高等概率论讨论班 (第 0 次)
符号测度、条件期望与离散时间鞅

应数 91 陈昱坤　　　　统计 91 董晟渤

西安交通大学数学与统计学院

日期：2022 年 3 月

目录

Part I

测度论部分

1 符号测度与 Radon-Nikodym 定理

1.1 符号测度

定义 1.1. 设 $(X,\mathscr{F})$ 是一个可测空间, 实值函数 $\nu:\mathscr{F}\to\mathbb{R}\cup\{-\infty,\infty\}$ 称为一个**符号测度** (Signed Measure), 如果

(1) $\nu(\varnothing)=0$;

(2) ν 至多达到 $-\infty$ 和 ∞ 中的一个;

(3) ν 具有可数可加性, 即对任意两两互不相交的集列 $\{A_i\}_{i=1}^{\infty}$, 有

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

命题 1.1. 设 (X,μ) 是一个符号测度空间. 如果 $\{E_n\}$ 是一列上升的可测集列, 则

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\lim_{n\to\infty}\mu(E_n).$$

如果 $\{E_n\}$ 是一列下降的可测集列, 并且 $\mu(E_1)<\infty$, 那么

$$\mu\left(\bigcap_{n=1}^{\infty}E_n\right)=\lim_{n\to\infty}\mu(E_n).$$

评论. 为了区别符号测度与通常的测度, 常常称过去定义的测度为**正测度**.

符号测度常常以如下两种方式出现:

- 设 μ_1,μ_2 是 $(X,\mathscr{F})$ 上的正测度, 且至少有其中一个是有限测度时,

  $$\nu(E):=\mu_1(E)-\mu_2(E),\quad \forall E\in\mathscr{F}$$

  定义了一个符号测度.

- 设 f 是测度空间 $(X,\mathscr{F},\mu)$ 上的 L^1 函数, 那么由 f 诱导了符号测度

  $$\nu_f(E):=\int_E f\mathrm{d}\mu=\int_X\chi_E f\mathrm{d}\mu.$$

  特别地, 当 $(\Omega,\mathscr{F},\mathbb{P})$ 是概率空间时, 随机变量 $X\in L^1(\Omega)$ 定义了一个 $\mathscr{F}$(或者其子 σ-域) 上的符号测度:

  $$\varphi(C):=\int_C X\mathrm{d}\mathbb{P}.$$

我们将说明这两种情形是典范的: 前者对应着 Hahn 分解 & Jordan 分解; 后者对应着我们本节的主题——Radon-Nikodym 定理.

1.2 Hahn 分解与 Jordan 分解

为了方便理解, 可以首先我们所研究的测度理解成为可积函数 f 所诱导出来的符号测度 ν_f.

这里, 我们令

$$f^+(x) = \begin{cases} f(x), & f(x) > 0, \\ 0, & f(x) \leq 0. \end{cases}$$

且令 $f^- = f^+ - f$, 那么 $f = f^+ - f^-$. 这里 f^+, f^- 都是非负函数.

那么, 由 f^+ 和 f^- 可以诱导出正测度

$$\mu^+(E) = \int_E f^+ \mathrm{d}\mu,$$

以及

$$\mu^-(E) = \int_E f^- \mathrm{d}\mu,$$

于是

$$\mu(E) = \int_E f \mathrm{d}\mu = \int_E f^+ - f^- \mathrm{d}\mu = \mu^+(E) - \mu^-(E).$$

我们可以观察到上面的分解具有如下性质:

- μ^+ 和 μ^- 存在于不相交的集合上: 即

$$\mu^+(E) \neq 0 \implies \mu^-(E) = 0, \mu^-(E) \neq 0 \implies \mu^+(E) = 0.$$

- f^+ 的支撑集 $\mathrm{supp} f^+$ 满足: 对任意的可测集 $E \subset \mathrm{supp} f^+, \nu_f(E) \geq 0$;$f^-$ 的支撑集 $\mathrm{supp} f^-$ 满足: 对任意的可测集 $E \subset \mathrm{supp} f^-, \nu_f(E) \leq 0$.

由此, 可以引导我们给出下列定义:

定义 1.2. 设 μ 是 $(X, \mathscr{F})$ 上的符号测度, 如果存在集合 $A \in \mathscr{F}$, 使得对任意的 $E \in \mathscr{F}$, 都有 $\mu(E) = \mu(A \cap E)$, 则称 μ **集中于** A. 设 μ_1, μ_2 是 $(X, \mathscr{F})$ 上的不同测度,$A, B \in \mathscr{F}$ 且 $A \cap B = \varnothing$, 若 μ_1 集中于 A,μ_2 集中于 B, 则说 μ_1 和 μ_2 **相互奇异**, 记为 $\mu_1 \perp \mu_2$.

定义 1.3. 设 (X, ν) 是符号测度空间, 集合 $E \in \mathscr{F}$ 称为是**正集**, 如果 $\nu(E) \geq 0$, 并且对任意的 $F \subset E, F \in \mathscr{F}$, 都有 $\nu(F) \geq 0$; 类似地, 可以定义**负集**.

定理 1.2 (Hahn 分解)**.** *设 (X, ν) 是符号测度空间, 那么存在正集 P 和负集 N, 使得 $P \cap N = \varnothing, P \cup N = X$. 并且这一分解在相差一零测集的意义下是唯一的, 即: 如果存在另一对正集 P' 和负集 N', 满足上述条件, 那么 $P \Delta P' = N \Delta N'$ 是零测集.*

定理 1.3 (Jordan 分解)**.** 设 (X,ν) 是符号测度空间, 那么存在唯一的正测度 ν^+ 和正测度 ν^-, 使得

$$\nu = \nu^+ - \nu^-, \quad \nu^+ \perp \nu^-.$$

从上文的讨论中可以发现, 这两个分解定理在本质上是相同的. 下面我们着手证明这两个定理.

引理 1.4. 正集的可测子集也是正集, 正集的可数并也是正集.

证明. 设 E 是正集,F 是 E 的可测子集, 那么对于任意 F 的可测子集 G, 都有 $G \subset E$, 因而具有非负测度. 所以 F 也是正集.

设 $\{E_n\}$ 是正集列, 令 $E = \bigcup_{n=1}^\infty E_n$. 设 $P_n = E_n \setminus \bigcup_{i=1}^{n-1} E_n$, 则 P_n 是正集, 并且

$$\bigcup_{n=1}^\infty P_n = E, P_n \cap P_m = \varnothing, n \neq m.$$

若 F 是 E 的子集, 那么

$$\bigcup_{n=1}^\infty (F \cap P_n) = F \cap \left(\bigcup_{n=1}^\infty P_n \right) = F.$$

所以

$$\nu(F) = \nu\left(\bigcup_{n=1}^\infty (F \cap P_n) \right) = \sum_{n=1}^\infty \mu(F \cap P_n) > 0.$$

□

Hahn 分解定理的证明. 不妨假设 ν 不能取到 ∞. 那么取 $\mathscr{F}$ 中所有正集构成的集合为 $\mathscr{P}$, 令

$$m := \sup_{P \in \mathscr{P}} \nu(P).$$

注意到 $\varnothing \in \mathscr{P}$, 所以 $m \geq 0$. 根据定义, 存在正集列 $\{P_j\}$, 使得

$$\lim_{j \to \infty} \nu(P_j) = m.$$

因此, 定义 $P = \bigcup_{j=1}^\infty P_j$, 则有 P 是正集, 且 $\nu(P) = m$.

我们证明 $N := X \setminus P$ 是负集.

我们宣称 N 中不存在非空的正集, 否则将其并入 P 则得到一个测度大于 m 的正集, 矛盾. 基于这一结果, 我们证明 N 确实是一个负集.

如果 N 中有一个非空集合 A, 使得 $\nu(A) \geq 0$. 那么 A 不是正集, 因此存在可测子集 C, 使得 $\nu(E) < 0$. 令 $B = A \setminus C$, 则

$$\nu(B) + \nu(C) = \nu(A),$$

所以 $\nu(B) = \nu(A) - \nu(C) > \nu(A)$.

现取最小的自然数 n_1, 使得存在集合 B_1, 满足

$$\mu(B_1) > \nu(A) + \frac{1}{n_1},$$

令 $A_{n_1} = B$. 依此类推, 取最小的自然数 n_j, 使得存在集合 B_j, 满足

$$\mu(B_j) > \nu(A_{j-1}) + \frac{1}{n_j},$$

令 $A_{n_j} = B_j$. 令

$$\mathcal{A} = \bigcap_{j=1}^{\infty} A_{n_j},$$

那么 $\nu(\mathcal{A}) > \nu(A) + \sum_{j=1}^{\infty} \frac{1}{n_j}$. 又因为 $\mathcal{A}$ 不是正集, 所以存在 $\mathcal{B} \subset \mathcal{A}$, 使得存在 $n \in \mathbb{N}$, 满足

$$\nu(\mathcal{B}) > \nu(\mathcal{A}) + \frac{1}{n} > \nu(A_{n_j}) + \frac{1}{n}, \forall j \in \mathbb{N}.$$

因为 $\nu(\mathcal{A})$ 测度有限, 所以 $\lim\limits_{j\to\infty} \frac{1}{n_j} = 0$, 因此, 存在 $n_J > n$. 此时

$$\nu(\mathcal{B}) > \nu(A_{n_{J-1}}) + \frac{1}{n}.$$

这与 A_{n_J} 的构造矛盾. 因此 N 是一个负集.

下面我们说明分解的“唯一性”. 设 P', N' 是另一对分解, 那么 $P\Delta P'$ 是正集. 注意到

$$P\Delta P' = (P \cup P') \setminus (P \cap P') = (P \setminus (P \cap P')) \cup (P' \setminus (P \cap P')),$$

因为 $P\setminus(P\cap P')\cap P' = \varnothing$, 所以 $P\setminus(P\cap P') \subset N'$, 故 $\nu(P\setminus(P\cap P')) = 0$. 同理,$\nu(P'\setminus(P\cap P')) = 0$. 因此,$\nu(P\Delta P') = 0$.

□

Jordan 分解的证明. 设 P, N 是空间 (X, ν) 的 Hahn 分解, 其中 P 是正集,N 是负集. 定义 $\nu^+(E) = \nu(E \cap P), \nu^-(E) = -\nu(E \cap N)$, 那么

$$\nu(E) = \nu^+(E) - \nu^-(E).$$

又因为 ν^+ 集中在 P 上,ν^- 集中在 N 上, 且 $P \cap N = \varnothing, P \cup N = X$, 所以 $\nu^+ \perp \nu^-$. □

定义 1.4. 设 (X, ν) 是符号测度空间,$\nu = \nu^+ - \nu^-$ 是符号测度 ν 的 Jordan 分解. 则分别称 ν^+ 和 ν^- 为 ν 的**正变差**和**负变差**, 并称 $|\nu| := \nu^+ + \nu^-$ 为 ν 的**全变差** (测度).

评论. 显然有

$$|\nu(E)| \leq |\nu|(E), \quad \forall E \in \mathscr{F}.$$

1.3 Lebesgue-Radon-Nikodym 定理

本小节包括两部分, 即 Lebesgue 分解定理和 Radon-Nikodym 定理.

现在, 让我们再次考虑由可积函数诱导的测度 ν_f. 注意到当我们在讨论这一对象时, 我们事先在可测空间 $(X, \mathscr{F})$ 给定了一个正测度 μ. 当可测集 E 是 μ- 零测集时, 根据定义有 $\nu_f(E) = 0$. 这样的性质被称为**绝对连续性**.

定义 1.5. 设 $(X, \mathscr{F})$ 是一个可测空间,μ 是其上的正测度,ν 是其上的符号测度. 我们说 ν 关于 μ 是**绝对连续的**, 是指

$$\forall E \in \mathscr{F} \, [\mu(E) = 0 \implies \nu(E) = 0] \, ,$$

记为 $\nu \ll \mu$.

在 ν 是有限测度的情形, 有如下的等价刻画:

命题 1.5 (绝对连续性的等价描述)**.** *设 $(X, \mathscr{F})$ 是一个可测空间,μ 是其上的正测度,ν 是其上的有限的符号测度. 那么 ν 关于 μ 是**绝对连续的**, 当且仅当对任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得对任意的可测集 E 且 $\mu(E) < \delta$ 时, 必有 $|\nu(E)| < \varepsilon$.*

证明. 充分性是显然的, 我们只证明必要性. 如果不然, 存在 $\varepsilon_0 > 0$, 以及对任意的正整数 n, 存在 E_n, 使得

$$\mu(E_n) < \frac{1}{2^n}, \quad |\nu(E_n)| \geq \varepsilon_0.$$

令

$$A_n = \bigcup_{i=n}^{\infty} E_i, \quad A = \bigcap_{n=1}^{\infty} A_n.$$

则 $\mu(A_n) < 2^{-n+1}$,$A_n \supset A_{n+1}$, 因此 $\mu(A) = 0$. 并且由 $|\nu|(A_n) \geq |\nu|(E_n)$ 可知,

$$|\nu|(A) = \lim_{n \to \infty} |\nu|(A_n) \geq \varepsilon > 0.$$

于是 $|\nu| \ll \mu$ 不成立, 矛盾. $\square$

对于任意的符号测度 ν, 我们已经证明了可以将其分解为正测度 ν^+ 与正测度 ν^- 之差. 所以, 我们考虑正测度的情形就足够了.

引理 1.6. *若 μ 是集 X 的 σ-域 $\mathscr{F}$ 上的正 $\sigma-$ 有限测度, 则存在一个函数 $w \in L^1(\mu)$, 使得对于每个 $x \in X$, 有 $0 < w(x) < 1$.*

证明. 由于 μ 是 σ-有限的, 所以存在一列可测集 $\{E_i\}_{i=1}^{\infty}$, 使得 $\mu(E_i) < \infty$ 且 $\bigcup_i E_i = X$. 定义

$$w_n(x) := \begin{cases} 2^{-n}/(1 + \mu(E_n)) & , x \in E_n, \\ 0 & , x \notin E_n. \end{cases}$$

取 $w=\sum_{n=1}^{\infty} w_n$, 则满足要求, 此处的收敛为点态收敛. 这是因为,

$$0<w(x)\leq\sum_{n=1}^{\infty}\frac{2^{-n}}{1+\mu(E_n)}<\sum_{n=1}^{\infty}2^{-n}=1.$$

且

$$\int_X\sum_{n=1}^{\infty}w_n\mathrm{d}\mu=\sum_{i=n}^{\infty}\int_X w_n(x)\mathrm{d}\mu=\sum_{n=1}^{\infty}\frac{2^{-n}\mu(E_n)}{1+\mu(E_n)}<1.$$

其中积分和求和交换次序是使用了单调收敛定理. 因此 $w\in L^1(\mu)$. □

通过这一引理, 我们定义测度 $\widetilde{\mu}=w\mu$, 即

$$\widetilde{\mu}(E)=\int_E w(x)\mathrm{d}\mu,$$

则 $\widetilde{\mu}$ 是有限测度.

定理 1.7 (Lebesgue-Radon-Nikodym 定理)**.** *设 μ 是集 X 的 σ-域 $\mathscr{F}$ 上的正 σ-有限测度, 并设 ν 是其上的 σ-有限的符号测度.*

(a)(Lebsugue 分解) 在 $\mathscr{F}$ 上存在唯一的一对符号测度 ν_a 和 ν_s, 使得

$$\nu=\nu_a+\nu_s,\quad \nu_a\ll\mu,\quad \nu_s\perp\mu.$$

若 ν 是正有限测度, 则 ν_a 和 ν_s 也是正有限测度.

(b)(Radon-Nikodym 定理) 存在唯一一个 μ-可测函数 h, 使得

$$\nu_a(E)=\int_E h\mathrm{d}\mu(E\in\mathscr{F}).$$

证明. 我们先证明唯一性. 假设有另一对复测度 (ν_a',ν_s'), 满足定理条件. 那么

$$\nu_a-\nu_a'=\nu_s'-\nu_s,$$

而 $\nu_a-\nu_a'\ll\mu,\nu_s'-\nu_s\perp\mu$, 所以

$$\nu_a-\nu_a'=\nu_s'-\nu_s=0.$$

即 $\nu_a=\nu_a'$,$\nu_s=\nu_s'$. 下面证明存在性.

(1) 首先假设 ν 是正有限测度. 那么取引理中的 w, 则 $\varphi=\nu+w\mu$ 也是正有限测度. 于是, 对于特征函数 $f=\chi_E$,

$$\int_X f\mathrm{d}\varphi=\int_X f\mathrm{d}\nu+\int_X fw\mathrm{d}\mu.$$

因此对于简单函数, 进而对于非负可测函数 f 成立. 设 $f\in L^2(\varphi)$, 则

$$|\int_X f\mathrm{d}\nu|\leq\int_X|f|\mathrm{d}\nu\leq\int_X|f|\mathrm{d}\varphi\leq(\int_X|f|^2\mathrm{d}\varphi)^{\frac{1}{2}}(\varphi(X))^{\frac{1}{2}}.$$

因为 φ 是有限的, 所以映射 $f\mapsto\int_X f\mathrm{d}\nu$ 是 $L^2(\varphi)$ 上的有界线性泛函. 根据 Riesz 表现定理, 存在唯一的 $g\in L^2(\varphi)$, 使得

$$\int_X f\mathrm{d}\nu=\int_X fg\mathrm{d}\varphi. \tag{1}$$

注意到 g 是在相差一个与 φ-零测集上取非零值的函数的意义下唯一.

取 $E \in \mathscr{F}$, 使得 $\varphi(E) > 0$, 令 $f = \chi_E$, 则

$$\nu(E) = \int_E g \mathrm{d}\varphi \geq 0.$$

因此

$$0 \leq \frac{1}{\varphi(E)} \int_E g \mathrm{d}\varphi = \frac{\nu(E)}{\varphi(E)} \leq 1.$$

所以 $g(x) \in [0,1]$ 对几乎所有的 $x \in X$ 成立, 因此, 不妨令 $0 \leq g(x) \leq 1$. 将式 (1) 重写为

$$\int_X (1-g) f \mathrm{d}\nu = \int_X f g w \mathrm{d}\mu. \tag{2}$$

令 $A = \{x : 0 \leq g(x) < 1\}$,$B = \{x : g(x) = 1\}$, 定义

$$\nu_a(E) = \nu(A \cap E), \nu_s(E) = \nu(B \cap E) (E \in \mathscr{F}).$$

在式 (2) 中, 取 $f = \chi_B$, 则

$$0 = \int_B w \mathrm{d}\mu.$$

所以 $\mu(B) = 0$, 这说明了 $\nu_s \perp \mu$. 另一方面, 由于 $g(x) \in [0,1]$, 所以令 $f = (1+g+\cdots+g^n)\chi_E$, 则

$$\int_E (1-g^{n+1}) \mathrm{d}\nu = \int_E g(1+g+\cdots+g^n) w \mathrm{d}\mu.$$

考虑上式左边, 由于当 $x \in B$ 时,$g(x) = 1$, 故此时 $1-g^{n+1}(x) = 0$; 当 $x \in A$ 时,$0 \leq g(x) < 1$, 所以 $\lim_{n\to\infty} g^n(x) = 0$. 因此两边取极限, 则左边相当于 $\nu_a(E) = \nu(A \cap E)$, 而右边是单调增加的有界序列, 所以它收敛于某个非负可测的极限函数 h, 此时, 有

$$\nu_a(E) = \int_E h \mathrm{d}\mu.$$

这就证明了 (b), 并且该式也说明了 $\nu_a \ll \mu$. 所以该命题对正有限测度 ν 成立.

(2) 过渡到 ν 是 σ-有限的正测度的情况只需要注意到 X 可以表示为两两不相交的可数的测度有限子集的并即可. □

评论. 回忆起我们在数学分析中学到的变量替换公式:

$$\int \mathrm{d}f = \int f'(x) \mathrm{d}x.$$

形式上有

$$f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}.$$

我们也可以写

$$h = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

因此, 这里的 h 经常被称为 **Radon-Nikodym** 导数.

本节的最后, 我们给出 Radon-Nikodym 定理在分布函数分类中的一个应用.

设 X 是概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上的一个随机变量, 其概率分布函数是

$$\mathbb{P}X^{-1}:(\mathbb{R},\mathscr{B}_{\mathbb{R}})\to[0,1],B\mapsto\mathbb{P}(X^{-1}(B)),$$

这是 $(\mathbb{R},\mathscr{B}_{\mathbb{R}})$ 上的一个 σ-有限的正测度. 令 λ 是 $(\mathbb{R},\mathscr{B}_{\mathbb{R}})$ 上的另一个 σ-有限的正测度 (例如 Lebesgue 测度). 根据 Lebesgue 分解, 有

$$\mathbb{P}X^{-1}=\mu_a+\mu_s,\mu_a\ll\lambda,\mu_s\perp\lambda.$$

设

$$D:=\{x\in\mathbb{R}:\mu_s(\{x\})>0\},$$

定义

$$\mu_1(E)=\mu(E\cap D),\mu_2=\mu_a,\mu_3=\mu_s-\mu_1.$$

则有

$$\mathbb{P}X^{-1}=\mu_1+\mu_2+\mu_3.$$

但是此处的 μ_a,μ_1,μ_3 在全空间上的测度不一定为 1, 所以考虑如下的正规化:

$$\widetilde{\mu_i}(E)=\begin{cases}\mu_i(E)/\mu_i(\mathbb{R}), & \mu_i(\mathbb{R})>0,\\ \nu, & \mu_i(\mathbb{R})=0.\end{cases}$$

其中 ν 是任意的概率测度. 再设 $\alpha_i=\mu_i(\mathbb{R})$, 则有

$$\mathbb{P}X^{-1}=\alpha_1\widetilde{\mu_1}+\alpha_2\widetilde{\mu_2}+\alpha_3\widetilde{\mu_3}.$$

此时,$\alpha_1+\alpha_2+\alpha_3=1$.

- 当 $\alpha_1=1$ 时, 随机变量 X 称为**离散型随机变量**. 此时,D 是至多可数集 (为什么?). 设 $D=\{x_n\}_{n=1}^{\infty}$, 那么 $\mathbb{P}X^{-1}$ 在 x_n 上的取值

  $$\{p_n=\mathbb{P}(X=x_n):n\in\mathbb{N}\}$$

  唯一决定了 $\mathbb{P}X^{-1}$. 此时, 对于随机变量的函数 $f(X)$, 有

  $$\mathbb{E}(f(X))=\int_{\mathbb{R}}f(x)\mathrm{d}(\mathbb{P}X^{-1})=\sum_{n=1}^{\infty}f(x_n)p_n.$$

- 当 $\alpha_2=1$ 时, 随机变量 X 称为**连续型随机变量**. 此时,$\mathbb{P}X^{-1}$ 关于测度 λ 绝对连续, 所以存在 λ-可积函数 p, 使得

  $$\mathbb{P}X^{-1}(E)=\int_E p\mathrm{d}\lambda,$$

  此时的 p 称为随机变量 X 的**密度函数**. 对于随机变量的函数 $f(X)$, 有

  $$\mathbb{E}(f(X))=\int_{\mathbb{R}}f(x)p(x)\mathrm{d}\lambda(x).$$

- 当 $\alpha_3=1$ 时, 随机变量 X 称为**奇异型随机变量**.

定理 1.8. 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是一个概率空间,X 是其上的一个随机变量. 则 X 的概率分布 $\mathbb{P}X^{-1}$ 可以写为离散型概率分布函数、连续型概率分布函数以及奇异型概率分布函数的凸组合.

Part II

概率论部分

2 条件期望

2.1 条件期望与条件概率

设 $(\Omega, \mathscr{F}, \mathbb{P})$ 为概率空间, 条件期望首先是对于 $\mathscr{F}$ 的子 σ-域而言的.

定义 2.1 (子 σ-域)**.** 设 $\mathscr{G} \subset \mathscr{F}$ 是 σ-域, 则称 $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域.

设 $X \in \mathscr{L}^1(\Omega)$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, 定义

$$\varphi(C) := \int_C X \mathrm{d}\mathbb{P}, \quad \forall C \in \mathscr{G},$$

则 φ 是 $\mathscr{G}$ 上的符号测度, 且根据

$$\mathbb{P}(C) = 0 \implies \varphi(C) = \int_C X \mathrm{d}\mathbb{P} = 0$$

得知 $\varphi \ll \mathbb{P}$, 从而由 Radon-Nikodym 定理知, 存在概率空间 $(\Omega, \mathscr{G}, \mathbb{P})$ 上的随机变量 $Y \in \mathscr{L}^1(\Omega)$, 使得

$$\varphi(C) = \int_C X \mathrm{d}\mathbb{P} = \int_C Y \mathrm{d}\mathbb{P}, \quad \forall C \in \mathscr{G}.$$

需要注意的是, 上式中 X 是 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, 而 Y 是 $(\Omega, \mathscr{G}, \mathbb{P})$ 上的随机变量. 据此, 条件期望可以得到定义:

定义 2.2 (条件期望)**.** 设 $\mathscr{G} \subset \mathscr{F}$ 是子 σ-域, $X \in \mathscr{L}^1(\Omega)$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, $\mathbb{E}(X|\mathscr{G}) \in \mathscr{L}^1(\Omega)$ 是概率空间 $(\Omega, \mathscr{G}, \mathbb{P})$ 上的随机变量, 且满足

$$\int_C X \mathrm{d}\mathbb{P} = \int_C \mathbb{E}(X|\mathscr{G}) \mathrm{d}\mathbb{P}, \quad \forall C \in \mathscr{G},$$

则称 $\mathbb{E}(X|\mathscr{G})$ 是 X 关于子 σ-域 $\mathscr{G}$ 的条件期望.

根据定义, 取 $C = \Omega \in \mathscr{G}$, 则

$$\mathbb{E}X = \int_\Omega X \mathrm{d}\mathbb{P} = \int_\Omega \mathbb{E}(X|\mathscr{G}) \mathrm{d}\mathbb{P} = \mathbb{E}[\mathbb{E}(X|\mathscr{G})],$$

这是条件期望的一个重要性质. 另外, 在初等概率论中, 设 $A \in \mathscr{F}$, I_A 是 A 的示性函数, 则 $\mathbb{E}I_A = \mathbb{P}(A)$. 受此启发, 可以定义条件概率:

定义 2.3 (条件概率)**.** 设 $\mathscr{G} \subset \mathscr{F}$ 是子 σ-域, $A \in \mathscr{F}$, 称 $\mathbb{P}(A|\mathscr{G}) = \mathbb{E}(I_A|\mathscr{G})$ 是 A 关于子 σ-域 $\mathscr{G}$ 的条件概率.

需要注意, 条件概率实际上也是随机变量.

2.2 条件期望与条件概率的例子

为了加强对条件概率与条件期望的定义的理解, 在这里考虑一些简单的例子.

例 2.1. 设 $B \in \mathscr{F}, \mathscr{G} = \sigma(B) = \{\varnothing, B, B^C, \Omega\}$, 且 $\mathbb{P}(B) \notin \{0, 1\}$. 设 $A \in \mathscr{G}$, 则 $\mathbb{P}(A|\mathscr{G})$ 是 $\mathscr{G}$ 上的可测函数, 从而

$$\mathbb{P}(A|\mathscr{G}) = aI_B + bI_{B^C};$$

另外, 根据定义得

$$\int_C \mathbb{P}(A|\mathscr{G})\mathrm{d}\mathbb{P} = \int_C I_A \mathrm{d}\mathbb{P} = \mathbb{P}(AC),$$

取 $C = B$ 得 $a = \dfrac{\mathbb{P}(AB)}{\mathbb{P}(B)} = \mathbb{P}(A|B)$, 取 $C = B^C$ 得 $b = \dfrac{\mathbb{P}(AB^C)}{\mathbb{P}(B^C)} = \mathbb{P}(A|B^C)$, 从而

$$\mathbb{P}(A|\mathscr{G}) = \mathbb{P}(A|B) \cdot I_B + \mathbb{P}(A|B^C) \cdot I_{B^C}.$$

例 2.2. 设 $B_1, B_2, \cdots, B_m \in \mathscr{F}$ 互不相交, 且 $B_1 \cup B_2 \cup \cdots \cup B_m = \Omega, \mathscr{G} = \sigma(B_1, B_2, \cdots, B_m)$. 设 $A \in \mathscr{F}$, 类似例2.1, 可以求出

$$\mathbb{P}(A|\mathscr{G}) = \sum_{j=1}^{m} \mathbb{P}(A|B_j) \cdot I_{B_j}.$$

对于2.2例的结果, 如果 $\omega \in B_j$, 也即给定条件 B_j, 则

$$\mathbb{P}(A|\mathscr{G})(\omega) = \mathbb{P}(A|B_j),$$

右边是在条件 B_j 下, 事件 A 发生的概率, 这说明了抽象定义的合理性.

例 2.3. 设 $A_1, A_2, \cdots, A_n \in \mathscr{F}$ 互不相交, 且 $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$, 令

$$X = \sum_{i=1}^{n} a_i \cdot I_{A_i},$$

再设 $B_1, B_2, \cdots, B_m \in \mathscr{F}$ 互不相交, 且 $B_1 \cup B_2 \cup \cdots \cup B_m = \Omega, \mathscr{G} = \sigma(B_1, B_2, \cdots, B_m)$, 类似例2.1, 可以求出

$$\mathbb{E}(X|\mathscr{G}) = \sum_{j=1}^{m} \sum_{i=1}^{n} a_i \cdot \mathbb{P}(A_i|B_j) \cdot I_{B_j}.$$

例2.3的结果非常有趣. 如果 $\omega \in B_j$, 也即给定条件 B_j, 我们知道

$$\mathbb{E}(X|\mathscr{G})(\omega) = \sum_{i=1}^{n} a_i \cdot \mathbb{P}(A_i|B_j),$$

这事实上是将样本空间作划分 $\Omega = A_1 \cup A_2 \cup \cdots \cup A_n$, 随机变量 X 在第 i 个集合 A_i 上的取值为 a_i, 而 $\mathbb{P}(A_i|B_j)$ 是在条件 B_j 下, 事件 A_i 发生的概率, 它们相乘求和, 就得到了初等概率论中定义的条件期望.

2.3 条件期望的性质

本节中将探讨条件期望的一些基本性质.

命题 2.1 (特殊的条件期望)**.** *设 X 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域.*

(1) 如果 X 关于 $\mathscr{G}$ 可测, 则 $\mathbb{E}(X|\mathscr{G}) = X, \text{a.s.}$;

(2) 如果 X 与 $\mathscr{G}$ 独立, 则 $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}X, \text{a.s.}$.

证明. (1) X 是 $(\Omega, \mathscr{G}, \mathbb{P})$ 上的可测函数, 且

$$\int_C X \mathrm{d}\mathbb{P} = \int_C X \mathrm{d}\mathbb{P}, \quad \forall C \in \mathscr{G},$$

根据定义知 $\mathbb{E}(X|\mathscr{G}) = X, \text{a.s.}$.

(2) $\mathbb{E}X$ 是 $(\Omega, \mathscr{G}, \mathbb{P})$ 上的可测函数, 且[1]

$$\int_C X \mathrm{d}P = \mathbb{E}(X \cdot I_C) = \mathbb{E}X \cdot \mathbb{P}(C) = \int_C \mathbb{E}X \mathrm{d}\mathbb{P},$$

根据定义知 $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}X, \text{a.s.}$. □

命题 2.2 (复合条件期望)**.** *设 X 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$ 是 $\mathscr{F}$ 的子 σ-域, 则*

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

证明. 一方面, 根据 $\mathbb{E}(X|\mathscr{G}_1)$ 关于 $\mathscr{G}_1 \subset \mathscr{G}_2$ 可测, 应用命题2.1的 (1) 得

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

另外一方面, 根据 $\mathbb{E}[\mathbb{E}(X|\mathscr{G}_2)|\mathscr{G}_1]$ 关于 $\mathscr{G}_1$ 可测, 且对任意的 $C \in \mathscr{G}_1 \subset \mathscr{G}_2$, 都有

$$\int_C \mathbb{E}[\mathbb{E}(X|\mathscr{G}_2)|\mathscr{G}_1] \mathrm{d}\mathbb{P} = \int_C \mathbb{E}(X|\mathscr{G}_2) \mathrm{d}\mathbb{P} = \int_C X \mathrm{d}\mathbb{P} = \int_C \mathbb{E}(X|\mathscr{G}_1) \mathrm{d}\mathbb{P},$$

由 C 的任意性得

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

综合以上两条可知命题成立. □

命题 2.3 (单调性)**.** *设 X, Y 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, $X \leq Y, \text{a.s.}$, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域, 则*

$$\mathbb{E}(X|\mathscr{G}) \leq \mathbb{E}(Y|\mathscr{G}), \quad \text{a.s.}.$$

---

[1]在这里应用了结论: 若 X 与 C 独立, 则 $\mathbb{E}(X \cdot I_C) = \mathbb{E}X \cdot \mathbb{P}(C)$, 证明需要使用典型方法.

特别地, $|\mathbb{E}(X|\mathscr{G})| \le \mathbb{E}(|X||\mathscr{G})$, a.s..

证明. 对任意的 $C \in \mathscr{G}$, 都有

$$\int_C \mathbb{E}(X|\mathscr{G})\mathrm{d}\mathbb{P} = \int_C X\mathrm{d}\mathbb{P} \le \int_C Y\mathrm{d}\mathbb{P} = \int_C \mathbb{E}(Y|\mathscr{G})\mathrm{d}\mathbb{P},$$

由 C 的任意性得

$$\mathbb{E}(X|\mathscr{G}) \le \mathbb{E}(Y|\mathscr{G}), \quad \text{a.s..}$$

另外, 注意到 $-X \le |X|, X \le |X|$, 因此

$$\begin{cases} -\mathbb{E}(X|\mathscr{G}) \le \mathbb{E}(|X||\mathscr{G}), & \text{a.s.,} \\ \mathbb{E}(X|\mathscr{G}) \le \mathbb{E}(|X||\mathscr{G}), & \text{a.s..} \end{cases}$$

综合以上两式, 有 $|\mathbb{E}(X|\mathscr{G})| \le \mathbb{E}(|X||\mathscr{G})$, a.s.. □

命题 2.4 (可加性)**.** *设 X, Y 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域, $a, b \in \mathbb{R}$, 且 $a\mathbb{E}X + b\mathbb{E}Y$ 存在, 则*

$$\mathbb{E}(aX + bY|\mathscr{G}) = a\mathbb{E}(X|\mathscr{G}) + b\mathbb{E}(Y|\mathscr{G}), \quad \text{a.s..}$$

证明. 由 $\mathbb{E}(aX + bY) = a\mathbb{E}X + b\mathbb{E}Y$ 存在, 知 $\mathbb{E}(aX + bY|\mathscr{G})$ 有定义. 对任意的 $C \in \mathscr{G}$, 都有

$$\begin{aligned} \int_C (aX + bY)\mathrm{d}\mathbb{P} &= a\int_C X\mathrm{d}\mathbb{P} + b\int_C Y\mathrm{d}\mathbb{P} \\ &= a\int_C \mathbb{E}(X|\mathscr{G})\mathrm{d}\mathbb{P} + b\int_C \mathbb{E}(Y|\mathscr{G})\mathrm{d}\mathbb{P} \\ &= \int_C (a\mathbb{E}(X|\mathscr{G}) + b\mathbb{E}(Y|\mathscr{G}))\mathrm{d}\mathbb{P}. \end{aligned}$$

因此, $\mathbb{E}(aX + bY|\mathscr{G}) = a\mathbb{E}(X|\mathscr{G}) + b\mathbb{E}(Y|\mathscr{G})$, a.s.. □

命题2.3和2.4都是期望 $\mathbb{E}(\cdot)$ 的性质向条件期望 $\mathbb{E}(\cdot|\mathscr{G})$ 的推广. 联想到期望所满足的单调收敛定理、Fatou 引理及 Lebesgue 控制收敛定理, 我们接下来证明这些定理对于条件期望也是成立的.

定理 2.5 (单调收敛定理)**.** *设 $\{X_n\}$ 和 X 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的积分存在的随机变量, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域, 若 $0 \le X_n \uparrow X$, a.s., 则*

$$0 \le \mathbb{E}(X_n|\mathscr{G}) \uparrow \mathbb{E}(X|\mathscr{G}), \quad \text{a.s..}$$

证明. 由命题2.3得 $\{\mathbb{E}(X_n|\mathscr{G})\}$ 单调递增, 且根据 $\mathbb{E}(X_n|\mathscr{G})$ 关于 $\mathscr{G}$ 可测知 $\lim\limits_{n\to\infty} \mathbb{E}(X_n|\mathscr{G})$ 关于 $\mathscr{G}$ 可测. 对任意的 $C \in \mathscr{G}$, 应用单调收敛定理得

$$\int_C \lim_{n\to\infty} \mathbb{E}(X_n|\mathscr{G})\mathrm{d}\mathbb{P} = \lim_{n\to\infty}\int_C \mathbb{E}(X_n|\mathscr{G})\mathrm{d}\mathbb{P} = \lim_{n\to\infty}\int_C X_n\mathrm{d}\mathbb{P} = \int_C X\mathrm{d}\mathbb{P},$$

这便说明了 $\mathbb{E}(X_n|\mathscr{G}) \uparrow \mathbb{E}(X|\mathscr{G})$, a.s.. □

定理 2.6 (Fatou 引理)**.** 设 $\{X_n\}$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的积分存在的随机变量序列, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域, 若 $X_n \geq 0$, a.s., 则

$$\mathbb{E}\left(\liminf_{n\to\infty} X_n|\mathscr{G}\right) \leq \liminf_{n\to\infty} \mathbb{E}(X_n|\mathscr{G}), \quad \text{a.s..}$$

证明. 对任意的 $C \in \mathscr{G}$, 应用 Fatou 引理得

$$\int_C \mathbb{E}\left(\liminf_{n\to\infty} X_n|\mathscr{G}\right) \mathrm{d}\mathbb{P} = \int_C \liminf_{n\to\infty} X_n \mathrm{d}\mathbb{P} \leq \liminf_{n\to\infty} \int_C X_n \mathrm{d}\mathbb{P},$$

这便说明了 $\mathbb{E}\left(\liminf_{n\to\infty} X_n|\mathscr{G}\right) \leq \liminf_{n\to\infty} \mathbb{E}(X_n|\mathscr{G})$, a.s.. □

定理 2.7 (Lebesgue 控制收敛定理)**.** 设 $\{X_n\}$ 和 X 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的积分存在的随机变量, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域, 若 $\lim_{n\to\infty} X_n = X$, 且存在 $Y \in L^1(\Omega)$, 使得对任意的 $n \geq 1$, 都有 $|X_n| \leq Y$, a.s., 则

$$\lim_{n\to\infty} \mathbb{E}(X_n|\mathscr{G}) = \mathbb{E}(X|\mathscr{G}), \quad \text{a.s..}$$

证明. 由命题2.3知

$$|\mathbb{E}(X_n|\mathscr{G})| \leq \mathbb{E}(|X_n||\mathscr{G}) \leq \mathbb{E}(Y|\mathscr{G}) \in L^1(\Omega).$$

对任意的 $C \in \mathscr{G}$, 应用 Lebesgue 控制收敛定理得

$$\int_C \lim_{n\to\infty} \mathbb{E}(X_n|\mathscr{G})\mathrm{d}\mathbb{P} = \lim_{n\to\infty} \int_C \mathbb{E}(X_n|\mathscr{G})\mathrm{d}\mathbb{P} = \lim_{n\to\infty} \int_C X_n \mathrm{d}\mathbb{P} = \int_C X \mathrm{d}\mathbb{P},$$

这便说明了 $\lim_{n\to\infty} \mathbb{E}(X_n|\mathscr{G}) = \mathbb{E}(X|\mathscr{G})$, a.s.. □

定理2.5、2.6和2.7的证明过程说明了, 条件期望是离不开期望的. 在处理条件期望的问题时, 常常需要借助条件期望的定义转化为期望. 最后, 作为定理2.5的应用, 我们引出如下的重要结论, 其在随机过程中有应用.

命题 2.8. 设 X, Y 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-域, X 和 XY 的积分存在, 且 Y 关于 $\mathscr{G}$ 可测, 则

$$\mathbb{E}(XY|\mathscr{G}) = Y \cdot \mathbb{E}(X|\mathscr{G}), \quad \text{a.s..}$$

证明. 用典型方法. □

例 2.4. 设 $\{(X_n, \mathscr{F}_n), n \geq 0\}$ 是带流的随机过程, 也即对任意的 $n \geq 0$, X_n 是关于 $\mathscr{F}_n$ 可测的随机变量, 且 $\mathscr{F}_n \subset \mathscr{F}_{n+1}$. 设其期望为 0, 方差为 $\sigma^2 < \infty$. 若其是一个鞅差过程, 也即对任意的 $n \geq 0$, 都有

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = 0, \quad \text{a.s.,}$$

则其一定是一个白噪声过程, 也即对任意的 $n \geq 0$, 都有

$$\mathrm{Cov}(X_n, X_{n+1}) = 0.$$

这是因为, X_n 是关于 $\mathscr{F}_n$ 可测的, 从而

$$\mathbb{E}(X_n X_{n+1}|\mathscr{F}_n) = X_n \cdot \mathbb{E}(X_{n+1}|\mathscr{F}_n) = 0, \quad \text{a.s.},$$

对上式两边取期望得

$$\mathrm{Cov}(X_n, X_{n+1}) = \mathbb{E}[\mathbb{E}(X_n X_{n+1}|\mathscr{F}_n)] = 0.$$

3 离散时间鞅

3.1 离散时间鞅的定义

鞅 (Martingale) 的概念来源于公平赌博. 在上一节的2.4中, 所提到的鞅差过程也与鞅有一定的联系.

定义 3.1 (离散时间鞅)**.** 设 $\{(X_n, \mathscr{F}_n), n \geq 0\}$ 是带流的随机过程, 也即对任意的 $n \geq 0$, X_n 是关于 $\mathscr{F}_n$ 可测的随机变量, 且 $\mathscr{F}_n \subset \mathscr{F}_{n+1}$. 如果对任意的 $n \geq 0$, X_n 的积分存在, 且

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = X_n, \quad \text{a.s.},$$

则称 $\{X_n\}$ 为鞅. 如果将上式改为

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) \leq X_n, \quad \text{a.s.}, \quad 或 \quad \mathbb{E}(X_{n+1}|\mathscr{F}_n) \geq X_n, \quad \text{a.s.},$$

则分别称 $\{X_n\}$ 为上鞅 或下鞅.

根据以上定义, 若 $\{-X_n\}$ 为上鞅, 则 $\{X_n\}$ 为下鞅, 反之亦成立.

例 3.1 (公平赌博)**.** 考虑简单的赌博模型, 设本金为 X_0, 对 $n \geq 1$, 每次赌博的金额为 1 元, 收益为 δ_n, 且

$$\mathbb{P}(\delta_n = -1) = \mathbb{P}(\delta_n = 1) = \frac{1}{2},$$

设经过 n 次赌博后的本金为 X_n, 则有

$$X_{n+1} = X_0 + \sum_{i=1}^{n+1} \delta_i = X_n + \delta_{n+1}.$$

记 $\mathscr{F}_n = \sigma(X_0, \delta_1, \cdots, \delta_n)$, 则 X_n 是关于 $\mathscr{F}_n$ 可测的, 且 $\mathscr{F}_n \subset \mathscr{F}_{n+1}$, 并且

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = \mathbb{E}(X_n|\mathscr{F}_n) + \mathbb{E}(\delta_{n+1}|\mathscr{F}) = X_n + \mathbb{E}\delta_{n+1} = X_n,$$

其中 X_n 关于 $\mathscr{F}_n$ 可测, δ_{n+1} 与 $\mathscr{F}_n$ 独立. 从而 $\{(X_n, \mathscr{F}_n), n \geq 0\}$ 是鞅.

在大致了解了鞅之后, 现在需要进一步研究鞅的性质. 以下通常用 $\{X_n\}$ 表示鞅, 并且记使得 $X_0, X_1, \cdots, X_n$ 可测的最小 σ-代数

$$\mathscr{F}_n = \sigma(X_0, X_1, \cdots, X_n).$$

命题 3.1 (期望). 设 $\{X_n\}$ 为鞅, 则 $\mathbb{E}X_n$ 为常数; 设 $\{X_n\}$ 为上鞅, 则 $\{\mathbb{E}X_n\}$ 单调递减; 设 $\{X_n\}$ 为下鞅, 则 $\{\mathbb{E}X_n\}$ 单调递增.

证明. 对于鞅的情形, 对任意的 $n \geq 0$, 都有

$$\mathbb{E}X_{n+1} = \mathbb{E}[\mathbb{E}(X_{n+1}|\mathscr{F}_n)] = \mathbb{E}X_n,$$

这便说明了 $\mathbb{E}X_n$ 为常数. 上鞅和下鞅的情形可以类似证明. □

命题 3.2 (可加性). 设 $\{X_n\}$, $\{Y_n\}$ 为鞅 (上鞅、下鞅), 则 $\{X_n + Y_n\}$ 也为鞅 (上鞅、下鞅).

证明. 对于鞅的情形, 注意到

$$\mathbb{E}(X_{n+1} + Y_{n+1}|\mathscr{F}_n) = \mathbb{E}(X_{n+1}|\mathscr{F}_n) + \mathbb{E}(Y_{n+1}|\mathscr{F}_n) = X_n + Y_n,$$

因此 $\{X_n + Y_n\}$ 为鞅, 同理也可以证明上鞅和下鞅的情形. □

命题 3.3 (最大值与最小值). 设 $\{X_n\}$ 和 $\{Y_n\}$ 为鞅 (上鞅), 则 $\{X_n \wedge Y_n\}$ 为上鞅; 设 $\{X_n\}$ 和 $\{Y_n\}$ 为鞅 (下鞅), 则 $\{X_n \vee Y_n\}$ 为下鞅.

证明. 首先, 设 $\{X_n\}$ 和 $\{Y_n\}$ 为鞅 (上鞅), 计算得

$$\mathbb{E}(X_{n+1} \wedge Y_{n+1}|\mathscr{F}_n) \leq \mathbb{E}(X_{n+1}|\mathscr{F}_n) \wedge \mathbb{E}(Y_{n+1}|\mathscr{F}_n) \leq X_n \wedge Y_n,$$

因此 $\{X_n \wedge Y_n\}$ 为上鞅; 接下来, 设 $\{X_n\}$ 和 $\{Y_n\}$ 为鞅 (下鞅), 计算得

$$\mathbb{E}(X_{n+1} \vee Y_{n+1}|\mathscr{F}_n) \geq \mathbb{E}(X_{n+1}|\mathscr{F}_n) \vee \mathbb{E}(Y_{n+1}|\mathscr{F}_n) \geq X_n \vee Y_n,$$

因此 $\{X_n \vee Y_n\}$ 为下鞅. □

命题 3.4 (凸函数). 设 $\{X_n\}$ 为鞅, f 为 $\mathbb{R}$ 上的连续凸函数, 如果每个 $f(X_n)$ 可积, 则 $\{f(X_n)\}$ 为下鞅; 设 $\{X_n\}$ 为下鞅, f 是连续非降的凸函数, 如果每个 $f(X_n)$ 可积, 则 $\{f(X_n)\}$ 为下鞅.

证明. 设 $\{X_n\}$ 为鞅, 根据 Jensen 不等式得

$$\mathbb{E}(f(X_{n+1})|\mathscr{F}_n) \geq f(\mathbb{E}(X_{n+1}|\mathscr{F}_n)) = f(X_n),$$

因此 $\{f(X_n)\}$ 为下鞅; 再设 $\{X_n\}$ 为下鞅, 同样根据 Jensen 不等式得

$$\mathbb{E}(f(X_{n+1})|\mathscr{F}_n) \geq f(\mathbb{E}(X_{n+1}|\mathscr{F}_n)) \geq f(X_n),$$

因此 $\{f(X_n)\}$ 为下鞅. □

回忆起在例2.4中定义了鞅差.

定义 3.2 (鞅差). 设 $\{(X_n, \mathscr{F}_n), n \geq 0\}$ 是带流的随机过程, 也即对任意的 $n \geq 0$, X_n 是关于 $\mathscr{F}_n$ 可测的随机变量, 且 $\mathscr{F}_n \subset \mathscr{F}_{n+1}$. 如果对任意的 $n \geq 0$, X_n 的积分存在, 且

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = 0, \quad \text{a.s.},$$

则称 $\{X_n\}$ 为鞅差.

事实上, 设 $\{X_n\}$ 为鞅, 若记 $x_0 = X_0, x_{n+1} = X_{n+1} - X_n (n \geq 0)$, 则有

$$X_n = \sum_{i=0}^{n} x_i, \quad 且 \quad \mathbb{E}(x_{n+1}|\mathscr{F}_n) = X_n - X_n = 0,$$

这便说明了 $\{x_n\}$ 是鞅差. 反之, 如果 $\{x_n\}$ 是鞅差, 也即对任意的 $n \geq 0$, 都有 $\mathbb{E}(x_{n+1}|\mathscr{F}_n) = 0$, 记

$$X_n = \sum_{i=0}^{n} x_i,$$

则根据 $x_1, x_2, \cdots, x_n$ 关于 $\mathscr{F}_n$ 是可测的, 知 X_n 关于 $\mathscr{F}_n$ 是可测的, 从而

$$\mathbb{E}(X_{n+1}|\mathscr{F}_n) = \mathbb{E}(X_n + x_{n+1}|\mathscr{F}_n) = X_n,$$

这便说明了 $\{X_n\}$ 是鞅. 上面的过程说明了鞅和鞅差是可以一一对应的. 下面的定理就需要用到这样的技巧.

定理 3.5 (Doob 分解)**.** 设 $\{X_n\}$ 是下鞅, 则存在鞅 $\{Y_n\}$ 及增过程 $\{Z_n\}$, 使得

$$X_n = Y_n + Z_n.$$

证明. 记 $x_0 = X_0, x_{n+1} = X_{n+1} - X_n$, 则

$$\mathbb{E}(x_{n+1}|\mathscr{F}_n) = \mathbb{E}(X_{n+1}|\mathscr{F}_n) - X_n \geq 0.$$

记

$$z_0 = 0, \quad z_n = \mathbb{E}(x_n|\mathscr{F}_{n-1}) \geq 0, \quad Z_n = \sum_{i=0}^{n} z_i,$$

则 $\{Z_n\}$ 是增过程, 再记

$$y_0 = x_0, \quad y_n = x_n - z_n, \quad Y_n = \sum_{i=0}^{n} y_i,$$

则 $X_n = Y_n + Z_n$, 并且根据

$$\mathbb{E}(y_{n+1}|\mathscr{F}_n) = \mathbb{E}(x_{n+1}|\mathscr{F}_n) - \mathbb{E}(x_{n+1}|\mathscr{F}_n) = 0$$

知 $\{y_n\}$ 是鞅差, 从而 $\{Y_n\}$ 是鞅. □

3.2 停时与上穿不等式

定义 3.3 (停时)**.** 设 $\{\mathscr{F}_n\}$ 是上升的 σ-域, 也即对任意的 $n \geq 0$, 都有 $\mathscr{F}_n \subset \mathscr{F}_{n+1}$. 随机变量 $T : \Omega \to \{0, 1, 2, \cdots\}$, 且对任意的 $n \geq 0$, 都有

$$\{T = n\} \in \mathscr{F}_n,$$

则称 T 是 $\{\mathscr{F}_n\}$ 的停时.

事实上, 注意到 $\{\mathscr{F}_n\}$ 是上升的, 在这里 $\{T = n\} \in \mathscr{F}_n$ 对任意的 $n \geq 0$ 成立, 等价于 $\{T \leq n\} \in \mathscr{F}_n$ 对任意的 $n \geq 0$ 成立.

例 3.2 (赌徒停止赌博的时间)**.** 设 T 表示赌徒停止赌博的时间, $T=n$ 表示在经过 n 次赌博后赌徒停止赌博, 则事件 $\{T=n\}$ 取决于前 n 次赌博中所获得的信息 $\mathscr{F}_n$.

定义 3.4 (T 前 σ-域)**.** 设 T 是停时, 记

$$\mathscr{F}_\infty=\sigma\left(\bigcup_{n=1}^{\infty}\mathscr{F}_n\right),\quad \mathscr{F}_T=\{A\in\mathscr{F}_\infty: A\cap\{T=n\}\in\mathscr{F}_n,\forall n\geq 0\},$$

称 $\mathscr{F}_T$ 为 T **前 σ-域**.

关于停时和 T 前 σ-域, 有一些基本的性质.

命题 3.6 (可加性)**.** 设 S,T 是停时, 则 $S+T$ 是停时.

证明. 根据

$$\{S+T=n\}=\bigcup_{i=0}^{n}\{S=i,T=n-i\}=\bigcup_{i=0}^{n}\{S=i\}\{T=n-i\}\in\mathscr{F}_n,$$

知 $S+T$ 是停时. □

命题 3.7 (单调性)**.** 设 S,T 是停时, $S\leq T$, 则 $\mathscr{F}_S\subset\mathscr{F}_T$.

证明. 任取 $A\in\mathscr{F}_S$, 则对任意的 $n\geq 0$, 都有

$$A\cap\{S\leq n\}=A\cap\{S\leq T\}\cap\{T=n\}\in\mathscr{F}_n,$$

从而 $A=A\cap\{S\leq T\}\in\mathscr{F}_T$. □

定理 3.8 (Doob 停止定理)**.** *设 S,T 是有界停时, 且 $S\leq T$,*

(1) 设 $\{X_n\}$ 是鞅, 则 $\mathbb{E}(X_T|\mathscr{F}_S)=X_S,\text{a.s.}$;

(2) 设 $\{X_n\}$ 是上鞅, 则 $\mathbb{E}(X_T|\mathscr{F}_S)\leq X_S,\text{a.s.}$.

证明. 对于上鞅的情形, 设 $\{X_n\}$ 是上鞅, $S\leq T\leq N$, 任取 $A\in\mathscr{F}_S$. 首先设 $T=S+1$, 对任意的 $0\leq i\leq N$, 记

$$A_i=A\cap[S=i]\cap[T\geq i]\in\mathscr{F}_i.$$

利用命题3.1, 有

$$\int_A(X_S-X_T)\mathrm{d}\mathbb{P}=\sum_{i=0}^{N}\int_{A_i}(X_i-X_{i+1})\mathrm{d}\mathbb{P}\geq 0\implies\int_A X_S\mathrm{d}\mathbb{P}\geq\int_A X_T\mathrm{d}\mathbb{P};$$

接下来设 $T>S+1$, 令 $R_i=S+i,1\leq i\leq T-S-1$, 则 $S\leq R_1\leq R_2\leq\cdots\leq T$, 且它们之间间隔为 1. 应用上式可得

$$\int_A X_S\mathrm{d}\mathbb{P}\geq\int_A X_{R_1}\mathrm{d}\mathbb{P}\geq\int_A X_{R_2}\mathrm{d}\mathbb{P}\geq\cdots\geq\int_A X_T\mathrm{d}\mathbb{P}.$$

从而, 对任意的 $S \le T$, 及对任意的 $A \in \mathscr{F}_S$, 都有

$$\int_A X_S \mathrm{d}\mathbb{P} \ge \int_A X_T \mathrm{d}\mathbb{P} = \int_A \mathbb{E}(X_T|\mathscr{F}_S)\mathrm{d}\mathbb{P},$$

这便说明了 $\mathbb{E}(X_T|\mathscr{F}_S) \le X_S$, a.s.. 鞅的情形同理可证. □

最后, 我们将介绍鞅的上穿不等式, 为此引入一些记号. 对于上鞅 $\{X_n\}$, 记 $U_a^b(k)$ 表示 $\{X_1, X_2, \cdots, X_k\}$ 上穿区间 $[a,b]$ 的次数. 为了更精确地描述 $U_a^b(k)$, 记

$$T_0 = \inf\{n \ge 0, X_n \le a\}, \quad T_1 = \inf\{n > T_0, X_n \ge b\},$$

以及

$$T_{2i} = \inf\{n > T_{2i-1} | X_n \le a\}, \quad T_{2i+1} = \inf\{n > T_{2i} | X_n \ge b\},$$

则 T_{2i} 表示第 $i+1$ 上穿区间 $[a,b]$ 的开始时间, 而 T_{2i+1} 表示第 $i+1$ 上穿区间 $[a,b]$ 的结束时间, 并且 $\{T_n, n \ge 0\}$ 为递增的停时序列. 从而,

$$\{U_a^b(k) = i\} = \{T_{2i-1} < k < T_{2i+1}\} \in \mathscr{F}_k,$$

也即 $U_a^b(k)$ 是关于 $\mathscr{F}_k$ 可测的随机变量.

定理 3.9 (上穿不等式)**.** 设 $\{X_n\}$ 是上鞅, $U_a^b(k)$ 表示 $\{X_1, X_2, \cdots, X_k\}$ 上穿区间 $[a,b]$ 的次数, 则

$$\mathbb{E}U_a^b(k) \le \frac{1}{b-a} \cdot \mathbb{E}[(X_k - a)^-].$$

证明. 由 $\{X_n\}$ 是上鞅知 $\mathbb{E}X_{T_{2i+1}\wedge k} \le \mathbb{E}X_{T_{2i}\wedge k}$, 从而

$$\begin{aligned}
0 &\ge \mathbb{E}(X_{T_{2i+1}\wedge k} - X_{T_{2i}\wedge k}) \\
&= \mathbb{E}(X_{T_{2i+1}\wedge k} - X_{T_{2i}\wedge k}) \cdot \left(I_{\{T_{2i} \le k \le T_{2i+1}\}} + I_{\{k \ge T_{2i+1}\}}\right) \\
&\ge \mathbb{E}(X_k - a) \cdot I_{\{T_{2i} \le k \le T_{2i+1}\}} + (b-a) \cdot \mathbb{E}I_{\{k \ge T_{2i+1}\}}.
\end{aligned}$$

由于 $\{U_a^b(k) \ge i+1\} \subset \{k \ge T_{2i+1}\}$, 并且 $\{2i \le k \le 2i+1\} \subset \{U_a^b(k) = i\}$, 因此

$$(b-a) \cdot \mathbb{P}(U_a^b(k) \ge i+1) \le \mathbb{E}(X_k - a)^- \cdot I_{\{U_a^b(k)=i\}},$$

在上式中对 i 进行求和, 可得

$$\mathbb{E}U_a^b(k) = \sum_{i=0}^{\infty} \mathbb{P}(U_a^b(k) \ge i+1) \le \frac{1}{b-a} \cdot \mathbb{E}(X_k - a)^-.$$

这便完成了定理的证明. □

3.3 鞅收敛定理

本节将介绍 Doob 鞅收敛定理及其推论.

定理 3.10 (Doob 收敛定理)**.** 设 $\{X_n\}$ 为上鞅, 如果

$$\sup_{n\geq 0}\mathbb{E}|X_n| < +\infty,$$

则当 $n\to\infty$ 时, X_na.s. 收敛于一可积随机变量 X_∞. 进一步, 若 $\{X_n\}$ 为非负上鞅, 则对任意的 $n\geq 0$, 有

$$\mathbb{E}(X_\infty|\mathscr{F}_n)\leq X_n,\quad \text{a.s.}.$$

证明. 设 $a,b\in\mathbb{Q}$, 记

$$U_a^b = \lim_{k\to\infty} U_a^b(k)$$

表示 $\{X_n, n\geq 0\}$ 上穿区间 $[a,b]$ 的次数, 由定理3.9知

$$\mathbb{E}U_a^b \leq \frac{1}{b-a}\cdot\sup_{n\geq 0}\mathbb{E}(X_n-a)^- \leq \frac{1}{b-a}\cdot\left(a+\sup_{n\geq 1}\mathbb{E}X_n^-\right) < +\infty,$$

从而 U_a^b 可积, 并且 $U_a^b < +\infty$, a.s.. 记

$$W_{a,b} = \left\{\liminf_{n\to\infty} X_n < a\right\}\cup\left\{\limsup_{n\to\infty} X_n > b\right\},\quad W = \bigcup_{a,b\in\mathbb{Q}} W_{a,b}.$$

对于 $W_{a,b}$, 其表示 $\{X_n\}$ 的下极限小于 a 且上极限大于 b 的部分, 在这两个条件下, $\{X_n\}$ 极限不存在. 而对于 W, 其表示了 X_n 极限不存在的所有点. 由 $W_{a,b}\subset\{U_a^b=+\infty\}$ 以及 $U_a^b<+\infty$, a.s., 知 $\mathbb{P}(W_{a,b})=0$, 从而 $\mathbb{P}(W)=0$. 令

$$X_\infty(\omega)=\begin{cases}0, & \omega\in W,\\ \lim\limits_{n\to\infty} X_n(\omega), & \omega\in\Omega\setminus W,\end{cases}$$

则 $X_n\to X_\infty$, a.s., 且由 $\{X_n\}$ 可积知 X_∞ 可积. 进一步, 设 $\{X_n\}$ 是非负上鞅, 对任意的 $n\geq 0$, 由定理2.6得

$$\mathbb{E}(X_\infty|\mathscr{F}_n)=\mathbb{E}\left(\lim_{k\to\infty} X_k\middle|\mathscr{F}_n\right)\leq\liminf_{k\to\infty}\mathbb{E}(X_k|\mathscr{F}_n)\leq\mathbb{E}(X_{n+1}|\mathscr{F}_n)\leq X_n,\quad\text{a.s.}.$$

也即 $\mathbb{E}(X_\infty|\mathscr{F}_n)\leq X_n$, a.s.. □

命题 3.11. 设 $\{X_n\}$ 为上鞅, 如果 $\{X_n\}$ 一致可积, 也即

$$\lim_{c\to\infty}\sup_{n\geq 0}\mathbb{E}(|X_n|\cdot I_{|X_n|\geq c})=0,$$

则 X_na.s. 且 L^1 收敛于一可积随机变量 X_∞, 且

$$\mathbb{E}(X_\infty|\mathscr{F}_n)\leq X_n,\quad\text{a.s.}.$$

证明. 由 $\{X_n\}$ 一致可积知 $\sup\limits_{n\geq 0}\mathbb{E}|X_n|<+\infty$, 应用定理3.10知, 存在可积随机变量 X_∞, 使得 $X_n\to X_\infty$, a.s., 且

$$\mathbb{E}(X_\infty|\mathscr{F}_n)\leq X_n,\quad\text{a.s.}.$$

并且由一致可积性知 $X_n\xrightarrow{L^1}X_\infty$. □