2\right) + \cdots.\end{aligned}$$

解答 1.11. 若 $P(B)\neq 0$，则 A 和 B 独立当且仅当

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

反过来，这个等式成立当且仅当 $P(A)=P(A|B)$.

解答 1.12. σ-域 $\sigma(\xi)$ 和 $\sigma(\eta)$ 分别由形如

$$\{\xi\in A\} \quad \text{和} \quad \{\eta\in B\}$$

的事件组成，其中 A 和 B 是 $\mathbb{R}$ 中的 *Borel* 集. 因此，$\sigma(\xi)$ 和 $\sigma(\eta)$ 独立当且仅当对任意 *Borel* 集 A 和 B，事件 $\{\xi\in A\}$ 和 $\{\eta\in B\}$ 独立，而这又等价于 ξ 和 η 独立.

第二章
条件期望

条件期望是研究随机过程的关键工具. 因此，培养对这一概念的直观理解至关重要，尽管其定义初看起来可能有些抽象. 本章旨在通过若干复杂度逐渐增加的特例，引导读者逐步理解条件期望的一般定义. 此外，本章还提供了大量例子和习题来帮助读者加深理解.

第 1 节　基于事件的条件化

首先考虑最简单的情形：给定事件 B 时随机变量 ξ 的条件期望 $E(\xi|B)$.

定义 2.1. *对于任意可积随机变量 ξ 和满足 $P(B) \neq 0$ 的事件 $B \in \mathcal{F}$，给定 B 时 ξ 的条件期望定义为*

$$E(\xi|B) = \frac{1}{P(B)} \int_B \xi \, \mathrm{d}P.$$

例 2.1. *抛掷三枚硬币，面值分别为 10 便士、20 便士和 50 便士. 将正面朝上的硬币面值相加得到总金额 ξ. 已知有两枚硬币正面朝上，求期望总金额.*

设 B 表示有两枚硬币正面朝上的事件. 下面求 $E(\xi|B)$. 显然，B 包含三个元素

$$B = \{\mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\},$$

每个元素的概率都是 $\frac{1}{8}$.（这里 H 表示正面，T 表示反面.）ξ 的相应取值为

$$\xi(\mathrm{HHT}) = 10 + 20 = 30,$$
$$\xi(\mathrm{HTH}) = 10 + 50 = 60,$$
$$\xi(\mathrm{THH}) = 20 + 50 = 70.$$

因此

$$E(\xi|B) = \frac{1}{P(B)} \int_B \xi \, \mathrm{d}P = \frac{1}{\frac{3}{8}} \left(\frac{30}{8} + \frac{60}{8} + \frac{70}{8} \right) = 53\frac{1}{3}.$$

练习 2.1. 证明 $E(\xi|\Omega) = E(\xi)$.

提示 $E(\xi)$ 的定义涉及一个积分，$E(\xi|\Omega)$ 的定义也涉及一个积分. 这两个积分之间有什么关系？

练习 2.2. 证明若

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{当}\omega \in A, \\ 0 & \text{当}\omega \notin A \end{cases}$$

(A 的指示函数)，则

$$E(\mathbf{1}_A|B) = P(A|B),$$

其中

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

是给定 B 时 A 的条件概率.

提示 将 $\int_B \mathbf{1}_A \, \mathrm{d}P$ 写成 $P(A \cap B)$ 的形式.

第 2 节　基于离散随机变量的条件化

迈向条件期望一般定义的下一步涉及基于离散随机变量 η 的条件化，其可能取值为 $y_1, y_2, \ldots$，且对每个 n 有 $P\{\eta = y_n\} \neq 0$. 获知 η 的值等同于获知 $\{\eta = y_n\}$ 中哪个事件发生. 因此，基于 η 的条件化应该与基于事件 $\{\eta = y_n\}$ 的条件化相同. 由于事先不知道这些事件中哪个会发生，需要考虑所有可能性，这就涉及一列条件期望

$$E\left(\xi \mid \{\eta = y_1\}\right), E\left(\xi \mid \{\eta = y_2\}\right), \ldots.$$

一种方便的做法是构造一个新的离散随机变量，在每个集合 $\{\eta = y_n\}$ 上恒等于 $E\left(\xi \mid \{\eta = y_n\}\right)$. 这引导我们得到下面的定义.

定义 2.2. 设 ξ 是可积随机变量，η 是如上所述的离散随机变量. 则给定 η 时 ξ 的条件期望定义为一个随机变量 $E(\xi|\eta)$，使得

$$E(\xi|\eta)(\omega) = E(\xi|\{\eta = y_n\}) \quad \text{若}\eta(\omega) = y_n$$

对任意 $n = 1, 2, \ldots$ 成立.

例 2.2. 如例 *2.1* 中抛掷三枚 *10* 便士、*20* 便士和 *50* 便士的硬币. 已知仅由 *10* 便士和 *20* 便士硬币显示的总金额 η，求三枚硬币显示的总金额 ξ 的条件期望 $E(\xi|\eta)$.

显然，η 是一个离散随机变量，有四个可能取值：*0*、*10*、*20* 和 *30*. 下面用与例 *2.1* 类似的方式求出四个相应的条件期望：

$$E(\xi|\{\eta=0\})=25,\quad E(\xi|\{\eta=10\})=35,$$
$$E(\xi|\{\eta=20\})=45,\quad E(\xi|\{\eta=30\})=55.$$

因此

$$E(\xi|\eta)(\omega)=\begin{cases}25 & 若\eta(\omega)=0,\\ 35 & 若\eta(\omega)=10,\\ 45 & 若\eta(\omega)=20,\\ 55 & 若\eta(\omega)=30.\end{cases}$$

例 2.3. 取 $\Omega=[0,1]$，其上的 σ-域为 *Borel* 集，P 为 $(0,1]$ 上的 *Lebesgue* 测度. 对于

$$\xi(x)=2x^2,\qquad \eta(x)=\begin{cases}1 & 若x\in\left[0,\frac{1}{3}\right],\\ 2 & 若x\in\left(\frac{1}{3},\frac{2}{3}\right),\\ 0 & 若x\in\left(\frac{2}{3},1\right],\end{cases}$$

下面求 $E(\xi|\eta)$.

显然，η 是离散的，有三个可能取值 *1*、*2*、*0*. 相应的事件为

$$\{\eta=1\}=\left[0,\frac{1}{3}\right],$$
$$\{\eta=2\}=\left(\frac{1}{3},\frac{2}{3}\right),$$
$$\{\eta=0\}=\left(\frac{2}{3},1\right].$$

对于 $x\in\left[0,\frac{1}{3}\right]$

$$E(\xi|\eta)(x)=E\left(\xi\mid\left[0,\frac{1}{3}\right]\right)=\frac{1}{\frac{1}{3}}\int_0^{\frac{1}{3}}2x^2\,\mathrm{d}x=\frac{2}{27}.$$

对于 $x\in\left(\frac{1}{3},\frac{2}{3}\right)$

$$E(\xi|\eta)(x)=E\left(\xi\mid\left(\frac{1}{3},\frac{2}{3}\right)\right)=\frac{1}{\frac{1}{3}}\int_{\frac{1}{3}}^{\frac{2}{3}}2x^2\,\mathrm{d}x=\frac{14}{27}.$$

对于 $x \in \left(\frac{2}{3}, 1\right]$

$$E(\xi|\eta)(x) = E\left(\xi \mid \left(\frac{2}{3}, 1\right]\right) = \frac{1}{\frac{1}{3}} \int_{\frac{2}{3}}^{1} 2x^2 \, \mathrm{d}x = \frac{38}{27}.$$

$E(\xi|\eta)$ 的图像如图 *2.1* 所示，同时给出了 ξ 和 η 的图像.

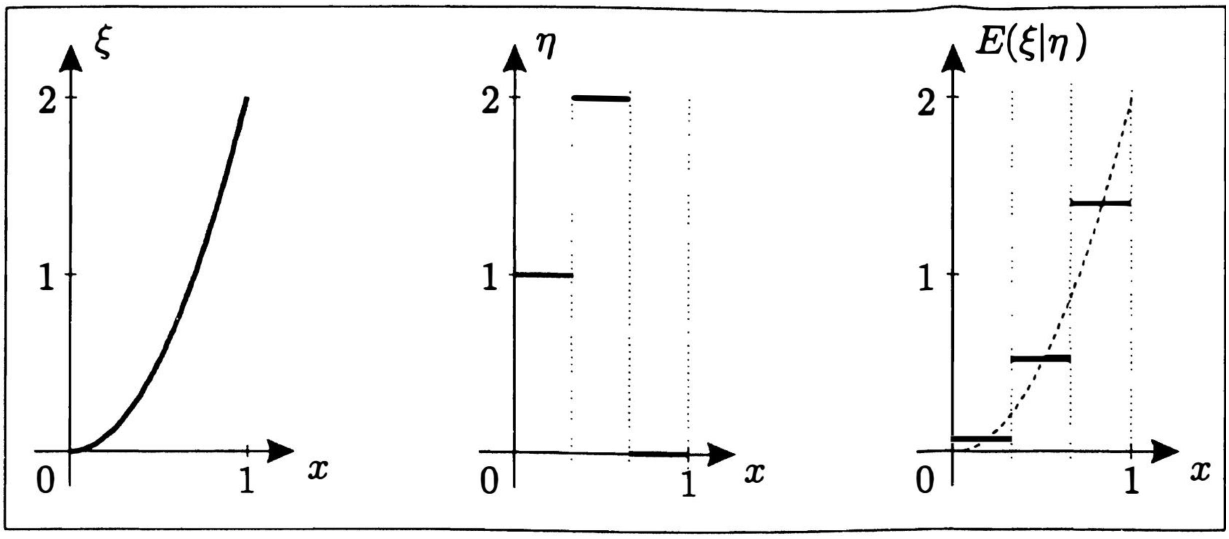

图 2.1: 例 2.3 中 $E(\xi|\eta)$ 的图像

练习 2.3. 证明若 η 是常值函数，则 $E(\xi|\eta)$ 是常数且等于 $E(\xi)$.

提示 对任意 $c \in \mathbb{R}$，事件 $\{\eta = c\}$ 必为 $\varnothing$ 或 Ω.

练习 2.4. 证明对任意满足 $1 \neq P(B) \neq 0$ 的 B 有

$$E(\mathbf{1}_A|\mathbf{1}_B)(\omega) = \begin{cases} P(A|B) & \text{若}\omega \in B, \\ P(A|\Omega \setminus B) & \text{若}\omega \notin B. \end{cases}$$

提示 $\mathbf{1}_B$ 取多少个不同的值？这些值分别在什么集合上取到？

练习 2.5. 假设 η 是离散随机变量，证明

$$E(E(\xi|\eta)) = E(\xi).$$

提示 观察到对任意 η 在其上为常数的事件 B 有

$$\int_B E(\xi|\eta) \, \mathrm{d}P = \int_B \xi \, \mathrm{d}P.$$

所求等式可以通过用可数个这种互不相交的事件覆盖 Ω 而得到.

命题 2.1. 若 ξ 是可积随机变量且 η 是离散随机变量，则

(1) $E(\xi|\eta)$ 是 $\sigma(\eta)$-可测的；

(2) 对任意 $A \in \sigma(\eta)$

$$\int_A E(\xi|\eta)\,\mathrm{d}P = \int_A \xi\,\mathrm{d}P. \tag{2.1}$$

证明. 假设 η 有互不相同的两两不同的值 $y_1, y_2, \ldots$. 则事件

$$\{\eta = y_1\}, \{\eta = y_2\}, \ldots$$

两两不交且覆盖 Ω. σ-域 $\sigma(\eta)$ 由这些事件生成，事实上每个 $A \in \sigma(\eta)$ 都是形如 $\{\eta = y_n\}$ 的可数个集合的并. 因为 $E(\xi|\eta)$ 在每个这些集合上都是常数，所以它必是 $\sigma(\eta)$-可测的.

对每个 n 有

$$\begin{aligned}\int_{\{\eta=y_n\}} E(\xi|\eta)\,\mathrm{d}P &= \int_{\{\eta=y_n\}} E(\xi|\{\eta=y_n\})\,\mathrm{d}P \\ &= \int_{\{\eta=y_n\}} \xi\,\mathrm{d}P.\end{aligned}$$

因为每个 $A \in \sigma(\eta)$ 都是形如 $\{\eta = y_n\}$ 的可数个两两不交集合的并，所以

$$\int_A E(\xi|\eta)\,\mathrm{d}P = \int_A \xi\,\mathrm{d}P,$$

如所要求. □

第 3 节　基于任意随机变量的条件化

命题 2.1 中的性质 (1) 和 (2) 为给定任意随机变量 η 时条件期望的定义提供了关键.

定义 2.3. 设 ξ 是可积随机变量，η 是任意随机变量. 则给定 η 时 ξ 的条件期望定义为一个随机变量 $E(\xi|\eta)$ 使得

(1) $E(\xi|\eta)$ 是 $\sigma(\eta)$-可测的；

(2) 对任意 $A \in \sigma(\eta)$

$$\int_A E(\xi|\eta)\,\mathrm{d}P = \int_A \xi\,\mathrm{d}P.$$

注记 2.1. 也可以通过

$$P(A|\eta) = E(\mathbf{1}_A|\eta)$$

来定义给定 η 时事件 $A \in \mathcal{F}$ 的条件概率，其中 $\mathbf{1}_A$ 是 A 的指示函数.

定义 2.3 的条件是否唯一刻画了 $E(\xi|\eta)$？下面的引理意味着 $E(\xi|\eta)$ 在满测度集合上至多有定义. 即

$$若\xi = \xi' \ a.s.\square 则 E(\xi|\eta) = E(\xi'|\eta) \ a.s. \tag{2.2}$$

$E(\xi|\eta)$ 的存在性将在本章稍后讨论.

引理 2.1. 设 $(\Omega, \mathcal{F}, P)$ 是概率空间，$\mathcal{G}$ 是包含于 $\mathcal{F}$ 的 σ-域. 若 ξ 是 $\mathcal{G}$-可测随机变量且对任意 $B \in \mathcal{G}$

$$\int_B \xi \, \mathrm{d}P = 0,$$

则 $\xi = 0$ *a.s.*.

证明. 注意对任意 $\varepsilon > 0$ 有 $P\{\xi \geqslant \varepsilon\} = 0$，因为

$$0 \leqslant \varepsilon P\{\xi \geqslant \varepsilon\} = \int_{\{\xi \geqslant \varepsilon\}} \varepsilon \, \mathrm{d}P \leqslant \int_{\{\xi \geqslant \varepsilon\}} \xi \, \mathrm{d}P = 0.$$

最后一个等式成立是因为 $\{\xi \geqslant \varepsilon\} \in \mathcal{G}$. 类似地，对任意 $\varepsilon > 0$ 有 $P\{\xi \leqslant -\varepsilon\} = 0$. 因此，对任意 $\varepsilon > 0$

$$P\{-\varepsilon < \xi < \varepsilon\} = 1.$$

记

$$A_n = \left\{-\frac{1}{n} < \xi < \frac{1}{n}\right\}.$$

则 $P(A_n) = 1$ 且

$$\{\xi = 0\} = \bigcap_{n=1}^{\infty} A_n.$$

因为 A_n 构成一个递减序列的事件，所以

$$P\{\xi = 0\} = \lim_{n \to \infty} P(A_n) = 1,$$

如所要求. □

定义 2.3 的一个困难是没有给出 $E(\xi|\eta)$ 的显式公式. 如果已知这样的公式，那么通常很容易验证条件 (1) 和 (2). 但首先如何找到它呢？下面的例子和习题旨在展示如何在具体情况下解决这个问题.

例 2.4. 取 $\Omega = [0,1]$，其上的 σ-域为 *Borel* 集，P 为 $[0,1]$ 上的 *Lebesgue* 测度. 对于

$$\xi(x) = 2x^2, \qquad \eta(x) = \begin{cases} 2 & 若 x \in \left[0, \frac{1}{2}\right), \\ x & 若 x \in \left[\frac{1}{2}, 1\right], \end{cases}$$

下面求 $E(\xi|\eta)$.

这里 η 不再是离散的，应使用一般的定义 *2.3*.

首先描述 σ-域 $\sigma(\eta)$. 对任意 *Borel* 集 $B\subset\left[\frac{1}{2},1\right]$ 有

$$B=\{\eta\in B\}\in\sigma(\eta)$$

且

$$\left[0,\frac{1}{2}\right)\cup B=\{\eta\in B\}\cup\{\eta=2\}\in\sigma(\eta).$$

事实上，这两类集合穷尽了 $\sigma(\eta)$ 的所有元素. 任意 *Borel* 集 $C\subset\mathbb{R}$ 的逆像 $\{\eta\in C\}$ 若 $2\notin C$ 则为第一类，若 $2\in C$ 则为第二类.

若 $E(\xi|\eta)$ 是 $\sigma(\eta)$-可测的，它必在 $\left[0,\frac{1}{2}\right)$ 上为常数，因为 η 在此区间上是常数. 若对任意 $x\in\left[0,\frac{1}{2}\right)$

$$\begin{aligned}E(\xi|\eta)(x)&=E\left(\xi\mid\left[0,\frac{1}{2}\right)\right)\\&=\frac{1}{P\left(\left[0,\frac{1}{2}\right)\right)}\int_{\left[0,\frac{1}{2}\right)}\xi(x)\,\mathrm{d}x\\&=\frac{1}{\frac{1}{2}}\int_0^{\frac{1}{2}}2x^2\,\mathrm{d}x\\&=\frac{1}{6},\end{aligned}$$

则

$$\int_{\left[0,\frac{1}{2}\right)}E(\xi|\eta)(x)\,\mathrm{d}x=\int_{\left[0,\frac{1}{2}\right)}\xi(x)\,\mathrm{d}x,$$

即定义 *2.3* 的条件 *(2)* 对 $A=\left[0,\frac{1}{2}\right)$ 成立.

此外，若 $E(\xi|\eta)=\xi$ 在 $\left[\frac{1}{2},1\right]$ 上，则当然对任意 *Borel* 集 $B\subset\left[\frac{1}{2},1\right]$ 有

$$\int_B E(\xi|\eta)(x)\,\mathrm{d}x=\int_B\xi(x)\,\mathrm{d}x.$$

因此，得到

$$E(\xi|\eta)(x)=\begin{cases}\frac{1}{6} & 若x\in\left[0,\frac{1}{2}\right),\\ 2x^2 & 若x\in\left[\frac{1}{2},1\right].\end{cases}$$

因为 $\sigma(\eta)$ 的每个元素都是形如 B 或 $\left[0,\frac{1}{2}\right)\cup B$ 的集合，其中 $B\subset\left[\frac{1}{2},1\right]$ 是 *Borel* 集，所以定义 *2.3* 的条件 *(1)* 和 *(2)* 立即得到满足. $E(\xi|\eta)$ 的图像如图 *2.2* 所示，同时给出了 ξ 和 η 的图像.

练习 2.6. 设 $\Omega=[0,1]$，其上的测度与例 *2.4* 相同. 若

$$\xi(x)=2x^2,\qquad\eta(x)=1-|2x-1|.$$

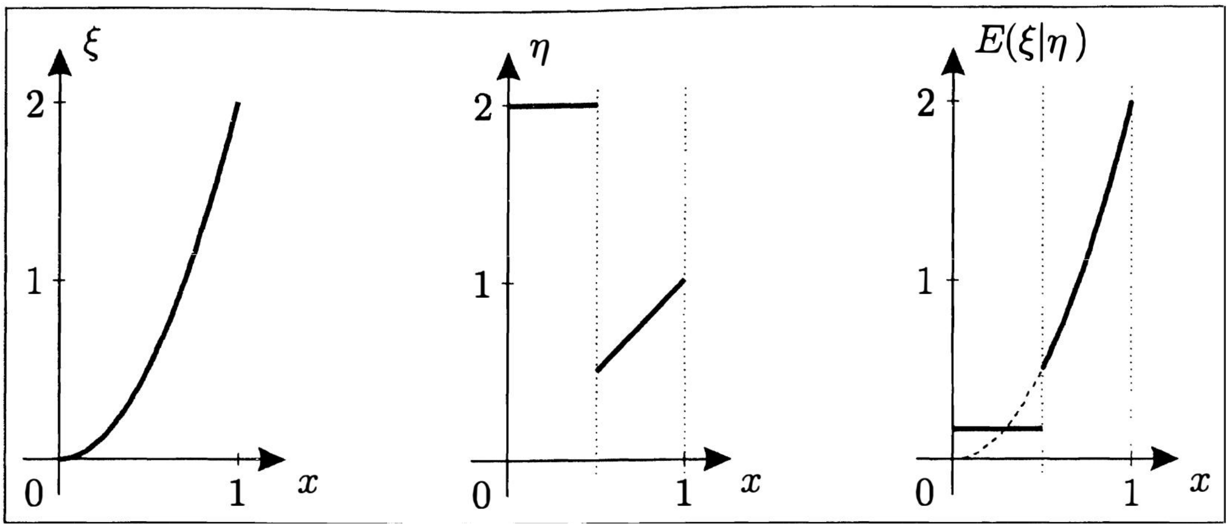

图 2.2: 例 2.4 中 $E(\xi|\eta)$ 的图像

求条件期望 $E(\xi|\eta)$.

提示 首先描述由 η 生成的 σ-域. 观察到 η 关于 $\frac{1}{2}$ 对称. 这告诉你关于 $\sigma(\eta)$ 中的集合什么信息？这告诉你关于 $E(\xi|\eta)$ 的什么信息，如果它要是 $\sigma(\eta)$-可测的？它需要对称吗？对任意 $A \in \sigma(\eta)$ 尝试变换 $\int_A \xi \, \mathrm{d}P$ 使得被积函数对称.

练习 2.7. 设 Ω 是单位正方形 $[0,1] \times [0,1]$，其上的 σ-域为 *Borel* 集，P 为 $[0,1] \times [0,1]$ 上的 *Lebesgue* 测度. 假设 ξ 和 η 是 Ω 上的随机变量，具有联合密度

$$f_{\xi,\eta}(x,y) = x + y$$

对任意 $x, y \in [0,1]$，且 $f_{\xi,\eta}(x,y) = 0$ 在其他地方. 证明

$$E(\xi|\eta) = \frac{2+3\eta}{3+6\eta}.$$

提示 只需（为什么？）证明对任意 *Borel* 集 B

$$\int_{\{\eta \in B\}} \xi \, \mathrm{d}P = \int_{\{\eta \in B\}} \frac{2+3\eta}{3+6\eta} \, \mathrm{d}P.$$

尝试将等式两边都表示为使用联合密度 $f_{\xi,\eta}(x,y)$ 在单位正方形 $[0,1] \times [0,1]$ 上的积分.

练习 2.8. 设 Ω 是单位正方形 $[0,1] \times [0,1]$，其上的测度与练习 *2.7* 相同. 若 ξ 和 η 是 Ω 上的随机变量，具有联合密度

$$f_{\xi,\eta}(x,y) = \frac{3}{2}\left(x^2 + y^2\right)$$

对任意 $x, y \in [0,1]$，且 $f_{\xi,\eta}(x,y) = 0$ 在其他地方，求 $E(\xi|\eta)$.

提示 这比练习 *2.7* 稍难，因为这里需要推导条件期望的公式. 研究练习 *2.7* 的解答以找到获得这种公式的方法.

练习 2.9. 设 Ω 是单位圆盘 $\{(x,y) : x^2+y^2 \leqslant 1\}$，其上的 σ-域为 *Borel* 集，P 为圆盘上的 *Lebesgue* 测度，归一化使得 $P(\Omega)=1$，即

$$P(A)=\frac{1}{\pi}\iint_A \mathrm{d}x\,\mathrm{d}y$$

对任意 *Borel* 集 $A\subset\Omega$. 假设 ξ 和 η 分别是投影到 x 轴和 y 轴上

$$\xi(x,y)=x,\quad \eta(x,y)=y$$

对任意 $(x,y)\in\Omega$. 求 $E(\xi^2|\eta)$.

提示 ξ 和 η 的联合密度是什么？使用这个密度变换积分

$$\int_{\{\eta\in B\}}\xi^2\,\mathrm{d}P$$

对任意 *Borel* 集 B，使得被积函数成为 η 的函数. 这个 η 的函数与 $E(\xi^2|\eta)$ 有何关系？

第 4 节　基于 σ-域的条件化

现在可以迈向条件期望的一般定义的最后一步. 它基于观察：$E(\xi|\eta)$ 只依赖于 η 生成的 σ-域 $\sigma(\eta)$，而不是 η 的实际取值.

命题 2.2. 若 $\sigma(\eta)=\sigma(\eta')$，则 $E(\xi|\eta)=E(\xi|\eta')$ *a.s.*（与 *(2.2)* 比较.）

证明. 这是引理 2.1 的直接推论. □

由于命题 2.2，谈论给定 σ-域的条件期望是合理的. 下面的定义与定义 2.3 的唯一区别是用任意 σ-域 $\mathcal{G}$ 代替由随机变量 η 生成的 σ-域 $\sigma(\eta)$.

定义 2.4. 设 ξ 是概率空间 $(\Omega,\mathcal{F},P)$ 上的可积随机变量，$\mathcal{G}$ 是包含于 $\mathcal{F}$ 的 σ-域. 则给定 $\mathcal{G}$ 时 ξ 的条件期望定义为一个随机变量 $E(\xi|\mathcal{G})$ 使得

(1) $E(\xi|\mathcal{G})$ 是 $\mathcal{G}$-可测的；

(2) 对任意 $A\in\mathcal{G}$

$$\int_A E(\xi|\mathcal{G})\,\mathrm{d}P=\int_A \xi\,\mathrm{d}P. \tag{2.3}$$

注记 2.2. 给定 σ-域 $\mathcal{G}$ 时事件 $A\in\mathcal{F}$ 的条件概率可以通过

$$P(A|\mathcal{G})=E(\mathbf{1}_A|\mathcal{G})$$

来定义，其中 $\mathbf{1}_A$ 是 A 的指示函数.

关于 σ-域的条件期望的概念推广了关于随机变量 η 的条件化，其意义在于

$$E(\xi|\sigma(\eta)) = E(\xi|\eta),$$

其中 $\sigma(\eta)$ 是由 η 生成的 σ-域.

命题 2.3. $E(\xi|\mathcal{G})$ 存在且在如下意义下唯一：若 $\xi = \xi'$ *a.s.*，则 $E(\xi|\mathcal{G}) = E(\xi'|\mathcal{G})$ *a.s.*.

证明. 存在性和唯一性分别由下面的定理 2.1 和引理 2.1 得出. □

定理 2.1（Radon-Nikodym）. 设 $(\Omega, \mathcal{F}, P)$ 是概率空间，$\mathcal{G}$ 是包含于 $\mathcal{F}$ 的 σ-域. 则对任意随机变量 ξ 存在 $\mathcal{G}$-可测随机变量 ζ 使得

$$\int_A \xi \, \mathrm{d}P = \int_A \zeta \, \mathrm{d}P$$

对每个 $A \in \mathcal{G}$ 成立.

Radon-Nikodym 定理从理论观点看是重要的. 然而，在实践中通常有其他方法可以建立条件期望的存在性，例如通过寻找显式公式，如上一节中的例子和习题所示. Radon-Nikodym 定理的证明超出了本课程的范围，在此省略.

练习 2.10. 证明若 $\mathcal{G} = \{\emptyset, \Omega\}$，则 $E(\xi|\mathcal{G}) = E(\xi)$ *a.s.*.

提示 若 $\mathcal{G} = \{\emptyset, \Omega\}$，什么随机变量是 $\mathcal{G}$-可测的？

练习 2.11. 证明若 ξ 是 $\mathcal{G}$-可测的，则 $E(\xi|\mathcal{G}) = \xi$ *a.s.*.

提示 若 ξ 是 $\mathcal{G}$-可测的，则定义 *2.4* 的条件显然由 ξ 满足.

练习 2.12. 证明若 $B \in \mathcal{G}$，则

$$E\left(E(\xi|\mathcal{G}) \mid B\right) = E(\xi|B).$$

提示 等式两边的条件期望都涉及在一个积分. 这些积分之间有什么关系？

第 5 节 一般性质

命题 2.4. 条件期望具有以下性质：

(1) $E(a\xi + b\zeta|\mathcal{G}) = aE(\xi|\mathcal{G}) + bE(\zeta|\mathcal{G})$（线性性）；

(2) $E(E(\xi|\mathcal{G})) = E(\xi)$；

(3) $E(\xi\zeta|\mathcal{G}) = \xi E(\zeta|\mathcal{G})$ 若 ξ 是 $\mathcal{G}$-可测的（取出已知量）；

(4) $E(\xi|\mathcal{G}) = E(\xi)$ 若 ξ 独立于 $\mathcal{G}$（独立条件消失）；

(5) $E(E(\xi|\mathcal{G})|\mathcal{H}) = E(\xi|\mathcal{H})$ 若 $\mathcal{H} \subset \mathcal{G}$（塔性质）；

(6) 若 $\xi \geqslant 0$，则 $E(\xi|\mathcal{G}) \geqslant 0$（正性）.

这里 a, b 是任意实数，ξ, ζ 是概率空间 $(\Omega, \mathcal{F}, P)$ 上的可积随机变量，$\mathcal{G}, \mathcal{H}$ 是 Ω 上包含于 $\mathcal{F}$ 的 σ-域. 在 *(3)* 中还假设乘积 $\xi\zeta$ 是可积的. 所有等式和 *(6)* 中的不等式都是 P-*a.s.* 成立的.

证明. (1) 对任意 $B \in \mathcal{G}$

$$\begin{aligned}\int_B (aE(\xi|\mathcal{G}) + bE(\zeta|\mathcal{G}))\,\mathrm{d}P &= a\int_B E(\xi|\mathcal{G})\,\mathrm{d}P + b\int_B E(\zeta|\mathcal{G})\,\mathrm{d}P \\ &= a\int_B \xi\,\mathrm{d}P + b\int_B \zeta\,\mathrm{d}P \\ &= \int_B (a\xi + b\zeta)\,\mathrm{d}P.\end{aligned}$$

由唯一性，这证明了所求的等式.

(2) 这由 (2.3) 中令 $A = \Omega$ 得到. 同时，(2) 也是 (5) 当 $\mathcal{H} = \{\emptyset, \Omega\}$ 时的特例.

(3) 首先对 $\xi = \mathbf{1}_A$（其中 $A \in \mathcal{G}$）验证结果. 此时对任意 $B \in \mathcal{G}$

$$\begin{aligned}\int_B \mathbf{1}_A E(\eta|\mathcal{G})\,\mathrm{d}P &= \int_{A\cap B} E(\eta|\mathcal{G})\,\mathrm{d}P \\ &= \int_{A\cap B} \eta\,\mathrm{d}P \\ &= \int_B \mathbf{1}_A \eta\,\mathrm{d}P,\end{aligned}$$

这意味着由唯一性

$$\mathbf{1}_A E(\eta|\mathcal{G}) = E(\mathbf{1}_A \eta|\mathcal{G}).$$

类似地，若 ξ 是 $\mathcal{G}$-可测的阶梯函数

$$\xi = \sum_{j=1}^m a_j \mathbf{1}_{A_j},$$

其中 $A_j \in \mathcal{G}$ 对 $j = 1, \ldots, m$，结果也成立. 最后，一般情形的结果通过用 $\mathcal{G}$-可测阶梯函数逼近 ξ 得到.

(4) 因为 ξ 独立于 $\mathcal{G}$，所以随机变量 ξ 和 $\mathbf{1}_B$ 对任意 $B \in \mathcal{G}$ 是独立的. 由命题 1.1（独立随机变量是不相关的）可得

$$\begin{aligned}\int_B E(\xi)\,\mathrm{d}P &= E(\xi)E(\mathbf{1}_B) \\ &= E(\xi\mathbf{1}_B) \\ &= \int_B \xi\,\mathrm{d}P,\end{aligned}$$

这证明了断言.

(5) 由定义 2.4

$$\int_B E(\xi|\mathcal{G})\,\mathrm{d}P = \int_B \xi\,\mathrm{d}P$$

对每个 $B \in \mathcal{G}$ 成立，且

$$\int_B E(\xi|\mathcal{H})\,\mathrm{d}P = \int_B \xi\,\mathrm{d}P$$

对每个 $B \in \mathcal{H}$ 成立. 因为 $\mathcal{H} \subset \mathcal{G}$，所以对每个 $B \in \mathcal{H}$

$$\int_B E(\xi|\mathcal{G})\,\mathrm{d}P = \int_B E(\xi|\mathcal{H})\,\mathrm{d}P.$$

再次应用定义 2.4，得到

$$E(E(\xi|\mathcal{G})|\mathcal{H}) = E(\xi|\mathcal{H}).$$

(6) 对任意 n 记

$$A_n = \left\{E(\xi|\mathcal{G}) \leqslant -\frac{1}{n}\right\}.$$

则 $A_n \in \mathcal{G}$. 若 $\xi \geqslant 0$ a.s.，则

$$0 \leqslant \int_{A_n} \xi\,\mathrm{d}P = \int_{A_n} E(\xi|\mathcal{G})\,\mathrm{d}P \leqslant -\frac{1}{n}P(A_n),$$

这意味着 $P(A_n) = 0$. 因为

$$\{E(\xi|\mathcal{G}) < 0\} = \bigcup_{n=1}^{\infty} A_n$$

所以

$$P\{E(\xi|\mathcal{G}) < 0\} = 0,$$

证毕. □

下一个定理将不予证明地陈述，它涉及凸函数的概念，例如 $\max(1, x)$ 或 $e^{|x|}$. 在本课程中，该定理主要用于 $|x|$，它也是凸函数. 一般地，称函数 $\varphi : \mathbb{R} \to \mathbb{R}$ 是凸的，如果对任意 $x, y \in \mathbb{R}$ 和任意 $\lambda \in [0, 1]$

$$\varphi(\lambda x + (1-\lambda)y) \leqslant \lambda\varphi(x) + (1-\lambda)\varphi(y).$$

这个条件意味着 φ 的图像位于连接点 $(x, \varphi(x))$ 和 $(y, \varphi(y))$ 的弦的下方.

定理 2.2（Jensen 不等式）. 设 $\varphi : \mathbb{R} \to \mathbb{R}$ 是凸函数，ξ 是概率空间 $(\Omega, \mathcal{F}, P)$ 上的可积随机变量，使得 $\varphi(\xi)$ 也是可积的. 则对任意 Ω 上包含于 $\mathcal{F}$ 的 σ-域 $\mathcal{G}$

$$\varphi(E(\xi|\mathcal{G})) \leqslant E(\varphi(\xi)|\mathcal{G}) \quad a.s..$$

第 6 节 条件期望的各种练习

练习 2.13. 琼斯太太为她的两个儿子做了一个牛排腰子饼. 吃掉超过一半会让任何人消化不良. 当她去邻居家喝茶时，大儿子先拿了一块饼吃. 然后小儿子来了，吃了他哥哥剩下的部分. 假设两个儿子吃的每一块饼的大小都是随机的，且在当前可用份额上均匀分布. 已知两个儿子都没有消化不良，问剩下饼的期望大小是多少？

提示 所有可能的结果可以用一对数表示，即两个儿子消耗的饼的份额. 因此 Ω 可以选为平面的一个子集. 观察到大儿子只受饼的大小限制，而小儿子受他哥哥剩下的份额限制. 这将决定 Ω 的形状. 接下来在 Ω 上引入与习题条件一致的概率测度. 这可以通过在 Ω 上的适当密度来实现. 现在可以计算两个儿子都不会消化不良的概率. 对应的 Ω 的子集是什么？最后，在 Ω 上定义一个表示儿子们剩下的饼的份额的随机变量，并计算条件期望.

练习 2.14. 作为概率空间，取 $\Omega = [0,1)$，其上的 σ-域为 *Borel* 集，$(0,1)$ 上的测度为 *Lebesgue* 测度. 求 $E(\xi|\eta)$ 若

$$\xi(x) = 2x^2, \qquad \eta(x) = \begin{cases} 2x & \text{当} 0 \leqslant x < \frac{1}{2}, \\ 2x-1 & \text{当} \frac{1}{2} \leqslant x < 1. \end{cases}$$

提示 $\sigma(\eta)$ 中的事件看起来如何？$\sigma(\eta)$-可测随机变量看起来如何？如果能想出一种简洁的描述方式，将使寻找 $E(\xi|\eta)$ 的任务变得容易得多. 需要变换定义 *2.3* 条件 *(2)* 中的积分以找到条件期望的公式.

练习 2.15. 取 $\Omega = [0,1]$，其上的 σ-域为 *Borel* 集，P 为 $[0,1]$ 上的 *Lebesgue* 测度. 设

$$\eta(x) = x(1-x)$$

对 $x \in [0,1]$. 证明

$$E(\xi|\eta)(x) = \frac{\xi(x) + \xi(1-x)}{2}$$

对任意 $x \in [0,1]$.

提示 观察到 $\eta(x) = \eta(1-x)$. 这告诉你关于 η 生成的 σ-域什么信息？$\frac{1}{2}(\xi(x)+\xi(1-x))$ 关于这个 σ-域是否可测？如果是，则只需验证定义 *2.3* 的条件 *(2)*.

练习 2.16. 设 ξ, η 是具有联合密度 $f_{\xi,\eta}(x,y)$ 的可积随机变量. 证明

$$E(\xi|\eta) = \frac{\int_{\mathbb{R}} x f_{\xi,\eta}(x,\eta)\,\mathrm{d}x}{\int_{\mathbb{R}} f_{\xi,\eta}(x,\eta)\,\mathrm{d}x}.$$

提示 研究练习 *2.7* 和 *2.8* 的解答.

注记 2.3. 若记

$$f_{\xi,\eta}(x|y) = \frac{f_{\xi,\eta}(x,y)}{f_\eta(y)},$$

其中

$$f_\eta(y) = \int_{\mathbb{R}} f_{\xi,\eta}(x,y)\,\mathrm{d}x$$

是 η 的密度，则由练习 *2.16* 的结果

$$E(\xi|\eta) = \int_{\mathbb{R}} x f_{\xi,\eta}(x|\eta)\,\mathrm{d}x.$$

称 $f_{\xi,\eta}(x|y)$ 为给定 η 时 ξ 的条件密度.

练习 2.17. 考虑 $L^2(\mathcal{F}) = L^2(\Omega,\mathcal{F},P)$ 作为 *Hilbert* 空间，其标量积为

$$L^2(\mathcal{F}) \times L^2(\mathcal{F}) \ni (\xi,\zeta) \mapsto E(\xi\zeta) \in \mathbb{R}.$$

证明若 ξ 是 $L^2(\mathcal{F})$ 中的随机变量且 $\mathcal{G}$ 是包含于 $\mathcal{F}$ 的 σ-域，则 $E(\xi|\mathcal{G})$ 是 ξ 在 $L^2(\mathcal{F})$ 的子空间 $L^2(\mathcal{G})$ 上的正交投影，该子空间由 $\mathcal{G}$-可测随机变量组成.

提示 观察到定义 *2.4* 的条件 *(2)* 意味着 $\xi - E(\xi|\mathcal{G})$（在上述标量积的意义下）与任意 $A \in \mathcal{G}$ 的指示函数 $\mathbf{1}_A$ 正交.

第 7 节 解答

解答 2.1. 因为 $P(\Omega) = 1$ 且 $\int_\Omega \xi\,\mathrm{d}P = E(\xi)$，

$$E(\xi|\Omega) = \frac{1}{P(\Omega)} \int_\Omega \xi\,\mathrm{d}P = E(\xi).$$

解答 2.2. 由定义 *2.1*

$$\begin{aligned} E(\mathbf{1}_A|B) &= \frac{1}{P(B)} \int_B \mathbf{1}_A\,\mathrm{d}P \\ &= \frac{1}{P(B)} \int_{A\cap B} \mathrm{d}P \\ &= \frac{P(A\cap B)}{P(B)} \\ &= P(A|B). \end{aligned}$$

解答 2.3. 因为 η 是常数，它只有一个值 $c \in \mathbb{R}$，对此

$$\{\eta = c\} = \Omega.$$

因此 $E(\xi|\eta)$ 是常数且等于

$$E(\xi|\eta)(\omega) = E(\xi|\{\eta = c\}) = E(\xi|\Omega) = E(\xi)$$

对每个 $\omega \in \Omega$. 最后一个等式已在练习 *2.1* 中验证.

解答 2.4. 指示函数 $\mathbf{1}_B$ 取两个值 *1* 和 *0*. 这些值分别在集合

$$\{\mathbf{1}_B = 1\} = B, \quad \{\mathbf{1}_B = 0\} = \Omega \setminus B$$

上取到.

因此，对 $\omega \in B$

$$E(\mathbf{1}_A|\mathbf{1}_B)(\omega) = E(\mathbf{1}_A|B) = P(A|B),$$

如同练习 *2.2*. 类似地，对 $\omega \in \Omega \setminus B$

$$E(\mathbf{1}_A|\mathbf{1}_B)(\omega) = E(\mathbf{1}_A|\Omega \setminus B) = P(A|\Omega \setminus B).$$

解答 2.5. 首先观察到对任意事件 B

$$\int_B E(\xi|B)\,\mathrm{d}P = \int_B \left(\frac{1}{P(B)}\int_B \xi\,\mathrm{d}P\right)\mathrm{d}P = \int_B \xi\,\mathrm{d}P. \tag{2.4}$$

因为 η 是离散的，它有可数个值 $y_1, y_2, \ldots$. 可以假设这些值是两两不同的，即若 $i \neq j$ 则 $y_i \neq y_j$. 则集合 $\{\eta = y_1\}, \{\eta = y_2\}, \ldots$ 两两不交且覆盖整个空间 Ω. 因此，由 *(2.4)*

$$\begin{aligned}
E(E(\xi|\eta)) &= \int_\Omega E(\xi|\eta)\,\mathrm{d}P \\
&= \sum_n \int_{\{\eta = y_n\}} E(\xi|\{\eta = y_n\})\,\mathrm{d}P \\
&= \sum_n \int_{\{\eta = y_n\}} \xi\,\mathrm{d}P \\
&= \int_\Omega \xi\,\mathrm{d}P \\
&= E(\xi).
\end{aligned}$$

解答 2.6. 首先描述由 η 生成的 σ-域 $\sigma(\eta)$. 观察到 η 关于 $\frac{1}{2}$ 对称，

$$\eta(x) = \eta(1-x)$$

对任意 $x \in [0,1]$. 断言 $\sigma(\eta)$ 由所有关于 $\frac{1}{2}$ 对称的 *Borel* 集 $A \subset [0,1]$ 组成，即满足

$$A = 1 - A. \tag{2.5}$$

事实上，若 A 是这样的集合，则 $A = \{\eta \in B\}$，其中

$$B = \{2x : x \in A \cap [0, \frac{1}{2}]\}$$

是 *Borel* 集，所以 $A \in \sigma(\eta)$. 另一方面，若 $A \in \sigma(\eta)$，则存在 $\mathbb{R}$ 中的 *Borel* 集 B 使得 $A = \{\eta \in B\}$. 则

$$\begin{aligned} x \in A &\Leftrightarrow \eta(x) \in B \\ &\Leftrightarrow \eta(1-x) \in B \\ &\Leftrightarrow 1 - x \in A, \end{aligned}$$

所以 A 满足 *(2.5)*.

下面求 $E(\xi|\eta)$. 如果它要是 $\sigma(\eta)$-可测的，它必须关于 $\frac{1}{2}$ 对称，即

$$E(\xi|\eta)(x) = E(\xi|\eta)(1-x)$$

对任意 $x \in [0,1]$. 对任意 $A \in \sigma(\eta)$ 将变换下面的积分以使被积函数关于 $\frac{1}{2}$ 对称：

$$\begin{aligned} \int_A 2x^2 \,\mathrm{d}x &= \int_A x^2 \,\mathrm{d}x + \int_A x^2 \,\mathrm{d}x \\ &= \int_A x^2 \,\mathrm{d}x + \int_{1-A} (1-x)^2 \,\mathrm{d}x \\ &= \int_A x^2 \,\mathrm{d}x + \int_A (1-x)^2 \,\mathrm{d}x \\ &= \int_A \left(x^2 + (1-x)^2\right) \,\mathrm{d}x. \end{aligned}$$

由此可得

$$E(\xi|\eta)(x) = x^2 + (1-x)^2.$$

ξ, η 和 $E(\xi|\eta)$ 的图像如图 *2.3* 所示.

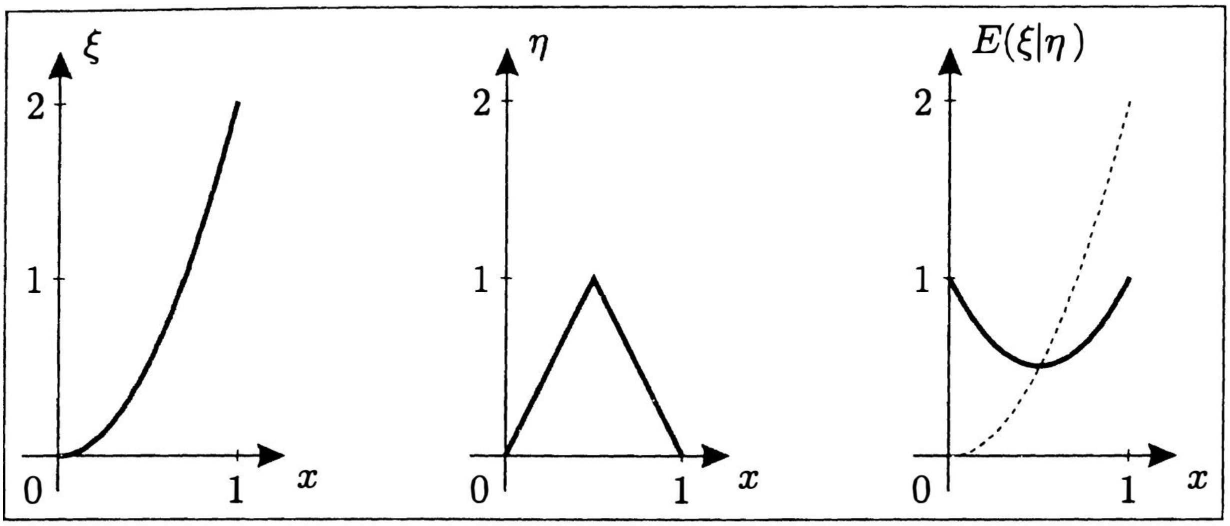

图 2.3: 练习 2.6 中 $E(\xi|\eta)$ 的图像

解答 2.7. 因为对任意 *Borel* 集 B 有 $\{\eta \in B\} = [0,1] \times B$，有

$$\begin{aligned}\int_{\{\eta\in B\}} \xi \, \mathrm{d}P &= \int_B \int_{\mathbb{R}} x f_{\xi,\eta}(x,y) \, \mathrm{d}x \, \mathrm{d}y \\ &= \int_B \left(\int_{[0,1]} x(x+y) \, \mathrm{d}x \right) \mathrm{d}y \\ &= \int_B \left(\frac{1}{3} + \frac{1}{2}y \right) \mathrm{d}y\end{aligned}$$

和

$$\begin{aligned}\int_{\{\eta\in B\}} \frac{2+3\eta}{3+6\eta} \, \mathrm{d}P &= \int_B \int_{\mathbb{R}} \frac{2+3y}{3+6y} f_{\xi,\eta}(x,y) \, \mathrm{d}x \, \mathrm{d}y \\ &= \int_B \frac{2+3y}{3+6y} \left(\int_{[0,1]} (x+y) \, \mathrm{d}x \right) \mathrm{d}y \\ &= \int_B \left(\frac{1}{3} + \frac{1}{2}y \right) \mathrm{d}y.\end{aligned}$$

因为 $\sigma(\eta)$ 中的每个事件都是形如 $\{\eta \in B\}$ 的集合对某个 *Borel* 集 B，这意味着定义 *2.3* 的条件 *(2)* 被满足. 随机变量 $\frac{2+3\eta}{3+6\eta}$ 是 $\sigma(\eta)$-可测的，所以条件 *(1)* 也成立. 由此可得

$$E(\xi|\eta) = \frac{2+3\eta}{3+6\eta}.$$

解答 2.8. 要找 *Borel* 函数 $F : \mathbb{R} \to \mathbb{R}$ 使得对任意 *Borel* 集 B

$$\int_{\{\eta\in B\}} \xi \, \mathrm{d}P = \int_{\{\eta\in B\}} F(\eta) \, \mathrm{d}P. \tag{2.6}$$

则由定义 2.3 有 $E(\xi|\eta)=F(\eta)$.

将用联合密度 $f_{\xi,\eta}(x,y)$ 以与练习 2.7 的解答非常相似的方式变换上面的两个积分，只是这里不知道 $F(y)$ 的确切形式. 即

$$\begin{aligned}\int_{\{\eta\in B\}}\xi\,\mathrm{d}P&=\int_B\int_{\mathbb{R}}xf_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\&=\frac{3}{2}\int_B\left(\int_{[0,1]}x(x^2+y^2)\,\mathrm{d}x\right)\,\mathrm{d}y\\&=\frac{3}{2}\int_B\left(\frac{1}{4}+\frac{1}{2}y^2\right)\,\mathrm{d}y\end{aligned}$$

和

$$\begin{aligned}\int_{\{\eta\in B\}}F(\eta)\,\mathrm{d}P&=\int_B\int_{\mathbb{R}}F(y)f_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\&=\frac{3}{2}\int_B F(y)\left(\int_{[0,1]}(x^2+y^2)\,\mathrm{d}x\right)\,\mathrm{d}y\\&=\frac{3}{2}\int_B F(y)\left(\frac{1}{3}+y^2\right)\,\mathrm{d}y.\end{aligned}$$

则，(2.6) 将对任意 Borel 集 B 成立，如果

$$F(y)=\frac{\frac{1}{4}+\frac{1}{2}y^2}{\frac{1}{3}+y^2}=\frac{3+6y^2}{4+12y^2}.$$

由此可得

$$E(\xi|\eta)=F(\eta)=\frac{3+6\eta^2}{4+12\eta^2}.$$

解答 2.9. 要找 Borel 函数 $F:\mathbb{R}\to\mathbb{R}$ 使得对任意 Borel 集 $B\subset\mathbb{R}$

$$\int_{\{\eta\in B\}}\xi^2\,\mathrm{d}P=\int_{\{\eta\in B\}}F(\eta)\,\mathrm{d}P. \tag{2.7}$$

则由定义 2.3 有 $E(\xi^2|\eta)=F(\eta)$.

变换 (2.7) 的两边. 为此观察到随机变量 ξ 和 η 在单位圆盘 $\Omega=\{(x,y):x^2+y^2\leqslant 1\}$ 上有均匀联合分布，密度为

$$f_{\xi,\eta}(x,y)=\frac{1}{\pi}$$

若 $x^2+y^2 \leqslant 1$，且 $f_{\xi,\eta}(x,y)=0$ 在其他地方. 由此可得

$$\begin{aligned}\int_{\{\eta\in B\}} \xi^2 \,\mathrm{d}P &= \int_B \int_{\mathbb{R}} x^2 f_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{\pi}\int_B \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x^2 \,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{2}{3\pi}\int_B (1-y^2)^{3/2}\,\mathrm{d}y\end{aligned}$$

和

$$\begin{aligned}\int_{\{\eta\in B\}} F(\eta) \,\mathrm{d}P &= \int_B \int_{\mathbb{R}} F(y) f_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{\pi}\int_B F(y) \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{2}{\pi}\int_B F(y)(1-y^2)^{1/2}\,\mathrm{d}y.\end{aligned}$$

若 *(2.7)* 要对所有 *Borel* 集 B 满足，则

$$F(y) = \frac{1}{3}(1-y^2).$$

这意味着

$$E(\xi^2|\eta)(x,y) = F(\eta(x,y)) = F(y) = \frac{1}{3}(1-y^2)$$

对 Ω 中任意 (x,y).

解答 2.10. 若 $\mathcal{G}=\{\emptyset,\Omega\}$，则任何常值随机变量都是 $\mathcal{G}$-可测的. 因为

$$\int_\Omega \xi \,\mathrm{d}P = E(\xi) = \int_\Omega E(\xi)\,\mathrm{d}P$$

和

$$\int_\varnothing \xi \,\mathrm{d}P = 0 = \int_\varnothing E(\xi)\,\mathrm{d}P,$$

由此可得 $E(\xi|\mathcal{G}) = E(\xi)$ *a.s.*，如所要求.

解答 2.11. 因为显然恒等式

$$\int_A \xi \,\mathrm{d}P = \int_A \xi \,\mathrm{d}P$$

对任意 $A\in\mathcal{G}$ 成立且 ξ 是 $\mathcal{G}$-可测的，由此可得 $E(\xi|\mathcal{G}) = \xi$ *a.s.*.

解答 2.12. 由定义 *2.3*

$$\int_B E(\xi|\mathcal{G})\,\mathrm{d}P = \int_B \xi\,\mathrm{d}P,$$

因为 $B \in \mathcal{G}$. 由此可得

$$\begin{aligned} E(E(\xi|\mathcal{G})|B) &= \frac{1}{P(B)} \int_B E(\xi|\mathcal{G}) \,\mathrm{d}P \\ &= \frac{1}{P(B)} \int_B \xi \,\mathrm{d}P \\ &= E(\xi|B). \end{aligned}$$

解答 2.13. 整个饼用区间 $[0,1]$ 表示. 设 $x \in [0,1]$ 为大儿子消耗的份额. 则 $[0,1-x]$ 将对小儿子可用，他取的份额大小为 $y \in [0,1-x]$. 所有可能结果的集合是

$$\Omega = \{(x,y) : x, y \geqslant 0, x+y \leqslant 1\}.$$

琼斯太太的儿子都不会消化不良的事件是

$$A = \left\{(x,y) \in \Omega : x, y < \frac{1}{2}\right\}.$$

这些集合如图 *2.4* 所示. 若 x 在 $[0,1]$ 上均匀分布且 y 在 $[0,1-x]$ 上均匀分布，则 Ω 上的概率测度 P 具有密度

$$f(x,y) = \frac{1}{1-x}$$

将描述结果 (x,y) 的联合分布，见图 *2.5*.

现在可以计算

$$\begin{aligned} P(A) &= \int_A f(x,y) \,\mathrm{d}x \,\mathrm{d}y \\ &= \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} \frac{1}{1-x} \,\mathrm{d}x \,\mathrm{d}y \\ &= \ln \sqrt{2}. \end{aligned}$$

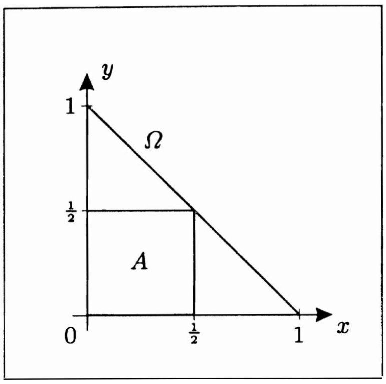

图 2.4: 练习 2.13 中的集合 Ω 和 A

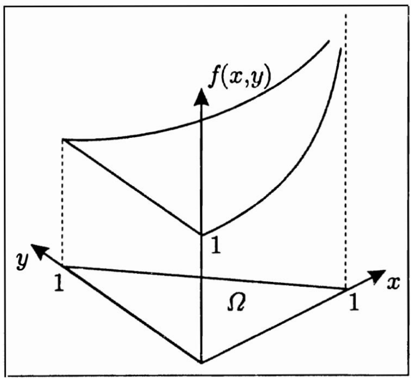

图 2.5: 练习 2.13 中的密度 $f(x,y)$

定义在 Ω 上的随机变量

$$\xi(x,y)=1-x-y$$

表示琼斯太太的儿子们剩下的饼的大小. 最后，得到

$$\begin{aligned} E(\xi|A) &= \frac{1}{P(A)}\int_A (1-x-y)f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{\ln\sqrt{2}}\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}}\frac{1-x-y}{1-x}\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1-\ln\sqrt{2}}{\ln 4}. \end{aligned}$$

解答 2.14. 由 η 生成的 σ-域 $\sigma(\eta)$ 由形如 $B\cup(B+\frac{1}{2})$ 的集合组成，其中 $B\subset[0,\frac{1}{2})$ 是某个 *Borel* 集. 因此，要找 $\sigma(\eta)$-可测随机变量 ζ 使得对每个 *Borel* 集 $B\subset[0,\frac{1}{2})$

$$\int_{B\cup(B+\frac{1}{2})}\xi(x)\,\mathrm{d}x=\int_{B\cup(B+\frac{1}{2})}\zeta(x)\,\mathrm{d}x. \tag{2.8}$$

则由定义 *2.3* 有 $E(\xi|\eta)=\zeta$.

变换左边的积分，得到

$$\begin{aligned} \int_{B\cup(B+\frac{1}{2})}\xi(x)\,\mathrm{d}x &= \int_B 2x^2\,\mathrm{d}x+\int_{B+\frac{1}{2}}2x^2\,\mathrm{d}x \\ &= \int_B 2x^2\,\mathrm{d}x+\int_B 2\left(x+\frac{1}{2}\right)^2\mathrm{d}x \\ &= 2\int_B\left(x^2+\left(x+\frac{1}{2}\right)^2\right)\mathrm{d}x. \end{aligned}$$

为使 ζ 是 $\sigma(\eta)$-可测的，它必须满足

$$\zeta(x)=\zeta\left(x+\frac{1}{2}\right) \tag{2.9}$$

对每个 $x\in[0,\frac{1}{2})$. 则

$$\begin{aligned}\int_{B\cup(B+\frac{1}{2})}\zeta(x)\,\mathrm{d}P&=\int_B\zeta(x)\,\mathrm{d}x+\int_{B+\frac{1}{2}}\zeta(x)\,\mathrm{d}x\\&=\int_B\zeta(x)\,\mathrm{d}x+\int_B\zeta\left(x+\frac{1}{2}\right)\,\mathrm{d}x\\&=\int_B\zeta(x)\,\mathrm{d}x+\int_B\zeta(x)\,\mathrm{d}x\\&=2\int_B\zeta(x)\,\mathrm{d}x.\end{aligned}$$

若 *(2.8)* 要对任意 *Borel* 集 $B\subset[0,\frac{1}{2})$ 成立，则

$$\zeta(x)=x^2+\left(x+\frac{1}{2}\right)^2$$

对每个 $x\in[0,\frac{1}{2})$. $\zeta(x)$ 对 $x\in[\frac{1}{2},1)$ 的值可由 *(2.9)* 得到. 由此可得

$$E(\xi|\eta)(x)=\zeta(x)=\begin{cases}x^2+\left(x+\frac{1}{2}\right)^2 & \text{当}\,0\leqslant x<\frac{1}{2},\\ \left(x-\frac{1}{2}\right)^2+x^2 & \text{当}\,\frac{1}{2}\leqslant x<1.\end{cases}$$

ξ,η 和 $E(\xi|\eta)$ 的图像如图 *2.6* 所示.

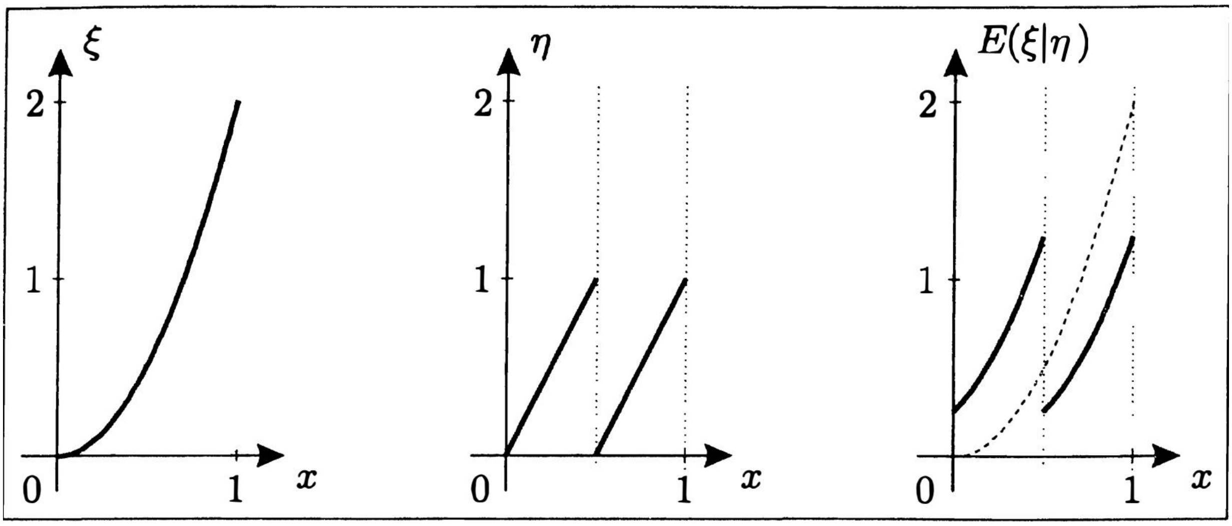

图 2.6: 练习 2.14 中 $E(\xi|\eta)$ 的图像

解答 2.15. 因为 $\eta(x)=\eta(1-x)$，σ-域 $\sigma(\eta)$ 由满足

$$B=1-B$$

的 *Borel* 集 $B\subset[0,1]$ 组成，其中 $1-B=\{1-x:x\in B\}$. 对任意这样的 B

$$\begin{aligned}\int_B \xi(x)\,\mathrm{d}x &= \frac{1}{2}\int_B \xi(x)\,\mathrm{d}x+\frac{1}{2}\int_B \xi(x)\,\mathrm{d}x\\ &= \frac{1}{2}\int_B \xi(x)\,\mathrm{d}x+\frac{1}{2}\int_{1-B} \xi(1-x)\,\mathrm{d}x\\ &= \frac{1}{2}\int_B \xi(x)\,\mathrm{d}x+\frac{1}{2}\int_B \xi(1-x)\,\mathrm{d}x\\ &= \int_B \frac{\xi(x)+\xi(1-x)}{2}\,\mathrm{d}x.\end{aligned}$$

因为 $\frac{1}{2}(\xi(x)+\xi(1-x))$ 是 $\sigma(\eta)$-可测的，由此可得

$$E(\xi|\eta)(x)=\frac{\xi(x)+\xi(1-x)}{2}.$$

解答 2.16. 要找 *Borel* 函数 $F(y)$ 使得

$$\int_{\{\eta\in B\}}\xi\,\mathrm{d}P=\int_{\{\eta\in B\}}F(\eta)\,\mathrm{d}P$$

对 $\mathbb{R}$ 中任意 *Borel* 集 B. 因为 $F(\eta)$ 是 $\sigma(\eta)$-可测的且 $\sigma(\eta)$ 中的每个事件都可以写成 $\{\eta\in B\}$ 的形式对某个 *Borel* 集 B，这意味着 $E(\xi|\eta)=F(\eta)$.

用 ξ 和 η 的联合密度变换上面的两个积分

$$\begin{aligned}\int_{\{\eta\in B\}}\xi\,\mathrm{d}P &= \int_B\int_{\mathbb{R}} x f_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_B\left(\int_{\mathbb{R}} x f_{\xi,\eta}(x,y)\,\mathrm{d}x\right)\mathrm{d}y\end{aligned}$$

和

$$\begin{aligned}\int_{\{\eta\in B\}}F(\eta)\,\mathrm{d}P &= \int_B\int_{\mathbb{R}} F(y) f_{\xi,\eta}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_B F(y)\left(\int_{\mathbb{R}} f_{\xi,\eta}(x,y)\,\mathrm{d}x\right)\mathrm{d}y.\end{aligned}$$

若这两个积分要对每个 *Borel* 集 B 相等，则

$$F(y)=\frac{\int_{\mathbb{R}} x f_{\xi,\eta}(x,y)\,\mathrm{d}x}{\int_{\mathbb{R}} f_{\xi,\eta}(x,y)\,\mathrm{d}x}.$$

由此可得

$$E(\xi|\eta)=F(\eta)=\frac{\int_{\mathbb{R}} x f_{\xi,\eta}(x,\eta)\,\mathrm{d}x}{\int_{\mathbb{R}} f_{\xi,\eta}(x,\eta)\,\mathrm{d}x}.$$

解答 2.17. 用 ζ 表示 ξ 在子空间 $L^2(\mathcal{G})\subset L^2(\mathcal{F})$ 上的正交投影，该子空间由 $\mathcal{G}$-可测随机变量组成. 因此，$\xi-\zeta$ 与 $L^2(\mathcal{G})$ 正交，即

$$E[(\xi-\zeta)\gamma]=0$$

对每个 $\gamma\in L^2(\mathcal{G})$. 但对任意 $A\in\mathcal{G}$ 指示函数 $\mathbf{1}_A$ 属于 $L^2(\mathcal{G})$，所以

$$E[(\xi-\zeta)\mathbf{1}_A]=0.$$

因此

$$\int_A \xi\,\mathrm{d}P=E(\xi\mathbf{1}_A)=E(\zeta\mathbf{1}_A)=\int_A \zeta\,\mathrm{d}P$$

对任意 $A\in\mathcal{G}$. 这意味着 $\zeta=E(\xi|\mathcal{G})$.

第三章
离散时间鞅

第 1 节　随机变量序列

随机变量序列 $\xi_1, \xi_2, \ldots$ 通常用作一系列随机现象结果的数学模型，例如抛硬币或伦敦证券交易所连续工作日 FTSE 全股指数的取值. 这种序列中的随机变量由整数索引，这些整数通常称为 **离散时间**. 需要注意的是，这些整数未必与序列所建模事件的实际发生物理时间相关. 离散时间用于记录事件的顺序，而这些事件在物理时间上可能是均匀分布的，也可能不是. 例如，股票指数仅在工作日记录，周六、周日及其他假日则不予记录. 与其反复抛掷一枚硬币，不如同时抛掷 100 枚硬币并记录结果.

定义 3.1. 对任意固定的 $\omega \in \Omega$，数列 $\xi_1(\omega), \xi_2(\omega), \ldots$ 称为 **样本路径**.

图 3.1 展示了一列抛硬币的样本路径（+1 表示正面，-1 表示反面）. 图 3.2 展示了截至 1997 年 FTSE 全股指数的样本路径. 严格来说，这些图应由点组成，分别代表取值 $\xi_1(\omega), \xi_2(\omega), \ldots$，但出于说明目的，通常用折线将其连接.

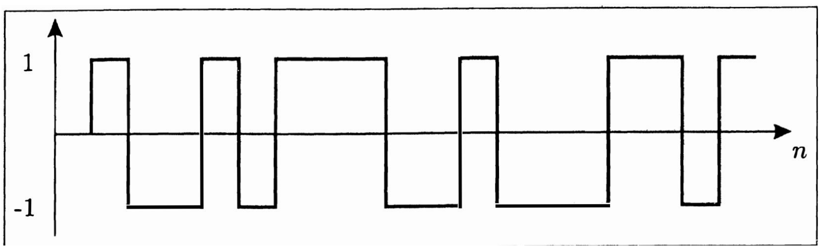

图 3.1: 一列抛硬币的样本路径

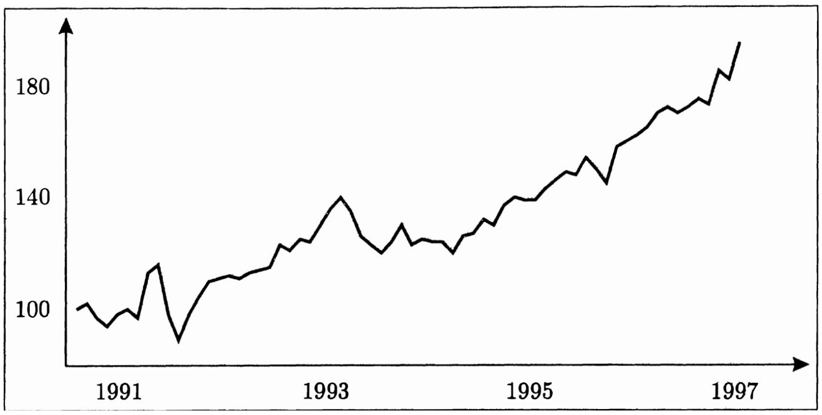

图 3.2: 截至 1997 年的 FTSE 全股指数的样本路径

第 2 节 滤子

随着时间推移，对历史事件的信息不断累积. 这可用下面定义的滤子来建模.

定义 3.2. Ω 上的一列 σ-域 $\mathcal{F}_1, \mathcal{F}_2, \ldots$ 如果满足

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}$$

则称为一个**滤子**.

这里 $\mathcal{F}_n$ 代表时刻 n 的知识状态. 它包含所有事件 A，使得在时刻 n 能够判断 A 是否已经发生. 随着 n 的增加，这样的事件 A 会更多，即代表信息量的族 $\mathcal{F}_n$ 会变大. (经验随时间增长!)

例 3.1. 对于抛硬币序列 $\xi_1, \xi_2, \ldots$，取 $\mathcal{F}_n$ 为由 $\xi_1, \ldots, \xi_n$ 生成的 σ-域，

$$\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n).$$

设

$$A = \{\text{前 5 次抛掷中至少出现 2 次正面}\}.$$

在离散时刻 $n = 5$，即硬币已被抛掷五次后，就能够判断 A 是否已经发生. 这意味着 $A \in \mathcal{F}_5$. 然而，在 $n = 4$ 时，尚无法判断 A 是否已经发生. 若前四次抛掷的结果为

$$\text{反面，反面，正面，反面，}$$

则事件 A 仍无法确定. 需待第五次抛掷后才能知晓结果. 因此 $A \notin \mathcal{F}_4$.

此例说明了另一相关问题. 假设前四次抛硬币的结果为

$$\text{反面，正面，反面，正面.}$$

在此情形下，无论第五次抛掷结果如何，在 $n=4$ 时就已经能够判断 A 已经发生. 但这并不意味着 A 属于 $\mathcal{F}_4$. 关键在于，要使 A 属于 $\mathcal{F}_4$，必须在前四次抛掷之后，无论前四个结果是什么，都能判断 A 是否已经发生. 在本例中，这显然不成立.

练习 3.1. 设 $\xi_1, \xi_2, \ldots$ 是一个抛硬币序列，$\mathcal{F}_n$ 是由 $\xi_1, \ldots, \xi_n$ 生成的 σ-域. 对下列每个事件，找出使得该事件属于 $\mathcal{F}_n$ 的最小 n:

$$A = \{\text{第一次出现正面之前至多出现 10 次反面}\},$$

$$B = \{\text{在序列}\xi_1, \xi_2, \ldots \text{中至少出现 1 次正面}\},$$

$$C = \{\text{前 100 次抛掷产生相同的结果}\}.$$

提示 要找出一个数集中的最小元素，首先需要确保该集合非空.

假设 $\xi_1, \xi_2, \ldots$ 是一个随机变量序列，$\mathcal{F}_1, \mathcal{F}_2, \ldots$ 是一个滤子. 先验而言，二者可能毫无关联. 然而，在实际应用中，滤子通常包含观测随机序列所累积的信息，如例 3.1 所示. 下面定义中的条件意味着 $\mathcal{F}_n$ 包含 $\xi_1, \ldots, \xi_n$ 能提供的全部信息. 一般来说，其往往包含更多信息.

定义 3.3. 如果随机变量序列 $\xi_1, \xi_2, \ldots$ 满足对每个 $n = 1, 2, \ldots$，ξ_n 都是 $\mathcal{F}_n$-可测的，则称该序列**适应于**滤子 $\mathcal{F}_1, \mathcal{F}_2, \ldots$.

例 3.2. 如果 $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ 是由 $\xi_1, \ldots, \xi_n$ 生成的 σ-域，则 $\xi_1, \xi_2, \ldots$ 适应于 $\mathcal{F}_1, \mathcal{F}_2, \ldots$.

练习 3.2. 证明

$$\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$$

是使得序列 $\xi_1, \xi_2, \ldots$ 适应于 $\mathcal{F}_1, \mathcal{F}_2, \ldots$ 的最小滤子. 也就是说，如果 $\mathcal{G}_1, \mathcal{G}_2, \ldots$ 是另一个滤子使得 $\xi_1, \xi_2, \ldots$ 适应于 $\mathcal{G}_1, \mathcal{G}_2, \ldots$，则对每个 n 都有 $\mathcal{F}_n \subset \mathcal{G}_n$.

提示 要使 $\sigma(\xi_1, \ldots, \xi_n)$ 包含于 $\mathcal{G}_n$，需要证明 $\xi_1, \ldots, \xi_n$ 都是 $\mathcal{G}_n$-可测的.

第 3 节 鞅

鞅的概念起源于赌博，描述了一种公平的赌博游戏，下一节将对此展开详细讨论. 类似地，下鞅和上鞅（将在下文定义）的概念分别与有利和不利机会的赌博游戏相关. 赌博的某些方面本质上存在于金融数学中，尤其是期权等金融衍生品理论中. 鞅在其中发挥着至关重要的作用，这一点不足为奇. 事实上，鞅不仅限于博弈论，还出现在现代概率与随机分析的诸多领域，尤其是扩散理论中. 下面首先在离散时间情形下介绍其基本定义与性质.

定义 3.4. 随机变量序列 $\xi_1, \xi_2, \ldots$ 称为关于滤子 $\mathcal{F}_1, \mathcal{F}_2, \ldots$ 的**鞅**，如果

(i) 对每个 $n=1,2,\ldots$，ξ_n 是**可积的**；

(ii) $\xi_1,\xi_2,\ldots$ 是**适应的**于 $\mathcal{F}_1,\mathcal{F}_2,\ldots$；

(iii) 对每个 $n=1,2,\ldots$，$E(\xi_{n+1}\mid\mathcal{F}_n)=\xi_n$ *a.s.*

例 3.3. 设 $\eta_1,\eta_2,\ldots$ 是一列独立的可积随机变量，满足对所有 $n=1,2,\ldots$，$E(\eta_n)=0$. 记

$$\begin{aligned}&\xi_n=\eta_1+\cdots+\eta_n,\\&\mathcal{F}_n=\sigma(\eta_1,\ldots,\eta_n).\end{aligned}$$

则 ξ_n 适应于滤子 $\mathcal{F}_n$，且可积，因为

$$\begin{aligned}&E(|\xi_n|)=E(|\eta_1+\cdots+\eta_n|)\\&\leqslant E\left(|\eta_1|\right)+\cdots+E\left(|\eta_n|\right)\\&<\infty.\end{aligned}$$

此外，

$$\begin{aligned}&E\left(\xi_{n+1}\mid\mathcal{F}_n\right)=E\left(\eta_{n+1}\mid\mathcal{F}_n\right)+E\left(\xi_n\mid\mathcal{F}_n\right)\\&=E\left(\eta_{n+1}\right)+\xi_n\\&=\xi_n,\end{aligned}$$

其中用到 η_{n+1} 独立于 $\mathcal{F}_n$（“独立的条件消失”）且 ξ_n 是 $\mathcal{F}_n$-可测的（“拿出已知的东西”）. 这表明 ξ_n 是关于 $\mathcal{F}_n$ 的鞅.

例 3.4. 设 ξ 是一个可积随机变量，$\mathcal{F}_1,\mathcal{F}_2,\ldots$ 是一个滤子. 记

$$\xi_n=E(\xi\mid\mathcal{F}_n)$$

对于 $n=1,2,\ldots$

则 ξ_n 是 $\mathcal{F}_n$-可测的，

$$|\xi_n|=|E(\xi\mid\mathcal{F}_n)|\leqslant E(|\xi|\mid\mathcal{F}_n),$$

这意味着

$$E(|\xi_n|)\leqslant E(E(|\xi|\mid\mathcal{F}_n))=E(|\xi|)<\infty,$$

并且

$$E\left(\xi_{n+1}\mid\mathcal{F}_n\right)=E\left(E\left(\xi\mid\mathcal{F}_{n+1}\right)\mid\mathcal{F}_n\right)=E\left(\xi\mid\mathcal{F}_n\right)=\xi_n,$$

因为 $\mathcal{F}_n\subset\mathcal{F}_{n+1}$（条件期望的塔性质）. 因此 ξ_n 是关于 $\mathcal{F}_n$ 的鞅.

练习 3.3. 证明若 ξ_n 是关于 $\mathcal{F}_n$ 的鞅，则

$$E(\xi_1)=E(\xi_2)=\ldots.$$

提示：$E(\xi_{n+1} \mid \mathcal{F}_n)$ 的期望是多少？

练习 3.4. 假设 ξ_n 是关于滤子 $\mathcal{F}_n$ 的鞅. 证明 ξ_n 是关于滤子

$$\mathcal{G}_n = \sigma(\xi_1, \ldots, \xi_n)$$

的鞅.

提示：注意到 $\mathcal{G}_n \subset \mathcal{F}_n$ 并使用条件期望的塔性质.

练习 3.5. 设 ξ_n 是一个**对称随机游走**，即

$$\xi_n = \eta_1 + \cdots + \eta_n,$$

其中 $\eta_1, \eta_2, \ldots$ 是一列独立的**同分布的**随机变量，满足

$$P\{\eta_n = 1\} = P\{\eta_n = -1\} = \frac{1}{2}$$

（例如，一系列掷硬币）. 证明 $\xi_n^2 - n$ 是关于滤子

$$\mathcal{F}_n = \sigma(\eta_1, \ldots, \eta_n)$$

的鞅.

提示：需将 $E\left(\xi_{n+1}^2 - (n+1) \mid \mathcal{F}_n\right)$ 变换为 $\xi_n^2 - n$. 写出

$$\begin{aligned} \xi_{n+1}^2 &= \left(\xi_n + \eta_{n+1}\right)^2 \\ &= \eta_{n+1}^2 + 2\eta_{n+1}\xi_n + \xi_n^2 \end{aligned}$$

并注意到 ξ_n 是 $\mathcal{F}_n$-可测的，而 η_{n+1} 独立于 $\mathcal{F}_n$. 要变换条件期望，可“拿出已知的东西”并使用“独立的条件消失”这一事实. 不要忘记验证 $\xi_n^2 - n$ 是可积的且适应于 $\mathcal{F}_n$.

练习 3.6. 设 ξ_n 是一个对称随机游走，$\mathcal{F}_n$ 是练习 *3.5* 中定义的滤子. 证明

$$\zeta_n = (-1)^n \cos(\pi \xi_n)$$

是关于 $\mathcal{F}_n$ 的鞅.

提示：需将 $E\big((-1)^{n+1}\cos(\pi\xi_{n+1}) \mid \mathcal{F}_n\big)$ 变换为 $(-1)^n \cos(\pi\xi_n)$. 使用与练习 *3.5* 中类似的论证来实现这一点. 但是，首先确保 ζ_n 是可积的且适应于 $\mathcal{F}_n$.

定义 3.5. 若满足以下条件，则称 $\xi_1, \xi_2, \ldots$ 是关于滤子 $\mathcal{F}_1, \mathcal{F}_2, \ldots$ 的**上鞅**（**下鞅**）:

(i) 对每个 $n = 1, 2, \ldots$，ξ_n 是可积的；

(ii) $\xi_1, \xi_2, \ldots$ 是适应于 $\mathcal{F}_1, \mathcal{F}_2, \ldots$;
(iii) 对每个 $n = 1, 2, \ldots$, $E(\xi_{n+1} \mid \mathcal{F}_n) \leqslant \xi_n$（相应地，$E(\xi_{n+1} \mid \mathcal{F}_n) \geqslant \xi_n$）*a.s.*

练习 3.7. 设 ξ_n 是一列平方可积随机变量. 证明若 ξ_n 是关于滤子 $\mathcal{F}_n$ 的鞅，则 ξ_n^2 是关于同一滤子的下鞅.

提示：使用 *Jensen* 不等式，其中凸函数 $\varphi(x) = x^2$.

第 4 节 赌博博弈

假设你参与一种赌博游戏，例如轮盘赌. 设 $\eta_1, \eta_2, \ldots$ 为一列可积随机变量，其中 η_n 表示你在第 n 场游戏中每单位赌注的赢利（或亏损）. 若每场游戏的赌注为一，则你在 n 场游戏后的总赢利为

$$\xi_n = \eta_1 + \cdots + \eta_n.$$

取滤子

$$\mathcal{F}_n = \sigma(\eta_1, \ldots, \eta_n)$$

且为记号简便，记 $\xi_0 = 0$ 及 $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

若截至目前已进行 $n-1$ 轮游戏，所积累的信息由 σ-域 $\mathcal{F}_{n-1}$ 表示. 若

$$E(\xi_n \mid \mathcal{F}_{n-1}) = \xi_{n-1},$$

则该游戏是公平的，即第 n 步时的财富期望与第 $n-1$ 步时相同. 若

$$E\left(\xi_n \mid \mathcal{F}_{n-1}\right) \geqslant \xi_{n-1},$$

则该游戏对你有利；若

$$E\left(\xi_n \mid \mathcal{F}_{n-1}\right) \leqslant \xi_{n-1}$$

对 $n = 1, 2, \ldots$ 成立，则该游戏对你不利. 这三种情形分别对应于 ξ_n 关于 $\mathcal{F}_n$ 是鞅、下鞅或上鞅，参见定义 3.4 和 3.5.

假设你可以在每场游戏中改变赌注为 α_n.（特别地，若放弃第 n 场游戏，α_n 可取零；若你经营赌场并接受他人赌注，α_n 甚至可为负.）决定赌注 α_n 时，你已知晓前 $n-1$ 场游戏的结果. 因此，合理假设 α_n 是 $\mathcal{F}_{n-1}$-可测的，其中 $\mathcal{F}_{n-1}$ 表示你在第 $n-1$ 场游戏及之前积累的信息. 特别地，由于在第一场游戏前一无所知，取 $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

定义 3.6. 一个赌博策略 $\alpha_1, \alpha_2, \ldots$(关于滤子 $\mathcal{F}_1, \mathcal{F}_2, \ldots$)是指一列随机变量,使得对每个 $n = 1, 2, \ldots$, α_n 是 $\mathcal{F}_{n-1}$-可测的,其中 $\mathcal{F}_0 = \{\emptyset, \Omega\}$.(在赌博的语境之外,这样的随机变量序列 α_n 称为**可预见的**.)

若遵循策略 $\alpha_1, \alpha_2, \ldots$,则你在 n 场游戏后的总赢利为

$$\begin{aligned} \zeta_n &= \alpha_1 \eta_1 + \cdots + \alpha_n \eta_n \\ &= \alpha_1 (\xi_1 - \xi_0) + \cdots + \alpha_n (\xi_n - \xi_{n-1}). \end{aligned}$$

为方便起见,还令 $\zeta_0 = 0$.

下面的命题对赌徒有着重要的意义. 它表明:公平的游戏无论采用何种赌博策略都永远是公平的. 若你不能投注负金额(例如经营赌场),则不可能把对你不利的游戏变成有利的,反之亦然. 你无法战胜系统!序列 α_n 的有界性意味着你的可用资金和信用额度均是有界的.

命题 3.1. 设 $\alpha_1, \alpha_2, \ldots$ 为一个赌博策略.

(1) 若 $\alpha_1, \alpha_2, \ldots$ 是**有界的**序列且 $\xi_0, \xi_1, \xi_2, \ldots$ 是鞅,则 $\zeta_0, \zeta_1, \zeta_2, \ldots$ 也是鞅(无论你做什么,公平的游戏仍是公平的);

(2) 若 $\alpha_1, \alpha_2, \ldots$ 是**非负的**、**有界的**序列且 $\xi_0, \xi_1, \xi_2, \ldots$ 是上鞅,则 $\zeta_0, \zeta_1, \zeta_2, \ldots$ 也是上鞅(不利的游戏仍是不利的).

(3) 若 $\alpha_1, \alpha_2, \ldots$ 是**非负的**、**有界的**序列且 $\xi_0, \xi_1, \xi_2, \ldots$ 是下鞅,则 $\zeta_0, \zeta_1, \zeta_2, \ldots$ 也是下鞅(有利的游戏仍是有利的).

证明. 由于 α_n 和 ζ_{n-1} 是 $\mathcal{F}_{n-1}$-可测的,可将它们从关于 $\mathcal{F}_{n-1}$ 的条件期望中提出(“取出已知量”,命题 2.4). 于是得到

$$\begin{aligned} E(\zeta_n \mid \mathcal{F}_{n-1}) &= E(\zeta_{n-1} + \alpha_n (\xi_n - \xi_{n-1}) \mid \mathcal{F}_{n-1}) \\ &= \zeta_{n-1} + \alpha_n (E(\xi_n \mid \mathcal{F}_{n-1}) - \xi_{n-1}). \end{aligned}$$

若 ξ_n 是鞅,则

$$\alpha_n (E(\xi_n \mid \mathcal{F}_{n-1}) - \xi_{n-1}) = 0,$$

这就证明了断言 (1). 若 ξ_n 是上鞅且 $\alpha_n \geqslant 0$,则

$$\alpha_n (E(\xi_n \mid \mathcal{F}_{n-1}) - \xi_{n-1}) \leqslant 0,$$

证明了断言 (2). 最后,若 ξ_n 是下鞅且 $\alpha_n \geqslant 0$,则

$$\alpha_n (E(\xi_n \mid \mathcal{F}_{n-1}) - \xi_{n-1}) \geqslant 0,$$

这就证明了断言 (3). □

第 5 节 停时

在轮盘赌及诸多赌博游戏中，参与者往往可随时选择退出. 退出前已进行的轮数记为 τ. 它可预先固定，例如若某人决定无论结果如何都在 10 轮后停止，则可设 $\tau = 10$. 但一般而言，是否退出的决定通常基于每轮结束后积累的信息做出. 因此，τ 假定为取值于 $\{1, 2, \ldots\} \cup \{\infty\}$ 的随机变量. 引入无穷大是为了涵盖理论上游戏永不终止的情形（这也是某些赌场梦寐以求的场景）. 在每一步 n，都应能判定是否停止游戏，即 $\tau = n$ 是否成立. 故事件 $\{\tau = n\}$ 应属于刻画时刻 n 所掌握信息的 σ-域 $\mathcal{F}_n$. 由此引出如下定义.

定义 3.7. 取值于 $\{1, 2, \ldots\} \cup \{\infty\}$ 的随机变量 τ 称为**停时**（关于滤子 $\mathcal{F}_n$），若对每个 $n = 1, 2, \ldots$

$$\{\tau = n\} \in \mathcal{F}_n.$$

练习 3.8. 证明以下条件等价：

(1) $\{\tau \leqslant n\} \in \mathcal{F}_n$ 对每个 $n = 1, 2, \ldots$ 成立；

(2) $\{\tau = n\} \in \mathcal{F}_n$ 对每个 $n = 1, 2, \ldots$ 成立.

提示 能否用事件 $\{\tau = k\}$（其中 $k = 1, \ldots, n$）表示 $\{\tau \leqslant n\}$？能否用事件 $\{\tau \leqslant k\}$（其中 $k = 1, \ldots, n$）表示 $\{\tau = n\}$？

例 3.5（首中时）. 假设反复抛掷一枚硬币，根据落地结果赢得或输掉 £1. 若初始资金为 £5，并决定玩到拥有 £10 或输光为止. 设 ξ_n 为第 n 步时的资金，则停止游戏的时刻为

$$\tau = \min\{n : \xi_n = 10 \text{ 或} 0\},$$

称为（随机序列 ξ_n 首次到达 10 或 0 的）**首中时**. 它关于滤子 $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ 是停时. 这是因为

$$\{\tau = n\} = \{0 < \xi_1 < 10\} \cap \cdots \cap \{0 < \xi_{n-1} < 10\} \cap \{\xi_n = 10 \text{ 或} 0\}.$$

右边每个集合均属于 $\mathcal{F}_n$，故其交也属于 $\mathcal{F}_n$. 这证明了

$$\{\tau = n\} \in \mathcal{F}_n$$

对每个 n 成立，因此 τ 是停时.

练习 3.9. 设 ξ_n 为适应于滤子 $\mathcal{F}_n$ 的随机变量序列，$B \subset \mathbb{R}$ 为 *Borel* 集. 证明 ξ_n 首次进入 B 的时刻

$$\tau = \min\{n : \xi_n \in B\}$$

是停时.

提示 例 3.5 涵盖了 $B=(-\infty,0]\cup[10,\infty)$ 的情形. 将论证推广至任意 *Borel* 集 B.

设 ξ_n 为适应于滤子 $\mathcal{F}_n$ 的随机变量序列，τ 为（关于同一滤子的）停时. 若 ξ_n 表示 n 轮游戏后的盈亏，而决定在 τ 轮后退出，则总盈亏为 ξ_τ. 此时，n 轮后的实际盈亏为 $\xi_{\tau\wedge n}$. 这里 $a\wedge b$ 表示 a 与 b 中较小者，

$$a\wedge b=\min(a,b).$$

定义 3.8. 称 $\xi_{\tau\wedge n}$ 为在 τ 处**停止的序列**. 通常记为 ξ_n^τ. 因此，对每个 $\omega\in\Omega$

$$\xi_n^\tau(\omega)=\xi_{\tau(\omega)\wedge n}(\omega).$$

练习 3.10. 证明若 ξ_n 为适应于滤子 $\mathcal{F}_n$ 的随机变量序列，则序列 $\xi_{\tau\wedge n}$ 也是适应的.

提示 对任意 *Borel* 集 B，将 $\{\xi_{\tau\wedge n}\in B\}$ 用事件 $\{\xi_k\in B\}$ 与 $\{\tau=k\}$（其中 $k=1,\ldots,n$）表示.

已知使用赌博策略无法将公平游戏变为不公平，也无法将不利游戏变为有利，反之亦然. 下述命题表明，使用停时同样无法实现这一点（本质上是因为停止本身也是一种赌博策略）.

命题 3.2. 设 τ 为停时.

(1) 若 ξ_n 是鞅，则 $\xi_{\tau\wedge n}$ 也是鞅；

(2) 若 ξ_n 是上鞅，则 $\xi_{\tau\wedge n}$ 也是上鞅；

(3) 若 ξ_n 是下鞅，则 $\xi_{\tau\wedge n}$ 也是下鞅.

证明. 这实际上是命题 3.1 的推论. 给定停时 τ，记

$$\alpha_n=\begin{cases}1 & 若 \tau\geqslant n,\\ 0 & 若 \tau<n.\end{cases}$$

断言 α_n 是一个赌博策略（即 α_n 是 $\mathcal{F}_{n-1}$-可测的）. 这是因为对任意 Borel 集 $B\subset\mathbb{R}$，逆像 $\{\alpha_n\in B\}$ 等于

$$\varnothing\in\mathcal{F}_{n-1}$$

当 $0,1\notin B$ 时；或等于

$$\Omega\in\mathcal{F}_{n-1}$$

当 $0,1\in B$ 时；或等于

$$\{\alpha_n=1\}=\{\tau\geqslant n\}=\{\tau>n-1\}\in\mathcal{F}_{n-1}$$

当 $1\in B$ 且 $0\notin B$ 时；或等于

$$\{\alpha_n=0\}=\{\tau<n\}=\{\tau\leqslant n-1\}\in\mathcal{F}_{n-1}$$

当 $1 \notin B$ 且 $0 \in B$ 时. 对该赌博策略有

$$\xi_{\tau\wedge n} = \alpha_1(\xi_1 - \xi_0) + \cdots + \alpha_n(\xi_n - \xi_{n-1}).$$

因此命题 3.1 蕴含了上述断言 (1)、(2) 和 (3). □

例 3.6. *(若有无限资金与无限时间，便可尝试击败系统.）以下赌博策略称为"the martingale". （请勿与本节前面鞅的一般定义混淆.）假设反复抛掷一枚硬币. 用 $\eta_1, \eta_2, \ldots$ 表示结果，取值为 $+1$（正面）或 -1（反面）. 在正面上下注 £1. 若获胜，则退出. 若失败，则将赌注加倍再玩一次. 若此次获胜，则退出. 否则再次加倍赌注，依此类推. 因此，赌博策略为*

$$\alpha_n = \begin{cases} 2^{n-1} & \text{若}\eta_1 = \cdots = \eta_{n-1} = \text{反面}, \\ 0 & \text{其他情况}. \end{cases}$$

记

$$\zeta_n = \eta_1 + 2\eta_2 + \cdots + 2^{n-1}\eta_n$$

并考虑停时

$$\tau = \min\{n : \eta_n = \text{正面}\}.$$

则 $\zeta_{\tau\wedge n}$ 为 n 轮后的盈亏. 它是鞅（请验证!）.

可以证明 $P\{\tau < \infty\} = 1$（在序列 $\eta_1, \eta_2, \ldots$ 中，正面最终将以概率 1 出现）. 因此考虑 ζ_τ 是有意义的. 若能一直玩下去，无论第一次正面出现需要多久，这将成为总盈亏. 这需要无限的时间与资金. 若能负担得起，最终必将获胜，因为 $\zeta_\tau = 1$ 恒成立，这是由于对任意 n 有

$$-1 - 2 - \cdots - 2^{n-1} + 2^n = 1.$$

练习 3.11. *证明若赌徒采用"the martingale"策略，其在最终获胜前的期望损失为无穷大，即*

$$E(\zeta_{\tau-1}) = -\infty.$$

提示 *游戏在第 n 步终止（即 $\tau = n$）的概率是多少？若 $\tau = n$，则 $\zeta_{\tau-1}$ 等于多少？这将给出 $\zeta_{\tau-1}$ 的所有可能取值及其概率. 现在计算 $\zeta_{\tau-1}$ 的期望.*

第 6 节　可选停时定理

若 ξ_n 是一个鞅，那么特别地，

$$E(\xi_n) = E(\xi_1)$$

对每个 n 成立. 例 3.6 表明，对于停时 τ，$E(\xi_\tau)$ 不一定等于 $E(\xi_1)$. 然而，若等式

$$E(\xi_\tau) = E(\xi_1)$$

确实成立，它将非常有用. 可选停时定理为此提供了充分条件.

定理 3.1（可选停时定理）. 设 ξ_n 是关于滤子 $\mathcal{F}_n$ 的鞅，τ 是停时，且满足以下条件：

(1) $\tau < \infty$ *a.s.*,

(2) ξ_τ 是可积的，

(3) $E(\xi_n 1_{\{\tau>n\}}) \to 0$ 当 $n \to \infty$.

则

$$E(\xi_\tau) = E(\xi_1).$$

证明. 因为

$$\xi_\tau = \xi_{\tau\wedge n} + (\xi_\tau - \xi_n)1_{\{\tau>n\}},$$

所以

$$E(\xi_\tau) = E\left(\xi_{\tau\wedge n}\right) + E\left(\xi_\tau 1_{\{\tau>n\}}\right) - E\left(\xi_n 1_{\{\tau>n\}}\right).$$

由于由命题 3.2 知 $\xi_{\tau\wedge n}$ 是鞅，右边第一项等于

$$E\left(\xi_{\tau\wedge n}\right) = E\left(\xi_1\right).$$

最后一项由假设 3) 趋于零. 中间项

$$E\left(\xi_\tau 1_{\{\tau>n\}}\right) = \sum_{k=n+1}^{\infty} E\left(\xi_k 1_{\{\tau=k\}}\right)$$

当 $n \to \infty$ 时趋于零，因为级数

$$E(\xi_\tau) = \sum_{k=1}^{\infty} E(\xi_k 1_{\{\tau=k\}})$$

由 2) 是收敛的. 于是得到 $E(\xi_\tau) = E(\xi_1)$，证毕. □

例 3.7（随机游走首中时的期望）

设 ξ_n 是如练习 3.5 中的对称随机游走，K 是正整数. 记 ξ_n 首次到达 $\pm K$ 的时间为

$$\tau = \min\{n : |\xi_n| = K\}.$$

由练习 3.9 知 τ 是停时. 由练习 3.5 知 $\xi_n^2 - n$ 是鞅. 若可选停时定理可以应用，则

$$E\left(\xi_\tau^2 - \tau\right) = E\left(\xi_1^2 - 1\right) = 0.$$

由此可求出期望

$$E(\tau) = E\left(\xi_\tau^2\right) = K^2,$$

因为 $|\xi_\tau| = K$.

下面验证可选停时定理的条件 1)–3).

1) 下面证明 $P\{\tau = \infty\} = 0$. 为此估计 $P\{\tau > 2Kn\}$. 可将 $2Kn$ 次抛硬币看作 n 次各含 $2K$ 次抛掷的序列. $\tau > 2Kn$ 的一个必要条件是这 n 个序列中没有一个全是正面. 因此

$$P\{\tau > 2Kn\} \leqslant \left(1 - \frac{1}{2^{2K}}\right)^n \to 0$$

当 $n \to \infty$. 因为 $\{\tau > 2Kn\}$ 对于 $n = 1, 2, \ldots$ 是递减的集合序列（即 $\{\tau > 2Kn\} \supset \{\tau > 2K(n+1)\}$），所以

$$\begin{aligned}
&P\{\tau = \infty\} = P\left(\textstyle\bigcap_{n=1}^{\infty}\{\tau > 2Kn\}\right) \\
&= \lim_{n\to\infty} P\{\tau > 2Kn\} = 0,
\end{aligned}$$

证毕.

2) 下面证明

$$E\left(\left|\xi_\tau^2 - \tau\right|\right) < \infty.$$

事实上，

$$\begin{aligned}
&E(\tau) = \sum_{n=1}^{\infty} nP\{\tau = n\} \\
&= \sum_{n=0}^{\infty}\sum_{k=1}^{2K}(2Kn + k)P\{\tau = 2Kn + k\} \\
&\leqslant \sum_{n=0}^{\infty}\sum_{k=1}^{2K} 2K(n+1)P\{\tau > 2Kn\} \\
&\leqslant 4K^2\sum_{n=0}^{\infty}(n+1)\left(1 - \frac{1}{2^{2K}}\right)^n \\
&< \infty,
\end{aligned}$$

因为对任意 $q \in (-1, 1)$，级数 $\sum_{n=1}^{\infty}(n+1)q^n$ 是收敛的. 这里重复使用了 1) 中对 $P\{\tau > 2Kn\}$

的估计. 此外，$\xi_\tau^2 = K^2$，所以

$$\begin{aligned} &E\left(\left|\xi_\tau^2 - \tau\right|\right) \leqslant E\left(\xi_\tau^2\right) + E(\tau) \\ &= K^2 + E(\tau) \\ &< \infty. \end{aligned}$$

3) 因为在 $\{\tau > n\}$ 上 $\xi_n^2 \leqslant K^2$，

$$E\left(\xi_n^2 1_{\{\tau>n\}}\right) \leqslant K^2 P\{\tau > n\} \to 0$$

当 $n \to \infty$. 此外，

$$E\left(n 1_{\{\tau>n\}}\right) \leqslant E\left(\tau 1_{\{\tau>n\}}\right) \to 0$$

当 $n \to \infty$. 收敛到 0 成立是因为由 2) 知 $E(\tau) < \infty$，且 $\{\tau > n\}$ 是递减的集合序列，其交 $\{\tau = \infty\}$ 的测度为零. 于是得到

$$E\left(\left(\xi_n^2 - n\right) 1_{\{\tau>n\}}\right) \to 0,$$

证毕.

练习 3.12. 设 ξ_n 是对称随机游走，$\mathcal{F}_n$ 是练习 *3.5* 中定义的滤子. 记 τ 为如例 *3.7* 中满足 $\xi_n = K$ 的最小 n. 验证

$$\zeta_n = (-1)^n \cos[\pi(\xi_n + K)]$$

是鞅（见练习 *3.6*）. 然后证明 ζ_n 和 τ 满足可选停时定理的条件，并应用该定理求 $E[(-1)^\tau]$.

提示：等式 $\zeta_\tau = (-1)^\tau$ 是利用可选停时定理计算 $E[(-1)^\tau]$ 的关键. 对于本例，该定理的前两个条件要么是显然的，要么在本章其他地方已经验证过. 为确保条件 *3)* 成立，证明以下不等式可能有所帮助：

$$\left|E\left(\zeta_n 1_{\{\tau>n\}}\right)\right| \leqslant P\{\tau > n\}.$$

使用凸函数 $\varphi(x) = |x|$ 的 *Jensen* 不等式来估计左边. 首先验证 ζ_n 是鞅.

第 7 节　解答

解答 3.1. A 属于 $\mathcal{F}_{11}$，但不属于 $\mathcal{F}_{10}$. 最小的 n 是 *11*.

B 对于任何 n 都不属于 $\mathcal{F}_n$. 不存在最小的 n 使得 $B \in \mathcal{F}_n$. C 属于 $\mathcal{F}_{100}$，但不属于 $\mathcal{F}_{99}$. 最小的 n 是 *100*.

由于 $D = \emptyset$，它对于每个 $n = 1, 2, \ldots$ 都属于 $\mathcal{F}_n$. 这里最小的 n 是 *1*.

解答 3.2. 因为随机变量序列 $\xi_1, \xi_2, \ldots$ 适应于滤子 $\mathcal{G}_1, \mathcal{G}_2, \ldots$，所以对每个 n，ξ_n 是 $\mathcal{G}_n$-可测的. 但是

$$\mathcal{G}_1 \subset \mathcal{G}_2 \subset \ldots,$$

因此对每个 n，$\xi_1, \ldots, \xi_n$ 都是 $\mathcal{G}_n$-可测的. 于是

$$\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n) \subset \mathcal{G}_n$$

对每个 n 成立.

解答 3.3. 对等式两边取期望

$$\xi_n = E\left(\xi_{n+1} \mid \mathcal{F}_n\right),$$

得到

$$E(\xi_n) = E\left(E\left(\xi_{n+1} \mid \mathcal{F}_n\right)\right) = E\left(\xi_{n+1}\right)$$

对每个 n 成立. 证毕.

解答 3.4. 随机变量 ξ_n 是可积的，因为 ξ_n 是关于 $\mathcal{F}_n$ 的鞅. 由于 $\mathcal{G}_n$ 是由 $\xi_1, \ldots, \xi_n$ 生成的 σ-域，因此 ξ_n 适应于 $\mathcal{G}_n$. 最后，因为 $\mathcal{G}_n \subset \mathcal{F}_n$,

$$\begin{aligned}
\xi_n &= E(\xi_n \mid \mathcal{G}_n) \\
&= E\left(E\left(\xi_{n+1} \mid \mathcal{F}_n\right) \mid \mathcal{G}_n\right) \\
&= E\left(\xi_{n+1} \mid \mathcal{G}_n\right)
\end{aligned}$$

根据条件期望的塔性质（命题 *2.4*）. 这就证明了 ξ_n 是关于 $\mathcal{G}_n$ 的鞅.

解答 3.5. 因为

$$\xi_n^2 - n = \left(\eta_1 + \cdots + \eta_n\right)^2 - n$$

是 $\eta_1, \ldots, \eta_n$ 的函数，它关于由 $\eta_1, \ldots, \eta_n$ 生成的 σ-域 $\mathcal{F}_n$ 可测，即 $\xi_n^2 - n$ 适应于 $\mathcal{F}_n$. 由于

$$|\xi_n| = |\eta_1 + \cdots + \eta_n| \leqslant |\eta_1| + \cdots + |\eta_n| = n,$$

因此

$$E\left(\left|\xi_n^2 - n\right|\right) \leqslant E\left(\xi_n^2\right) + n \leqslant n^2 + n < \infty,$$

所以对每个 n，$\xi_n^2 - n$ 是可积的. 因为

$$\xi_{n+1}^2 = \eta_{n+1}^2 + 2\eta_{n+1}\xi_n + \xi_n^2,$$

其中 ξ_n 和 ξ_n^2 是 $\mathcal{F}_n$-可测的，且 η_{n+1} 与 $\mathcal{F}_n$ 独立，故可用命题 *2.4*（“拿出已知量” 和 “独立条件失效”）得到

$$
\begin{aligned}
& E\left(\xi_{n+1}^2 \mid \mathcal{F}_n\right)=E\left(\eta_{n+1}^2 \mid \mathcal{F}_n\right)+2 E\left(\eta_{n+1} \xi_n \mid \mathcal{F}_n\right)+E\left(\xi_n^2 \mid \mathcal{F}_n\right) \\
& =E\left(\eta_{n+1}^2\right)+2 \xi_n E\left(\eta_{n+1}\right)+\xi_n^2 \\
& =1+\xi_n^2 .
\end{aligned}
$$

这意味着

$$
E\left(\xi_{n+1}^2-n-1 \mid \mathcal{F}_n\right)=\xi_n^2-n,
$$

所以 ξ_n^2-n 是一个鞅.

解答 3.6. 作为 ξ_n 的函数，随机变量 ζ_n 对每个 n 都是 $\mathcal{F}_n$-可测的，因为 ξ_n 是. 由于 $|\zeta_n| \leqslant 1$，显然 ζ_n 是可积的. 因为 η_{n+1} 与 $\mathcal{F}_n$ 独立且 ξ_n 是 $\mathcal{F}_n$-可测的，因此

$$
\begin{aligned}
& E\left(\zeta_{n+1} \mid \mathcal{F}_n\right)=E\left((-1)^{n+1} \cos \left[\pi\left(\xi_n+\eta_{n+1}\right)\right] \mid \mathcal{F}_n\right) \\
& =(-1)^{n+1} E\left(\cos (\pi \xi_n) \cos (\pi \eta_{n+1}) \mid \mathcal{F}_n\right) \\
& -(-1)^{n+1} E\left(\sin \left(\pi \xi_n\right) \sin \left(\pi \eta_{n+1}\right) \mid \mathcal{F}_n\right) \\
& =(-1)^{n+1} \cos (\pi \xi_n) E(\cos (\pi \eta_{n+1})) \\
& -(-1)^{n+1} \sin (\pi \xi_n) E(\sin (\pi \eta_{n+1})) \\
& =(-1)^n \cos (\pi \xi_n) \\
& =\zeta_n,
\end{aligned}
$$

利用公式

$$
\cos (\alpha+\beta)=\cos \alpha \cos \beta-\sin \alpha \sin \beta .
$$

为计算 $E\left(\cos (\pi \eta_{n+1})\right)$ 和 $E\left(\sin (\pi \eta_{n+1})\right)$，注意 $\eta_{n+1}=1$ 或 -1 且

$$
\begin{aligned}
& \cos \pi=\cos (-\pi)=-1, \\
& \sin \pi=\sin (-\pi)=0 .
\end{aligned}
$$

因此 ζ_n 是关于滤子 $\mathcal{F}_n$ 的鞅.

解答 3.7. 若 ξ_n 适应于 $\mathcal{F}_n$，则 ξ_n^2 亦然. 由于对每个 n 有 $\xi_n=E\left(\xi_{n+1} \mid \mathcal{F}_n\right)$ 且 $\varphi(x)=x^2$ 是凸函数，应用 *Jensen* 不等式（定理 *2.2*）可得

$$
\xi_n^2=\left[E\left(\xi_{n+1} \mid \mathcal{F}_n\right)\right]^2 \leqslant E\left(\xi_{n+1}^2 \mid \mathcal{F}_n\right)
$$

对每个 n 成立. 这意味着 ξ_n^2 是关于 $\mathcal{F}_n$ 的下鞅.

解答 3.8. 1) $\Rightarrow$ 2). 若 τ 具有性质 *1)*，则

$$\{\tau \leqslant n\} \in \mathcal{F}_n$$

且

$$\{\tau \leqslant n-1\} \in \mathcal{F}_{n-1} \subset \mathcal{F}_n,$$

所以

$$\{\tau = n\} = \{\tau \leqslant n\} \setminus \{\tau \leqslant n-1\} \in \mathcal{F}_n.$$

2) $\Rightarrow$ 1). 若 τ 具有性质 *2)*，则对每个 $k = 1, \ldots, n$ 有

$$\{\tau = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$$

因此

$$\{\tau \leqslant n\} = \{\tau = 1\} \cup \cdots \cup \{\tau = n\} \in \mathcal{F}_n.$$

解答 3.9. 若

$$\tau = \min\{n : \xi_n \in B\},$$

则对任何 n

$$\{\tau = n\} = \{\xi_1 \notin B\} \cap \cdots \cap \{\xi_{n-1} \notin B\} \cap \{\xi_n \in B\}.$$

因为 B 是 *Borel* 集，右边每个集合都属于 σ-域 $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$，它们的交集也属于该域. 这就证明了对每个 n 有 $\{\tau = n\} \in \mathcal{F}_n$，所以 τ 是一个停时.

解答 3.10. 记 $B \subset \mathbb{R}$ 是一个 *Borel* 集. 可写

$$\{\xi_{\tau\wedge n} \in B\} = \{\xi_n \in B, \tau > n\} \cup \bigcup_{k=1}^{n} \{\xi_k \in B, \tau = k\},$$

其中

$$\{\xi_n \in B, \tau > n\} = \{\xi_n \in B\} \cap \{\tau > n\} \in \mathcal{F}_n$$

且对每个 $k = 1, \ldots, n$

$$\{\xi_k \in B, \tau = k\} = \{\xi_k \in B\} \cap \{\tau = k\} \in \mathcal{F}_k \subset \mathcal{F}_n.$$

因此对每个 n

$$\{\xi_{\tau\wedge n} \in B\} \in \mathcal{F}_n,$$

符合要求.

解答 3.11. 鞅在第 n 步终止的概率是

$$P\{\tau=n\}=\frac{1}{2^n}$$

($n-1$ 次反面后跟第 n 次正面). 因此

$$\begin{aligned}&E\left(\zeta_{\tau-1}\right)=\textstyle\sum_{n=1}^{\infty}\zeta_{n-1}P\{\tau=n\}\\&=\textstyle\sum_{n=1}^{\infty}(-1-2-\cdots-2^{n-2})\frac{1}{2^n}\\&=-\textstyle\sum_{1}^{\infty}\frac{2^{n-1}-1}{2^n}=-\infty.\end{aligned}$$

解答 3.12. 证明 ζ_n 是鞅的过程与练习 *3.6* 几乎相同. 下面验证 ζ_n 和 τ 满足可选停时定理的条件 *1)–3)*.

条件 *1)* 实际上已在例 *3.7* 中得到验证.

条件 *2)* 成立，因为 $|\zeta_\tau|\leqslant 1$，所以 $E(|\zeta_\tau|)\leqslant 1<\infty$.

为验证条件 *3)*，注意对所有 n 有 $|\zeta_n|\leqslant 1$，所以

$$\begin{aligned}&\left|E\left(\zeta_n 1_{\{\tau>n\}}\right)\right|\leqslant E\left(|\zeta_n|1_{\{\tau>n\}}\right)\\&\leqslant E(1_{\{\tau>n\}})\\&=P\{\tau>n\}.\end{aligned}$$

事件族 $\{\tau>n\}$，$n=1,2,\ldots$ 是一个递减序列，其交集为 $\{\tau=\infty\}$. 因此

$$\left|E\left(\zeta_n 1_{\{\tau>n\}}\right)\right|\leqslant P\{\tau>n\}\searrow P\{\tau=\infty\}$$

当 $n\to\infty$. 但是

$$P\{\tau=\infty\}=0$$

根据 *1)*，证毕.

可选停时定理意味着

$$E\left(\zeta_\tau\right)=E\left(\zeta_1\right).$$

因为 $\xi_\tau=K$ 或 $-K$，有

$$\zeta_\tau=(-1)^\tau\cos[\pi(K+\xi_\tau)]=(-1)^\tau.$$

下面计算

$$\begin{aligned}&E\left(\zeta_1\right)=-\tfrac{1}{2}\left(\cos[\pi(1+K)]+\cos[\pi(-1+K)]\right)\\&=\cos(\pi K)=(-1)^K.\end{aligned}$$

因此

$$E[(-1)^\tau]=(-1)^K.$$

术语表

外文术语	中文术语
absolutely continuous distribution	绝对连续分布.
adapted	适应的.
almost surely	几乎必然.
Borel function	Borel 函数.
Borel set	Borel 集.
Borel σ-field	Borel σ-域.
Borel-Cantelli lemma	Borel-Cantelli 引理.
bounded	有界的.
composition	复合.
conditional density	条件密度.
conditional expectation	条件期望.
conditional probability	条件概率.
conditioning	条件化.
contracting sequence	递减序列.
convex function	凸函数.
De Morgan’s law	De Morgan 法则.
density	密度函数.
discrete distribution	离散分布.
distribution	分布.
distribution function	分布函数.
Doob-Dynkin lemma	Doob-Dynkin 引理.
events	事件.
expanding sequence	递增序列.
expectation	期望.
filtration	滤子.

外文术语	中文术语
first hitting time	首中时.
gambling strategy	赌博策略.
generated	生成的.
identically distributed	同分布的.
independence	独立性.
independent	独立的.
indicator function	指示函数.
integrable	可积的.
inverse image	逆像.
Jensen's inequality	Jensen 不等式.
joint density	联合密度.
joint distribution	联合分布.
Lebesgue measure	Lebesgue 测度.
linearity	线性性.
martingale	鞅.
measurable	可测的.
monotone convergence	单调收敛.
non-decreasing	单调递增的.
non-negative	非负的.
Optional Stopping Theorem	可选停时定理.
orthogonal projection	正交投影.
pairwise disjoint	两两不交.
pairwise distinct	两两不同.
positivity	正性.
previsible	可预见的.
probability measure	概率测度.
probability space	概率空间.
Radon-Nikodym theorem	Radon-Nikodym 定理.
random variable	随机变量.
right-continuous	右连续的.
sample path	样本路径.
scalar product	标量积.
Schwarz inequality	Schwarz 不等式.
square integrable	平方可积的.
step function	阶梯函数.
stopping time	停时.
subadditivity	次可加性.
submartingale	下鞅.

外文术语	中文术语
supermartingale	上鞅.
symmetric random walk	对称随机游走.
total probability formula	全概率公式.
tower property	塔性质.
uncorrelated	不相关的.
variance	方差.
σ-field	σ-域.