概率论笔记

Fiddie

2021 年 4 月 23 日

Contents

2020-2021学年宋玉林老师开的《概率论(续)》课程.

【特别说明】此笔记未经审核, 难免有笔误, 如果有的话请在微信公众号【数学兔的极大理想】私信指出!

教材:

- 严加安《测度论讲义》, 内容高度抽象.
- 程士宏《测度论与概率论基础》, 北大本科教材.
- 《现代概率论基础》, 北师大的教材.

大部分时间都讲测度论, 还有现代概率论、条件期望、鞅. 对于抽象的东西, 找具体的东西类比会更轻松, 或者一步到位去理解.

没有期中考试, 30%平时分、70%期末考试, 正常来说可以拿80%到90%的分. 考试的证明都不会超过六七行.

授课内容:

$$\begin{cases} \text{测度论}\begin{cases}\text{测度及扩张理论}\\ \text{可测函数(构造、性质)}\\ \text{积分理论}\end{cases}\\ \text{概率论初步}\begin{cases}\text{独立性}\\ \text{条件期望(大篇幅, 现代概率论基础)}\\ \text{鞅论初步}\end{cases}\end{cases}$$

期末考试一共有8道题, 第一道题30分(叙述题), 第八道题6分.

叙述考过的内容有:

- σ-代数的条件期望. (两个条件)
- 离散时间鞅的概念.(适应、可积)
- $\{\mu_n, \mu\} \subset \mathscr{M}(E)$, $\mu_n \xrightarrow{W} \mu$与等价条件.
- 叙述三级数定理.
- 可测空间中测度的概念.
- 鞅、停时、有界停时的Doob停止定理.
- 胎紧.
- Wasserstein距离.
- Doob不等式与极大值不等式.

第 1 章　集类与测度

§ 1.1　几个重要的集类

1.1.1　集合的有关概念

复习一下集合的一些概念与性质: 设Ω是抽象集合, $A,B\subset\Omega$. 我们定义

- 差集: $A\setminus B=A\cap B^c$.
- 对称差: $A\triangle B=(A\setminus B)\cup(B\setminus A)$.
- Ω上的集类: 指每个元素都为Ω的子集构成的集合.

还有de Morgan律:

定理 1.1.1 (de Morgan律) 设I是指标集, $\{A_i\}_{i\in I}$是Ω上的集类. 则

$$\left(\bigcup_{i\in I}A_i\right)^c=\bigcap_{i\in I}A_i^c,\ \left(\bigcap_{i\in I}A_i\right)^c=\bigcup_{i\in I}A_i^c.$$

定义 1.1.1 (极限) 设$\{A_n\}$为一列集合.

(1)若$A_n\nearrow$, 则称$A\triangleq\bigcup\limits_{i=1}^{\infty}A_i$为其极限.

(2)若$A_n\searrow$, 则称$A\triangleq\bigcap\limits_{i=1}^{\infty}A_i$为其极限.

定义 1.1.2 (上极限集与下极限集) 设$\{A_n\}$为一列集合.

(1)把

$$\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n=\{\omega:\omega\text{在无穷多个}A_n\text{中}\}=\{\omega:\forall j\in\mathbb{N},\exists k\geq j(x\in A_k)\}$$

称为A_n的上极限集, 记为$\limsup\limits_{n\to\infty}A_n$.

(2)把

$$\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}A_n=\{\omega:\omega\text{不在至多有限个}A_n\text{中}\}=\{\omega:\exists j_0\in\mathbb{N},\forall k\geq j_0(x\in A_k)\}$$

称为A_n的下极限集, 记为$\liminf\limits_{n\to\infty}A_n$.

注: 显然$\liminf\limits_{n\to\infty}A_n\subset\limsup\limits_{n\to\infty}A_n$. 若$\limsup\limits_{n\to\infty}A_n=\limsup\limits_{n\to\infty}A_n$, 则称$A_n$**极限存在**, 记为$A=\lim\limits_{n\to\infty}A_n$.

定义 1.1.3 (可数分割) 设$\{A_n\}_{n\geq1}$为一列集合, 满足$A_n\cap A_m=\varnothing(m\neq n)$, 且$\bigcup\limits_{n=1}^{+\infty}A_n=\Omega$, 称$\{A_n\}$为$\Omega$的一个**可数分割**, 记为$\Omega=\sum\limits_{n=1}^{+\infty}A_n$.

注: 对任一集合序列$\{B_n\}_{n\geq 1}$, $\bigcup\limits_{n=1}^{+\infty} B_n = \Omega$, 可以用**不交并处理**来构造可数分割: 令

$$A_1 = B_1, A_n = B_n \setminus (B_1 \cup \cdots \cup B_{n-1}), \forall n \geq 2,$$

这样$\{A_n\}$两两不交且$\bigcup\limits_{n=1}^{+\infty} A_n = \Omega$, 从而是$\Omega$的一个可数分割.

1.1.2 几个重要的集类

下面要介绍的是几个非常重要的集类概念, 每天都要唠叨, 需要很熟练. 下面要介绍的π类、代数、单调类、λ类都是为σ-代数服务的, 为了让你更好认识σ-代数. 因为测度定义在σ-代数上.

设$\mathscr{C}$是一个集类.

定义 1.1.4 (π类) *称$\mathscr{C}$为π**类**, 若$\mathscr{C}$关于**有限交运算**封闭, 即若$A, B \in \mathscr{C}$, 则$A \cap B \in \mathscr{C}$.*

定义 1.1.5 (代数) *称$\mathscr{C}$为**代数**, 若$\mathscr{C}$满足:*
(1)$\Omega \in \mathscr{C}$;【等价于$\varnothing \in \mathscr{C}$】
(2)关于取余运算封闭, 即$A \in \mathscr{C} \Rightarrow A^c \in \mathscr{C}$;
(3)关于有限交运算封闭, 即$A, B \in \mathscr{C} \Rightarrow A \cap B \in \mathscr{C}$. 【等价于关于有限并运算封闭】

定义 1.1.6 (单调类) *称$\mathscr{C}$为**单调类**, 若$\mathscr{C}$关于**单调序列的极限运算**封闭, 即$\{A_n\} \subset \mathscr{C}, A_n \searrow A$或$A_n \nearrow A$, 则$A \in \mathscr{C}$.*

定义 1.1.7 (λ类) *称$\mathscr{C}$为λ**类**, 若$\mathscr{C}$满足:*
(1)$\Omega \in \mathscr{C}$;【等价于$\varnothing \in \mathscr{C}$】
(2)关于真差运算封闭, 即$A, B \in \mathscr{C}, B \subset A \Rightarrow A \setminus B \in \mathscr{C}$;【可以推出$\lambda$类关于取余运算封闭】
(3)关于单调增序列的极限运算封闭, 即$A_n \nearrow A, \{A_n\} \subset \mathscr{C} \Rightarrow A \in \mathscr{C}$. 【等价于单调减封闭】

定义 1.1.8 (σ-代数) *称$\mathscr{C}$为σ**-代数**, 若$\mathscr{C}$满足:*
(1)$\Omega \in \mathscr{C}$;【等价于$\varnothing \in \mathscr{C}$】
(2)关于取余运算封闭, 即$A \in \mathscr{C} \Rightarrow A^c \in \mathscr{C}$;
(3)关于可列交运算封闭, 即$\{A_n\} \subset \mathscr{C} \Rightarrow \bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{C}$. 【等价于关于可列并运算封闭】

例 1.1.1 $\{A | A \subset \Omega\}$是σ-代数（最大的σ-代数）.

例 1.1.2 $\{\varnothing, \Omega\}$是σ-代数（最小的σ-代数）.

例 1.1.3 $A \subset \Omega, \{\varnothing, A, A^c, \Omega\}$是$\sigma$-代数.

定理 1.1.2 我们有
(1)$\mathscr{C}$是σ-代数$\Leftrightarrow \mathscr{C}$为代数且$\mathscr{C}$为单调类.
(2)$\mathscr{C}$是σ-代数$\Leftrightarrow \mathscr{C}$为$\pi$类且$\mathscr{C}$为$\lambda$类.

证明: 两个命题的其中一个方向都是显然的.

(1)设$\mathscr{C}$为代数且为单调类, 下证$\mathscr{C}$关于可列并运算封闭.

设$\{A_n\} \subset \mathscr{C}$, 则对任意$k \geq 1$, $\bigcup_{n=1}^{k} A_n \in \mathscr{C}$. 而$\bigcup_{n=1}^{k} A_n \nearrow$, 由于$\mathscr{C}$为单调类, 则$\bigcup_{n=1}^{+\infty} A_n = \lim_{k\to\infty} \bigcup_{n=1}^{k} A_n \in \mathscr{C}$, 因此$\mathscr{C}$关于可列并运算封闭. 于是$\mathscr{C}$是$\sigma$-代数.

(2)设$\mathscr{C}$为π类也是λ类, $\{A_n\} \subset \mathscr{C}$, 下证$\bigcup_{n=1}^{\infty} A_n \in \mathscr{C}$. 由$\mathscr{C}$为$\lambda$类, 只需证$\bigcap_{n=1}^{\infty} A_n^c \in \mathscr{C}$. 显然$\{A_n^c\} \subset \mathscr{C}$, 又$\mathscr{C}$是$\lambda$类, 则$\bigcap_{n=1}^{k} A_n^c \in \mathscr{C}(\forall k \geq 1)$. 而$\mathscr{C}$关于单调减封闭, 则$\bigcap_{n=1}^{\infty} A_n^c = \lim_{k\to\infty} \bigcap_{n=1}^{k} A_n^c \in \mathscr{C}$. □

定理 1.1.3 设$\{\mathscr{C}_i\}_{i\in I}$是一族集类, 若对每个$i \in I$, $\mathscr{C}_i$为σ-代数, 则$\bigcap_{i\in I} \mathscr{C}_i$也是$\sigma$-代数.

证明: 用定义验证. □

1.1.3 单调类定理

定义 1.1.9 设$\mathscr{C}$为集类, 称$\sigma(\mathscr{C}), \lambda(\mathscr{C}), m(\mathscr{C})$分别表示包含$\mathscr{C}$的最小$\sigma$-代数、$\lambda$类、单调类. 也称为**由$\mathscr{C}$生成的$\sigma$-代数**.

包含$\mathscr{C}$的最小σ-代数(λ类、单调类) 就是由所有包含$\mathscr{C}$的$\sigma-$代数（λ类、单调类）之交.

注: 容易验证

$$m(\mathscr{C}) \subset \lambda(\mathscr{C}) \subset \sigma(\mathscr{C}).$$

定理 1.1.4 (单调类定理) 我们有:

(1)若$\mathscr{C}$为代数, 则$m(\mathscr{C}) = \sigma(\mathscr{C})$.

(2)若$\mathscr{C}$为π类, 则$\lambda(\mathscr{C}) = \sigma(\mathscr{C})$.

证明: (1)只需证$\sigma(\mathscr{C}) \subset m(\mathscr{C})$, 只需证$m(\mathscr{C})$是$\sigma$-代数! 而$m(\mathscr{C})$是单调类, 那么只需证$m(\mathscr{C})$是代数. 显然$\Omega \in \mathscr{C} \subset m(\mathscr{C})$, 下面验证另外两个条件. 首先看有限交运算封闭:

[分析]显然有$A, B \in \mathscr{C} \Rightarrow A \cap B \in \mathscr{C}$. 我们想证明$A, B \in m(\mathscr{C}) \Rightarrow A \cap B \in \mathscr{C}$. 但不能一步到位, 可以考虑先证明$A \in m(\mathscr{C}), B \in \mathscr{C} \Rightarrow A \cap B \in m(\mathscr{C})$. 要在$m(\mathscr{C})$中找比较好的集合.

令

$$\boxed{\mathscr{G}_1 = \{A \in m(\mathscr{C}) : \forall B \in \mathscr{C}(A \cap B \in m(\mathscr{C})), \text{且} A^c \in m(\mathscr{C})\}}$$

显然, $\mathscr{C} \subset \mathscr{G}_1$, 下证$\mathscr{G}_1$为单调类:

若$\{A_n\} \subset \mathscr{G}_1$, 且$A_n \nearrow A$, 则根据$\mathscr{G}_1$的定义, $\forall B \in \mathscr{C}$, $\{A_n \cap B\} \subset m(\mathscr{C})$且$A_n^c \in m(\mathscr{C})$. 利用$m(\mathscr{C})$是个单调类, 关于单调增运算封闭可知

$$A \cap B = \bigcup_{n=1}^{\infty} (A_n \cap B) \in m(\mathscr{C}), A^c = \bigcup_{n=1}^{\infty} A_n^c \in m(\mathscr{C}).$$

则$A \in \mathscr{G}_1$, 从而$\mathscr{G}_1$是单调类. 由于$m(\mathscr{C})$是包含$\mathscr{C}$的最小的单调类, 因此$m(\mathscr{C}) \subset \mathscr{G}_1$. 显然有$\mathscr{G}_1 \subset m(\mathscr{C})$, 则$m(\mathscr{C}) = \mathscr{G}_1$.

再令

$$\boxed{\mathscr{G}_2=\{A\in m(\mathscr{C}):\forall B\in m(\mathscr{C}),A\cap B\in m(\mathscr{C})\}}$$

由$\mathscr{G}_1=m(\mathscr{C})$可知$\mathscr{C}\subset\mathscr{G}_2$. (这是因为当$A\in\mathscr{C}$, 对任意$B\in m(\mathscr{C})=\mathscr{G}_1$, 都有$B\cap A\in m(\mathscr{C})$), 下证$\mathscr{G}_2$是单调类: (其实与$\mathscr{G}_1$是单调类的证明一样)

若$\{A_n\}\subset\mathscr{G}_2$且$A_n\nearrow A$, 根据$\mathscr{G}_2$的定义, 对任意$B\in m(\mathscr{C}),\{A_n\cap B\}\subset m(\mathscr{C})$. 由于$m(\mathscr{C})$是单调类, 关于单调增运算封闭, 则

$$A\cap B=\bigcup_{n=1}^{\infty}(A_n\cap B)\in m(\mathscr{C}).$$

所以$A\in\mathscr{G}_2$, 从而$\mathscr{G}_2$是单调类. 由于$m(\mathscr{C})$是包含$\mathscr{C}$的最小的单调类, 因此$m(\mathscr{C})\subset\mathscr{G}_2$. 显然有$\mathscr{G}_2\subset m(\mathscr{C})$, 则$m(\mathscr{C})=\mathscr{G}_2$. 即$m(\mathscr{C})$关于有限交封闭. 由于$\mathscr{G}_1,m(\mathscr{C})$关于取余运算封闭, 则$m(\mathscr{C})$是代数, 从而是$\sigma$-代数.

注: 证单调类的时候需要对$A_n\nearrow A$与$A_n\searrow A$都说明, 上面忽略了一部分.

(2)只需证$m(\mathscr{C})\subset\lambda(\mathscr{C})$, 只需证$\lambda(\mathscr{C})$是$\sigma$-代数, 只需证它是$\pi$类(关于有限交封闭). 思路: 先证明$\forall A\in\lambda(\mathscr{C}),B\in\mathscr{C}\Rightarrow A\cap B\in\lambda(\mathscr{C})$, 再证明$\forall A,B\in\lambda(\mathscr{C})\Rightarrow A\cap B\in\lambda(\mathscr{C})$.

令

$$\mathscr{G}_1=\{A\in\lambda\in(\mathscr{C}):\forall B\in\mathscr{C},A\cap B\in\lambda(\mathscr{C})\}.$$

则$\mathscr{C}\subset\mathscr{G}_1\subset\lambda(\mathscr{C})$, 且$\mathscr{G}_1$是$\lambda$类:

事实上, ①显然$\Omega\in\mathscr{G}_1$. ②关于真差封闭: 若$A_1,A_2\in\mathscr{G}_1,A_1\subset A_2$, 则

$$(A_2\setminus A_1)\cap B=(A_2\cap B)\cap(A_1^c\cap B)\in\lambda(\mathscr{C}).$$

③单调增封闭: 若$\{A_n\}\searrow A$, 则$\forall B\in\mathscr{C},A_n\cap B\in\lambda(\mathscr{C})$, 所以$\bigcap_{n=1}^{\infty}(A_n\cap B)\in\lambda(\mathscr{C})$, 则$A\cap B\in\lambda(\mathscr{C})\Rightarrow A\in\mathscr{G}_1$.

因此$\lambda(\mathscr{C})=\mathscr{G}_1$. 令

$$\mathscr{G}_2=\{A\in\lambda(\mathscr{C}):\forall B\in\lambda(\mathscr{C}),A\cap B\in\lambda(\mathscr{C})\}.$$

由$\mathscr{G}_1=\lambda(\mathscr{C})\Rightarrow\mathscr{C}\subset\mathscr{G}_2$, 易证$\mathscr{G}_2$是$\lambda$-类, 从而$\lambda(\mathscr{C})=\mathscr{G}_2$. □

注: 证明也有其他方法, 比如用单调序列的极限构成一个集合. 以后经常会用单调类定理.

小结: 为了证明σ-代数$\mathscr{F}$中元素有某种性质, 可以证明: (1)生成$\mathscr{F}$的代数(或π类)$\mathscr{C}$有此性质; (2)具有此性质的元素全体构成单调类(或λ类). 若$\mathscr{C}$不是代数也不是π类, 则要证明由$\mathscr{C}$构成的满足某种性质的集合构成σ-代数.

定理 1.1.5 *设$A\subset\Omega$且$A,A^c\neq\varnothing$. 设$\mathscr{C}$为Ω上的集类, 则$\sigma_A(A\cap\mathscr{C})=A\cap\sigma(\mathscr{C})$. 在这里, $A\cap\mathscr{C}$指$\{A\cap B:B\in\mathscr{C}\}$; $\sigma_A(A\cap\mathscr{C})$指以$A$为全集, 用$A\cap\mathscr{C}$生成的$\sigma$-代数.*

证明: (1) “$\sigma_A(A\cap\mathscr{C})\subset A\cap\sigma(\mathscr{C})$”: 只需证$A\cap\sigma(\mathscr{C})$构成$\sigma$-代数. 用定义验证即可.
(2) “$\sigma_A(A\cap\mathscr{C})\supset A\cap\sigma(\mathscr{C})$”: 令$\mathscr{G}=\{B\in\sigma(\mathscr{C}):A\cap B\in\sigma_A(A\cap\mathscr{C})\}$, 则$\mathscr{C}\subset\mathscr{G}$且$\mathscr{G}$是$\sigma$-代数.
综上, $\sigma(\mathscr{C})=\mathscr{G}$. □

例 1.1.4 (Borel σ-代数) *设E是拓扑空间,*

$$\mathscr{B}(E)=\sigma(\{A|A\text{为}E\text{上开集}\})=\sigma(\{A|A\text{为}E\text{上闭集}\}).$$

注: 开集的补是闭集.

基于此,

$$\mathscr{B}(\mathbb{R}^d)=\sigma(\{(-\infty,x):x\in\mathbb{R}^d\})=\cdots$$

有很多等价定义;

$$\mathscr{B}([0,T])=[0,T]\cap\mathscr{B}(\mathbb{R}).$$

定义 1.1.10 设$\mathscr{F}$为Ω上的σ-代数, 称$(\Omega,\mathscr{F})$为**可测空间**, $\mathscr{F}$中元素叫做$\mathscr{F}$-**可测集**. 若存在可数子类$\mathscr{C}$使得$\mathscr{F}=\sigma(\mathscr{C})$, 称$\mathscr{F}$**可分**, 此时$(\Omega,\mathscr{F})$叫**可分可测空间**.

例 1.1.5 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$是可分可测空间, 因为直线上开集可以写成至多可数个开区间的并.

注: 可分σ-代数元素个数不可数! 但是可以是有限个.

注: 一定要知道测度定义在可测空间上, 随机变量也定义在可测空间(样本空间)上(注意随机变量就是可测空间上的可测函数), 概率定义在σ-代数上(而不是可测空间上)(概率是$\sigma-$代数上的集函数).

定理 1.1.6 设$\mathscr{C}$是一集类.

(1)为使$m(\mathscr{C})=\sigma(\mathscr{C})$, 必须且仅需

$$A\in\mathscr{C}\Rightarrow A^c\in m(\mathscr{C}); A,B\in\mathscr{C}\Rightarrow A\cap B\in m(\mathscr{C}).$$

(2)为使$\lambda(\mathscr{C})=\sigma(\mathscr{C})$, 必须且仅需

$$A,B\in\mathscr{C}\Rightarrow A\cap B\in\lambda(\mathscr{C}).$$

注: 此定理的交可以改成并.

定理 1.1.7 设$\mathscr{C}$是一集类. 下列两个条件等价:

(1)$\lambda(\mathscr{C})=m(\mathscr{C})$.

(2)$A\in\mathscr{C}\Rightarrow A^c\in m(\mathscr{C}); A,B\in\mathscr{C}, A\cap B=\varnothing\Rightarrow A\cup B\in m(\mathscr{C})$.

§ 1.2 测度与测度扩张定理

1.2.1 与测度有关的定义

定义 1.2.1 设$\mathscr{C}$为Ω上的集类. 称定义在$\mathscr{C}$上的取值于$(-\infty,+\infty]$上的映射为**集函数**, 记为μ.

定义 1.2.2 (测度) 设$(\Omega,\mathscr{F})$是可测空间. μ为$\mathscr{F}$上的集函数. 称μ为**测度**, 若

(1)非负性: $\mu(A)\geq 0,\forall A\in\mathscr{F}$.

(2)可列可加性: $\forall\{A_n\}\subset\mathscr{F}, A_n\cap A_m=\varnothing(n\neq m)$, 有

$$\mu\left(\sum_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mu(A_n).$$

(3)空集为0: $\mu(\varnothing)=0$.

若$\mu(\Omega)<+\infty$, 则称μ是**有限测度**, $(\Omega,\mathscr{F},\mu)$叫**有限测度空间**.

若$(\Omega,\mathscr{F},\mu)$是有限测度空间，$\mu(\Omega)=1$，则称μ为**概率测度**，$(\Omega,\mathscr{F},\mu)$是**概率测度空间**.

若存在$\{A_n\}\subset\mathscr{F}$使得$\bigcup\limits_{n=1}^{\infty}A_n=\Omega$且$\mu(A_n)<+\infty$，则称$\mu$是**$\boldsymbol{\sigma}$-有限测度**，$(\Omega,\mathscr{F},\mu)$是**$\boldsymbol{\sigma}$-有限测度空间**.

注: 测度是定义在σ-代数上的.

例 1.2.1 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$上的有限测度举例.

(1)Dirac测度: $\sigma_{x_0}(A)=\begin{cases}1, x_0\in A\\ 0, x_0\notin A\end{cases}$.

(2)Bernoulli随机变量的分布形如$p\delta_1+(1-p)\delta_0$.

(3)离散型随机变量的分布是Dirac测度的凸组合.

(4)Gauss分布: $\mu(\mathrm{d}x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x$. 那么$\mu(A)=\displaystyle\int_A\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x$.

(5)均匀分布: $([0,T],\mathscr{B}(\mathbb{R})\cap[0,T])$.

(6)连续型随机变量有密度函数.

定义 1.2.3 设$(\Omega,\mathscr{F},\mu)$为一测度空间，若$A\in\mathscr{F}$，且$\mu(A)=0$，称A为**$\boldsymbol{\mu}$-零测集**. 如果任何μ-零测集的子集都属于$\mathscr{F}$，称$\mathscr{F}$关于μ是**完备的**，$(\Omega,\mathscr{F},\mu)$叫**完备测度空间**.

例 1.2.2 (代数上的测度) 设$\mathscr{F}$为代数，μ为$\mathscr{F}$上的集函数，满足:

(1)非负性: $\mu(A)\geq 0,\forall A\in\mathscr{F}$.

(2)可列可加性: $\forall\{A_n\}\subset\mathscr{F},A_n\cap A_m=\varnothing(n\neq m)$, $\underline{\text{如果}\sum\limits_{n=1}^{\infty}A_n\in\mathscr{F}}$，那么

$$\mu\left(\sum_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mu(A_n).$$

(3)空集为0: $\mu(\varnothing)=0$.

称μ为**代数上的测度**.

性质 1.2.1 设$\mathscr{F}$为代数，μ为其上测度，则:

(1)**单调性**: $A,B\in\mathscr{F},A\subset B\Rightarrow\mu(A)\leq\mu(B)$.

(2)**可减性**: $A,B\in\mathscr{F},A\subset B,\underline{\mu(B)<+\infty}$, 则$\mu(B\setminus A)=\mu(B)-\mu(A)$.

(3)**从下连续性**: $\{A_n\}\subset\mathscr{F},A_n\nearrow A,\underline{A\in\mathscr{F}},\Rightarrow\mu(A)=\lim\limits_{n\to\infty}\mu(A_n)$.

(4)**从上连续性**: $\{A_n\}\subset\mathscr{F},A_n\searrow A,\underline{A\in\mathscr{F}},\mu(A_1)<\infty,\Rightarrow\mu(A)=\lim\limits_{n\to\infty}\mu(A_n)$.

(5)**在$\varnothing$处连续**: $\{A_n\}\subset\mathscr{F},A_n\searrow\varnothing,\mu(A_1)<\infty,\Rightarrow\lim\limits_{n\to\infty}\mu(A_n)=0$.

(6)**次σ可加性**: $\{A_n\}\subset\mathscr{F},\bigcup\limits_{n=1}^{+\infty}A_n\in\mathscr{F},\Rightarrow\mu\left(\bigcup\limits_{n=1}^{+\infty}A_n\right)\leq\sum\limits_{n=1}^{+\infty}\mu(A_n)$.

(7)**半σ可加性**: $\{A_n\}\subset\mathscr{F},A\in\mathscr{F},A\subset\bigcup\limits_{n=1}^{+\infty}A_n,\Rightarrow\mu(A)\leq\sum\limits_{n=1}^{+\infty}\mu(A_n)$.

注: 可减性$\mu(B)<+\infty$是保证没有无穷减无穷的情况; 当$\mathscr{F}$是σ-代数时, 从下连续性与从上连续性的条件可以不加横线部分.

注: 从上连续性的条件也可以改为存在某个k使得$\mu(A_k) < \infty$. 去掉这个条件不成立: 比如取$A_n = [n, +\infty)$, 取Lebesgue测度. (取计数测度(集合元素个数)也可以)

证明: (3)作不交并处理, 取$B_n = A_n \setminus A_{n-1}$, $B_1 = A_1$, 则

$$\begin{aligned}\mu(A) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\sum_{n=1}^{\infty} B_n\right)\\ &= \sum_{n=1}^{\infty}\mu(B_n) = \lim_{k\to\infty}\sum_{n=1}^{k}\mu(B_n)\\ &= \lim_{k\to\infty}\sum_{n=2}^{k}(\mu(A_n)-\mu(A_{n-1}))+\mu(A_1)\\ &= \lim_{k\to\infty}\mu(A_k).\end{aligned}$$

(4)取$B_n = A_1 \setminus A_n$, 则$B_n \nearrow A_1 \setminus A$, 利用(3),

$$\mu(A_1)-\mu(A_n)=\mu(B_n)\to\mu(B)=\mu(A_1)-\mu(A)(n\to\infty).$$

从而$\mu(A_n)\to\mu(A)$.

(6)取$B_n = A_n \setminus (A_1\cup A_2\cup\cdots\cup A_{n-1})$, 则

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right)=\mu\left(\bigcup_{n=1}^{+\infty} B_n\right)=\sum_{n=1}^{+\infty}\mu(B_n)\le\sum_{n=1}^{+\infty}\mu(A_n).$$

(7)只需注意到

$$\mu(A)=\mu\left(A\cap\bigcup_{n=1}^{\infty}A_n\right)\le\sum_{n=1}^{\infty}\mu(A\cap A_n)\le\sum_{n=1}^{\infty}\mu(A_n).$$

注: 下周就可以友好地把半σ可加性忘记掉了.

注: 如果$A=\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{F}$, 那么半$\sigma$可加性就是次$\sigma$可加性.

接下来看看代数上非负集函数的性质.

定理 1.2.1 *设$\mathscr{C}$是代数, μ是$\mathscr{C}$上满足有限可加性的非负集函数, 而且$\mu(\varnothing)=0$. 则μ有单调性与可减性, 此外, 下面几个命题是等价的:*

(1)μ是σ-可加的.

(2)μ有从下连续性. (可以推出从上连续, 进一步可推出在$\varnothing$ 连续.)

(3)μ有半σ可加性.

证明: μ有单调性与可减性都是显然的. 只需看后面的等价关系.

"$(1)\Rightarrow(2)$" 与 "$(1)\Rightarrow(3)$": 前面已证.

"$(2)\Rightarrow(1)$": 取$\{A_n\}$两两不交, 则

$$P\left(\sum_{n=1}^{\infty}A_n\right)=P\left(\lim_{m\to\infty}\bigcup_{n=1}^{m}A_n\right)=\lim_{m\to\infty}P\left(\bigcup_{n=1}^{m}A_n\right)=\lim_{m\to\infty}\sum_{n=1}^{m}P(A_n)=\sum_{n=1}^{\infty}P(A_n).$$

“(3) $\Rightarrow$ (2)”：设$\{A_n\} \subset \mathscr{F}$满足$A_n \nearrow A \in \mathscr{F}$. 利用$\mu$的单调性可知

$$\lim_{n\to\infty} \mu(A_n) \le \mu(A).$$

另一方面,

$$\begin{aligned}\mu(A) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=2}^{\infty} (A_n \setminus A_{n-1}) + A_1\right) \\ &\le \sum_{n=2}^{\infty} [\mu(A_n) - \mu(A_{n-1})] + \mu(A_1) = \lim_{n\to\infty} \mu(A_n).\end{aligned}$$

注意$A = \bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$, 所以小于等于号处可以直接用次$\sigma$可加性. □

注: 此定理不需要特别关注, 用到的时候查一下即可. 另外“(3) $\Rightarrow$ (1)”相对来说更简单, 读者可以尝试一下.

注: 如果$\mu(\Omega) < +\infty$, 则从上连续与从下连续等价.

§ 1.3 外测度与测度的扩张

后面会经常用到下面的扩张与限制的概念.

定义 1.3.1 (扩张) 设$\mathscr{C}_1 \subset \mathscr{C}_2$为集类, μ_1, μ_2分别是定义在$\mathscr{C}_1, \mathscr{C}_2$上的集函数. 若对于定义在$\mathscr{C}_1$上的集合$A$有$\mu_1(A) = \mu_2(A)$, 称$\mu_2$是$\mu_1$在$\mathscr{C}_2$上的**扩张**, μ_1叫做μ_2在μ_1上的**限制**, 记为$\mu_1 = \mu_2|_{\mathscr{C}_1}$.

观察: $\mathscr{B}(\mathbb{R})$可以用一条线段的长度来推广, 我们要采用**外测度**来推广, 将**半环**$\mathscr{C}$上的非负集函数扩张成$\sigma(\mathscr{C})$上的测度. 步骤如下: ①用覆盖的手法在2^{Ω}上定义外测度μ^*. ②在2^{Ω}上挑选一些集合构成σ-代数$\mathcal{U}$, 使得$\mu^*|_{\mathcal{U}}$是测度(用卡式条件来找). ③$\mathscr{C} \subset \mathcal{U}$, 从而$\sigma(\mathscr{C}) \subset \mathcal{U}$, 这样就完成了扩张的过程.

μ	$\Rightarrow$	μ^* (外测度)	$\Rightarrow$	μ^* (测度)	$\Rightarrow$	μ^* (测度)
$\mathscr{C}$	$\Rightarrow$	2^{Ω}	$\Rightarrow$	$\mathcal{U}$	$\Rightarrow$	$\sigma(\mathscr{C})$

下面记

$$\mathscr{C}_{\Sigma f} = \left\{\sum_{n=1}^{m} A_n : A_n\text{两两不交}, A_n \in \mathscr{C}, \forall n \ge 1\right\}$$

表示$\mathscr{C}$中所有集合的**有限不交并**构成的集类. 这里的f表示finite(有限的). 显然$\mathscr{C} \subset \mathscr{C}_{\Sigma f}$.

定义 1.3.2 (半环) 设$\mathscr{C}$是集类, 满足:
(1)有限交封闭: $\forall A, B \in \mathscr{C} \Rightarrow A \cap B \in \mathscr{C}$;
(2)$\forall A, B \in \mathscr{C}, A \setminus B \in \mathscr{C}_{\Sigma f}$;
(3)（默认的）$\varnothing \in \mathscr{F}$.
则称$\mathscr{C}$是一个**半环**.

例 1.3.1 $\mathscr{C} = \{(a, b] | a \le b\}$是个半环.

定义 1.3.3 (半代数) 称$\mathscr{C}$为**半代数**, 若$\mathscr{C}$为半环且$\Omega \in \mathscr{C}$.

注: 如果$\mathscr{C}$是半环, 不能说$\mathscr{C}\cup\{\Omega\}$是个半代数, 比如说

$$\mathscr{C}_1=\{(a,b]|a\le b\}\cup\{\mathbb{R}\}$$

不是半代数, 因为$(-\infty,a]$不在$\mathscr{C}_1$里!

例 1.3.2 $\mathscr{C}=\{(a,b]|a\le b\}\cup\{(-\infty,x]|x\in\mathbb{R}\}\cup\{(y,+\infty)|y\in\mathbb{R}\}\cup\{\mathbb{R}\}$是个半代数.

定义 1.3.4 (外测度) 设μ^*为定义在2^Ω上的非负集函数, 若μ^*满足单调性与次σ-可加性, 则称μ^*是**外测度**. 约定$\mu^*(\varnothing)=0$.

注: 这里的定义没有σ-可加性! 注意2^Ω不是σ-代数.

下面要把半环$\mathscr{C}$上的集函数变成测度, 要加上半σ-可加性.

命题 1.3.1 设$\mathscr{C}$是集类, μ为$\mathscr{C}$上的非负集函数, 且具有半σ-可加性. 令

$$\mu^*(A)=\inf\left\{\sum_{n=1}^\infty\mu(A_n):A_n\in\mathscr{C},A\subset\bigcup_{n=1}^\infty A_n\right\}$$

(“最小的覆盖”), 约定如果没有集合可以覆盖A, 此时$\mu^*(A)=\inf\varnothing\triangleq+\infty$. 则$\mu^*$为$\Omega$的**由$\mu$诱导出来的外测度**, 且$\mu^*|_{\mathscr{C}}=\mu$.

证明: (1)先验证μ^*是外测度. 单调性是显然的. 下面验证次σ-可加性. 对任意$\{A_n\}\subset\mathscr{C}$, 由外测度定义,

$$\forall\varepsilon>0,\exists\{A_{n,k}\}_{k\ge1}\subset\mathscr{C},\text{使得}A_n\subset\bigcup_{k=1}^\infty A_{n,k}(\text{覆盖}),\text{且}\mu^*(A_n)+\frac{\varepsilon}{2^n}\ge\sum_{n=1}^\infty\mu(A_{n,k}).$$

注意

$$\bigcup_{n=1}^\infty A_n\subset\bigcup_{n=1}^\infty\bigcup_{k=1}^\infty A_{n,k}.$$

因此根据外测度定义,

$$\mu^*\left(\bigcup_{n=1}^\infty A_n\right)\le\sum_{n=1}^\infty\sum_{k=1}^\infty\mu(A_{n,k})\le\sum_{n=1}^\infty\mu^*(A_n)+\varepsilon.$$

因此μ^*是外测度.

(2)对任意$A\in\mathscr{C}$, 显然$\mu^*(A)\le\mu(A)$. (A可以覆盖自己). 而用半σ-可加性可以推出不等号反向:

$$\forall\{A_n\}\subset\mathscr{C},\text{若}A\subset\bigcup_{n=1}^\infty A_n,\text{则}\mu(A)\le\sum_{n=1}^\infty\mu(A_n).$$

由$\{A_n\}$的任意性, $\mu(A)\le\mu^*(A)$. □

引理 1.3.2 设$\mathscr{C}$是半环, μ是$\mathscr{C}$上的集函数, 满足$\mu(\varnothing)=0$. 则

$$\mu\text{在}\mathscr{C}\text{上}\sigma\text{-可加}\Leftrightarrow\text{有限可加且半}\sigma\text{-可加}.$$

证明: 留作作业. □

注: 这个引理把$\mathscr{C}$的条件由代数变成了半环!

定理 1.3.3 设μ^*是外测度, 定义

$$\mathcal{U}=\{A\subset\Omega|\mu^*(D)=\mu^*(A\cap D)+\mu^*(A^c\cap D),\forall D\subset\Omega\}.$$

则$\mathcal{U}$为σ-代数且$\mu^*|_{\mathcal{U}}$为测度.

注: $\mathcal{U}$中元素称为满足**Carathéodory条件**. 这里找A使得在外测度意义下分割D保持加性. D也叫"检验集合".

证明: (1)先证明$\mathcal{U}$是代数. 显然$\Omega\subset\mathcal{U}$, 且$\mathcal{U}$关于取余运算封闭, 下证$\mathcal{U}$关于有限并运算封闭.

对任意$A,B\in\mathcal{U},D\in\Omega$, 有

$$\begin{aligned}\mu^*(D)&=\mu^*(A\cap D)+\mu^*(A^c\cap D)\\&=\mu^*(A\cap D\cap B)+\mu^*(A^c\cap D\cap B)+\mu^*(A^c\cap D\cap B^c)+\mu^*(A\cap D\cap B^c)\\&\geq\mu^*((A\cup B)\cap D)+\mu^*((A\cup B)^c\cap D).\qquad(\text{次}\sigma\text{-可加性})\end{aligned}$$

又根据次σ-可加性, "$\leq$"方向也对, 因此$A\cup B\in\mathcal{U}$.

(2)下证$\mathcal{U}$是单调类. 只需证明对于$\{A_n\}\subset\mathcal{U},A_n\cap A_m=\varnothing,\forall n\neq m$, 则$\sum\limits_{n=1}^{\infty}A_n\in\mathcal{U}$. (作不交并处理即可得到单调类!) 由上一步可知

$$\begin{aligned}\mu^*(D)&=\mu^*(A_1\cap A_2\cap D)+\mu^*(A_1\cap A_2^c\cap D)+\mu^*(A_2\cap A_1^c\cap D)+\mu^*(A_1^c\cap A_2^c\cap D)\\&=\mu^*(A_1\cap D)+\mu^*(A_2\cap D)+\mu^*((A_1\cup A_2)^c\cap D)\\&=\cdots=\sum_{i=1}^{n}\mu^*(A_i\cap D)+\mu^*\left(\left(\sum_{i=1}^{n}A_i\right)^c\cap D\right),\forall n.\\&\geq\sum_{i=1}^{n}\mu^*(A_i\cap D)+\mu^*\left(\left(\sum_{i=1}^{\infty}A_i\right)^c\cap D\right),\forall n.\end{aligned}$$

对上式让$n\to\infty$, 可得

$$\begin{aligned}\mu^*(D)&\geq\sum_{i=1}^{\infty}\mu^*(A_i\cap D)+\mu^*\left(\left(\sum_{i=1}^{\infty}A_i\right)^c\cap D\right)\\&\geq\mu^*\left(\left(\sum_{i=1}^{\infty}A_i\right)\cap D\right)+\mu^*\left(\left(\sum_{i=1}^{\infty}A_i\right)^c\cap D\right).\qquad(\text{次}\sigma\text{-可加性})\end{aligned}$$

而"$\leq$"方向是显然的, 因此$\sum\limits_{i=1}^{\infty}A_i\in\mathcal{U}$. 而单调递减封闭用取余封闭即可.

综上, $\mathcal{U}$是σ-代数.

(3)要证μ^*有σ-可加性, 取上面的$D=\sum\limits_{i=1}^{\infty}A_i$, 则上式变为

$$\mu^*\left(\sum_{i=1}^{\infty}A_i\right)\geq\sum_{i=1}^{\infty}\mu^*(A_i).$$

而"$\leq$"用次σ-可加性即可, 因此$\mu^*|_{\mathcal{U}}$是测度. □

注: 称$\mathcal{U}$中元素为**μ^*-可测集**.

命题 1.3.4 设$\mathscr{C}$是集类, $\varnothing \in \mathscr{C}$, μ是$\mathscr{C}$上满足半σ-可加性的集合, 且$\mu(\varnothing)=0$. 令μ^*表示由μ诱导出的外测度, 则

$$A\text{为}\mu^*\text{-可测集} \Leftrightarrow \forall D \in \mathscr{C}, \mu(D) \geq \mu^*(A\cap D)+\mu^*(A^c \cap D).$$

注: D不需要对所有Ω都成立也可以! 且$\mu^*|_{\mathscr{C}}=\mu$. 注意外测度有次σ-可加性, 故"$\leq$"是显然的.

证明: "$\Rightarrow$": 显然. "$\Leftarrow$": 设$C\subset\Omega$是任意子集, 下证$\mu^*(C)\geq\mu^*(A\cap C)+\mu^*(A^c\cap C)$. 不妨假定$\mu^*(C)<+\infty$(等于正无穷的情况是显然的), 由外测度的定义,

$$\mu^*(C)=\inf\left\{\sum_{n=1}^{\infty}C_n : C\subset\bigcup_{n=1}^{\infty}\mu(C_n), C_n\in\mathscr{C}\right\}.$$

(把Ω中的元素"变成"$\mathscr{C}$中的元素).

对任意的$\varepsilon>0$, 存在$\{C_n\}\subset\mathscr{C}$, 使得$C\subset\bigcup\limits_{n=1}^{\infty}C_n$, 且于是

$$\begin{aligned}\mu^*(C) &\geq \sum_{n=1}^{\infty}\mu(C_n)-\varepsilon, \geq \sum_{n=1}^{\infty}(\mu^*(A\cap C_n)+\mu^*(A^c\cap C_n))-\varepsilon\\ &\geq \mu^*\left(\bigcup_{n=1}^{\infty}(A\cap C_n)\right)+\mu^*\left(A^c\cap\bigcup_{n=1}^{\infty}C_n\right)-\varepsilon\\ &\geq \mu^*(A\cap C)+\mu^*(A^c\cap C)-\varepsilon. \qquad \text{(测度的单调性)}\end{aligned}$$

由ε的任意性, 欲证式子成立. □

引理 1.3.5 (Powerful) 设μ_1^*, μ_2^*是$\mathscr{C}$上两个有限测度, 假定$\Omega\in\mathscr{C}$, 且$\mathscr{C}$是<u>π类</u>, 若μ_1, μ_2在$\mathscr{C}$上一致, 即

$$\mu_1(A)=\mu_2(A), \forall A\in\mathscr{C}.$$

那么μ_1, μ_2在$\sigma(\mathscr{C})$上也一致.

证明: 令$\mathscr{G}=\{A:\mu_1(A)=\mu_2(A), \forall A\in\sigma(\mathscr{G})\}$, 则$\mathscr{G}$是$\pi$类, 下证$\mathscr{G}$是$\lambda$类.

显然$\Omega\in\mathscr{G}$, 用测度的可减性可以推真差运算封闭(有限测度保证了可减性), 用测度的从下连续性可以推对单调增封闭. 因此$\mathscr{G}$是λ类. □

注: 如果两个随机变量的分布函数一样, 则它的分布也一样. 事实上, 设$(\Omega, \mathscr{F}, P)$中的两个随机变量ξ与η. (注意随机变量都是可测函数, $\xi^{-1}(B)\in\mathscr{F}, \forall B\in\mathscr{B}(\mathbb{R})$) 分布函数(distribution function)为

$$F_\xi(x)=P(\xi\leq x)=P(\xi\in(-\infty,x)).$$

分布(distribution)是定义在$(\mathbb{R},\mathscr{B}(\mathbb{R}))$的概率:

$$P\circ\xi^{-1}(B)\triangleq P(\xi\in B), \forall B\in\mathscr{B}(\mathbb{R}),$$

这是个概率测度, 叫做ξ在P下的分布.

分布决定分布函数: 取B为$(-\infty,x)$即可; 而所有$(-\infty,x)$构成的集类为π类, 所以分布函数决定分布.

定理 1.3.6 (Carathéodory测度扩张定理) 设$\mathscr{C}$为Ω上的半环，μ是$\mathscr{C}$上的一个σ可加的非负集函数，则μ可以扩张成为$\sigma(\mathscr{C})$上的测度. 若进一步μ在$\mathscr{C}$上是σ有限的，且

$$\Omega \in C_\sigma = \left\{ \bigcup_{n=1}^{\infty} C_n : C_n \in \mathscr{C}, n \geq 1 \right\},$$

则这一扩张是唯一的，并且扩张所得的测度在$\sigma(\mathscr{C})$上也是σ有限的.

证明: (1)由引理1.3.2, μ在$\mathscr{C}$上有半σ-可加性. 令μ^*为μ诱导的外测度, 即

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A_n \in \mathscr{C}, A \subset \bigcup_{n=1}^{\infty} A_n \right\},$$

则μ^*是2^Ω上的外测度. 令$\mathcal{U}$为μ^*-可测集全体, 即

$$\mathcal{U} = \{A \subset \Omega | \mu^*(D) = \mu^*(A \cap D) + \mu^*(A^c \cap D), \forall D \in \mathscr{C}\}.$$

则根据定理1.3.3, $\mathcal{U}$是σ-代数且$\mu^*|_{\mathcal{U}}$是测度.

下证$\mathscr{C} \subset \mathcal{U}$, 设$A \in \mathscr{C}$, 对任意$D \in \mathscr{C}$, 有

$$D = (A \cap D) \cup \underbrace{(A^c \cap D)}_{\in \mathscr{C}_{\Sigma f}} = (A \cap D) \cup \sum_{i=1}^{n} A_i, \text{其中} A_i \text{两两不交}.$$

由有限可加性,

$$\mu(D) = \mu(A \cap D) + \sum_{i=1}^{n} \mu(A_i) \geq \mu^*(A \cap D) + \mu^*(A^c \cap D).$$

（不能写$\mu(A^c \cap D)$，因为有可能不在$\mathscr{C}$里，此外ΣA_i覆盖了$A^c \cap D$）. 所以$A \in \mathcal{U}$. 所以$\mathscr{C} \subset \mathcal{U}$, 进一步, $\sigma(\mathscr{C}) \subset \mathcal{U}$, (因为$\mathcal{U}$是包含了$\mathscr{C}$的$\sigma$-代数).

(2)由于$\Omega \in \mathscr{C}_\sigma$，$\mu$在$\mathscr{C}$上$\sigma$-有限，则存在$\Omega$的一个分割$\{A_n\} \subset \mathscr{C}$，使得$\Omega = \sum_{n=1}^{\infty} A_n$(两两不交)且$\mu(A_n) < +\infty$. **(WHY?)**

令$\mathscr{C}_n = A_n \cap \sigma(\mathscr{C}) \subset \sigma(\mathscr{C})$, 此时$\mu^*|_{\mathscr{C}_n}$是有限测度. 下证唯一性. 设$\mu_1, \mu_2$为$\mu$在$\sigma(\mathscr{C})$上扩张所得的测度, 则$\mu_1|_{A_n \cap \mathscr{C}}$与$\mu_2|_{A_n \cap \mathscr{C}}$一致. 又$\sigma(A_n \cap \mathscr{C}) = A_n \cap \sigma(\mathscr{C})$且$A_n \cap \mathscr{C}$是$\pi$类, 从而$\mu_1|_{A_n \cap \sigma(\mathscr{C})} = \mu_2|_{A_n \cap \sigma(\mathscr{C})}$.

因此对任意$C \in \sigma(\mathscr{C})$, 都有

$$\mu_1(C) = \mu_1 \left(C \cap \sum_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu_1(C \cap A_n) = \sum_{n=1}^{\infty} \mu_2(C \cap A_n) = \mu_2 \left(C \cap \sum_{n=1}^{\infty} A_n \right) = \mu_2(C).$$

唯一性证完. □

命题 1.3.7 设$(\Omega,\mathscr{F},\mu)$是测度空间, 令

$$\begin{aligned}\mathscr{N}&=\{N\subset\Omega|\exists A\in\mathscr{F},\text{满足}N\subset A\text{且}\mu(A)=0\},\\ \overline{\mathscr{F}}&=\{A\cup N|A\in\mathscr{F},N\in\mathscr{N}\}.\end{aligned}$$

在$\overline{\mathscr{F}}$上定义集函数$\bar{\mu}(\overline{A})=\mu(A)$, 其中$\overline{A}=A\cup N,A\in\mathscr{F},N\in\mathscr{N}$. 则$(\Omega,\overline{\mathscr{F}},\bar{\mu})$是完备测度空间.

证明: (1)下证$\overline{\mathscr{F}}$是σ-代数.
①显然, $\Omega\in\overline{\mathscr{F}}$.
②对于$\{\overline{A_n}\}\subset\overline{\mathscr{F}}$, 存在$\{B_n\}\subset\mathscr{F}$以及$\{N_n\}\subset\mathscr{F}$, 使得$\overline{A_n}=B_n\cup N_n$, 从而

$$\bigcup_{n=1}^{\infty}\overline{A_n}=\underbrace{\bigcup_{n=1}^{\infty}B_n}_{\in\mathscr{F}}\cup\underbrace{\bigcup_{n=1}^{\infty}N_n}_{\mu(\cdot)=0}\in\overline{\mathscr{F}}.$$

③对$\overline{A}\in\overline{\mathscr{F}}$, 存在$A\in\mathscr{F},N\in\mathscr{N}$, 使得$\overline{A}=A\cup N$, 且存在$B\in\mathscr{F}$使得$\mu(B)=0,N\subset B$. 于是

$$\overline{A}^c=(A\cup N)^c=A^c\cap N^c=A^c\cap((N^c\cap B^c)\cup B^c)=\underbrace{(A^c\cap B^c)}_{\in\mathscr{F}}\cup\underbrace{(A^c\cap N^c\cap B)}_{\subset B\in\mathscr{N}}\in\overline{\mathscr{F}}.$$

(2)下证$\bar{\mu}$是$\overline{\mathscr{F}}$上的测度. 设$\{\overline{A_n}\}\subset\overline{\mathscr{F}}$两两不交, 则存在$\{A_n\}\subset\mathscr{F},\{N_n\}\subset\mathscr{N}$, 使得$\overline{A_n}=A_n\cup N_n(n\geq 1)$. 显然$\{A_n\}$两两不交且$\sum\limits_{n=1}^{\infty}N_n\in\mathscr{N}$, 因此

$$\bar{\mu}\left(\sum_{n=1}^{\infty}\overline{A_n}\right)=\bar{\mu}\left(\sum_{n=1}^{\infty}A_n\right)+\bar{\mu}\left(\sum_{n=1}^{\infty}N_n\right)=\mu\left(\sum_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mu(A_n)=\sum_{n=1}^{\infty}\bar{\mu}(\overline{A_n}).$$

(3)下证完备. 设$\overline{A}\in\overline{\mathscr{F}}$, 且$\bar{\mu}(\overline{A})=0$. 对任意$\overline{B}\subset\overline{A}$, 有$\overline{B}\in\overline{\mathscr{F}}$. 注意$\overline{A}\in\overline{\mathscr{F}}$, 所以存在$B\in\mathscr{F},N\in\mathscr{N}$, 使得$\overline{A}=B\cup N\in\mathscr{N}$(这是因为$\mu(B)=0$, 存在$C\in\mathscr{F}$使得$\mu(C)=0,N\subset C$, 于是$B\cup N\subset B\cup C\in\mathscr{F}$满足$\mu(B\cup C)=0$), 因此$\overline{B}\in\mathscr{N}$, 特别地, $\overline{B}\in\overline{\mathscr{F}}$. □

定理 1.3.8 设μ是半环$\mathscr{C}$上的σ-有限的非负集函数, 若$\Omega\in\mathscr{C}_\sigma$且$\mu$在$\mathscr{C}$上$\sigma$-有限, 则$(\Omega,\mathcal{U},\mu^*|_{\mathcal{U}})$是$(\Omega,\sigma(\mathscr{C}),\mu^*|_{\sigma(\mathscr{C})})$的完备化.

§ 1.4 Lebesgue-Stieltjes测度

对于$a,b\in\mathbb{R}^n$, 定义

$$(a,b]\triangleq\{(x_1,x_2,\cdots,x_n)|a_i<x_i\leq b_i,i=1,2,\cdots,n\}.$$

令$\mu((a,b])=\prod\limits_{i=1}^{n}(b_i-a_i)$.

命题 1.4.1 若设$\mathscr{C}$是半环, 则上述μ是$\mathscr{C}$上σ-可加的集函数.

注: 由Carathéodory扩张定理, μ可以扩张成$\sigma(\mathscr{C})$上的σ-有限的测度, 而且此扩张是唯一的, 叫**Lebesgue测度**.

$\mathbb{R}^n = \bigcup_{m=1}^{\infty}(-m,m]$, 那么Borel集类$\mathscr{B}(\mathbb{R}^n)$的完备化如下:

$$(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n) \underbrace{, m(\mathrm{d}x)}_{\text{Lebesgue测度}}) \Rightarrow (\mathbb{R}^n, \underbrace{\overline{(\mathscr{B}(\mathbb{R}^n))}}_{\text{Lebesgue可测集}}, \overline{m}(\mathrm{d}x)).$$

Lebesgue可测集与Borel可测集只差一个零测集.

定义 1.4.1 设$F:\mathbb{R}^n \to \mathbb{R}$是右连续函数, 对任意$a, b \in \mathbb{R}^n$, 令

$$\Delta_{b,a}F = \Delta^{(n)}_{b_n,a_n}\Delta^{(n-1)}_{b_{n-1},a_{n-1}}\cdots\Delta^{(2)}_{b_2,a_2}\underbrace{\Delta^{(1)}_{b_1,a_1}F}_{n-1\text{元函数}}.$$

其中, $\Delta^{(i)}_{b_i,a_i}G(x) = G(x_1,\cdots,x_{i-1},b_i,x_{i+1},\cdots,x_n) - G(x_1,\cdots,x_{i-1},a_1,x_{i+1},\cdots,x_n)$.

若$\Delta_{b,a}F \ge 0$, 称F是**增函数**.

例 1.4.1 $F(x,y) = P(X \le x, Y \le y)$是增函数. $\Delta_{b,a}F = P((X,Y) \in (a,b])$.

定义 1.4.2 设μ是$(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$上的$\sigma$-有限测度. 若对任意$A \in \mathscr{C} \triangleq \{(a,b] | a \le b, a, b \in \mathbb{R}^n\}$, $\mu(A) < \infty$, 则称μ为$(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$上的*Lebesgue-Stieltjes(L-S)*测度.

注: Lebesgue测度是L-S测度, 分布是L-S测度.

下面看$\mathscr{B}(\mathbb{R}^n)$上的L-S测度与$\mathbb{R}^n$上的右连续增函数的关系.

定理 1.4.2 设F是$\mathbb{R}^n$上的右连续增函数. 令

$$\mu_F(\varnothing) = 0, \mu_F((a,b]) = \Delta_{b,a}F, a \le b, a, b \in \mathbb{R}^n.$$

则μ_F可以唯一扩张成$\mathscr{B}(\mathbb{R}^n)$上的*L-S*测度. 反之, 设$\mu$是$\mathscr{B}(\mathbb{R}^n)$的*L-S*测度, 则存在$\mathbb{R}^n$上的右连续增函数$F$, 使得$\mu$是$\mu_F$从$\mathscr{C}$到$\mathscr{B}(\mathbb{R}^n)$上的唯一扩张.

注: 不唯一: 考虑

$$F(x_1,\cdots,x_n) = x_1\cdots x_n, \mu_F((a,b]) = \Delta_{b,a}F = \prod_{i=1}^{n}(b_i - a_i).$$

于是F可以唯一扩张成$\mathscr{B}(\mathbb{R}^n)$上的L-S测度: 但是$F(x_1,\cdots,x_n) = x_1\cdots x_n + 1$ 扩张成的L-S测度与前面一样.

注: $F(x_1,\cdots,x_n) = P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n)$, 扩张成的L-S测度是

$$\mu_F((a,b]) = P(a_1 < X_1 \le b_1, \cdots, a_n < X_n \le b_n).$$

可以扩张成分布: $P \circ (X_1,\cdots,X_n)^{-1}$.

第 2 章　可测映射

§ 2.1　定义与基本性质

定义 2.1.1 设$(\Omega,\mathscr{F}),(E,\mathscr{E})$是两个可测空间, $f:\Omega\to E$是映射. 若对任意$A\in\mathscr{E}$, 如果

$$(原像)f^{-1}(A)\triangleq\{\omega\in\Omega|f(\omega)\in A\}\in\mathscr{F},$$

则称f是$\mathscr{F}/\mathscr{E}$**可测的**.

引理 2.1.1 设$\mathscr{C}$是集类, 则$\sigma(f^{-1}(\mathscr{C}))=f^{-1}(\sigma(\mathscr{C}))$.

证明: 留作作业. 先证明$f^{-1}(\sigma(\mathscr{C}))$是$\sigma$-代数, 再用单调类定理. □

下面设

$$\overline{\mathbb{R}}=\mathbb{R}\cup\{-\infty,+\infty\},$$

并设

$$\mathscr{B}(\overline{\mathbb{R}})=\sigma(\mathscr{B}(\mathbb{R})\cup\{\{-\infty\}\}\cup\{\{+\infty\}\})=\sigma(\{[-\infty,x]|x\in\mathbb{R}\}).$$

定义 2.1.2 设f是$(\Omega,\mathscr{F})\to(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$可测映射, 则称$f$是***Borel*可测函数**, 若$f$只取实值, 则称$f$为**实可测函数**, 如果$f$可取复值, 且实部与虚部同为实值可测, 称$f$为**复可测函数**.

下面来看可测映射的刻画. 以后为了方便, 记$f^{-1}([-\infty,a))$为$[f\le a]$.

命题 2.1.2 设$f:(\Omega,\mathscr{F})\to(E,\mathscr{E})$, $\mathscr{C}\subset\mathscr{E}$, 且$\sigma(\mathscr{C})=\mathscr{E}$. 则

$$f为\mathscr{F}/\mathscr{E}可测\Leftrightarrow\forall A\in\mathscr{C},f^{-1}(A)\in\mathscr{F}.$$

证明: "$\Rightarrow$": 由定义显然.

"$\Leftarrow$": 令$\mathscr{G}=\{A\in\mathscr{E}|f^{-1}(A)\in\mathscr{F}\}$, 显然$\mathscr{E}\subset\mathscr{G}$, 且根据前面引理可知$\mathscr{G}$是$\sigma$-代数. 所以$\sigma(\mathscr{C}\subset\mathscr{G}$. 又由于$\mathscr{G}\subset\mathscr{E}=\sigma(\mathscr{C})$, 因此$\mathscr{G}=\sigma(\mathscr{C})$. □

推论 2.1.3 设$f:\Omega\to\overline{\mathbb{R}}$是映射,则下列命题等价.

(1)f是$\mathscr{F}/\mathscr{B}(\overline{\mathbb{R}})$可测.

(2)$\forall a\in\mathbb{R},[f\le a]\in\mathscr{F}$.

(3)$\forall a\in\mathbb{R},[f<a]\in\mathscr{F}$.

(4)$\forall a\in\mathbb{R},[f\ge a]\in\mathscr{F}$.

(5)$\forall a\in\mathbb{R},[f>a]\in\mathscr{F}$.

证明: (2)只需要证: $[-\infty,a]$生成的σ-代数是不是$\mathscr{B}(\overline{\mathbb{R}})$. 根据$\mathscr{B}(\mathbb{R})$可以由$(-\infty,a]$生成即可.

(3)事实上$\sigma(\{[-\infty,a]|a\in\mathbb{R}\})=\sigma(\{[-\infty,a)|a\in\mathbb{R}\})$, 这是因为$[-\infty,a]=\bigcap_{n=1}^{\infty}[-\infty,a+1/n)$. □

关于无穷大的约定: 正无穷减正无穷、正无穷加负无穷、无穷除以无穷、任何数除以0都认为无意义.

命题 2.1.4 $(\Omega,\mathscr{F})$上的实值可测函数全体构成实数域上的线性空间.

证明: 设$f,g:\Omega\to\mathbb{R}$是实可测函数. 只需要注意到

$$[f+g<a]=\bigcup_{r\in\mathbb{R}}([f<a-r]\cap[g<r])=\bigcup_{r\in\mathbb{Q}}([f<a-r]\cap[g<r]).$$

容易验证af也是实可测函数. □

命题 2.1.5 设$f,g,\{f_n\}_{n\geq 1}$都是$(\Omega,\mathscr{F})$上的可测函数, 则

(1)若$f+g$处处有意义, 则$f+g$可测.

(2)$\forall a\in\mathbb{R}$, af可测.

(3)$f\wedge g, f\vee g$均可测.

(4)fg可测.

(5)若f/g处处有意义, 则f/g可测.

(6) $\inf\limits_n f_n, \sup\limits_n f_n, \liminf\limits_{n\to\infty} f_n, \limsup\limits_{n\to\infty} f_n$均可测.

证明: (3)$\forall x\in\mathbb{R}$, $[f\wedge g\leq x]=[f\leq x]\cup[g\leq x]\in\mathscr{F}, [f\vee g\leq x]=[f\leq x]\cap[g\leq x]\in\mathscr{F}$.

(4)$fg=(f^+-f^-)(g^+-g^-)=f^+g^++f^-g^--f^-g^+-f^+g^-$, 其中$g^+=g\vee 0, g^-=-g\wedge 0$, 所以可以不妨设$f,g$非负, 于是对任意$x\in\mathbb{R}$, 如果$x\leq 0$, 则$[f-g<x]=\varnothing$, 如果$x>0$, 则

$$[f-g<x]=[f=0]\cup[g=0]\cup\bigcup_{r\in\mathbb{Q}^+}([f<r]\cap[g<x/r])\in\mathscr{F}.$$

(5)$f/g=f\cdot(1/g)$. 对任意$x\in\mathbb{R}$, 如果$x<0$, 则$[1/g<x]=[1/x<g<0]\in\mathscr{F}$; 如果$x=0$, 则$[1/g<x]=[g<0]$; 如果$x>0$, 则$[1/g<x]=[g<0]\cup[g>1/x]$.

(6)$[\inf\limits_n f_n<x]=\bigcup\limits_{n=1}^{\infty}[f_n<x]$, 而$[\sup\limits_n f_n\leq x]=\bigcap\limits_{n=1}^{\infty}[f_n\leq x]$. $\limsup\limits_n f_n=\inf\limits_n\sup\limits_{k\geq n} f_k$. □

注: $[\sup\limits_n f_n>x]=\bigcup\limits_{n=1}^{\infty}[f_n>x]$, $[\inf\limits_n f_n\geq x]=\bigcap\limits_{n=1}^{\infty}[f_n\geq x]$.

命题 2.1.6 设$f:(\Omega,\mathscr{F})\to(E,\mathscr{E})$可测, $g:(E,\mathscr{E})\to(G,\mathscr{G})$可测, 则$f\circ g:(\Omega,\mathscr{F})\to(G,\mathscr{G})$可测.

证明: 只需要验证$(f\circ g)^{-1}(A)=f^{-1}(\underbrace{g^{-1}(A)}_{\in\mathscr{E}})\in\mathscr{F}$. □

注: 随机变量是$(\Omega,\mathscr{F})$到$(\mathbb{R},\mathscr{B}(\mathbb{R}))$的可测映射. 对随机变量$\xi:\Omega\to\mathbb{R}$, $\xi^{-1}(A)\in\mathscr{F},\forall A\in\mathscr{B}(\mathbb{R})$.

注: 设ξ是随机变量, f是可测函数, 那么根据这个命题, $f(\xi)$是随机变量.

§ 2.2 可测函数的构造

设$A\subset\Omega$, 记$I_A(\omega)=\begin{cases}1, & \omega\in A,\\ 0, & \omega\notin A.\end{cases}$ 为A的**示性函数**.

设f是Ω上的函数, f仅取有限个值$\{a_1,\cdots,a_n\}\subset\mathbb{R}$. $A_i=f^{-1}(\{a_i\})$, 把$f=\sum\limits_{i=1}^{n}a_iI_{A_i}$叫**简单函数**.

定理 2.2.1 设函数$f:\Omega\to\overline{\mathbb{R}}$.
(1)若$f\geq 0$, 则f为$\mathscr{F}/\mathscr{B}(\mathbb{R})$可测$\Leftrightarrow$存在非负可测简单函数序列$\{f_n\}$使得$f_n\leq f_{n+1}$且$f=\lim\limits_{n\to\infty}f_n$.
(2)f可测$\Leftrightarrow$存在非负可测函数g,h使得$f=g-h$.

证明: (1)仅证明充分性. 考虑

$$\boxed{f_n=\sum_{k=0}^{n\cdot 2^n-1}\frac{k}{2^n}I_{[\frac{k}{2^n}\leq f<\frac{k+1}{2^n}]}+nI_{[f\geq n]}}.$$

则$f_n\to f$（逐点）, 且f_n关于n递增. 特别地, 若f有界, 则这个收敛是一致收敛.
(2)注意$f=f^+-f^-$. □

示性函数在概率上的运用:
(1)设A_i是第i次Bernoulli独立重复试验成功, $1\leq i\leq n$. 则$\xi=\sum\limits_{i=1}^{n}I_{A_i}\sim B(n,p)$.
(2)Poisson分布: $A_i=[\xi=i],\xi=\sum\limits_{i=0}^{\infty}iI_{A_i}$, 则$P(A_i)=\dfrac{e^{-\lambda}}{i!}\lambda^i$.

设$\mathscr{H}$是Ω到E的一族映射, 令$\mathscr{F}=\sigma\left(\bigcup\limits_{f\in\mathscr{H}}f^{-1}(\mathscr{E})\right)$. 则$\mathscr{F}$是使得$\mathscr{H}$中所有映射均可测的最小$\sigma$-代数. 若$(E,\mathscr{E})=(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$, 则把$\mathscr{F}$记为$\sigma(\{f|f\in\mathscr{H}\})$.
注: 若$\mathscr{H}$只有一个点f, 记$\sigma(f)$表示$\{f^{-1}(B)|B\in\mathscr{B}(\mathbb{R})\}$. 这是$\sigma$-代数.

定理 2.2.2 设$f:\Omega\to E$是映射, $\varphi:\Omega\to\mathbb{R}$是函数, 则

φ为$\sigma(f)$可测$\Leftrightarrow$存在$\mathscr{E}/\mathscr{B}(\mathbb{R})$可测的$h:E\to\mathbb{R}$, 使得$\varphi=h\circ f$.

注: 用交换图表示为:

$$\begin{tikzcd}
(\Omega,\sigma(f)) \arrow[r, "f"] \arrow[dr, "\varphi"'] & (E,\mathscr{E}) \arrow[d, "\exists h(\mathscr{E}/\mathscr{B}(\mathbb{R})\text{可测})"] \\
& (\mathbb{R},\mathscr{B}(\mathbb{R}))
\end{tikzcd}$$

证明: "$\Rightarrow$": 可测函数的复合依然是可测的.
"$\Leftarrow$": (标准步骤: 示性函数$\Rightarrow$非负简单函数$\Rightarrow$可测函数)
(i)设$\varphi=I_A,A\in\sigma(f)$, 则存在$B\in\mathscr{B}(\mathbb{R})$, 使得$A=f^{-1}(B)$. 所以$I_A=I_{f^{-1}(B)}=I_B(f(\cdot))$, 取$h=I_B$即可.
(ii)$\varphi=\sum\limits_{i=1}^{n}a_iI_{A_i},a_i\geq 0$, 且$A_i\in\sigma(f)$两两不交, $a_i\neq a_j$, 由(i), 存在$B_i\in\mathscr{E},1\leq i\leq n$, 使得$I_{A_i}=I_{B_i}\circ f$, 从而$\varphi=\sum\limits_{i=1}^{n}a_iI_{B_i}\circ f$.
(iii)$\varphi\geq 0$, 且φ关于$\sigma(f)$可测, 则有$\sigma(f)$可测的非负简单函数序列$\{\varphi_n\}$, 使得$\varphi_n\nearrow\varphi$. 由(ii), 对固定的n, 存在$\sigma(f)/\mathscr{E}$可测函数h_n, 使得$\varphi_n=h_n\circ f$, 对两边取极限并令$h=\limsup\limits_{n\to\infty}h_n$.（这里$h_n$定义在$\mathscr{E}$上，$h_n\circ f$单调，但是$h_n$不一定单调，因为在$f$值域之外的情况未知.）所以$\varphi=h\circ f$. (若$f$有界, 则$h$可以有界, 取$h|_{R(f)^c}=0$即可.
(iv)φ一般可测, 取$\varphi=\varphi^+-\varphi^-,\cdots$. □

注: 示性函数I_B可测$\Leftrightarrow$ B是可测集.

必须牢牢记住单调类定理与简单函数逼近定理. 下面定理把这两个定理结合起来, 可以作为考试题, 但是不忍心考, 因为后面内容很丰富.

定理 2.2.3 设$\mathscr{C}$是Ω上的π类, $\mathscr{H}$是Ω上由实值函数构成的线性空间. 若以下条件成立:

(1)$1\in\mathscr{H}$.

(2)$\{f_n\}\subset\mathscr{H}, 0\le f_n\nearrow f$, 且$f$有限（相应地, 有界）, 则$f\in\mathscr{H}$.

(3)$\forall A\in\mathscr{C}, I_A\in\mathscr{H}$.

则$\mathscr{H}$包含Ω上所有$\sigma(\mathscr{C})$可测实值（相应地, 有界）函数

注: 条件(3)暗示用单调类定理, 以$I_A\in\mathscr{H}$为性质的集合全体构成λ类.

证明: (i)设$A\in\sigma(\mathscr{C})$, 下证$I_A\in\mathscr{H}$. 令

$$\mathscr{G}=\{A\in\sigma(\mathscr{C})|I_A\in\mathscr{H}.$$

由(3), $\mathscr{C}\subset\mathscr{G}$, 且$\mathscr{G}$是$\lambda$类, 所以$\sigma(\mathscr{C})=\mathscr{G}$.

(ii)非负简单函数;(iii)一般非负可测函数;(iv)一般可测函数. □

定义 2.2.1 设$\mathscr{H}$是Ω上的非负有界函数类, 满足:

(1)$1\in\mathscr{H}$;

(2)$f\in\mathscr{H},\alpha\in\mathbb{R}_+\Rightarrow\alpha f\in\mathscr{H}$;

(3)**(真差运算封闭)**$f,g\in\mathscr{H}, f\ge g\Rightarrow f-g\in\mathscr{H}$;

(4)**(单调增封闭)**$f_n\in\mathscr{H}, n\ge 1, f_n\nearrow f, f$有界$\Rightarrow f\in\mathscr{H}$.

则称$\mathscr{H}$为**Λ族**.

设$\mathscr{C}$是Ω上的非负有界函数类, 用$\Lambda(\mathscr{C})$表示包含$\mathscr{C}$的最小Λ族.

注: 所有非负有界函数构成Λ族, 所以一定存在最小Λ族.

注: 若$f,g\in\mathscr{H}$且$\mathscr{H}$是Λ族, 则$f+g\in\mathscr{H}$. 这是因为, f,g均有界, 则存在$C>0$使得$f+g\le C$, 所以$f+g=C-((C-f)-g)\in\mathscr{H}$.

定义 2.2.2 设$(E,\mathscr{E})$是可测空间, $\mathscr{C}$是Ω到E的一族映射. 令

$$\mathscr{F}=\sigma\Big\{\bigcup_{f\in\mathscr{C}}f^{-1}(\mathscr{E})\Big\},$$

则$\mathscr{F}$是使得$\mathscr{C}$中所有元素为可测的最小σ代数, 叫函数类$\mathscr{C}$在Ω上**生成的σ-代数**.

特别地, 若$(E,\mathscr{E})=(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$, 常用$\sigma(f:f\in\mathscr{C})$来表示这个$\sigma$代数$\mathscr{F}$.

记$\mathscr{L}_b^+(\mathscr{C})$是非负有界$\sigma(f:f\in\mathscr{C})$可测函数全体.

定理 2.2.4 设$\mathscr{C}$是Ω上非负有界函数, 则

$$\mathscr{L}_b^+(\mathscr{C})=\Lambda(\mathscr{C})\Leftrightarrow\forall f,g\in\mathscr{C}, fg\in\Lambda(\mathscr{C}).$$

证明: "$\Rightarrow$": 显然.

"$\Leftarrow$": 容易证$\mathscr{L}_b^+(\mathscr{C})$是$\Lambda$族, 则$\Lambda(\mathscr{C}) \subset \mathscr{L}_b^+(\mathscr{C})$. 下证"$\supset$"方向.

先证$\mathscr{L}_b^+(\mathscr{C})$中的示性函数在$\Lambda(\mathscr{C})$中.

(i)下证$\Lambda(\mathscr{C})$关于乘积运算封闭. 令

$$\mathscr{H}_1 = \{f \in \Lambda(\mathscr{C}) | \forall g \in \mathscr{C}, fg \in \mathscr{C}\},$$

则$\mathscr{C} \subset \mathscr{H}_1$且$\mathscr{H}_1$是$\Lambda$族, 则$\Lambda(\mathscr{C}) = \mathscr{H}_1$. 令

$$\mathscr{H}_2 = \{f \in \Lambda(\mathscr{C}) | \forall g \in \Lambda(\mathscr{C}), fg \in \mathscr{C}\},$$

则$\mathscr{C} \subset \mathscr{H}_2$且$\mathscr{H}_2$是$\Lambda$族, 则$\Lambda(\mathscr{C}) = \mathscr{H}_2$.

(ii)下证$\Lambda(\mathscr{C})$关于取下端运算封闭, 即$f, g \in \Lambda(\mathscr{C}) \Rightarrow f \wedge g \in \Lambda(\mathscr{C})$. 令

$$P_0(x) = 0, P_{n+1}(x) = P_n(x) + \frac{1}{2}(x^2 - P_n(x)^2), \forall n \geq 1, |x| \leq 1.$$

则$P_n(x) \nearrow |x|$, 不妨设$f, g \in \Lambda(\mathscr{C})$且$0 \leq f - g \leq 1$, 则$|f - g| \leq 1$, 所以

$$P_1(|f-g|) = \frac{1}{2}(|f-g|^2) = \frac{1}{2}(f^2 - 2fg + g^2) \in \Lambda(\mathscr{C}).$$

由归纳法可知$P_n(f-g) \in \Lambda(\mathscr{C})$, 根据$\Lambda$族的定义, $|f-g| \in \Lambda(\mathscr{C})$, 则

$$f \wedge g = \frac{1}{2}(f + g - |f-g|) \in \Lambda(\mathscr{C}).$$

(iii)对任意$A \in \sigma(f : f \in \mathscr{C})$, 下证$I_A \in \Lambda(\mathscr{C})$. 令

$$\mathscr{F} = \{A \subset \Omega | I_A \in \Lambda(\mathscr{C})\},$$

($\mathscr{F}$联系了λ类与Λ族), 则$\mathscr{F}$既是π类(注意$I_{A\cap B} = I_A \wedge I_B = I_A I_B$)又是$\lambda$类, 从而是$\sigma$-代数.

下证$\sigma(f : f \in \mathscr{C}) \subset \mathscr{F}$, 从而只需证$[f < a] \in \mathscr{F} (a > 0)$. 事实上,

$$I_{[f<a]} = \lim_{n\to\infty} 1 - \left(\frac{f}{a} \wedge 1\right)^n \in \Lambda(\mathscr{C}).$$

因此$[f < a] \in \mathscr{F}$, 进一步, $\sigma(f : f \in \mathscr{C}) = \sigma\Big(\bigcup_{f\in\mathscr{C}} f^{-1}(\mathscr{B}(\mathbb{R}))\Big) \in \mathscr{F}$. 即对任意$A \in \sigma(f : f \in \mathscr{C})$都有$I_A \in \Lambda(\mathscr{C})$.

(iv)再根据可测函数构造的步骤即可得到$\mathscr{L}_b^+(\mathscr{C}) \subset \Lambda(\mathscr{C})$. □

定义 2.2.3 设$\mathscr{H}$是Ω上一族有界函数, 若其关于单调一致有界序列极限运算封闭, 则称之为**单调族**.

设$\mathscr{C}$是一族有界函数, 用$M(\mathscr{C})$表示包含$\mathscr{C}$的**最小单调族**, 记

$$\mathscr{L}_b(\mathscr{C}) = \{g : \Omega \to \mathbb{R} | g\text{有界, 关于}\sigma(f : f \in \mathscr{C})\text{可测}\}.$$

定理 2.2.5 设$\mathscr{C}$是有界函数类, 则

$$\mathscr{L}_b(\mathscr{C})=M(\mathscr{C})\Leftrightarrow\begin{cases}1\in M(\mathscr{C}),\\ f\in\mathscr{C}\Rightarrow af\in M(\mathscr{C}),\forall a\in\mathbb{R},\\ f,g\in\mathscr{C}\Rightarrow f+g,f\wedge g\in M(\mathscr{C}).\end{cases}$$

证明: "$\Leftarrow$": 仅证$\mathscr{L}_b(\mathscr{C})\subset M(\mathscr{C})$.

(1)先证明$M(\mathscr{C})$关于线性运算与取下端运算封闭. 令

$$\mathscr{H}_1=\{f\in M(\mathscr{C})|\forall g\in\mathscr{C},\forall a\in\mathbb{R},af\in M(\mathscr{C}),f+g\in M(\mathscr{C}),f\wedge g\in M(\mathscr{C}),$$

则$\mathscr{C}\subset\mathscr{H}_1$, 且$\mathscr{H}_1$是单调族, 从而$\mathscr{H}_1=M(\mathscr{C})$. 令

$$\mathscr{H}_2=\{f\in M(\mathscr{C})|\forall g\in M(\mathscr{C}),f+g\in M(\mathscr{C}),f\wedge g\in M(\mathscr{C}),$$

则$\mathscr{C}\subset\mathscr{H}_2$, 且$\mathscr{H}_2$是单调族, 从而$\mathscr{H}_2=M(\mathscr{C})$.

(2)令$\mathscr{F}=\{A\subset\Omega|I_A\in M(\mathscr{C})\}$, 则$\mathscr{F}$是$\sigma$-代数($\lambda$类+$\pi$类), 下证$\sigma(f:f\in\mathscr{C})\subset\mathscr{F}$. 只需证对任意$f\in\mathscr{C},\forall a\in\mathbb{R}$, 都有$I_{[f>a]}\in M(\mathscr{C})$, 事实上, $I_{[f>a]}=\lim\limits_{n\to\infty}(n(f-a)^+)\wedge 1\in M(\mathscr{C})$. (注意$(f-a)^+=(f-a)\vee 0$, 而$f\vee g=-((-f)\wedge(-g))\in M(\mathscr{C})$.)

(3)简单函数逼近可测函数, 设$M(\mathscr{C})$是线性空间可知$\forall f\in\mathscr{L}_b(\mathscr{C})$, 有$f\in M(\mathscr{C})$. □

§ 2.3 可测函数的收敛

设$(\Omega,\mathscr{F},\mu)$是测度空间. 若某一性质在除掉零测集以后在Ω上逐点成立, 称该性质是"**几乎处处成立**"(a.e.成立, almost everywhere).

定义 2.3.1 设$\{f_n;f\}$为$(\Omega,\mathscr{F},\mu)$上的实可测函数.

(1)称$\{f_n\}$**几乎处处收敛**于f, 若存在$N\in\mathscr{F},\mu(N)=0$, 满足$f_n(\omega)\to f(\omega),\forall\omega\in N^c$. 记为$f_n\xrightarrow{a.e.}f$.

(2)称$\{f_n\}$**几乎一致收敛**于f, 若对任意$\delta>0$, 存在$F\in\mathscr{F},\mu(F)<\delta$, 使得$\{f_n\}$在$F^c$上一致收敛于$f$, 即$\lim\limits_{n\to\infty}\sup\limits_{\omega\in F^c}|f_n(\omega)-f(\omega)|=0$, 记为$f_n\xrightarrow{a.un.}f$.

(3)称$\{f_n\}$**依测度收敛**于f, 若对任意$\varepsilon>0$, 都有$\lim\limits_{n\to\infty}\mu(|f_n-f|\geq\varepsilon)=0$. 即对任意$\varepsilon>0$与任意$\delta>0$, 存在$N>0$, 当$n>N$时, $\mu([|f_n-f|\geq\varepsilon])<\delta$. 记为$f_n\xrightarrow{\mu}f$.

注: 上述极限都有一定的唯一性.

(1)如果$f_n\xrightarrow{a.e.}f$且$f_n\xrightarrow{a.e.}g$, 则$f=g(a.e.)$.

(2)如果$f_n\xrightarrow{\mu}f$且$f_n\xrightarrow{\mu}g$, 则

$$\mu([|f-g|\geq\varepsilon])\leq\mu\left(\left[|f_n-f|\geq\frac{\varepsilon}{2}\right]\right)+\mu\left(\left[|f_n-g|\geq\frac{\varepsilon}{2}\right]\right)\to 0(n\to\infty).$$

(根据$\varepsilon\leq|f(x)-g(x)|\leq|f_n(x)-f(x)|+|f_n(x)-g(x)|$, 可知必有$|f_n(x)-f(x)|\geq\varepsilon/2$或$|f_n(x)-g(x)|\geq\varepsilon/2$, 所以$[|f-g|\geq\varepsilon]\subset[|f_n-f|\geq\varepsilon/2]\cup[|f_n-g|\geq\varepsilon/2]$.)

由次σ可加性与$[|f-g|>0]=\bigcup\limits_{k=1}^{\infty}[|f-g|\geq 1/k]$即可得$f=g(a.e.)$.

定理 2.3.1 (等价刻画) 设$\{f_n:f\}$均为$(\Omega,\mathscr{F},\mu)$上的实值函数.

(1)$f_n\xrightarrow{a.e.}f\Leftrightarrow\forall\varepsilon>0,\mu\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}[|f_k-f|\ge\varepsilon]\right)=0.$

(2)$f_n\xrightarrow{a.un.}f\Leftrightarrow\forall\varepsilon>0,\lim_{n\to\infty}\mu\left(\bigcup_{k=n}^{\infty}[|f_k-f|\ge\varepsilon]\right)=0.$

(3)$f_n\xrightarrow{\mu}f\Leftrightarrow\{f_n\}$的任一子列$\{f_n'\}$都有几乎一致收敛序列$\{f_{n_k}\}\xrightarrow{a.un.}f.$

证明: (1)注意到$\{\omega|f_n(\omega)\to f(\omega)\}=\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left[|f_k-f|<\frac{1}{m}\right]$. 则

$$\mu(\{\omega|f_n(\omega)\nrightarrow f(\omega)\})=\mu\left(\bigcup_{m=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left[|f_k-f|\ge\frac{1}{m}\right]\right).$$

所以

$$\begin{aligned}f_n\xrightarrow{a.e.}f&\Leftrightarrow\mu(\{\omega|f_n(\omega)\nrightarrow f(\omega)\})=0\\&\Leftrightarrow\mu\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left[|f_k-f|\ge\frac{1}{m}\right]\right)=0,\forall m\ge1,\\&\Leftrightarrow\mu\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}[|f_k-f|\ge\varepsilon]\right)=0,\forall\varepsilon>0.\qquad\square\end{aligned}$$

(2)“$\Rightarrow$”: 假定$f_n\xrightarrow{a.un.}f$, 即对任意$\delta>0$, 存在$F\in\mathscr{F},\mu(F)<\delta$, 使得在$F^c$上$f_n\rightrightarrows f$. 对任意$\varepsilon>0$, 存在$n_0\in\mathbb{N}$, 当$n>n_0$时, $\sup_{\omega\in F^c}|f_n(\omega)-f(\omega)|<\varepsilon$ 所以$F^c\subset\bigcap_{n=1}^{\infty}[|f_n-f|<\varepsilon]$. 从而

$$\begin{aligned}\bigcup_{n=1}^{\infty}[|f_n-f|\ge\varepsilon]\subset F\Rightarrow\mu\left(\bigcup_{n=n_0}^{\infty}[|f_n-f|\ge\varepsilon]\right)&\le\mu(F)<\delta\\\Rightarrow\lim_{n\to\infty}\mu\left(\bigcup_{k=n}^{\infty}[|f_n-f|\ge\varepsilon]\right)&\le\delta.\end{aligned}$$

由δ的任意性, $\lim_{n\to\infty}\mu\left(\bigcup_{k=n}^{\infty}[|f_k-f|\ge\varepsilon]\right)=0.$

“$\Leftarrow$”: (把F找出来.)对任意$\delta>0$与任意$\varepsilon>0$, 存在$n_i>0$, 使得$\mu\left(\bigcup_{k=n_i}^{\infty}[|f_k-f|\ge\varepsilon]\right)\le\frac{\delta}{2^i}$.

令$F=\bigcup_{i=1}^{\infty}\bigcup_{k=n_i}^{\infty}[|f_k-f|\ge\varepsilon]$, 则$\mu(F)\le\delta$. 且在$F^c$上, $F^c=\bigcap_{i=1}^{\infty}\bigcap_{k=n_i}^{\infty}[|f_k-f|<\varepsilon]$.

(3)“$\Rightarrow$”(注意可以用(2)的结论!)假定$f_n\xrightarrow{\mu}f$, $\{f_{n'}\}\subset\{f_n\}$, 则$\{f_{n'}\}$依测度收敛于f. 即$\forall i\ge1$, 存在$n_i>0$, 使得$\mu\left([|f_{n_i}-f|\ge\frac{1}{i}]\right)\le\frac{1}{2^i}$. 从而对充分大的正整数$m$有

$$\mu\left(\bigcup_{i=m+1}^{\infty}\left[|f_{n_i}-f|\ge\frac{1}{i}\right]\right)\le\sum_{i=m+1}^{\infty}\frac{1}{2^i}=\frac{1}{2^m}.$$

对任意$\varepsilon>0,m'=[1/\varepsilon]\vee(m+1)$, 有$\mu\left(\bigcup_{i=m'}^{\infty}[|f_{n_i}-f|\ge\varepsilon]\right)\le\frac{1}{2^m}$. 从而由(2)可知$f_{n_i}\xrightarrow{a.un.}f(i\to\infty)$.

“$\Leftarrow$”: (反证)若$f_n\xrightarrow{\mu}f$不成立, 即存在$\varepsilon_0>0$与$\{f_{n'}\}\subset\{f_n\}$, 使得$\limsup_{n'\to\infty}\mu([|f_{n'}-f|\ge\varepsilon])=\delta_0>0$. 对充分大的$n'$, 有$\mu(|f_{n'}-f|\ge\varepsilon)>\frac{\delta_0}{2}$, 这与$\{f_{n'}\}$存在几乎一致收敛子列矛盾, 所以假设不成立. $\square$

注: (1)$f_n \xrightarrow{a.un.} f \Rightarrow f_n \xrightarrow{a.e.} f$.
(2)$f_n \xrightarrow{\mu} f \Rightarrow \forall\{f_{n'}\} \subset \{f_n\}, \exists\{f_{n'_k}\} \subset \{f_{n'}\}$使得$f_{n'_k} \xrightarrow{a.e.} f$.
(3)若μ是有限测度, 则$f_n \xrightarrow{a.e.} f \Leftrightarrow f_n \xrightarrow{a.un.} f$. (Egorov定理)

§ 2.4 第二章习题

例 2.4.1 设$f_n \xrightarrow{\mu} f, g_n \xrightarrow{\mu} g$, 则
(1)$f_n + g_n \xrightarrow{\mu} f + g$;
(2)$f_n g_n \xrightarrow{\mu} fg$.

证明: (1)可以用等价刻画, 也可以用

$$\mu([|f_n + g_n - f - g| \geq \varepsilon]) \leq \mu\left(\left[|f_n - f| \geq \frac{\varepsilon}{2}\right]\right) + \mu\left(\left[|g_n - g| \geq \frac{\varepsilon}{2}\right]\right) \to 0(n \to \infty).$$

(2)用等价刻画最快, $\{f_n\}$存在子列$\{f_{n_k}\}$使得$f_{n_k} \xrightarrow{a.un.} f$, 另一方面, $\{g_n\}$的子列$\{g_{n_k}\}$也存在子列$\{g_{n_{k_i}}\}$使得$g_{n_{k_i}} \xrightarrow{a.un.} g$, 所以结合$f_{n_{k_i}} \xrightarrow{a.un.} f$可得, $f_{n_{k_i}} g_{n_{k_i}} \xrightarrow{a.un.} fg$. (比较琐碎的步骤就在此省略). □

例 2.4.2 设$\xi_n \xrightarrow{P} \xi$, f是连续函数, 则$f(\xi_n) \xrightarrow{P} f(\xi)$. (提示: 用等价刻画).

例 2.4.3 设$f_n \xrightarrow{\mu} f$, 则$\liminf\limits_{n\to\infty} f_n \leq f \leq \limsup\limits_{n\to\infty} f_n$.

证明: 根据几乎一致收敛的定义, 以及等价刻画, 可知存在子列$\{f_{n_k}\}$使得

$$\lim_{k\to\infty} \sup_{\omega\in F^c} (f_{n_k}(\omega) - f(\omega)) = 0.$$

所以

$$0 = \lim_{k\to\infty} \sup_{\omega\in F^c} (f_{n_k}(\omega) - f(\omega)) \leq \lim_{n\to\infty} \sup_{\omega\in F^c} (f_n(\omega) - f(\omega)).$$

例 2.4.4 设$\{\xi_n\}$是$(\Omega, \mathscr{F}, P)$中的随机变量序列, 证明:

$$\xi_n \xrightarrow{P} \xi \Leftrightarrow \mathbb{E}\left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|}\right) \to 0.$$

证明: "$\Rightarrow$": 由条件, $\forall\varepsilon > 0$, $\lim\limits_{n\to\infty} P(|\xi_n - \xi| \geq \varepsilon) = 0$.
设$f(x) = \dfrac{x}{1+x}(x > 0)$, 则$f(x)$单调递增且有界.

$$\begin{aligned}\mathbb{E}\left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|}\right) &= \int_{[|\xi_n - \xi| \leq \varepsilon]} \frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} \mathrm{d}P + \int_{[|\xi_n - \xi| > \varepsilon]} \frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} \mathrm{d}P \\ &\leq \frac{\varepsilon}{1+\varepsilon} \underbrace{P(|\xi_n - \xi| \leq \varepsilon)}_{\leq 1} + P(|\xi_n - \xi| > \varepsilon).\end{aligned}$$

所以$\lim\limits_{n\to\infty} \mathbb{E}\left(\dfrac{|\xi_n - \xi|}{1 + |\xi_n - \xi|}\right) \leq \dfrac{\varepsilon}{1+\varepsilon}$. 由$\varepsilon$的任意性, 可得$\lim\limits_{n\to\infty} \mathbb{E}\left(\dfrac{|\xi_n - \xi|}{1 + |\xi_n - \xi|}\right) = 0$.

"$\Leftarrow$": 注意$\mathbb{E}\left(\dfrac{|\xi_n - \xi|}{1 + |\xi_n - \xi|}\right) \geq \displaystyle\int_{[|\xi_n - \xi| \geq \varepsilon]} \dfrac{|\xi_n - \xi|}{1 + |\xi_n - \xi|} \mathrm{d}P \geq \dfrac{\varepsilon}{1+\varepsilon} P(|\xi_n - \xi| \geq \varepsilon)$. □

第 3 章　积分的定义与性质

设$(\Omega,\mathscr{F},\mu)$是测度空间, f是可测函数. 记

$$\int_\Omega f(\omega)\mu(\mathrm{d}\omega)(\text{或}\int_\Omega f\mathrm{d}\mu,\ \text{或}\mu(f))$$

叫$\boldsymbol{f}$**关于**$\boldsymbol{\mu}$**的积分**.

再引入一些符号:

$\mathcal{S}^+$	$(\Omega,\mathscr{F})$上的非负简单函数全体
$\overline{\mathcal{L}}$	$(\Omega,\mathscr{F})$上的可测数值函数全体
$\mathcal{L}$	$(\Omega,\mathscr{F})$上的可测实值函数全体
$\overline{\mathcal{L}}^+$	$(\Omega,\mathscr{F})$上的非负可测数值函数全体
$\mathcal{L}^+$	$(\Omega,\mathscr{F})$上的非负可测实值函数全体

§ 3.1　积分的基本性质

定义 3.1.1 设$(\Omega,\mathscr{F},\mu)$是测度空间.

(1)设$A\in\mathscr{F}$, 则$\mu(I_A)\triangleq\mu(A)$.

(2)设$f=\sum\limits_{i=1}^n a_iI_{a_i}\in\mathcal{S}^+$, 则$\mu(f)\triangleq\int_\Omega f\mathrm{d}\mu\triangleq\sum\limits_{i=1}^n\alpha_i\mu(A_i)$.

命题 3.1.1 设$f,g,f_n,g_n\in\mathcal{S}^+$, 则

(1)$\mu(f+g)=\mu(f)+\mu(g)$.

(2)$\forall a\geq 0,\mu(af)=a\mu(f)$;

(3)$f\leq g\Rightarrow\mu(f)\leq\mu(g)$.

(4)若$f_n\searrow f$且$\mu(f_1)<+\infty$, 则$\mu(f_n)\searrow\mu(f)$.

(5)若$f_n\nearrow f$, 则$\mu(f_n)\nearrow\mu(f)$.

(6)若$\lim\limits_{n\to\infty}f_n\leq\lim\limits_{n\to\infty}g_n$且$f_n\nearrow,g_n\nearrow$, 则$\lim\limits_{n\to\infty}\mu(f_n)\leq\lim\limits_{n\to\infty}\mu(g_n)$.

证明: (1)(2)(3)显然.

(4)如果把f_n取为示性函数, 那么就相当于测度的从上连续性.

令$h_n=f_n-f$, 则$h_n\in S^+$且$h_n\leq f_1(\forall n\geq 1)$. 下证$\lim\limits_{n\to\infty}\mu(h_n)=0$. 令$\beta=\max\{h_1(\omega):\omega\in\Omega\}<+\infty$, 则对任意$\varepsilon>0$, 有

$$h_n=h_nI_{[h_n>\varepsilon]}+h_nI_{[h_n\leq\varepsilon]}\leq\beta I_{[h_n>\varepsilon]}+\varepsilon I_{[h_n\leq\varepsilon]}.$$

从而$\mu(h_n)\leq\beta\mu([h_n>\varepsilon])+\varepsilon\mu([h_n\leq\varepsilon])$. 当$n\to\infty$时, 由于$[h_n>\varepsilon]\searrow\varnothing$, 则$h_n\searrow 0$. 由

$$\mu([h_1>\varepsilon])\leq\frac{1}{\varepsilon}\mu(h_1)(\text{Chebyshev不等式!})$$

根据测度的从上连续性可知$\lim\limits_{n\to\infty}\mu([h_n>\varepsilon])=0$. 又由于$\mu([h_n\leq\varepsilon])\leq\mu([h_1>0])<+\infty$, 令$n\to\infty,\varepsilon\to 0$, 可得$\lim\limits_{n\to\infty}\mu(h_n)=0$, 即得$\lim\limits_{n\to\infty}\mu(f_n)=\mu(f)$.

(5)若$\mu(f)<+\infty$, 那么f的每个函数值都有限(简单函数的性质!), 所以$f-f_n\searrow 0$, 由于$\lim\limits_{n\to\infty}\mu(f_n)+\lim\limits_{n\to\infty}\mu(f-f_n)=\mu(f)$, 则$\lim\limits_{n\to\infty}\mu(f)=\mu(f)$.

若$\mu(f)=+\infty$, 则存在$a>0$使得$\mu([f=a])=+\infty$(f只取有限多个函数值!), 则

$$\mu(f_n)\geq\mu(f_nI_{[f_n>\frac{a}{2}]})\geq\frac{a}{2}\mu([f_n\geq\frac{a}{2}])\geq\frac{a}{2}\mu([f_n=a])=+\infty.$$

所以$\lim\limits_{n\to\infty}\mu(f)=\mu(f)=+\infty$.

(6)只需证对任意$n\geq 1$都有$\mu(f_n)\leq\lim\limits_{n\to\infty}\mu(g_n)$. 令$h_{n,m}=f_m\wedge g_n\in\mathcal{S}^+$, 则$h_{n,m}\nearrow f_m(n\to\infty)$, 利用(3)(5)可知

$$\mu(f_m)=\lim_{n\to\infty}\mu(h_{n,m})\leq\lim_{n\to\infty}\mu(g_n).$$

□

下面用非负简单函数逼近来定义非负可测函数的积分.

定义 3.1.2 设$f\in\overline{\mathcal{L}}^+$, 则$\mu(f)\triangleq\lim\limits_{n\to\infty}\mu(f_n)$, 其中$\{f_n\}\in\mathcal{S}^+$且$f_n\nearrow f$.

注: (1)f可以取$+\infty$; (2)积分的定义不依赖于$\{f_n\}$的选取.

命题 3.1.2 (1)设$f\in\overline{\mathcal{L}}^+$, $\alpha\geq 0$, 则$\mu(\alpha f)=\alpha\mu(f)$.

(2)$f,g\in\overline{\mathcal{L}}^+$, 则$\mu(f+g)=\mu(f)+\mu(g)$.

(3)$f\leq g\Rightarrow\mu(f)\leq\mu(g)$.

定义 3.1.3 设$f\in\overline{\mathcal{L}}$, 若$\mu(f^+)<+\infty$或$\mu(f^-)<+\infty$, 则称$f$关于$\mu$**积分存在**, 且$\mu(f)\triangleq\mu(f^+)-\mu(f^-)$.

若$\mu(f^+)<+\infty$且$\mu(f^-)<+\infty$, 则称f关于μ**可积**, 这等价于$\mu(|f|)<+\infty$.

特别地, 若$(\Omega,\mathscr{F},P)$是概率空间, X是随机变量, 则$EX\triangleq\int_\Omega X\mathrm{d}P$叫$f$的**数学期望**.

注: 这里给出了数学期望的最本质定义: 随机变量关于概率测度的积分.

注: 设$f\in\overline{\mathcal{L}}$, 若$f$积分存在, 则对任意$A\in\mathcal{F}$, fI_A积分存在. 用$\int_A f\mathrm{d}\mu$来表示$\int_\Omega fI_A\mathrm{d}\mu$.

定理 3.1.3 设f,g积分存在.

(1)对任意$a\in\mathbb{R}$, af积分存在, 且$\mu(af)=a\mu(f)$.

(2)若$f+g$处处有定义, 且$\mu(f)+\mu(g)$有意义(即不出现无穷减无穷), 则$f+g$积分存在, 且有$\mu(f+g)=\mu(f)+\mu(g)$.

(3)$|\mu(f)|\leq\mu(|f|)$.

(4)若N是零测集, 则$\mu(fI_N)=0$.

(5)若$f\leq g$(a.e.), 则$\mu(f)\leq\mu(g)$.

(6)若$f\in\overline{\mathcal{L}}^+$, 则$f=0$(a.e.)等价于$\mu(f)=0$.

(7)若$f\in\overline{\mathcal{L}}^+$且$\mu(f)<\infty$, 则$f<\infty$(a.e.), 且$[f>0]$关于$\mu$为$\sigma$有限的.

证明: (1)-(4)显然.

(5)令$N=[f>g]$, 则由条件, $\mu(N)=0$, 我们有

$$f=fI_{N^c}+fI_N,g=gI_{N^c}+gI_N,fI_{N^c}\leq gI_{N^c},$$

由(4)可知$\mu(f)=\mu(fI_{N^c}),\mu(g)=\mu(gI_{N^c})$. 但由积分的定义, $\mu(fI_{N^c})\le\mu(gI_{N^c})$, 从而有$\mu(f)\le\mu(g)$.

(6)"$\Rightarrow$": 由(5)立得. "$\Leftarrow$": 由

$$\mu([f>0])=\mu\left(\bigcup_{n=1}^{\infty}[f>\frac{1}{n}]\right)\le\sum_{n=1}^{\infty}\mu([f\ge\frac{1}{n}])\le\sum_{n=1}^{\infty}n\mu(f)=0.$$

(最后一个不等号是因为Chebyshev不等式!)

(7)设$f\in\overline{\mathcal{L}}^+$. (反证)若$\mu([f=+\infty])>0$, 则$f\ge\infty I_{f=\infty}$, 从而$\mu(f)=\infty$, 矛盾. 另外,

$$[f>0]=\bigcup_{n=1}^{\infty}[f\ge\frac{1}{n}],\mu([f\ge\frac{1}{n}])\le n\mu(f)<\infty,$$

所以$[f>0]$关于μ是σ有限的. □

推论 3.1.4 *设f,g积分存在, $f\le g(a.e.)$, 则对一切$A\in\mathscr{F}$, 有$\mu(fI_A)\le\mu(gI_A)$.*

下面证明: 在一定条件下, 上述推论的逆命题也成立.

命题 3.1.5 *设f,g积分存在, 且对一切$A\in\mathscr{F}$, 有$\mu(fI_A)\le\mu(gI_A)$.*
(1)若f,g可积, 则$f\le g(a.e.)$(即$\mu([f>g])=0$).
(2)若μ为σ有限测度, 则$f\le g(a.e.)$.

证明: (1)令$A=[f>g]$, 则$(f-g)I_A\ge0$. 由假定,

$$\mu((f-g)I_A)=\mu(fI_A)-\mu(gI_A)\le0.$$

所以$\mu((f-g)I_A)=0$, 则$(f-g)I_A=0$(a.e.). 而在A上有$f>g$, 则必定有$\mu(A)=0$, 所以$f\le g$(a.e.).

(2)设μ是σ有限测度, 下证$f\le g$(a.e.). (反证)假设$\mu([g<f])>0$, 令

$$A_n=[g<f-\frac{1}{n}]\cap[|f|<n],B_m=[g<m]\cap[f=+\infty].$$

则$[g<f]=\left(\bigcup\limits_{n=1}^{\infty}A_n\right)\cup\left(\bigcup\limits_{m=1}^{\infty}B_m\right)$. 于是存在某$n$或$m$, 使得$\mu(A_n)>0$或$\mu(B_m)>0$.

若$\mu(A_n)>0$, 由μ的σ有限性可知存在$A\subset A_n,A\in\mathscr{F}$, 使得$0<\mu(A)<\infty$. 此时有

$$\int_A g\mathrm{d}\mu\le\int_A\left(f-\frac{1}{n}\right)\mathrm{d}\mu=\int_A f\mathrm{d}\mu-\frac{1}{n}\mu(A)<\int_A f\mathrm{d}\mu.$$

这与假定$\int_A f\mathrm{d}\mu\le\int_A g\mathrm{d}\mu$矛盾. 若$\mu(B_m)>0$, 类似可以论证可导致矛盾, 则必定有$f\le g$(a.e.). □

命题 3.1.6 *设$(E,\mathscr{E}),(F,\mathscr{F})$是可测空间, $f:E\to F$是可测映射. 假定μ是$(E,\mathscr{E})$上的测度. 定义*

$$\mu\circ f^{-1}(A)=\mu(f^{-1}(A)),\forall A\in\mathscr{F}.$$

则$\mu\circ f^{-1}$是$(F,\mathscr{F})$上的测度. 假定g是F上的可测函数, 则g关于$\mu\circ f^{-1}$积分存在(或可积)$\Leftrightarrow$ $g\circ f$关于μ积分存在(或可积), 且

$$\int_E g\circ f\mathrm{d}\mu=\int_F g\mathrm{d}\mu\circ f^{-1}.$$

把$\mu\circ f^{-1}$叫**像测度**.

注: “像测度”不是说看起来像一个测度, 其实它就是个测度. 特别地, 取两个可测空间分别为$(\Omega,\mathscr{F},P)$与$(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$, 那么有

$$\mathbb{E}g(\xi)=\int_{\mathbb{R}^d} g(x)P\circ\xi^{-1}(\mathrm{d}x).$$

离散型随机变量:

$$P\circ\xi^{-1}(\mathrm{d}x)\triangleq\sum_i p_i\delta_{a_i}(\mathrm{d}x).$$

连续型随机变量:

$$P\circ\xi^{-1}(\mathrm{d}x)\triangleq p(x)\mathrm{d}x.$$

命题的证明方法依旧是先设g是示性函数, 再用简单函数逼近.

§ 3.2 积分号下取极限

引理 3.2.1 设$f_n\in\overline{\mathcal{L}}^+, n\geq 1, f\in\overline{\mathcal{L}}^+$.
(1)若$f_n\leq f_{n+1}$*(a.e.)*, $\forall n\geq 1$, 且$f_n\xrightarrow{a.e.}f$, 则$\lim\limits_{n\to\infty}\mu(f_n)=\mu(f)$;
(2)若$f_n\geq f_{n+1}$*(a.e.)*, $\forall n\geq 1$, $f_n\xrightarrow{a.e.}f$, 且$\mu(f_1)<\infty$, 则$\lim\limits_{n\to\infty}\mu(f_n)=\mu(f)$.

证明: (1)零测集对积分没有影响, 但是逼近没有a.e.的概念, 所以我们不妨设$\{f_n\}$处处单调递增, 且$f_n\nearrow f$处处成立. 对每个n, 令$f_{n,m}\in\mathcal{S}^+$, 使得$f_{n,m}\nearrow f_n(m\to\infty)$. 令$g_m=\bigvee\limits_{i=1}^m f_{i,m}$, 则$g_m\in\mathcal{S}^+, g_m\nearrow f$, 且$g_m\leq f_m$. 由积分的定义,

$$\mu(f)=\lim_{m\to\infty}\mu(g_m)\leq\lim_{m\to\infty}\mu(f_m).$$

但是恒有$\mu(f)\geq\mu(f_m)$, 因此$\lim\limits_{m\to\infty}\mu(f_m)=\mu(f)$.

$$\begin{array}{cccccc}
f_{n1} & f_{n2} & f_{n3} & \cdots & \to & f_n\\
\vdots & \vdots & \vdots & & \vdots & \vdots\\
f_{31} & f_{32} & \boldsymbol{f_{33}} & \cdots & \to & f_3\\
f_{21} & \boldsymbol{f_{22}} & f_{23} & \cdots & \to & f_2\\
\boldsymbol{f_{11}} & f_{12} & f_{13} & \cdots & \to & f_1
\end{array}$$

由对角线法则, $f_{mm}\to f$, 但要保持单调性, 故取$g_m=\bigvee\limits_{i=1}^m f_{i,m}$.

(2)不妨设$\{f_n\}$处处单调递减, 且$f_n\searrow f$处处成立. 由于$\mu(f_1)<\infty$, 则$\mu([f_1=\infty])=0$. (也就是说无穷减无穷的情况只会在零测集发生)

可以不妨设$f_1<\infty$处处成立, 并设$g_n=f_1-f_n$, 则$g_n\nearrow f_1-f$. 由(1), $\mu(g_n)\nearrow\mu(f_1)-\mu(f)$, 即$\mu(f_n)\searrow\mu(f)$. □

推论 3.2.2 设$f_n\in\overline{\mathcal{L}}^+$, $n\geq 1$. 则$\mu\left(\sum\limits_{n=1}^\infty f_n\right)=\sum\limits_{n=1}^\infty\mu(f_n)$.

证明: 令$g_n=\sum\limits_{i=1}^n f_i, g=\sum\limits_{i=1}^\infty f_i$. 则$g_n\nearrow g$, 故$\mu(g)=\lim\limits_{n\to\infty}\mu(g_n)=\lim\limits_{n\to\infty}\sum\limits_{i=1}^n\mu(f_i)=\sum\limits_{i=1}^\infty\mu(f_i)$. □

定理 3.2.3 (单调收敛定理) 设$f_n \in \overline{\mathcal{L}}, n \geq 1$. 又设$f_n$积分存在.

(1)设$\{f_n\}$$a.e.$单调增, 且$f_n \xrightarrow{a.e} f$, 若$\mu(f_1) > -\infty$, 则$f$积分存在, 且$\mu(f_n) \nearrow \mu(f)$.

(2)设$\{f_n\}$$a.e.$单调降, 且$f_n \xrightarrow{a.e} f$. 若$\mu(f_1) < \infty$, 则$f$积分存在, 且$\mu(f_n) \searrow \mu(f)$.

证明: (1)由假定, f_n^+a.e.单调增, f_n^-a.e.单调减, 且有$f_n^+ \xrightarrow{a.e} f^+$, $f_n^- \xrightarrow{a.e} f^-$. 由于$f_1^- \geq f^-$, 且$\mu(f_1) > -\infty, f_n^- > -\infty$, 故$\mu(f^-) \leq \mu(f_1^-) < \infty$, 从而$f$积分存在. 由引理3.2.1, $\mu(f_n^+) \nearrow \mu(f^+), \mu(f_n^-) \searrow \mu(f^-)$. 因此有$\mu(f_n) \nearrow \mu(f)$.

(2)对$-f_n$用(1)即可. □

定理 3.2.4 (Fatou引理) 设$f_n \in \overline{\mathcal{L}}, n \geq 1$, 且每个$f_n$的积分存在.

(1)若存在$g \in \overline{\mathcal{L}}, \mu(g) > -\infty$, 使得$f_n \geq g(a.e.)$, 则$\liminf\limits_{n\to\infty} f_n$积分存在, 且$\mu(\liminf\limits_{n\to\infty} f_n) \leq \liminf\limits_{n\to\infty} \mu(f_n)$.

(2)若存在$g \in \overline{\mathcal{L}}, \mu(g) < \infty$, 使得$f_n \leq g(a.e.)$, 则$\limsup\limits_{n\to\infty} f_n$积分存在, 且$\mu(\limsup\limits_{n\to\infty} f_n) \geq \limsup\limits_{n\to\infty} \mu(f_n)$.

证明: (1)注意$\liminf\limits_{n\to\infty} f_n = \lim\limits_{k\to\infty} \inf\limits_{k\geq n} f_k$. 我们令$g_n = \inf\limits_{k\geq n} f_k$. 则$g_n \nearrow \liminf\limits_{n\to\infty} f_n$, 且$g_1 \geq g$(a.e.). 于是$\mu(g_1) \geq \mu(g) > -\infty$. 由单调收敛定理, $\liminf\limits_{n\to\infty} f_n$积分存在, 且

$$\mu\left(\lim \inf_{n\to\infty} f_n\right) = \lim_{n\to\infty} \mu(g_n) \leq \liminf_{n\to\infty} \mu(f_n).$$

(2)对$-f_n$用(1)即可. □

定理 3.2.5 (控制收敛定理) 设$f_n \in \mathcal{L}, f \in \mathcal{L}, g \in \mathcal{L}^+$. 如果满足:

(1)$f_n \xrightarrow{a.e.} f$或$f_n \xrightarrow{\mu} f$;

(2)g可积;

(3)对任意$n \geq 1$, 都有$|f_n| \leq g(a.e.)$

则f可积, 且$\lim\limits_{n\to\infty} \mu(f_n) = \mu(f)$.

证明: (1)若$f_n \xrightarrow{a.e.} f$, 则由$|f_n| \leq g$(a.e.)可以推出$|f| \leq g$(a.e.). (可数个零测集之并也是零测集.)

若$f_n \xrightarrow{\mu} f$, 则存在子列$\{f_{n_k}\}$满足$f_{n_k} \xrightarrow{a.e.} f$, 所以$|f| \leq |f_{n_k} - f| + |f_{n_k}| \leq |f_{n_k} - f| + g$, 再让$k \to \infty$即可得$|f| \leq g$.

由于g可积, 则f可积.

(2)若$f_n \xrightarrow{a.e.} f$, 则结论可以由Fatou引理立得.

若$f_n \xrightarrow{\mu} f$, 则对$\{f_n\}$的任意子列$\{f_{n'}\}$, 存在子列$\{f_{n'_k}\}$ 使得$f_{n'_k} \xrightarrow{a.e.} f$. 于是

$$\lim_{n\to\infty} \mu(f_{n'_k}) = \mu(f) \qquad (*)$$

(反证)如果$\mu(f_n) \nrightarrow \mu(f)$, 则对某个$\varepsilon_0$, 存在子列$\{f_{n'}\}$使得$|\mu(f_{n'}) - \mu(f)| \geq \varepsilon_0$. 与(*)矛盾. □

注: 可以证明$\lim\limits_{n\to\infty} \mu(|f_n - f|) \to 0$. 对$g_n = |f_n - f|$用控制收敛定理即可.

下面推广Fatou引理与控制收敛定理.

定理 3.2.6 设$f_n \in \mathcal{L}, f \in \mathcal{L}$, 且$f_n \xrightarrow{a.e.} f$或$f_n \xrightarrow{\mu} f$. 又设每个$f_n$的积分存在.

(1)若存在$g \in \overline{\mathcal{L}}$, $\mu(g) > -\infty$, 使得$\forall n \geq 1, f_n \geq g(a.e.)$, 则$f$积分存在, 且$\mu(f) \leq \liminf\limits_{n\to\infty} \mu(f_n)$.

(2)若存在$g \in \overline{\mathcal{L}}$, $\mu(g) < \infty$, 使得$\forall n \geq 1, f_n \leq g(a.e.)$, 则$f$积分存在, 且$\mu(f) \geq \limsup\limits_{n\to\infty} \mu(f_n)$.

证明: (1)若$f_n \xrightarrow{a.e.} f$, 由Fatou引理立证.

下设$f_n \xrightarrow{\mu} f$, 则对$\{f_n\}$的任一子列$\{f_{n'}\}$, 存在子列$\{f_{n'_k}\}$使得$f_{n'_k} \xrightarrow{a.e.} f$. 于是由上所证有$\mu(f) \le \lim\limits_{k\to\infty} \mu(f_{n'_k})$. 但子列$\{f_{n'}\}$选取是任意的(或者用前一定理的反证法来说明), 则必有$\mu(f) \le \liminf\limits_{n\to\infty} \mu(f_n)$.

(2)对$-f_n$利用(1)即可. □

定理 3.2.7 (推广控制收敛定理) 设$f_n \in \mathcal{L}, g_n \in \mathcal{L}^+, g \in \mathcal{L}^+$, 若满足

(1)$f_n \xrightarrow{a.e.} f$或$f_n \xrightarrow{\mu} f$.

(2)$g_n \xrightarrow{a.e.} g$或$g_n \xrightarrow{\mu} g$.

(3)g与g_n都可积, $\mu(g_n) \to \mu(g)$.

(4)$|f_n| \le g_n(a.e.)$, $\forall n \ge 1$.

则$\lim\limits_{n\to\infty} \mu(|f_n - f|) = 0$. 特别有$\lim\limits_{n\to\infty} \mu(f_n) = \mu(f)$.

证明: 首先假定同时有$f_n \xrightarrow{a.e.} f$与$g_n \xrightarrow{a.e.} g$. 令

$$h_n = g_n + g - |f_n - f|.$$

则$h_n \ge 0$(a.e.), 且$h_n \xrightarrow{a.e.} 2g$. 由前一定理,

$$2\mu(g) \le \liminf_{n\to\infty} \mu(h_n) = 2\mu(g) - \limsup_{n\to\infty} \mu(|f_n - f|).$$

从而有$\lim\limits_{n\to\infty} \mu(|f_n - f|) = 0$. 特别地,

$$|\mu(f_n) - \mu(f)| \le \mu(|f_n - f|) \to 0.$$

若同时有$f_n \xrightarrow{\mu} f, g_m \xrightarrow{\mu} g$, 则$h_n \xrightarrow{\mu} 2g$. 由前一定理也可以得到本题结论.

若$f_n \xrightarrow{a.e.} f, g_m \xrightarrow{\mu} g$, 或$f_n \xrightarrow{\mu} f, g_m \xrightarrow{a.e.} g$, 与前一定理的证明类似可证$\lim\limits_{n\to\infty} \mu(|f_n - f|) = 0$. □

推论 3.2.8 设f_n, f是可积可测函数, $f_n \xrightarrow{a.e.} f$. 则

$$\text{“}\mu(|f_n - f|) \to 0\text{” 当且仅当 “}\mu(|f_n|) \to \mu(|f|)\text{”}.$$

证明: 必要性显然, 充分性对$g_n = |f_n|, g = |f|$用定理3.2.7. □

定理 3.2.9 设f_n, f是可积可测函数. 则

$$\text{“}\mu(|f_n - f|) \to 0\text{” 当且仅当 “}\mu(|f_n|) \to \mu(|f|)\text{, 且}f_n \xrightarrow{\mu} f\text{”}.$$

证明: “$\Leftarrow$”: 对$g_n = |f_n|, g = |f|$用定理3.2.7.

“$\Rightarrow$”: 对任意$\varepsilon > 0$, 有

$$\lim_{n\to\infty} \mu([|f_n - f| \ge \varepsilon]) \le \frac{1}{\varepsilon} \lim_{n\to\infty} \int_{[|f_n - f| \ge \varepsilon]} |f_n - f| \mathrm{d}\mu = \frac{1}{\varepsilon} \lim_{n\to\infty} \mu(|f_n - f|) = 0.$$

(Chebyshev不等式!) 所以$f_n \xrightarrow{\mu} f$. 用三角不等式即可得$\mu(|f_n|) \to \mu(|f|)$. □

§ 3.3 符号测度与不定积分

引理 3.3.1 设$\{f_n\}_{n\geq 1}$是可测函数列, f_n非负或$\sum_{n=1}^{\infty}\int_{\Omega}|f_n|\mathrm{d}\mu<+\infty$, 则$\sum_{n=1}^{\infty}f_n$积分存在, 且

$$\int_{\Omega}\sum_{n=1}^{\infty}f_n\mathrm{d}\mu=\sum_{n=1}^{\infty}\int_{\Omega}f_n\mathrm{d}\mu.$$

定义 3.3.1 设$f\in\overline{\mathcal{L}}$, 且$f$积分存在. 令

$$\nu(A)=\int_A f\mathrm{d}\mu,\forall A\in\mathscr{F},$$

则ν具有σ-可加性, 此时称ν为关于μ的**不定积分**, 记为$\nu=f.\mu$.

定义 3.3.2 设ν是定义在$(\Omega,\mathscr{F})$上的集函数, 若ν具有σ-可加性, 则称ν为$(\Omega,\mathscr{F})$上的**符号测度**.

注: (1)符号测度与测度的概念不一样, 应加以区分. (2)不定积分是符号测度.

命题 3.3.2 下面两者之一成立: $-\infty\leq\nu(A)<+\infty(\forall A\in\mathscr{F})$ 以及$-\infty<\nu(A)\leq+\infty(\forall A\in\mathscr{F})$.

证明: 若不然, 存在$A,B\in\mathscr{F}$使得$\nu(A)=+\infty,\nu(B)=-\infty$, 此时,

$$\begin{aligned}\nu(A\cup B)&=\nu(A)+\nu(B\setminus A)\\&=\nu(B)+\nu(A\setminus B).\end{aligned}$$

由σ-可加性, $\nu(B\setminus A)=\infty$或有限, $\nu(A\setminus B)=-\infty$或有限, 于是$\nu(A\cup B)=\infty$且$\nu(A\cup B)=-\infty$, 矛盾. (我们不讨论平凡的情况, 即$\nu(\cdot)$恒为$+\infty$或$-\infty$的情况). □

命题 3.3.3 设$\nu=f.\mu$, 令$\nu^+=f^+.\mu,\nu^-=f^-.\mu$, 则$\nu^+,\nu^-$为测度, 且其中之一为有限测度: $\nu=\nu^+-\nu^-$.

不定积分可以分成两个测度之差, 下面看一般的符号测度:

定理 3.3.4 (Jordan-Hahn分解定理) 设ν是$\mathscr{F}$上的符号测度, 对任意$A\in\mathscr{F}$, 令

$$\begin{aligned}\nu^+(A)&=\sup\{\nu(B)|B\subset A,B\in\mathscr{F}\},\\\nu^-(A)&=\sup\{-\nu(B)|B\subset A,B\in\mathscr{F}\}.\end{aligned}\tag{3.1}$$

则ν^+,ν^-分别是$(\Omega,\mathscr{F})$上的测度, 且其中之一是有限测度, 且$\nu=\nu^+-\nu^-$. 进一步, 存在可测集$D\in\mathscr{F}$, 使得

$$\nu^+(A)=\nu(A\cap D),\nu^-(A)=-\nu(A\cap D^c).$$

证明: 不妨设$\nu(A)>-\infty,\forall A\in\mathscr{F}$(符号测度取相反也是符号测度). 令$\nu^+,\nu^-$如(3.1)所示.
下证明存在$D\in\mathscr{F}$使得对$A\in\mathscr{F}$, ①如果$A\subset D$则$\nu(A)\geq 0$; ②如果$A\subset D^c$则$\nu(A)\leq 0$.
(1)令

$$\mathscr{B}=\{B\in\mathscr{F}|\nu^+(B)=0\}.$$

根据ν^+的定义, 可以改写$\mathscr{B}=\{B\in\mathscr{F}|\forall C\in\mathscr{F},C\subset B,\nu(C)\le 0\}$. $\mathscr{B}$对可列并运算封闭, 且$\mathscr{B}$中任一元素的可测子集仍在$\mathscr{B}$中. ($0=\nu^+(A)=\sup\{\nu(B)|B\subset A,B\in\mathscr{F}\}$, 对任意$A'\subset A$, $\nu(A')\le 0$, 但是$\varnothing\subset A'$满足$\nu(\varnothing)=0$, 故$\nu^+(A')=0$.)

此外, 设$B\in\mathscr{B},G\in\mathscr{F},G\subset B$, 则$G\in\mathscr{B}$. 令$B_n\in\mathscr{B},n\ge 1$, 使得

$$\lim_{n\to\infty}\nu(B_n)=\inf\{\nu(B)|B\in\mathscr{B}\}\triangleq\beta.$$

则由ν的σ-可加性可知$\bigcup\limits_{n=1}^{\infty}B_n\in\mathscr{B}$, 且对$m\ge 1$有

$$\beta\le\nu\left(\bigcup_n B_n\right)=\nu(B_m)+\nu\left(\bigcup_n B_n\setminus B_m\right)\le\nu(B_m).$$

取$m\to\infty$, 即可得到$\nu\Big(\bigcup\limits_n B_n\Big)=\beta$. 令$D=\Big(\bigcup\limits_n B_n\Big)^c$, 则$D^c\in\mathscr{B},\nu(D^c)=\beta$, 由$\beta$的定义可知②成立.

(2)下证①成立. (反证)假定存在$A\in\mathscr{F},A\subset D$, 使得$\nu(A)<0$, 我们断言: 必有$\nu^+(A)>0$. 事实上, 若$\nu^+(A)=0$, 则$A\in\mathscr{B}$, 故$A\cup D^c\in\mathscr{B}$, 但是

$$\nu(A\cup D^c)=\nu(A)+\nu(D^c)<\nu(D^c)=\beta,$$

与β定义矛盾. 所以必有$\nu^+(A)>0(\bigstar)$. 由ν^+的定义, 存在$A_1\in\mathscr{F},A_1\subset A$, 使得

$$\nu(A_1)\ge\frac{1}{2}(\nu^+(A)\wedge 1)>0$$

这时, $A\setminus A_1\subset D,\nu(A\setminus A_1)=\nu(A)-\nu(A_1)<0$, 因此由上所证可知$\nu^+(A\setminus A_1)>0$. 由归纳法, 存在$A_n\in\mathscr{F},A_n\subset D,n\ge 1$, 使得$A_n\subset A\setminus\sum\limits_{k=1}^{n-1}A_k$, 且

$$\nu(A_n)\ge\frac{1}{2}\left[\nu^+\left(A\setminus\sum_{k=1}^{n-1}A_k\right)\wedge 1\right]>0.$$

由于$\nu(A)<0$, 且

$$\nu(A)=\nu\left(A\setminus\sum_{k=1}^{\infty}A_k\right)+\sum_{k=1}^{\infty}\nu(A_k). \tag{3.2}$$

则$\sum\limits_{k=1}^{\infty}\nu(A_k)<\infty$(若不然, 就会出现$\infty-\infty$的情况), 特别有$\lim\limits_{k\to\infty}\nu(A_k)=0$. 所以

$$\lim_{n\to\infty}\nu^+\left(A\setminus\sum_{k=1}^{n-1}A_k\right)\wedge 1=0,\Rightarrow\lim_{n\to\infty}\nu^+\left(A\setminus\sum_{k=1}^{n-1}A_k\right)=0.$$

再利用$\nu^+\left(A\setminus\sum\limits_{k=1}^{\infty}A_k\right)\le\nu^+\left(A\setminus\sum\limits_{k=1}^{n-1}A_k\right)$, 可得$\nu^+\left(A\setminus\sum\limits_{k=1}^{\infty}A_k\right)=0$.

由($\bigstar$)的论述, 必有$\nu\left(A\setminus\sum\limits_{k=1}^{\infty}A_k\right)\ge 0$. 于是(3.2)式表明$\nu(A)>0$, 与$\nu(A)<0$矛盾. 故②成立.

(3)最后证明定理的结论, 设$A\in\mathscr{F}, B\in\mathscr{F}, B\subset A$, 则

$$\nu(B)+\nu((A\setminus B)\cap D)=\nu((A\cap D)\cup B)=\nu(A\cap D)+\nu(B\cap D^c).$$

由①②, $\nu(B)\le\nu(A\cap D)$, 所以$\nu^+(A)=\nu(A\cap D)$, 同理可证$\nu^-(A)=-\nu(A\cap D^c)$. 因此ν^+,ν^-都是$(\Omega,\mathscr{F})$上的测度, 且$\nu^-(\Omega)=-\nu(D^c)<\infty$, 此外有$\nu=\nu^+-\nu^-$, 定理证毕. □

注: 如果ν是不定积分, 取$D=[f\ge 0]$即可.

定义 3.3.3 *设ν,ν_1,ν_2是$(\Omega,\mathscr{F})$上的符号测度.*

*如果$\forall A\in\mathscr{F}$, $|\nu_2|(A)=0$可推出$|\nu_1|(A)=0$, 则称ν_1关于ν_2***绝对连续***, 记为$\nu_1\ll\nu_2$.*

*如果$\nu_1\ll\nu_2$且$\nu_2\ll\nu_1$, 则称ν_1与ν_2***等价***, 记为$\nu_1\sim\nu_2$.*

*如果存在$N\in\mathscr{F}$使得$|\nu_1|(N^c)=0, |\nu_2|(N)=0$, 则称$\nu_1$与$\nu_2$***相互奇异***, 记为$\nu_1\perp\nu_2$.*

*如果$N\in\mathscr{F}$满足$|\nu|(N^c)=0$, 则称N是ν的***支撑***(或者称ν***集中在N上***.)*

注: 一般来说, 支撑并非唯一确定.

注: (1)绝对连续的定义也可以改为$\forall A\in\mathscr{F}, |\nu_2|(A)=0\Rightarrow\nu_1(A)=0$. 这是因为, 当$|\nu_2|(A)=0$时, $|\nu_2|(A\cap D)=0\Rightarrow\nu_2(A\cap D)=0\Rightarrow\nu_1^+(A)=0$, 同理$\nu_1^-(A)=0$, 所以$|\nu_1|(A)=0$.

(2)设$\nu_1\ll\nu_2, \nu_1\perp\nu_2$, 则$\nu_1=0$.

例 3.3.1 *设$\mu_1(\mathrm{d}x)=\dfrac{1}{\sqrt{2\pi}\sigma}\exp\left(-\dfrac{(x-a)^2}{2\sigma^2}\right)\mathrm{d}x$, $\mu_2(\mathrm{d}x)=\mathrm{d}x$(Lebesgue测度), 则$\mu_1\sim\mu_2$.*

进一步, 若$p(x)\in(0,+\infty)$, 则$p(x)\mathrm{d}x$与$\mathrm{d}x$等价.

例 3.3.2 *设$\nu_1(\mathrm{d}x)=\delta_a(\mathrm{d}x), \nu_2(\mathrm{d}x)=\mathrm{d}x$, 则$\nu_1\perp\nu_2$. ($\delta_a(\mathrm{d}x)$: 若集合$A$含$a$, 则测度为1, 否则测度为0.)*

*进一步, $\sum\limits_{i=1}^{\infty}p_i\delta_{x_i}(\mathrm{d}x)$与$p(x)\mathrm{d}x$相互奇异, 即***离散型随机变量与连续型随机变量的分布是相互奇异的***. 因为$N=\{x_1,\cdots,x_n,\cdots\}$是两个测度的零测集.*

定理 3.3.5 (Lebesgue分解) *设μ,ν是$(\Omega,\mathscr{F})$上的两个有限符号测度, 则ν有如下的***Lebesgue分解***:*

$$\nu=\nu_s+\nu_c,$$

其中$\nu_s\perp\mu, \nu_c\ll\mu$, 且这样的分解是唯一的. 此外, 还满足:

(1)ν_s,ν_c均为σ有限的;

(2)存在$N\in\mathscr{F}, g\in\mathcal{L}$, 使得$|\mu|(N)=0, \nu_s(A)=\nu_s(A\cap N)$ (A是ν_s的支撑);

(3)g关于$|\mu|$的积分存在, $\nu_c=g.\mu$.

证明: 唯一性: 设$\nu_c+\nu_s=\nu_c'+\nu_s'$, 则$\nu_c-\nu_c'=\nu_s'-\nu_s$, 由于$\nu_c-\nu_c'\ll\mu, \nu_s'-\nu_s\perp\mu$, 则$\nu_c-\nu_c'=\nu_s'-\nu_s=0$.

存在性: 【待补充】. □

下面取$(\Omega,\mathscr{F})=(\mathbb{R},\mathscr{B}(\mathbb{R}))$, ν是随机变量ξ的概率分布, λ是$(\mathbb{R},\mathscr{B}(\mathbb{R}))$上的Lebesgue测度. 由Lebesgue分解定理, $\nu=\nu_s+\nu_c=\nu_s+\underbrace{g.\lambda}_{\text{连续}}$, 其中$\nu_s\perp\lambda, \nu_c\ll\lambda$. 令

$$D=\{x\in\mathbb{R}|\nu_s(\{x\})>0\}=\bigcup_{n=1}^{\infty}\{x\in\mathbb{R}|\nu_s(\{x\})\ge 1/n\},$$

则D是至多可数集(注意概率测度有限, 只能写成有限集合的可列并.)

令

$$\nu_{s,1}(A)=\nu_s(A\cap D),\forall A\in\mathscr{B}(\mathbb{R}),(\text{把}\nu_s\text{限制在}D\text{上})$$
$$\nu_{s,2}(A)=\nu_s(A)-\nu_{s,1}(A),\forall A\in\mathscr{B}(\mathbb{R}),$$

则$\nu_{s,1}(A),\nu_{s,2}(A)$是$(\mathbb{R},\mathscr{B}(\mathbb{R}))$上的概率测度. 从而

$$\nu=\nu_{s,1}+\nu_{s,2}+g.\lambda=\frac{\nu_{s,1}}{\nu_{s,1}(\mathbb{R})}\nu_{s,1}(\mathbb{R})+\frac{\nu_{s,2}}{\nu_{s,2}(\mathbb{R})}\nu_{s,2}(\mathbb{R})+\frac{g.\lambda}{g.\lambda(\mathbb{R})}g.\lambda(\mathbb{R}).$$

注意这里$\nu_{s,1}(\mathbb{R})+\nu_{s,2}(\mathbb{R})+g.\lambda(\mathbb{R})=1$, 因此我们把$\nu$写成了三个概率测度的凸组合, 且$\nu_{s,1}$是离散型随机变量的分布, $g.\lambda$是连续型随机变量的分布. 由于

$$\nu_{s,2}(\{x\})=\nu_s(\{x\})-\nu_s(\{x\}\cap D)=0,\forall x\in\Omega,$$

所以$\nu_{s,2}$不是离散型随机变量的分布, 又由于$\nu_{s,1}\perp\nu_{s,2}$, 所以$\nu_{s,2}$不是连续型随机变量的分布, 这样我们把$\nu_{s,2}$称为**奇异型**的, 它存在连续的分布函数但是没有密度函数.

例 3.3.3 *设$\xi\sim N(0,1)$, $\eta=\xi I_{(-\infty,0)}+I_{[0,+\infty)}(\xi)$, 则$\eta$的所有取值不可数, 故$\eta$不是离散型的. 注意*

$$P(\eta=1)=P(\xi\geq 0)=\frac{1}{2}.$$

则ξ不是连续型的.

例 3.3.4 *考虑$(\mathbb{R}^2,\mathscr{B}(\mathbb{R}^2))$. 设$\xi\sim N(0,1),\eta=\xi^2$, 则$(\xi,\eta)$的联合分布没有密度函数($(\xi,\eta)$取值在平面中为一条曲线, 此曲线在平面中测度为0), 但是$P\circ(\xi,\eta)^{-1}$关于$\mathrm{d}x$绝对连续.*

§ 3.4 Radon-Nikodym定理

定理 3.4.1 (Radon-Nikodym) *设$(\Omega,\mathscr{F})$是可测空间, μ是σ-有限测度, ν是符号测度(不必为σ-有限). 若ν关于μ绝对连续, 则存在关于μ积分存在的可测函数g, 使得$\nu=g.\mu$. 此外, g在μ等价的意义下唯一. 此外, g是μ-a.e.有限当且仅当ν是σ-有限的.*

注: g在μ等价的意义下唯一, 指如果g_1,g_2满足条件, 则$\mu([g_1\neq g_2])\neq 0$.

证明: 不妨设ν是测度(否则考虑ν^+和ν^-), 并设μ是有限测度(否则考虑$\tilde{\mu}=\sum_{n=1}^{\infty}\frac{1}{2^n}\frac{\mu(\cdot\cap A_n)}{\mu(A_n)}$.) 设$\nu(\Omega)=\infty$, 下面要找$g$.

设$\mathscr{B}=\{C\in\mathscr{F}|\nu(C)<\infty\}$, 则$\mathscr{B}$关于有限并运算封闭. 取$\{C_n\}\subset\mathscr{B}$使得$C_n\nearrow C$, 其中$\mu(C)=\sup\{\mu(B)|B\in\mathscr{F}\}$. 令$\nu_1(B)=\nu_1(B\cap C),\nu_2(B)=\nu(B\cap C^c),\forall B\in\mathscr{F}$, 则$\nu_1$是$\sigma$-有限测度. 由Lebesgue分解定理, 存在唯一非负实值可测函数g_1是恶$\nu_1=g_1.\mu$.

对任意$B\in\mathscr{F}$, 若$\mu(B\cap C^c)$, 则$\nu(B\cap C^c)=+\infty$. (若$\nu(B\cap C^c)<+\infty$, 则$B\cap C^c\in\mathscr{B}$, 此时$\mu((B\cap C^c)\cup C)=\mu(B\cap C^c)+\mu(C)>\mu(C)$, 矛盾.) 令$\nu_2=(+\infty)I_{C^c}.\mu$, (若$\int_B I_{C^c}\mathrm{d}\mu>0$, 则$\nu_2(B)=+\infty$). 则$\nu=\nu_1+\nu_2=(g_1I_c+(+\infty)I_{C^c}).\mu$. 由$g_1$唯一, 则此分解唯一.

若ν是σ-有限的, 则C^c有限, 从而g_1a.e.有限. □

例 3.4.1 *ν是σ-有限测度的条件不能去掉.*

反例: 取$\Omega=[0,1]$, $\boxed{\mathscr{F}=\{A\subset[0,1]|A\text{或}A^c\text{是至多可数集}\}}$, 则$\mathscr{F}$是$\sigma$-代数, 设

$$\mu(A)=\begin{cases}\sharp A, & A\text{至多可数},\\ +\infty, & A^c\text{至多可数},\end{cases},\nu(A)=\begin{cases}0, & A\text{至多可数},\\ 1, & A^c\text{至多可数},\end{cases}$$

则μ是$(\Omega,\mathscr{F})$上的测度, 但不是σ-有限测度; ν是测度, 且$\nu\ll\mu$.

(若μ是σ-有限测度, 则存在$\{A_n\}\subset\mathscr{F}$使得$\bigcup A_n=[0,1]$, 其中$A_n$是至多可数集, 但是$[0,1]$不可数, 这是不可能的.)

设$f\in\mathcal{L}^+$使得$\nu=f.\mu$, 则对任一$x\in[0,1]$, 有

$$0=\nu(\{x\})=\int_{\{x\}}f(y)\mu(\mathrm{d}y)=f(x)\mu(\{x\})=f(x),$$

但是$1=\nu([0,1])\neq\int_{[0,1]}f(y)\mu(\mathrm{d}y)=0$, 矛盾. 故不存在$f$使得$\nu=f.\mu$. □

定义 3.4.1 *把上述定理中的g记为$\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$, 称为$\nu$关于$\mu$的**Radon-Nikodym(RN)导数**.*

定理 3.4.2 *设$(\Omega,\mathscr{F},\mu)$是σ-有限测度空间, ν是$\mathscr{F}$上的符号测度, $\nu\ll\mu$. 令$g\in\overline{\mathcal{L}}$, 则$g$关于$\nu$积分存在当且仅当$g\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$关于$\mu$积分存在, 且此时有*

$$\int_A g\mathrm{d}\nu=\int_A\left(g\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\mathrm{d}\mu,\forall A\in\mathscr{F}.$$

注: 密度函数就是RN导数. $\mathbb{E}f(\xi)=\int_{\mathbb{R}}f(x)P\circ\xi^{-1}(\mathrm{d}x)=\int_{\mathbb{R}}f(x)p(x)\mathrm{d}x$.

定理 3.4.3 *设$(\Omega,\mathscr{F})$是可测空间, μ,ν是$\mathscr{F}$上的两个σ有限测度, φ是$\mathscr{F}$上的符号测度, 若$\varphi\ll\nu,\nu\ll\mu$, 则$\varphi\ll\mu$, 且*

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\mu}=\frac{\mathrm{d}\varphi}{\mathrm{d}\nu}\frac{\mathrm{d}\nu}{\mathrm{d}\mu},\mu-a.e..$$

注: 设$\tilde{\mu}=\sum_{n=1}^{\infty}\frac{1}{2^n}\frac{\mu(\cdot\cap A_n)}{\mu(A_n)}$, $\nu=g.\tilde{\mu}=g\frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu}.\mu$, 则$\nu\ll\tilde{\mu},\tilde{\mu}\ll\mu$, 且$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}=\frac{\mathrm{d}\nu}{\mathrm{d}\tilde{\mu}}\cdot\frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu}$.

如果μ,ν相互绝对连续(等价), 则$1=\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\cdot\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

*以上几个定理的证明: 待补充.

§ 3.5 L^p空间

§ 3.6 第三章习题

1. 设$(\Omega,\mathscr{F},\mu)$是有限测度空间, 在$\mathscr{F}$-可测函数全体构成的线性空间$\mathcal{L}$上, 定义距离$d(f,g)=\mu(|f-g|\wedge 1)$, 证明: $f_n\xrightarrow{d}f\Leftrightarrow f_n\xrightarrow{\mu}f$.

第 4 章　乘积可测空间

§ 4.1　乘积可测空间的定义

定义 4.1.1 设A_1, A_2是两个集合, 称

$$A_1 \times A_2 = \{(\omega_1, \omega_2) | \omega_1 \in A_1, \omega_2 \in A_2\},$$

为A_1, A_2的**乘积**. 若$(\Omega_1, \mathscr{F}_1)$与$(\Omega_2, \mathscr{F}_2)$是两个可测空间, 在$\Omega_1 \times \Omega_2$上定义如下$\sigma$-代数:

$$\mathscr{F}_1 \times \mathscr{F}_2 \triangleq \sigma(\{A_1 \times A_2 | A_1 \in \mathscr{F}_1, A_2 \times \mathscr{F}_2\}).$$

把$\mathscr{F}_1 \times \mathscr{F}_2$叫**乘积$\sigma$-代数**, $(\Omega_1 \times \Omega_2, \mathscr{F}_1 \times \mathscr{F}_2)$叫**乘积可测空间**.

容易推广到任意有限多个可测空间乘积的情形, 下面进一步定义一族可测空间的乘积.

定义 4.1.2 设$(\Omega_i)_{i \in I}$是一族集合, $\Omega = \bigcup\limits_{i \in I} \Omega_i$, 用$\Omega^I$表示从$I$到$\Omega$的映射全体, 把

$$\prod_{i \in I} \Omega_i \triangleq \{\omega \in \Omega^I | \omega(i) \in \Omega_i, i \in I\}$$

叫$(\Omega_i)_{i \in I}$的**乘积**. 此外, 对每个$i \in I$, 令

$$\pi_i(\omega) = \omega(i), \omega \in \prod_{i \in I} \Omega_i,$$

称π_i是$\prod\limits_{i \in I} \Omega_i$到$\Omega_i$的**投影映射**. 更一般地, 设$\varnothing \neq S \subset I$, 令

$$\pi_S(\omega) = (\omega(i), i \in S), \omega \in \prod_{i \in I} \Omega_i,$$

这里$(\omega(i), i \in S)$表示$\prod\limits_{i \in S} \Omega_i$中的元素, 在指标$i$处取值为$\omega(i)$. 称$\pi_S$是$\prod\limits_{i \in I} \Omega_i$到$\prod\limits_{i \in S} \Omega_i$的**投影映射**.

设$(\Omega_i, \mathscr{F}_i)_{i \in I}$是一族可测空间, 则在$\prod\limits_{i \in I} \Omega_i$上定义$\sigma$-代数如下:

$$\prod_{i \in I} \mathscr{F}_i \triangleq \sigma\Big(\bigcup_{i \in I} \pi_i^{-1}(\mathscr{F}_i)\Big),$$

称$\prod\limits_{i \in I} \mathscr{F}_i$为**乘积$\sigma$-代数**, $\Big(\prod\limits_{i \in I} \Omega_i, \prod\limits_{i \in I} \mathscr{F}_i\Big)$是**乘积可测空间**.

注: 若$I = \{1, 2\}$, 则相当于$\Omega_1 \times \Omega_2 = \{\omega \in \Omega^I | \omega(i) \in \Omega_i, i \in \{1, 2\}\}$, (x, y)可以看作$(1, 2) \mapsto (x, y)$的映射, 本质上与刚开始的有限个情形的定义是一样的. 这样$\pi_1^{-1}(A_1) = A_1 \times \Omega_2, \pi_2^{-1}(A_2) = \Omega_1 \times A_2$, $A_1 \times A_2 = \pi_1^{-1}(A_1) \cap \pi_2^{-1}(A_2)$.

显然, 乘积σ-代数是使得每个投影π_i为可测的最小σ-代数.

定理 4.1.1 设$\varnothing \neq S \subset I$, 则$\pi_S$是$\left(\prod_{i\in I}\Omega_i, \prod_{i\in I}\mathscr{F}_i\right)$ 到$\left(\prod_{i\in S}\Omega_i, \prod_{i\in S}\mathscr{F}_i\right)$的可测映射.

证明: 记π_i^S表示$\prod_{i\in S}\Omega_i$到Ω_i的投影. 由定义, $\prod_{i\in S}\mathscr{F}_i = \sigma\left(\bigcup_{i\in S}(\pi_i^S)^{-1}(\mathscr{F}_i)\right)$, 由命题2.1.2, 只需证

$$\pi_S^{-1}\left(\bigcup_{i\in S}(\pi_i^S)^{-1}(\mathscr{F}_i)\right) \subset \prod_{i\in I}\mathscr{F}_i.$$

这等价于

$$\pi_S^{-1}((\pi_i^S)^{-1}(\mathscr{F}_i)) \subset \prod_{i\in I}\mathscr{F}_i, \forall i \in S.$$

再根据$\pi_S^{-1}(\pi_i^S)^{-1}(\mathscr{F}_i) = \pi_i^{-1}(\mathscr{F}_i)$可得欲证结论. □

定理 4.1.2 令$\mathcal{P}_0$(resp. $\mathcal{P}$)表示I的非空有穷(resp. 至多可数)子集全体. 则

(1)**可测矩形**全体

$$\mathcal{I} = \left\{\pi_S^{-1}\left(\prod_{i\in S}A_i\right)\Big| A_i \in \mathscr{F}_i, i \in S; S \in \mathcal{P}_0\right\}$$

是$\prod_{i\in I}\Omega_i$上的半代数, 且$\sigma(\mathcal{I}) = \prod_{i\in I}\mathscr{F}_i$.

(2)**可测柱集**全体

$$\mathcal{Z} = \left\{\pi_S^{-1}\left(\prod_{i\in S}\mathscr{F}_i\right)\Big| S \in \mathcal{P}_0\right\}$$

是$\prod_{i\in I}\Omega_i$上的代数, 且$\sigma(\mathcal{Z}) = \prod_{i\in I}\mathscr{F}_i$.

(3) $\prod_{i\in I}\mathscr{F}_i = \left\{\pi_S^{-1}\left(\prod_{i\in S}\mathscr{F}_i\right)\Big| S \in \mathcal{P}\right\}$.

注: (1)$\mathcal{I}$中的元素形如$\prod_{i\in S}A_i \times \prod_{j\in I\setminus S}\Omega_j$, $\mathcal{Z}$中的元素形如$A \times \prod_{j\in I\setminus S}\Omega_j$, 其中$A \in \prod_{i\in S}\mathscr{F}_i$.

(2)$\prod_{i\in I}A_i$关于乘积可测空间未必可测, 因为$\bigcap_{i\in I}\pi_i^{-1}(A_i)$的$I$可能不是有限的.

§ 4.2 乘积测度与Fubini定理

4.2.1 乘积测度的定义

设$(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$是两个$\sigma$-有限测度空间, 下面要在乘积空间$(X \times Y, \mathcal{A} \times \mathcal{B})$上定义乘积测度$\mu \times \nu$与这个测度上的积分.

定义 4.2.1 设X, Y是两个集合, $E \subset X \times Y$, 令

$$E_x = \{y \in Y | (x, y) \in E\},$$
$$E_y = \{x \in X | (x, y) \in E\}.$$

分别称E_x, E^y为x, y处的**截口**.

设$f(x,y)$定义在$X\times Y$上, 记$f_x(y)=f(x,y), f^y(x)=f(x,y)$.

引理 4.2.1 设$(X,\mathcal{A}),(Y,\mathcal{B})$是可测空间.

(1)若$E\times\mathcal{A}\times\mathcal{B}$, 则$\forall x\in X, y\in Y$, 有$E_x\in\mathcal{B}, E^y\in\mathcal{A}$.

(2)若f是$X\times Y$上的$\mathcal{A}\times\mathcal{B}$可测函数, 则对一切$x\in X, y\in Y, f_x$是$Y$上的$\mathcal{B}$可测函数, f^y是X上的$\mathcal{A}$可测函数.

证明: (1)令$\mathcal{C}=\{A\times B|A\in\mathcal{A}, B\in\mathcal{B}\}$, 则对$E\in\mathcal{C}$, 引理结论成立. 再设

$$\mathcal{G}=\{E\in\mathcal{A}\times\mathcal{B}|\forall x\in X, y\in Y, E_x\in\mathcal{B}, E^y\in\mathcal{A}\}.$$

则$\mathcal{G}$是λ类且$\mathcal{C}\subset\mathcal{G}$. 由于$\mathcal{C}$是$\pi$类, 且$\sigma(\mathcal{C})=\mathcal{A}\times\mathcal{B}$, 由单调类定理, $\mathcal{G}=\sigma(\mathcal{C})$, 所以对$E\in\mathcal{A}\times\mathcal{B}$, 结论成立. □

引理 4.2.2 设$(X,\mathcal{A},\mu)$与$(Y,\mathcal{B},\nu)$是两个σ-有限的测度空间. 设$E\in\mathcal{A}\times\mathcal{B}$, 则函数$x\mapsto\nu(E_x)$是$\mathcal{A}$可测, 函数$y\mapsto\mu(E^y)$是$\mathcal{B}$可测.

证明: (1)设ν是有限测度. 令$\mathcal{C}=\{A\times B|A\in\mathcal{A}, B\in\mathcal{B}\}$, 设

$$\mathcal{G}=\{E\in\mathcal{A}\times\mathcal{B}|x\mapsto\nu(E_x)\text{是}\mathcal{A}\text{可测}\},$$

则$\mathcal{G}$是λ类且$\mathcal{C}\subset\mathcal{G}$(注意$\nu((A\times B)_x)=I_A(x)\nu(B)$). 由于$\mathcal{C}$是$\pi$类, 且$\sigma(\mathcal{C})=\mathcal{A}\times\mathcal{B}$, 由单调类定理, $\mathcal{G}=\sigma(\mathcal{C})=\mathcal{A}\times\mathcal{B}$, 故结论成立.

(2)下设ν是σ-有限测度. 取Y的可数划分$\{D_n\}$, 使得$D_n\in\mathcal{B}, \nu(D_n)<\infty$, 令$\nu_n(B)=\nu(B\cap D_n), B\in\mathscr{B}$, 则$\nu_n$是有限测度, $\nu=\sum\limits_{n=1}^{\infty}\nu_n$, 于是

$$\nu(E_x)=\sum_{n=1}^{\infty}\nu_n(E_x), E\in\mathcal{A}\times\mathcal{B}.$$

从而函数$x\mapsto\nu(E_x)$是$\mathcal{A}$可测. 同理可证$y\mapsto\mu(E^y)$是$\mathcal{B}$可测. □

定理 4.2.3 设$(X,\mathcal{A},\mu),(Y,\mathcal{B},\nu)$是两个$\sigma$有限测度空间, 则在$\mathcal{A}\times\mathcal{B}$上存在唯一的测度$\mu\times\nu$, 使得

(1)$(\mu\times\nu)(A\times B)=\mu(A)\nu(B), A\in\mathcal{A}, B\in\mathcal{B}$, 从而$\mu\times\nu$也是$\sigma$有限的. 此外,

(2)对任何$E\in\mathcal{A}\times\mathcal{B}$, 有$(\mu\times\nu)(E)=\displaystyle\int_X\nu(E_x)\mu(\mathrm{d}x)=\int_Y\mu(E^y)\nu(\mathrm{d}y)$.

把测度$\mu\times\nu$叫μ,ν的**乘积**.

注: (i)“存在唯一”: 我们仅在测度的扩张定理见过这样的话.

(ii)(1)可以联想随机变量的独立性, (2)把结论推广到对一般的集合成立.

(iii)$\displaystyle\int_X\nu(\cdot)\mu(\mathrm{d}x)$是测度, 用截口性质和单调收敛定理可以证明.

证明: 由前一引理, $x\mapsto\nu(E_x)$是$\mathcal{A}$可测的, $y\mapsto\mu(E^y)$是$\mathcal{B}$可测的, 所以可以定义

$$\begin{aligned}\lambda_1(E)&=\int_X\nu(E_x)\mu(\mathrm{d}x), \forall E\in\mathcal{A}\times\mathcal{B},\\ \lambda_2(E)&=\int_Y\nu(E^y)\nu(\mathrm{d}y), \forall E\in\mathcal{A}\times\mathcal{B}.\end{aligned}.$$

则λ_1,λ_2都是$\mathcal{A}\times\mathcal{B}$上的测度. 令$\mathcal{C}=\{A\times B|A\in\mathcal{A}, B\in\mathcal{B}\}$, 则$\lambda_1|_{\mathcal{C}}=\lambda_2|_{\mathcal{C}}$.

由Carathéodory测度扩张定理, 满足(1)的测度唯一, 特别有$\lambda_1=\lambda_2$, 令$\mu\times\nu=\lambda_1=\lambda_2$即可. □

4.2.2 乘积测度的积分

定理 4.2.4 令$(X,\mathcal{A},\mu),(Y,\mathcal{B},\nu)$是$\underline{\sigma\text{-有限}}$的测度空间，$f$是$X\times Y$上的$\underline{\text{非负}}\mathcal{A}\times\mathcal{B}$可测函数，则函数$x\mapsto\int_Y f_x\mathrm{d}\nu$是$\mathcal{A}$可测，$y\mapsto\int_X f^y\mathrm{d}\mu$是$\mathcal{B}$可测，而且

$$\int_{X\times Y}\mathrm{d}(\mu\times\nu)=\int_Y\Big(\int_X f^y\mathrm{d}\mu\Big)\nu(\mathrm{d}y)=\int_X\Big(\int_Y f^x\mathrm{d}\nu\Big)\mu(\mathrm{d}x).$$

证明: 不妨设μ,ν都是有限测度, 令$\mathcal{C}=\{A\times B|A\in\mathcal{A},B\in\mathcal{B}\}$, 则由定理4.2.3, $\mathcal{C}$中集合的示性函数满足定理的条件. 由可测函数的构造, 对一切有界的$\mathcal{A}\times\mathcal{B}$可测函数$f$, 定理结论成立. 因此对一切非负$\mathcal{A}\times\mathcal{B}$可测函数$f$, 定理结论成立. □

定理 4.2.5 (Fubini) 设$(X,\mathcal{A},\mu),(Y,\mathcal{B},\nu)$是$\sigma$-有限测度空间，$f$是$X\times Y$上的$\mathcal{A}\times\mathcal{B}$可测函数. 若$f$关于$\mu\times\nu$可积*(resp. 积分存在)*, 则有如下结论:

(1)对μ-*a.e.*x, f_x关于ν可积*(resp. 积分存在)*; 对ν-*a.e.*y, f_y关于μ可积*(resp. 积分存在)*.

(2)令

$$I_f(x)=\begin{cases}\displaystyle\int_Y f_x\mathrm{d}\nu, & \text{若}f_x\text{为}\nu\text{可积(resp. 积分存在)}\\ 0, & \text{其他}\end{cases}$$

$$I^f(y)=\begin{cases}\displaystyle\int_X f_y\mathrm{d}\mu, & \text{若}f_y\text{为}\mu\text{可积(resp. 积分存在)}\\ 0, & \text{其他}\end{cases}$$

则I_f是μ可积*(resp.积分存在)*; I^f是ν可积*(resp.积分存在)*, 且有

$$\int_{X\times Y}f\mathrm{d}(\mu\times\nu)=\int_X I_f(x)\mu(\mathrm{d}x)=\int_Y I^f(y)\nu(\mathrm{d}y).$$

例 4.2.1 累次积分存在但是重积分不一定存在的例子.

答: 设$x\in\mathbb{Q}$, 记为$\dfrac{p_x}{q_x}$, 其中p_x,q_x互素且$q_x>0$. 在正方形$D=[0,1]\times[0,1]$上定义函数f如下:

$$f(x,y)=\begin{cases}1, & (x,y)\text{是有理点且}q_x=q_y,\\ 0, & \text{其他}.\end{cases}$$

对固定的$y\in[0,1]$, 若y是无理数, 则$f(x,y)=0$, 从而$\int_0^1 f(x,y)\mathrm{d}x=0$. 若$y=\dfrac{p_y}{q_y}$是有理数, 分母等于$q$的有理数不超过$q$个, 所以$\int_0^1 f(x,y)\mathrm{d}x=0$. 这样对任一$y\in[0,1]$有$\int_0^1 f(x,y)\mathrm{d}x=0$, 从而$\int_0^1\mathrm{d}y\int_0^1 f(x,y)\mathrm{d}x=0$. 同理, $\int_0^1\mathrm{d}x\int_0^1 f(x,y)\mathrm{d}y=0$.

下证f二重积分不存在. 令$G=\{(x,y)|(x,y)\text{是}D\text{中有理点且}q_x=q_y\}$. 可以证明$G$在$D$内稠密. 若$(x,y)\in G$, 则$f(x,y)=1$; 若$(x,y)\notin G$, 则$f(x,y)=0$. 这样$f$在$G$上无处连续, 从而$f$在$D$上不可积.

这个例子可以见汪林《数学分析中的问题和反例》.

例 4.2.2 *Fubini定理中, σ-有限测度的条件不能去掉.*

答: 设$\sharp$是计数测度(它不是σ-有限的). $(X,\mathcal{A},\mu)=([0,1],[0,1]\cap\mathscr{B}(\mathbb{R}),\mathrm{d}x)$, $(Y,\mathcal{B},\nu)=([0,1],[0,1]\cap\mathscr{B}(\mathbb{R}),\sharp)$, $f(x,y)=I_D(x,y)$, 其中$D=\{(x,y)|x=y\}$, 则

$$\int_0^1\left[\int_0^1 I_D(x,y)\mathrm{d}x\right]\sharp(\mathrm{d}y)=\int_0^1\sharp(\mathrm{d}y)=0,$$
$$\int_0^1\left[\int_0^1 I_D(x,y)\sharp(\mathrm{d}y)\right]\mathrm{d}x=\int_0^1 1\mathrm{d}x=1.$$

注: Fubini定理中f关于$\mu\times\nu$积分存在(或者可积)条件也不可去掉.

§ 4.3　(*)转移测度

> 笔者注: 这里我并没有学明白, 所以笔记记得一塌糊涂. 如果有同学学明白这部分, 欢迎跟我分享你的笔记, 我的微信号是Fiddie_Math.

定义 4.3.1 *设$(X,\mathcal{A}),(Y,\mathcal{B})$是可测空间, $K:X\times\mathcal{B}\to[0,+\infty]$满足:*
(1)$\forall x\in X, K(x,\cdot)$是$\mathcal{B}$上的测度;
(2)$\forall B\in\mathcal{B}, x\mapsto K(x,B)$是$\mathcal{A}$-可测的函数,
*则称K是$(X,\mathcal{A})$到$(Y,\mathcal{B})$的**转移测度(核, kernel)**.*

*若$\forall x\in X, K(x,Y)<+\infty$, 则称$K$是**有限转移测度**. 若$\forall x\in X, K(x,Y)=1$, 称$K$是**转移概率**.*

若存在$\{B_n\}\subset\mathcal{B}$且$\{B_n\}$两两不交, 满足$\sum\limits_{n=1}^{\infty}B_n=Y$且$K(x,B_n)<+\infty$, $\forall x\in X, n\geq 1$, 则称K是σ-有限转移测度.

例 4.3.1 *设$(X,\mathcal{A})=(Y,\mathcal{B})=(\mathbb{R},\mathscr{B}(\mathbb{R}))$. 再设$p(y|x)=\dfrac{p(x,y)}{p_X(x)}$. 固定$x$, 则$p(y|x)\mathrm{d}y$是测度($K(x,\mathrm{d}y)=p(y|x)\mathrm{d}y$.)*

定义$K(x,B)=\displaystyle\int_B p(y|x)\mathrm{d}y$, 其中$B\in\mathscr{B}(\mathbb{R})$, 则$K$是转移测度.

注: $x\mapsto K(x,B)=\displaystyle\int_B p(y|x)\mathrm{d}y$是$\mathbb{R}$上的连续函数.

定理 4.3.1 *设K是$(X,\mathcal{A})$到$(Y,\mathcal{B})$的转移测度, μ是$(X,\mathcal{A})$上的测度, 则*

$$\nu(B)\triangleq\mu.K(B)=\int_X K(x,B)\mu(\mathrm{d}x)$$

是$(Y,\mathcal{B})$上的测度.

在例4.3.1中, $\mu.K(B)=\displaystyle\int_X\int_B p(y|x)\mathrm{d}y,\mu(\mathrm{d}x)=\int_X\underbrace{\int_B\frac{p(x,y)}{p_\xi(x)}\mathrm{d}y}_{=\int_B p_\eta(y)\mathrm{d}y}\cdot p_\xi(x)\mathrm{d}x$.

$\nu(I_B)=\displaystyle\int_X K(x,I_B)\mu(\mathrm{d}x)=\int_X\int_Y I_B K(x,\mathrm{d}y)\mu(\mathrm{d}x)$.

$K.\nu(A\times B)=\displaystyle\int_A\int_B p(y|x)\mathrm{d}y p_\xi(x)\mathrm{d}x=\int_{A\times B}p(x,y)\mathrm{d}x\mathrm{d}y$.

定理 4.3.2 设$(X,\mathcal{A},\mu)$是σ-有限测度空间, K是$(X,\mathcal{A})$到$(Y,\mathcal{B})$的σ-有限转移测度. 令$\mu.K(E)=\int_X K(x,E_x)\mu(\mathrm{d}x),\forall E\in\mathcal{A}\times\mathcal{B}$, 则$\mu.K$是$(X\times Y,\mathcal{A}\times\mathcal{B})$上的$\sigma$-有限测度.

(简证)设$\{A_n\}\subset\mathcal{A}$是可测划分, $\mu(A_n)<+\infty$, 并设$\{B_m\}\subset\mathcal{B}$使得$K(x,B_m)<+\infty,\forall x\in X,m\geq 1$. 证明各个划分$\{A_n\times B_m\}$的$\mu.K$测度均有限. □

注: 设$(X,\mathcal{A},\mu),(Y,\mathcal{B},\nu)$是$\sigma$-有限的测度空间, K是$(X,\mathcal{A})$到$(Y,\mathcal{B})$的σ-有限转移测度. 若$K(x,\cdot)=\nu,\forall x\in X$, 则$\mu\times\nu=\mu.K$.

§ 4.4 (*)无穷乘积上的概率测度

无穷次Bernoulli试验中, 设ξ是首次成功进行的试验次数, $\xi\in\{1,2,\cdots\}$. 定义

$$\Omega_i=\{0,1\},\mathscr{F}_i=\{\{1,0\},\{1\},\{0\},\varnothing\},P_i=p\delta_1+(1-p)\delta_0.$$

考虑$\left(\prod_{i=1}^{\infty}\Omega_i,\prod_{i=1}^{\infty}\mathscr{F}_i,?\right)$, 其中, $\prod_{i=1}^{\infty}\mathscr{F}_i=\sigma\left(\prod_{i=1}^{n}A_i:A_i\in\mathscr{F}_i;i=1,2,\cdots;n\geq 1\right)$.

只需要在$\prod_{i=1}^{n}A_i\times\prod_{j=n+1}^{\infty}\Omega_j$上考虑柱形集. (前$n$次发生之后, 后面不知道). 注意矩形集或柱形集全体都可以扩张为全空间上的测度.

由于$(\Omega_1,\mathscr{F}_1)\to(\Omega,\mathscr{F}_2)$上有转移概率$P(\omega_1,\mathrm{d}(\omega_1))$, 故可定义$(\Omega_1\times\Omega_2,\mathscr{F}_1\times\mathscr{F}_2,P(\omega_1,\mathrm{d}\omega_1)P(\mathrm{d}\omega_2))$. 进一步可以定义$P((\omega_1,\omega_2),\mathrm{d}\omega_3),\cdots,P((\omega_1,\omega_2,\cdots,\omega_{n-1})\mathrm{d}\omega_n)$等等.

§ 4.5 第四章习题

例 4.5.1 设F是连续的概率分布函数, 则$\int_{\mathbb{R}}F(x)\mathrm{d}F(x)=\frac{1}{2}$.

注: $F(x)=P(X\leq x)$是右连续的函数.

证明: 注意到

$$\begin{aligned}\int_{\mathbb{R}}F(x)\mathrm{d}F(x)&=\int_{\mathbb{R}}\int_{-\infty}^{x}\mathrm{d}F(y)\mathrm{d}F(x)\\&=\int_{\mathbb{R}}\int_{\mathbb{R}}I_{(-\infty,x]}(y)\mathrm{d}F(y)\mathrm{d}F(x)\\&=\int_{\mathbb{R}}\int_{\mathbb{R}}I_{[y,+\infty)}(x)\mathrm{d}F(x)\mathrm{d}F(y)\\&=\int_{\mathbb{R}}(1-F(y-0))\mathrm{d}F(y)\\ \text{【}x\text{连续】}\quad&=\int_{\mathbb{R}}(1-F(y))\mathrm{d}F(y)\\&=1-\int_{\mathbb{R}}F(y)\mathrm{d}F(y),\end{aligned}$$

所以$\int_{\mathbb{R}}F(x)\mathrm{d}F(x)=\frac{1}{2}$. □

例 4.5.2 (重要) 设ξ是非负随机变量, $p\geq 1$是常数, 则$\mathbb{E}\xi^p=p\int_0^{+\infty}x^{p-1}P(\xi\geq x)\mathrm{d}x$.

证明: 注意到

$$\begin{aligned}\mathbb{E}\xi^p &= \int_\Omega \xi^p \mathrm{d}P = \int_\Omega \int_0^\xi p x^{p-1} \mathrm{d}x \mathrm{d}P \\ &= p\int_0^{+\infty} \int_\Omega I_{[0,\xi]}(x) x^{p-1} \mathrm{d}P \mathrm{d}x \\ &= p\int_0^{+\infty} x^{p-1} \int_\Omega I_{[x,+\infty)}(\xi) \mathrm{d}P \mathrm{d}x \\ &= p\int_0^{+\infty} x^{p-1} P(\xi \geq x) \mathrm{d}x\end{aligned}$$

注: $\int_0^{+\infty} P(\xi \geq x)\mathrm{d}x = \int_0^{+\infty} P(\xi > x)\mathrm{d}x$. 因为$P(\xi \geq x)$与$P(\xi > x)$只在至多可数个点不一样(回顾单调函数性质), 而至多可数个点对积分无影响.

例 4.5.3 设F, G是$\mathbb{R}$上的右连续增函数, 则

$$\int_{(a,b]} F(x)\mathrm{d}G(x) = F(x)G(x)\Big|_a^b - \int_{(a,b]} G(x-0)\mathrm{d}F(x).$$

证明: 只需注意到

$$\begin{aligned}\text{LHS} &= \int_{(a,b]} \left[\int_{(a,b]} I_{(a,x]}(y)\mathrm{d}F(y) + F(a)\right] \mathrm{d}G(x) \\ &= \int_{(a,b]} \left[\int_{(a,b]} I_{(a,x]}(y)\mathrm{d}F(y)\right] \mathrm{d}G(x) + F(a)[G(b)-G(a)] \\ &= \int_{(a,b]} \left[\int_{(a,b]} I_{[y,b]}(x)\mathrm{d}G(x)\right] \mathrm{d}F(y) + F(a)[G(b)-G(a)] \\ &= \int_{(a,b]} [G(b) - G(y-0)]\, \mathrm{d}F(y) + F(a)[G(b)-G(a)] \\ &= \text{RHS}. \qquad \square\end{aligned}$$

例 4.5.4 设$c > 0$是常数, $F(x)$是概率分布函数, 则$\int_{\mathbb{R}} [F(x+c) - F(x)]\mathrm{d}x = c$.

证明: 只需注意到

$$\begin{aligned}\int_{\mathbb{R}} [F(x+c) - F(x)]\mathrm{d}x &= \int_{\mathbb{R}} \left(\int_{-\infty}^{x+c} \mathrm{d}F(y) - \int_{-\infty}^{x} \mathrm{d}F(y)\right) \mathrm{d}x \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} I_{(x,x+c]}(y)\mathrm{d}F(y)\mathrm{d}x \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} I_{[y-c,y)}(x)\mathrm{d}x\mathrm{d}F(y) \\ &= \int_{\mathbb{R}} c\mathrm{d}F(y) = c.\end{aligned}$$

(注意$x < y \leq x + c \Leftrightarrow y - c \leq x < y$.) $\square$

下面的引理很Nice, 所以叫Nice引理(后面也会经常用到).

引理 4.5.1 (Nice) 假定Y是非负$r.v.$, 则

$$\sum_{n=1}^{\infty} P(Y \geq n) \leq \mathbb{E}Y \leq \sum_{n=1}^{\infty} P(Y \geq n) + 1.$$

证明: 注意$P(Y \geq n)$关于n单调递减, 且由Y非负, 回顾: $\mathbb{E}Y = \int_0^{\infty} P(Y > y)\mathrm{d}y$. 所以

$$\mathbb{E}Y = \sum_{n=0}^{\infty} \int_n^{n+1} P(Y \geq y)\mathrm{d}y \leq \sum_{n=0}^{\infty} \int_n^{n+1} P(Y \geq n)\mathrm{d}y = \sum_{n=1}^{\infty} P(Y \geq n) + 1.$$

且

$$\mathbb{E}Y = \sum_{n=1}^{\infty} \int_{n-1}^{n} P(Y \geq y)\mathrm{d}y \geq \sum_{n=1}^{\infty} \int_{n-1}^{n} P(Y \geq n)\mathrm{d}y = \sum_{n=1}^{\infty} P(Y \geq n).$$

例 4.5.5 假定X是非负$r.v.$, 则$\mathbb{E}|X|^p < +\infty \Leftrightarrow \sum_{n=1}^{\infty} n^{p-1} P(|X| > n) < +\infty$.

证明: 注意到

$$\mathbb{E}|X|^p = p\int_0^{\infty} y^{p-1} P(|\xi| \geq y)\mathrm{d}y = p\sum_{n=1}^{\infty} \int_{n-1}^{n} y^{p-1} P(|\xi| \geq y)\mathrm{d}y.$$

所以

$$\mathbb{E}|X|^p \leq 2\sum_{n=1}^{\infty} n^{p-1} P(|\xi| \geq n-1) \leq 2 + 4^{p-1} \sum_{n=2}^{\infty} (n-1)^{p-1} P(|\xi| \geq n-1),$$

且

$$\mathbb{E}|X|^p \geq 2\sum_{n=1}^{\infty} (n-1)^{p-1} P(|\xi| \geq n) \leq \sum_{n=2}^{\infty} n^{p-1} P(|\xi| \geq n).$$

第 5 章　独立性、条件期望、一致可积

§ 5.1　独立性

定义 5.1.1 设$(\Omega,\mathscr{F},P)$是概率空间, 把$\mathscr{F}$中元称为**事件**, 把Ω称为**必然事件**, 把Ω上的$\mathscr{F}$可测函数叫**随机变量**.

设ξ是随机变量, 如果ξ关于P积分存在, 则称积分$\int_\Omega \xi \mathrm{d}P$为$\xi$的**数学期望**, 记为$\mathbb{E}\xi$.

关于独立的一些定义:

定义 5.1.2 设$(\Omega,\mathscr{F},P)$是概率空间.

(1)设$A,B\in\mathscr{F}$, 称A,B**相互独立**, 若$P(AB)=P(A)P(B)$.

(2)设$\{A_i\}\subset\mathscr{F}$, 称$A_1,\cdots,A_n$相互独立, 若$\forall\le i_1<i_2<\cdots<i_k\le n(2\le k\le n)$, 有

$$P\Big(\bigcap_{j=1}^k A_{i_j}\Big)=\prod_{j=1}^k P(A_{i_j}).$$

(3)设$\{A_t\}_{t\in T}\subset\mathscr{F}$, 称$\{A_t\}$**相互独立**, 若对$T$的任意有限子集$\{t_1,\cdots,t_n\}$都有

$$P(A_{t_1}\cap\cdots\cap A_{t_n})=\prod_{i=1}^n P(A_{t_i}).$$

(4)设$\{\mathscr{C}_t|t\in T\}$是一族事件类, $\mathscr{C}_t\subset\mathscr{F}(\forall t\in T)$, 称$\{\mathscr{C}_t\}_{t\in T}$为**独立事件类**, 若$\forall A_t\in\mathscr{C}_t$, 有$\{A_t\}_{t\in T}$相互独立.（每个集类中取一个，构成的新集类中事件相互独立.）

(5)设$\{X_t\}_{t\in T}$是一族随机变量, 称$\{X_t\}_{t\in T}$相互独立, 若$\{\sigma(X_t)\}_{t\in T}$为独立事件类.（特别地，若随机变量X,Y独立，指$\sigma(X),\sigma(Y)$独立.）

注: 回顾: 可测映射生成的σ-代数: $\sigma(X)=\sigma(X^{-1}(\mathscr{B}(\overline{\mathbb{R}})))$.

根据定义, 事件A,B独立等价于随机变量I_A,I_B独立, 这里

$$\sigma(I_A)=\{\Omega,A,A^c,\varnothing\},\sigma(I_B)=\{\Omega,B,B^c,\varnothing\}.$$

容易证明如果事件A,B独立, 则事件$\overline{A},B$也独立.

定理 5.1.1 (独立类扩张定理) 设$\{\mathscr{C}_t\}_{t\in T}$是独立事件类, 且$\forall t\in T,\mathscr{C}_t$是$\pi$类, 则$\{\sigma(\mathscr{C}_t)\}_{t\in T}$是独立事件类.

注: 当$|T|=2$时, $\mathscr{C}_1=\{[X<a]\},\mathscr{C}_2=\{[Y<a]\}$都是$\pi$类, 而$\sigma(\mathscr{C}_1)=\sigma(X),\sigma(\mathscr{C}_2)=\sigma(Y)$. 在矩形集上可以定义

$$P(X\le x,Y\le y)=P(X\le x)P(Y\le y),$$

而独立类扩张定理表明可以推出

$$P(X\in A,Y\in B)=P(X\in A)P(Y\in B),\Leftrightarrow P\circ(X,Y)^{-1}=P\circ X^{-1}\times P\circ Y^{-1}.$$

证明: 只需证$\forall A_t\in\sigma(\mathscr{C}_t)$, 得到$\{A_t\}_{t\in T}$, 对$t_1,\cdots,t_n\in T$, 都有$P\Big(\bigcap\limits_{i=1}^{n}A_{t_i}\Big)=\prod\limits_{i=1}^{n}P(A_{t_i})$. 于是$\{\sigma(\mathscr{C}_{t_1}),\cdots,\sigma(\mathscr{C}_{t_n})\}$是独立事件类. 只需证$\{\sigma(\mathscr{C}_{t_1}),\mathscr{C}_{t_2},\cdots,\mathscr{C}_{t_n}\}$是独立事件类.

设$\{t_1,\cdots,t_n\}$是T的任一有限子集, $A_{t_1}\in\sigma(\mathscr{C}_{t_1}),A_{t_i}\in\mathscr{C}_{t_i}(i=2,\cdots,n)$, 下证$P(A_{t_1}A_{t_2}\cdots A_{t_n})=\prod\limits_{i=1}^{n}P(A_{t_i})$. 但是$\Omega$存在与否不会影响独立性, 所以可以不妨假设$\Omega\in\mathscr{C}_t,\forall t\in T$. 令

$$\mathcal{H}=\{A\in\sigma(\mathscr{C}_{t_1})|P(A\cap A_{t_2}\cap\cdots\cap A_{t_n})=P(A)\prod_{j=2}^{n}P(A_{t_j}),\forall A_{t_i}\in\mathscr{C}_{t_i},2\le i\le n\}.$$

则$\mathscr{C}_{t_1}\subset\mathcal{H}$且$\mathscr{C}_{t_1}$是$\pi$类, $\mathcal{H}$是λ类, 所以由单调类定理, $\sigma(\mathscr{C}_{t_1})=\mathcal{H}$. 故$\{\sigma(\mathscr{C}_{t_1}),\mathscr{C}_{t_2},\cdots,\mathscr{C}_{t_n}\}$是独立事件类.

重复上面的步骤(对$k=2,3,\cdots,n$), 可得$\{\sigma(\mathscr{C}_{t_1}),\cdots,\sigma(\mathscr{C}_{t_n})\}$是独立事件类. □

推论 5.1.2 *设$\{\mathscr{F}_t\}_{t\in T}$是一族$\sigma$-代数, 且相互独立, $S\subset T$, 则$\bigvee\limits_{t\in S}\mathscr{F}_t$与$\bigvee\limits_{t\in T\setminus S}\mathscr{F}_t$相互独立. 这里$\bigvee\limits_{t\in S}\mathscr{F}_t$定义为$\sigma\Big(\bigcup\limits_{t\in S}\mathscr{F}_t\Big)$.*

推论 5.1.3 *设X,Y独立, 且X,Y与XY都期望存在, 则$\mathbb{E}XY=\mathbb{E}X\mathbb{E}Y$.*

证明: $\mathbb{E}XY=\int_{\mathbb{R}^2}xy\mathrm{d}F(x,y)=\int_{\mathbb{R}^2}xy\mathrm{d}F_X(x)\mathrm{d}F_Y(y)=\int_{\mathbb{R}}x\mathrm{d}F_X(x)\int_{\mathbb{R}}y\mathrm{d}F_Y(y)=\mathbb{E}X\mathbb{E}Y.$ □

注: 这里综合运用了独立性的定义和Fubini定理.

例 5.1.1 *设X,Y独立, f,g为可测函数, 则$f(X)$与$f(Y)$独立.*

证明: 注意$\sigma(f(X))\subset\sigma(X),\sigma(g(Y))\subset\sigma(Y)$, 这是因为$(f(X))^{-1}(A)=[x\in f^{-1}(A)]\in\sigma(X),\forall A$. □

下面的Borel-Cantelli引理需要刻在DNA里.

定理 5.1.4 (Borel-Cantelli) *(1)我们有*

$$\sum_{n=1}^{\infty}P(A_n)<+\infty\Longrightarrow P\left(A_n,i.o.\right)\triangleq P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)=0.$$

即A_n不可能发生无穷多次.

(2)若$\{A_n\}$相互独立, 则有

$$\sum_{n=1}^{\infty}P(A_n)=+\infty\Longrightarrow P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)=1.$$

即A_n必然发生无穷多次.

证明: (1)只需注意到

$$\begin{aligned}P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)&=\lim_{k\to\infty}P\left(\bigcup_{n=k}^{\infty}A_n\right)\qquad(\text{序列}\{\bigcup_{n=k}^{\infty}A_n\}_{k\ge1}\text{单调递减})\\&\le\lim_{k\to\infty}\sum_{n=k}^{\infty}P(A_n)\qquad(\text{次}\sigma\text{-可加性})\\&\le0.\qquad(\text{利用条件以及Cauchy准则})\end{aligned}$$

(2)欲证命题可以作如下转化：

$$
\begin{aligned}
P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)=1 &\Longleftrightarrow \lim_{k\to\infty}P\left(\bigcup_{n=k}^{\infty}A_n\right)=1\\
&\Longleftrightarrow P\left(\bigcup_{n=k}^{\infty}A_n\right)=1,\forall k\geq 1 \quad (\text{单调递减趋于}1,\text{ 只能为}1)\\
&\Longleftrightarrow P\left(\bigcap_{n=k}^{\infty}A_n^c\right)=0,\forall k\geq 1. \quad (\text{de Morgan})
\end{aligned}
$$

事实上,

$$
\begin{aligned}
0\leq P\left(\bigcap_{n=k}^{\infty}A_n^c\right) &= \lim_{m\to\infty}P\left(\bigcap_{n=k}^{m}A_n^c\right) \quad (\text{从下连续性})\\
&=\lim_{m\to\infty}\prod_{n=k}^{m}(1-P(A_n)) \quad (\text{独立性})\\
&\leq \lim_{m\to\infty}\prod_{n=k}^{m}(e^{-P(A_n)}) \quad (e^{-x}\geq -x+1)\\
&=\exp\left(-\lim_{m\to\infty}\sum_{n=k}^{m}P(A_n)\right)=0. \qquad \square
\end{aligned}
$$

推论 5.1.5 (Borel 0-1律) 若$\{A_n\}$相互独立, 则

$$
\begin{aligned}
\sum_{n=1}^{\infty}P(A_n)<\infty &\Longrightarrow P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)=0.\\
\sum_{n=1}^{\infty}P(A_n)=+\infty &\Longrightarrow P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)=1.
\end{aligned}
$$

例 5.1.2 设$\{X_n\}$相互独立, 则对任意$\varepsilon>0$, $X_n\xrightarrow{a.s.}0\Leftrightarrow\sum\limits_{n=1}^{+\infty}P(|X_n|\geq\varepsilon)<+\infty$.

证明: 取$A_n=[|X_n|\geq\varepsilon]$, 用Borel-Cantelli引理和a.s.收敛的等价刻画即可. $\square$

引理 5.1.6 (Chung-Erdös不等式) 设$\{A_k\}_{k\leq n}\subset\mathscr{F}$, 且$P\Big(\bigcup\limits_{k=1}^{n}A_k\Big)>0$, 则

$$
P\Big(\bigcup_{k=1}^{n}A_k\Big)\geq\frac{\Big(\sum\limits_{k=1}^{n}P(A_k)\Big)^2}{\sum\limits_{i,k=1}^{n}P(A_iA_k)}.
$$

证明: 只需注意到

$$
\begin{aligned}
\Big(\sum_{k=1}^{n}P(A_k)\Big)^2 &= \Big(\mathbb{E}\sum_{k=1}^{n}I_{A_k}\Big)^2=\Big(\mathbb{E}\sum_{k=1}^{n}I_{A_k}I_{\bigcup_{k=1}^{n}A_k}\Big)^2\\
&\leq \mathbb{E}\Big(\sum_{k=1}^{n}I_{A_k}\Big)^2\mathbb{E}\Big(I_{\bigcup_{k=1}^{n}A_k}\Big)^2=\mathbb{E}\sum_{i,k=1}^{n}I_{A_i}I_{A_k}P\Big(\bigcup_{k=1}^{n}A_k\Big)\\
&=\left(\sum_{i,k=1}^{n}P(A_iA_k)\right)P\Big(\bigcup_{k=1}^{n}A_k\Big). \qquad \square
\end{aligned}
$$

定理 5.1.7 (Kochen-Stone) 设$\{A_n\}\subset\mathscr{F}$, $\sum\limits_{n=1}^{\infty}P(A_n)=\infty$, 则

$$P(A_n,i.o.)\geq\limsup_{n\to\infty}\frac{\Big(\sum\limits_{k=1}^{n}P(A_k)\Big)^2}{\sum\limits_{i,k=1}^{n}P(A_iA_k)}=\limsup_{n\to\infty}\frac{\sum\limits_{1\leq i<k\leq n}P(A_i)P(A_k)}{\sum\limits_{1\leq i<k\leq n}P(A_iA_k)}.$$

证明: (1)设$a_n=\Big(\sum\limits_{k=1}^{n}P(A_k)\Big)^2, b_n=\sum\limits_{i,k=1}^{n}P(A_iA_k)$, 则$\lim\limits_{n\to\infty}a_n=\infty$. 由引理5.1.6, $\lim\limits_{n\to\infty}b_n=\infty$.

设$m\geq 1$. 由引理5.1.6, 以及$\sum\limits_{i,k=m+1}^{n}P(A_iA_k)\leq b_n-b_m$, 可得

$$P\Big(\bigcup_{k=m+1}^{\infty}A_k\Big)=\lim_{n\to\infty}P\Big(\bigcup_{k=m+1}^{\infty}A_k\Big)\geq\limsup_{n\to\infty}\frac{(\sqrt{a_n}-\sqrt{a_m})^2}{b_n-b_m}=\limsup_{n\to\infty}\frac{a_n}{b_n}.$$

由m的任意性与测度的从上连续性, 可得

$$P\Big(\bigcap_{m=1}^{\infty}\bigcup_{k=m+1}^{\infty}A_k\Big)=\lim_{m\to\infty}P\Big(\bigcup_{k=m+1}^{\infty}A_k\Big)=\limsup_{n\to\infty}\frac{a_n}{b_n}.$$

不等式证明完毕.

(2)由$a_n\to\infty$且

$$a_n=2\sum_{1\leq i<k\leq n}P(A_i)P(A_k)+\sum_{i=1}^{n}[P(A_i)]^2,$$

$$b_n=2\sum_{1\leq i<k\leq n}P(A_iA_k)+\sum_{i=1}^{n}P(A_i),$$

所以经过一系列放缩可得$\limsup\limits_{n\to\infty}\dfrac{a_n}{b_n}=\text{RHS}$. □

§ 5.2 条件期望

5.2.1 条件期望的定义

条件期望是现代概率论的基础.

回顾: 设$(\Omega,\mathscr{F},P)$是概率空间, 若$A\in\mathscr{F}, P(A)>0$, 定义$P(B|A)=\dfrac{P(AB)}{P(A)}, \forall B\in\mathscr{F}$.

则$P(\cdot|A):\mathscr{F}\to[0,1]$是个概率测度. $\mathbb{E}(\xi|A)=\int_\Omega\xi dP(\cdot|A)=\dfrac{\mathbb{E}(\xi I_A)}{P(A)}$.

最简单的情况: 设$\mathscr{C}=\{\Omega,\varnothing,A,A^c\}$, 满足$P(A)>0, P(A^c)>0$, 则定义条件期望$\mathbb{E}(\xi|\mathscr{C})\triangleq\mathbb{E}(\xi|A)I_A+\mathbb{E}(\xi|A^c)I_{A^c}$, 注意$I_A, I_{A^c}$都是r.v., 因此$\mathbb{E}(\xi|\mathscr{C})$也是r.v.

推广最简单的情况: 设$\{A_n\}_{n\geq1}\subset\mathscr{F}$是$\Omega$的划分, 即满足$\bigcup\limits_n A_n=\Omega$且诸$A_i$两两不交, $P(A_n)>0$, 记$\mathscr{C}=\sigma(\{A_n\}_{n\geq1})$为包含$\{A_n\}$的最小$\sigma-$代数, 定义$\mathbb{E}(\xi|\mathscr{C})\triangleq\sum\limits_{n=1}^{\infty}\mathbb{E}(\xi|A_n)I_{A_n}$,

满足: ①$\mathbb{E}(\xi|\mathscr{C})$关于$\mathscr{C}$可测, ②$\forall B\in\mathscr{C}$, $\int_B\mathbb{E}(\xi|\mathscr{C})dP=\int_B\xi dP$.(重要性质)

推广到更一般的sigma代数:

定义 5.2.1 (条件期望) 设$\mathscr{C} \subset \mathscr{F}$为$\sigma-$代数, ξ是可积$r.v.$, 称(关于)$\mathscr{C}$可测的$r.v.\eta$为ξ关于$\mathscr{C}$的条件期望, 若

$$\int_B \xi dP = \int_B \eta dP, \forall B \in \mathscr{C}.$$

此时记$\eta = \mathbb{E}(\xi|\mathscr{C})$.

注: ①存在性: 令$\nu(B) = \displaystyle\int_B \xi \mathrm{d}P, \forall B \in \mathscr{C}$, 则$\nu \ll P|_{\mathscr{C}}$, 故存在唯一的关于$\mathscr{C}$可测函数$g$使得$\nu = g.P|_{\mathscr{C}}$, 即$\nu(B) = \displaystyle\int_B g\mathrm{d}P|_{\mathscr{C}}$(Radon-Nikodym定理). 所以$\displaystyle\int_B \xi \mathrm{d}P = \int_B g\mathrm{d}P|_{\mathscr{C}} = \int_B g\mathrm{d}P$, 这里的$g$满足条件期望定义, g是关于P的R-N导数.

②唯一性成立(在$P|_{\mathscr{C}}$几乎处处定义, 即η, ξ差一个关于$\mathscr{C}$可测的零测集都成立)

③如果$E\xi$存在, 则ξ关于$\mathscr{C}$的$\sigma-$代数的条件期望存在(可测).

性质 5.2.1 若$\mathscr{C} = \{\varnothing, \Omega\}$, 则$\mathbb{E}(\xi|\mathscr{C}) = \mathbb{E}\xi$, $a.s.$.

性质 5.2.2 $\mathbb{E}(\mathbb{E}(\xi|\mathscr{C})) = E\xi$.

证明: 取$B = \Omega$即可. □

性质 5.2.3 若ξ关于$\mathscr{C}$可测, 则$\mathbb{E}(\xi|\mathscr{C}) = \xi$ $a.s.$, 即上面的唯一性.

性质 5.2.4 若$c_1, c_2 \in \mathbb{R}$, 且$c_1\xi_1 + c_2\xi_2, \xi_1, \xi_2$期望都存在, 则(期望存在意味着期望的正部或者负部都存在)

$$\mathbb{E}(c_1\xi_1 + c_2\xi_2|\mathscr{C}) = c_1\mathbb{E}(\xi_1|\mathscr{C}) + c_2\mathbb{E}(\xi_2|\mathscr{C}), a.s.$$

性质 5.2.5 (单调性) 若$X \geq Y$ $a.s.$(关于$\mathscr{F}$), 则$\mathbb{E}(X|\mathscr{C}) \geq \mathbb{E}(Y|\mathscr{C})$ $a.s.$(关于$P|_{\mathscr{C}}$).

证明: $\forall B \in \mathscr{C}, X \geq Y \Rightarrow \cdots \Rightarrow \displaystyle\int_B \mathbb{E}(X|\mathscr{C})dP \geq \int_B \mathbb{E}(Y|\mathscr{C})dP$. □

性质 5.2.6 $|\mathbb{E}(\xi|\mathscr{C})| \leq \mathbb{E}(|\xi||\mathscr{C})$ $a.s.$.

证明: 利用单调性.

性质 5.2.7 设$\{\xi_n\}$为非负$r.v.$列, 且$\xi_n \leq \xi_{n-1}$ $a.s.$, 则$\lim\limits_{n\to\infty} \mathbb{E}(\xi_n|\mathscr{C}) = E\left(\lim\limits_{n\to\infty} \xi_n \middle| \mathscr{C}\right)$, $a.s.$.

证明: 利用单调收敛定理. □

性质 5.2.8 (硬性质, 不能忘) 设$\xi, \xi\eta$的期望均存在且η关于$\mathscr{C}$可测, 则

$$\mathbb{E}(\xi\eta|\mathscr{C}) = \eta\mathbb{E}(\xi|\mathscr{C}) a.s.$$

(让$\xi = 1$可推出性质②)

证明: 证明思路: 证r.v.成立, 先证对示性函数成立, 再证对非负简单函数成立. 回顾$\mathbb{E}f(\xi) = \displaystyle\int_\Omega f(x)dF(x)$的证明过程.

先设$\eta = I_A, A \in \mathscr{C}$, 即证$\mathbb{E}(\xi I_A|\mathscr{C}) = I_A\mathbb{E}(\xi|\mathscr{C})$ a.s.成立. 由于$I_A\mathbb{E}(\xi|\mathscr{C})$是可测的, (这是因为$I_A$与$\mathbb{E}(\xi|\mathscr{C})$都关于$\mathscr{C}$可测, 它们的乘积也可测),

$\forall B \in \mathscr{C}$, 由于$\int_B \xi I_A dP = \int_{A\cap B} \xi dP = \int_{A\cap B} \mathbb{E}(\xi|\mathscr{C})dP = \int_B I_A \mathbb{E}(\xi|\mathscr{C})dP$,
根据性质②, 则$\mathbb{E}(\xi I_A|\mathscr{C}) = I_A\mathbb{E}(\xi|\mathscr{C})$ a.s.成立.

下面再证非负简单的情形$\eta = \sum_{i=1}^n a_i I_{A_i}, a_i \geq 0$, 且$A_i \in \mathscr{C}$两两不交, 用线性性(性质③)可知这是成立的. 然后证η是非负且关于$\mathscr{C}$可测成立, 用非负简单可测r.v.列$\{\eta_n\}$逼近即可(性质⑥). 对于一般的η, 记为$\eta = \eta^+ - \eta^-$(正部与负部)就OK了. □

性质 5.2.9 *若$\xi, \mathscr{C}$独立, 则$\mathbb{E}(\xi|\mathscr{C}) = E\xi$, a.s..*

证明: $\forall B \in \mathscr{C}$,

$$\begin{aligned}\int_B \mathbb{E}(\xi|\mathscr{C})dP &= \mathbb{E}(\mathbb{E}(\xi|\mathscr{C})I_B) = \mathbb{E}(\mathbb{E}(\xi I_B|\mathscr{C})) \qquad (\text{B是可测的})\\ &= \mathbb{E}(\xi I_B) \qquad (\xi, I_B\text{独立})\\ &= E\xi E I_B = E\xi \int_B 1dP = \int_B E\xi dP.\end{aligned}$$

这里$E\xi$是常数, 当然是可测的, 根据性质②, $\mathbb{E}(\xi|\mathscr{C}) = E\xi$. □

性质 5.2.10 (平滑性) *设$\mathscr{C}_1, \mathscr{C}_2 \subset \mathscr{F}$是$\sigma$-代数, 且$\mathscr{C}_1 \subset \mathscr{C}_2$, 则*

$$\mathbb{E}(\xi|\mathscr{C}_1) = E[\mathbb{E}(\xi|\mathscr{C}_2)|\mathscr{C}_1], a.s..$$
$$\mathbb{E}(\xi|\mathscr{C}_1) = E[\mathbb{E}(\xi|\mathscr{C}_1)|\mathscr{C}_2], a.s..$$

证明: 只证第一条. $\forall A \in \mathscr{C}_1$, 根据定义有

$$\int_A E[\mathbb{E}(\xi|\mathscr{C}_2)|\mathscr{C}_1]\mathrm{d}P = \int_A \mathbb{E}(\xi|\mathscr{C}_2)\mathrm{d}P = \int_A \xi \mathrm{d}P = \int_A \mathbb{E}(\xi|\mathscr{C}_1)\mathrm{d}P.$$

注: 这个性质是很好的. 根据证明过程不难知道, 类似可以证明如果$\mathscr{C}_2 \subset \mathscr{C}_1$, 则$\mathbb{E}(\xi|\mathscr{C}_2) = E[\mathbb{E}(\xi|\mathscr{C}_2)|\mathscr{C}_1]$, a.s.. 因此让$\xi$对$\mathscr{C}_1, \mathscr{C}_2$分别取条件期望(无论顺序), 得到的一定是较小sigma-代数的条件期望.

例 5.2.1 对$\Omega = \{a, b, c\}$, 给一个例子使得$\mathbb{E}(\mathbb{E}(X|\mathscr{C}_1)|\mathscr{C}_2) \neq \mathbb{E}(\mathbb{E}(X|\mathscr{C}_2)|\mathscr{C}_1)$.

下面把实变函数的部分定理推广到条件期望上来.

定理 5.2.1 (控制收敛定理) *设$\{X_n\}_{n\geq 1}$是r.v.序列, ξ是可积r.v.. 若$X_n \xrightarrow{a.s.} X$或$X_n \xrightarrow{P} X$, 且$|X_n| \leq \xi, \forall n \geq 1$, a.s., 则$\lim_{n\to\infty} \mathbb{E}(X_n|\mathscr{C}) = \mathbb{E}(\lim_{n\to\infty} X_n|\mathscr{C}) = \mathbb{E}(X|\mathscr{C})$.*

定理 5.2.2 (Fatou) *设$\{X_n\}$是r.v.序列, 且$EX_n (n = 1, 2, \cdots)$存在.*

(1)若存在r.v. Y, 使得$EY > -\infty$, 且对每个$n \geq 1$有$X_n \geq Y$, a.s., 则$\liminf_{n\to\infty} X_n$的期望存在, 且满足

$$\mathbb{E}(\liminf_{n\to\infty} X_n|\mathscr{C}) \leq \liminf_{n\to\infty} \mathbb{E}(X_n|\mathscr{C}).$$

(2)若存在r.v. Y, 使得$EY < +\infty$, 且对每个$n \geq 1$有$X_n \leq Y$, a.s., 则$\limsup_{n\to\infty} X_n$的期望存在, 且满足

$$\mathbb{E}(\limsup_{n\to\infty} X_n|\mathscr{C}) \geq \limsup_{n\to\infty} \mathbb{E}(X_n|\mathscr{C}).$$

定理 5.2.3 (Hölder不等式) $\forall p,q>1$满足$\dfrac{1}{p}+\dfrac{1}{q}=1$, 则

$$\mathbb{E}(|\xi\eta||\mathscr{C})\le \mathbb{E}\left(|\xi|^p|\mathscr{C}\right)^{\frac{1}{p}}\mathbb{E}\left(|\xi|^q|\mathscr{C}\right)^{\frac{1}{q}}.$$

注: 当$p=q=2$时为Cauchy-Schwarz不等式.

定理 5.2.4 (Jensen不等式) 设$f:\mathbb{R}\to\mathbb{R}$是连续凸函数,$r.v.\xi$满足$f(\xi)$积分存在, 则

$$f(E\xi)\le Ef(\xi), f(\mathbb{E}(\xi|\mathscr{C}))\le \mathbb{E}(f(\xi)|\mathscr{C}), a.s..$$

注: 特别地取$f(x)=|x|$显然成立(便于记忆). 取$f(x)=x^2$恰好是Cauchy-Schwarz不等式.

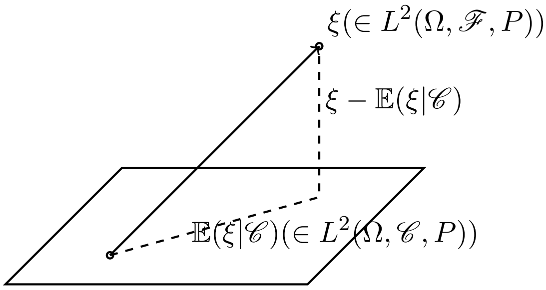
5.2.2 条件期望的几何意义

记

$$\begin{aligned}
&L^2(\Omega,\mathscr{F},P)=\{\xi:\xi\text{为}\mathscr{F}\text{中的可测r.v., 且}\mathbb{E}\xi^2<+\infty\},\\
&L^2(\Omega,\mathscr{C},P)=\{\xi:\xi\text{为}\mathscr{C}\text{中的可测r.v., 且}\mathbb{E}\xi^2<+\infty\}.\\
&\langle\xi,\eta\rangle_{L^2}=\mathbb{E}(\xi\eta).\\
&\|\xi\|_{L^2}=\langle\xi,\xi\rangle_{L^2}^{\frac{1}{2}}=(E\xi^2)^{\frac{1}{2}}.
\end{aligned}$$

则$L^2(\Omega,\mathscr{F},P)$和内积$\langle\cdot,\cdot\rangle_{L^2}$一起构成Hilbert空间, $L^2(\Omega,\mathscr{C},P)$是$L^2(\Omega,\mathscr{F},P)$的闭子空间(要证一下), $\mathbb{E}(\cdot|\mathscr{C})$是$L^2(\Omega,\mathscr{F},P)$到$L^2(\Omega,\mathscr{C},P)$的正交投影算子.

定理 5.2.5 $\forall\xi\in L^2(\Omega,\mathscr{F},P)$, 可对$\xi$做如下的正交分解: $\xi=\mathbb{E}(\xi|\mathscr{C})+(\xi-\mathbb{E}(\xi|\mathscr{C}))$.



注: $\mathbb{E}(\xi|\mathscr{C})\in L^2(\Omega,\mathscr{C},P)$是因为$\mathbb{E}(\xi|\mathscr{C})$可测, 且$\mathbb{E}(\mathbb{E}(\xi|\mathscr{C})^2)\le\mathbb{E}(\mathbb{E}(\xi^2|\mathscr{C}))=E\xi^2<+\infty$(Jensen不等式).

证明: 先验证垂直, 再证距离最短.

$$\begin{aligned}
\langle\mathbb{E}(\xi|\mathscr{C}),\xi-\mathbb{E}(\xi|\mathscr{C})\rangle_{L^2}=&\mathbb{E}\left[\mathbb{E}(\xi|\mathscr{C})\cdot(\xi-\mathbb{E}(\xi|\mathscr{C}))\right]\\
=&\mathbb{E}\left\{\mathbb{E}\left[\mathbb{E}(\xi|\mathscr{C})\cdot(\xi-\mathbb{E}(\xi|\mathscr{C}))|\mathscr{C}\right]\right\}\text{【性质①】}\\
=&\mathbb{E}\left\{\mathbb{E}(\xi|\mathscr{C})\mathbb{E}\left[\xi-\mathbb{E}(\xi|\mathscr{C})|\mathscr{C}\right]\right\}\text{【}\mathbb{E}(\xi|\mathscr{C})\text{可测, 性质⑦】}\\
=&\mathbb{E}\left\{\mathbb{E}(\xi|\mathscr{C})\left[\mathbb{E}(\xi|\mathscr{C})-\mathbb{E}(\xi|\mathscr{C})\right]\right\}\text{【性质③(线性性), 性质②】}\\
=&0.
\end{aligned}$$

下面验证距离最短: 即验证

$$\mathbb{E}(\xi-\mathbb{E}(\xi|\mathscr{C}))^2=\inf_{y\in L^2(\Omega,\mathscr{C},P)}\{\mathbb{E}(\xi-y)^2|y\in L^2(\Omega,\mathscr{C},P)\}.$$

事实上, $\forall y \in L^2(\Omega, \mathscr{C}, P)$,

$$\begin{aligned}
\mathbb{E}(\xi - \mathbb{E}(\xi|\mathscr{C}))^2 &= \mathbb{E}(\xi - \mathbb{E}(\xi|\mathscr{C}) - y + y)^2 \\
&= \mathbb{E}(\xi - y)^2 - 2\mathbb{E}[(\xi - y)(\mathbb{E}(\xi|\mathscr{C}) - y)] + E[\mathbb{E}(\xi|\mathscr{C}) - y]^2 \\
&= \mathbb{E}(\xi - y)^2 - 2\mathbb{E}\{\mathbb{E}[(\xi - y)(\mathbb{E}(\xi|\mathscr{C}) - y)|\mathscr{C}]\} + E[\mathbb{E}(\xi|\mathscr{C}) - y]^2 \text{【性质①】} \\
&= \mathbb{E}(\xi - y)^2 - E[\mathbb{E}(\xi|\mathscr{C}) - y]^2 \text{【}\mathbb{E}(\xi|\mathscr{C}) - y\text{可测, 性质⑦】} \\
&\le \mathbb{E}(\xi - y)^2.
\end{aligned}$$

根据y的任意性, 对上式取下确界即可. □

定理 5.2.6 *设$(\Omega, \mathscr{F}, P)$是概率空间, $\mathscr{C} \subset \mathscr{F}$是$\sigma-$代数. $(G, \mathscr{G}), (E, \mathscr{E})$是可测空间. 令$X : \Omega \to G$是$\mathscr{C}/\mathscr{G}$可测, $Y : \Omega \to E$是$\mathscr{F}/\mathscr{E}$可测. 若Y与$\mathscr{C}$独立, $g : G \times E \to \mathbb{R}$可测, 且$\mathbb{E}|g(X, Y)| < +\infty$, 则*

$$\mathbb{E}[g(X, Y)|\mathscr{C}] = \mathbb{E}(g(x, Y))|_{x=X}$$

证明: 设$f(x) = \mathbb{E}g(x, Y), x \in G$, 下证$\mathbb{E}(g(X, Y)|\mathscr{G}) = f(X)$, a.s.. 只需证$f(X)$关于$\mathscr{C}$可测(即$f$为$\mathscr{G}$-可测), 且$\int_A f(X)\mathrm{d}P = \int_A g(X, Y)\mathrm{d}P, \forall A \in \mathscr{C}$.

(1)下证f为$\mathscr{G}$-可测: 等价于$\forall B \in \mathscr{B}(\mathbb{R}), f^{-1}(B) \in \mathscr{G}$. 注意$\mathscr{G} \times \mathscr{E} = \sigma(A \times B | A \in \mathscr{G}, B \in \mathscr{C})$, 所以只需证矩形集的情形, 再用单调类定理即可. (步骤略). 所以$f(X)$是$\mathscr{C}$-可测.

(2)下证对任一$\mathscr{C}$-可测r.v.Z, 有$\mathbb{E}(g(X, Y)Z) = \mathbb{E}(f(X)Z)$. (等价于$\int_\Omega f(X)\mathrm{d}P = \int_\Omega g(X, Y)\mathrm{d}P$.) 注意$Z$有界且$g(X, Y)$可积, 所以$\mathbb{E}(g(X, Y)Z)$积分存在, 且

$$\begin{aligned}
\mathbb{E}(g(X, Y)Z) &= \int_{G \times E \times \mathbb{R}} g(x, y)z\mathrm{d}F_{(X,Y,Z)}(x, y, z) \\
&= \int_{G \times \mathbb{R}} \int_E g(x, y)z\mathrm{d}F_Y(y)\mathrm{d}F_{(X,Z)}(x, z) \qquad \text{(独立性)} \\
&= \int_{G \times \mathbb{R}} \int_E g(x, y)\mathrm{d}F_Y(y)z\mathrm{d}F_{(X,Z)}(x, z) \qquad \text{(Fubini)} \\
&= \int_{G \times \mathbb{R}} f(x)z\mathrm{d}F_{(X,Z)}(x, z) = \mathbb{E}(f(X)Z). \qquad \square
\end{aligned}$$

注: 取$g(x, y) = g_1(x)g_2(y)$, 则

$$\mathbb{E}[g_1(X)g_2(Y)|\mathscr{C}] = g_1(X)\mathbb{E}[g_2(Y)|\mathscr{C}] = g_1(X)\mathbb{E}g_2(Y) = g_1(x)\mathbb{E}g_2(Y)|_{x=X},$$

所以这个定理推广了条件期望的两条性质.

定理 5.2.7 *设$(\Omega, \mathscr{F}, P)$是概率空间, Q是概率测度, 且$Q \ll P$. 设$\mathbb{C}$是$\mathscr{F}$的子σ-代数, 则:*

(1)$Q\left(\mathbb{E}\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right) > 0\right) = 1$.

(2)*若X关于Q期望存在, 则*$\mathbb{E}_Q(X|\mathscr{C}) = \dfrac{\mathbb{E}\left(X\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)}{\mathbb{E}\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)}$. *(默认$\mathbb{E}$是关于$P$积分的期望.)*

证明: (1)注意到[①]

$$Q\left(\mathbb{E}\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)>0\right)=\int_\Omega I_{[\mathbb{E}(\frac{\mathrm{d}Q}{\mathrm{d}P}|\mathscr{C})>0]}\underbrace{\frac{\mathrm{d}Q}{\mathrm{d}P}}_{\text{是}r.v.}\mathrm{d}P=\int_\Omega I_{[\mathbb{E}(\frac{\mathrm{d}Q}{\mathrm{d}P}|\mathscr{C})>0]}\mathbb{E}(\frac{\mathrm{d}Q}{\mathrm{d}P}|\mathscr{C})\mathrm{d}P$$
$$=\int_\Omega \mathbb{E}(\frac{\mathrm{d}Q}{\mathrm{d}P}|\mathscr{C})\mathrm{d}P=\int_\Omega\frac{\mathrm{d}Q}{\mathrm{d}P}\mathrm{d}P=Q(\Omega)=1.$$

(2)下证$\mathbb{E}_Q(X|\mathscr{C})\mathbb{E}\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)=\mathbb{E}\left(X\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)$. 由条件期望的定义, 即证

$$\int_B \mathbb{E}_Q(X|\mathscr{C})\mathbb{E}\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)\mathrm{d}P=\int_B\mathbb{E}\left(X\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)\mathrm{d}P,\forall B\in\mathscr{C}.$$

事实上,

$$\begin{aligned}\mathrm{LHS}&=\int_\Omega\underbrace{I_B\mathbb{E}_Q(X|\mathscr{C})}_{\mathscr{C}\text{可测}}\mathbb{E}\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)\mathrm{d}P=\int_\Omega I_B\mathbb{E}_Q(X|\mathscr{C})\frac{\mathrm{d}Q}{\mathrm{d}P}\mathrm{d}P \qquad (\text{性质}5.2.8)\\&=\int_\Omega\mathbb{E}_Q(X|\mathscr{C})I_B\mathrm{d}Q=\int_\Omega XI_B\mathrm{d}Q \qquad (\text{性质}5.2.8)\\&=\int_B X\frac{\mathrm{d}Q}{\mathrm{d}P}\mathrm{d}P=\int_B\mathbb{E}\left(X\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathscr{C}\right)\mathrm{d}P \qquad (\text{条件期望的定义}) \qquad \square\end{aligned}$$

下设X积分存在, Y是r.v., 则$\mathbb{E}(X|Y)$关于$\sigma(Y)$可测. 由定理2.2.2, 存在可测函数$h:\mathbb{R}\to\mathbb{R}$使得$\mathbb{E}(X|Y)=h(Y)$. 若$h_1,h_2$满足此$h$, 则$P\circ Y^{-1}(h_1\neq h_2)=0$(条件期望的唯一性), 所以$h$仅在$Y$处a.s.唯一.

设$\mu(A)=P\circ Y^{-1}(A),\nu(A)=\mathbb{E}[\xi I_{Y^{-1}(A)}]$, 其中$A\in\mathscr{B}(\mathbb{R})$, 则$\nu$是符号测度且$\nu\ll\mu$, 从而$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$存在且$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}=h(\mu\text{-a.e.})\Leftrightarrow\forall A\in\mathscr{B}(\mathbb{R})$, 有

$$\begin{aligned}\nu(A)&=\int_A\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(y)\mu(\mathrm{d}y)=\int_A h(y)\mu(\mathrm{d}y)=\int_{\mathbb{R}}h(y)I_A(y)\mu(\mathrm{d}y)\\&=\mathbb{E}(h(Y)I_A(Y))=\mathbb{E}(\mathbb{E}(X|Y)I_A(Y))=\mathbb{E}XI_A(Y).\end{aligned}$$

注意$\sigma(Y)=\{Y^{-1}(A)|A\in\mathscr{B}(\mathbb{R})\}$, 则$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}=\mathbb{E}(X|Y)=h$, a.e.. (差$\sigma(Y)$中的零测集)

5.2.3 条件独立

定义 5.2.2 设$(\Omega,\mathscr{F},P)$是概率空间, $\mathscr{C}$是子σ-代数.
(1)称$P(B|\mathscr{C})\triangleq\mathbb{E}(I_B|\mathscr{C})$是$B$关于$\mathscr{C}$的**条件概率**.
(2)设$A,B\in\mathscr{F}$. 称A与B关于$\mathscr{C}$**条件独立**, 若$P(AB|\mathscr{C})=P(A|\mathscr{C})P(B|\mathscr{C})$.
(3)设$\mathscr{C}_1,\mathscr{C}_2\subset\mathscr{F}$. 称$\mathscr{C}_1,\mathscr{C}_2$关于$\mathscr{C}$**条件独立**, 若$P(AB|\mathscr{C})=P(A|\mathscr{C})P(B|\mathscr{C}),\forall A\in\mathscr{C}_1,B\in\mathscr{C}_2$.

命题 5.2.8 设$\mathscr{C}_1,\mathscr{C}_2$是$\mathscr{F}$的子$\sigma$-代数, 则$\mathscr{C}_1,\mathscr{C}_2$关于$\mathscr{C}$条件独立等价于对任意$B_2\in\mathscr{C}_2$, 有

$$P(B_2|\mathscr{C})=P(B_2|\mathscr{C}_1\vee\mathscr{C}),a.s.. \tag{5.1}$$

[①]若μ,ν是测度, 则$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\geq 0$, μ-a.s.. 要不然, 记$A=\left[\frac{\mathrm{d}\nu}{\mathrm{d}\mu}<0\right]$, 若$\mu(A)>0$, 则$\nu(A)=\int_A\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu<0$, 这与$\nu$非负矛盾.

注： (1)类比下述事实：若A,B独立，则$P(A|B)=P(A)$. (2)可以用单调类定理证明$\mathscr{C}_1\vee\mathscr{C}\triangleq\sigma(\mathscr{C}_1\cup\mathscr{C})=\sigma(\{A\cap B|A\in\mathscr{C}_1,B\in\mathscr{C}\})$.

证明： (5.1)式等价于对任意$B\in\mathscr{C},B_1\in\mathscr{C}_1$, 有$\displaystyle\int_{B\cap B_1}P(B_2|\mathscr{C})\mathrm{d}P=\int_B I_{B_1}I_{B_2}\mathrm{d}P$.

而$\mathscr{C}_1,\mathscr{C}_2$关于$\mathscr{C}$条件独立$\Leftrightarrow\forall B_1\in\mathscr{C}_1,B_2\in\mathscr{C}_2$, $\displaystyle\int_B P(B_1|\mathscr{C})P(B_2|\mathscr{C})\mathrm{d}P=\int_B I_{B_1\cap B_2}\mathrm{d}P,B\in\mathscr{C}$.

但是对$B\in\mathscr{C},B_1\in\mathscr{C}_1,B_2\in\mathscr{C}_2$, 有

$$\begin{aligned}\int_{B\cap B_1}P(B_2|\mathscr{C})\mathrm{d}P&=\int_B I_{B_1}P(B_1|\mathscr{C})\mathrm{d}P\\&=\int_B\mathbb{E}(I_{B_1}P(B_1|\mathscr{C})|\mathscr{C})\mathrm{d}P=\int_B P(B_1|\mathscr{C})P(B_2|\mathscr{C})\mathrm{d}P\end{aligned}$$

所以欲证命题成立. □

例 5.2.2 设$(\Omega,\mathscr{F},P)$为概率空间, $\mathscr{C}\subset\mathscr{F}$为子$\sigma$-代数, $\Gamma\in\mathscr{F}$是事件, 证明以下等价:
(1)$\Gamma,\mathscr{C}$独立;
*(2)*任一概率测度Q *on* $(\Omega,\mathscr{F})$, Q与P等价, 且$\dfrac{dQ}{dP}$为$\mathscr{C}$可测, 则$Q(\Gamma)=P(\Gamma)$.

证明: “(1)⇒(2)”: $Q(A)=\int_A\dfrac{dQ}{dP}(w)P(dw)$, 这里$\dfrac{dQ}{dP}(w)$是个r.v.. 记$E_Q$是关于$Q$积分的期望, E_P是关于P积分的期望.

$\Gamma,\mathscr{C}$独立$\Leftrightarrow P(\Gamma\cap B)=P(\Gamma)P(B),\forall B\in\mathscr{C}$, 而

$$\begin{aligned}Q(\Gamma)=\int_\Gamma dQ=\int_\Omega I_\Gamma dQ=\int_\Omega I_\Gamma\frac{dQ}{dP}dP\ \text{【形式上的转化】}\\ &=E_P\left(I_\Gamma\frac{dQ}{dP}\right)\\ &=E_P(I_\Gamma)E_P\left(\frac{dQ}{dP}\right)\ \text{【独立性】}\\ &=P(\Gamma)\int_\Omega\frac{dQ}{dP}dP=P(\Gamma)\int_\Omega dQ=P(\Gamma).\end{aligned}$$

“(1)⇐(2)”: 要证$P(\Gamma\cap B)=P(\Gamma)P(B),\forall B\in\mathscr{C}$. 事实上, 令$\boxed{dQ=\dfrac{I_B+1}{P(B)+1}dP}$(+1为了保证分子分母不为0, 除以$(P(B)+1)$这一常数是为了归一化). 下面验证$Q$是概率测度: 根据定义验证.

1°非负性: $Q(A)=\int_A\dfrac{I_B+1}{P(B)+1}=\int_\Omega I_A\dfrac{I_B+1}{P(B)+1}dP\geq0$, 这里$I_A\geq0$, $\dfrac{I_B+1}{P(B)+1}>0$.
2°可列可加性: $\forall\{A_n\}_{n\geq1}\subset\mathscr{F},A_n\cap A_m=\varnothing,\forall n\neq m$, 我们有:

$$\begin{aligned}Q\left(\sum_n A_n\right)&=\int_{\sum_n A_n}\frac{I_B+1}{P(B)+1}dP=\int_\Omega I_{\sum_n A_n}\frac{I_B+1}{P(B)+1}dP\\ &=\int_\Omega\sum_n\left(I_{A_n}\frac{I_B+1}{P(B)+1}\right)dP\text{【两两不交】}\\ &=\sum_n\int_{A_n}\frac{I_B+1}{P(B)+1}dP\text{【非负, 积分与求和可调换次序】}\\ &=\sum_n Q(A_n).\end{aligned}$$

3°规范性: $Q(\Omega)=\int_\Omega\dfrac{I_B+1}{P(B)+1}dP=\dfrac{1}{P(B)+1}(P(B)+1)=1$.
根据Radon-Nikodym定理, 因此$\dfrac{dQ}{dP}$是$\mathscr{C}$可测的, 根据条件,

$$\begin{aligned}Q(\Gamma)=\int_\Gamma\frac{I_B+1}{P(B)+1}dP=P(\Gamma)&\Longleftrightarrow\frac{\int_\Omega I_\Gamma(I_B+1)dP}{P(B)+1}=P(\Gamma)\\ &\Longleftrightarrow\int_\Omega I_{\Gamma\cap B}dP+P(\Gamma)=P(\Gamma)(P(B)+1)\\ &\Longleftrightarrow P(\Gamma\cap B)=P(\Gamma)P(B),\forall B\in\mathscr{C}.\end{aligned}$$

证完. □

例 5.2.3 设X,Y,Z是$r.v.$且Y可积, 证明若(X,Y)与Z独立, 则$\mathbb{E}(Y|X,Z)=\mathbb{E}(Y|X)$.

注: $\mathbb{E}(Y|X_1,X_2)$表示关于由X_1,X_2生成的sigma-代数$\sigma(X_1,X_2)=\sigma(\sigma(X_1)\cup\sigma(X_2))$的条件期望. (X,Y)与Z独立指$\sigma(Z),\sigma(X,Y)$独立.

证明: 只需证

$$\begin{aligned}
&\int_A YdP=\int_A \mathbb{E}(Y|X)dP, \forall A\in\sigma(X,Z)\\
\Longleftrightarrow &\int_A YdP=\int_A \mathbb{E}(Y|X)dP, \forall A\in\sigma(X)\cup\sigma(Z) \text{【单调类定理】}\\
&\qquad\text{【不需对所有都进行验证, 只需要看子类,}\\
&\qquad\qquad \sigma(X)\cup\sigma(Z)=\{X^{-1}(B),Z^{-1}(C):B,C\in\mathscr{B}(\mathbb{R})\}\text{】}\\
\Longleftrightarrow &\int_\Omega I_{X^{-1}(B)}I_{Z^{-1}(C)}YdP=\int_\Omega \mathbb{E}(Y|X)I_{X^{-1}(B)}I_{Z^{-1}(C)}dP, \forall B,C\in\mathscr{B}(\mathbb{R})\\
&\qquad\text{【}I_{X^{-1}(B)}\text{即}I_B(X)\text{】}
\end{aligned}$$

事实上,

$$\begin{aligned}
\int_\Omega I_B(X)I_C(Z)YdP&=P(Z\in C)\int_\Omega I_B(X)YdP \text{【}X,Z\text{独立】}\\
&=P(Z\in C)\mathbb{E}(I_B(X)Y)\\
&=P(Z\in C)\mathbb{E}(\mathbb{E}(I_B(X)Y|X)) \text{【条件期望的期望=无条件期望】}\\
&=EI_C(Z)\mathbb{E}(I_B(X)\mathbb{E}(Y|X)) \text{【}I_B(X)\text{关于}X\text{可测】}\\
&=\mathbb{E}(I_B(X)\mathbb{E}(Y|X)I_C(Z)) \text{【独立性】}\\
&=\int_\Omega \mathbb{E}(Y|X)I_B(X)I_C(Z)dP. \qquad\square
\end{aligned}$$

例 5.2.4 设一列$r.v.\{X_n\}$依分布收敛于一个$r.v.\ X$, 记$\{N_t\}_{t\geq 0}$是一列正整数$r.v.$集合, 与$\{X_n\}$独立且依概率收敛为$\infty(t\to\infty)$. 证明: $X_{N_t}\xrightarrow{d}X,(t\to\infty)$.

证明: 固定$c\in\mathbb{R}$, 记$a_n=Ee^{icX_n}, a=Ee^{icx}$, 由于$X_n\xrightarrow{d}X,(t\to\infty)$, 则$a_n\to a(n\to\infty)$. (依概率收敛与特征函数收敛是一一对应的!)

即$\forall\varepsilon>0,\exists M\in\mathbb{N}$, 使得当$n\geq M$时, 有$|a_n-a|\leq\varepsilon$. 因此

$$[|a_{N_t}-a|\leq\varepsilon]\supset[N_t\geq M],\Longleftrightarrow[|a_{N_t}-a|>\varepsilon]\subset[N_t<M].$$

则$P(|a_{N_t}-a|>\varepsilon)\leq P(N_t<M)\to 0(t\to\infty)$, 则$a_{N_t}\xrightarrow{d}a,(t\to\infty)$.

下面用条件期望: 注意到(把a_{N_t}看作关于r.v.N_t的随机函数. 根据后面“注”的定理,

$$\begin{aligned}
Ea_{N_t}&=E[\mathbb{E}(a_{N_t}|N_t)] \text{【取条件期望】}\\
&=E[(Ea_n)|_{n=N_t}]=Ea_{N_t} \text{【“注”的定理】}
\end{aligned}\tag{5.2}$$

则$Ea_{N_t}\to a=Ea$. 又由于$|a_{N_t}|\leq 1$, 根据控制收敛定理, $a_{N_t}\xrightarrow{L^1}a$, 特别地$Ea_{N_t}\to a(t\to\infty)$, 则$X_{N_t}\xrightarrow{d}X$.

§ 5.3　随机变量族的一致可积性

定义 5.3.1 设$(\Omega,\mathscr{F},P)$是概率空间, $\mathcal{H}$是一族随机变量. 若

$$\lim_{c\to\infty}\sup_{\xi\in\mathcal{H}}\mathbb{E}(|\xi|I_{|\xi|\geq c})=0,$$

则称$\mathcal{H}$是一致可积的.

例 5.3.1 (1)若存在$\eta\in L^1(\Omega,\mathscr{F},P)$, 使得$|\xi|\leq\eta,\forall\xi\in\mathcal{H}$, 则$\mathcal{H}$是一致可积的.
(2)若$\{\xi_1,\cdots,\xi_n\}\subset L^1(\Omega,\mathscr{F},P)$, 则$\{\xi_1,\cdots,\xi_n\}$是一致可积的.
(3)若$\mathcal{H}_1$是一致可积族, $\mathcal{H}_2$满足对任意$\xi\in\mathcal{H}_2$, 存在$\eta\in\mathcal{H}_1$, 使得$|\xi|\leq|\eta|$, 则$\mathcal{H}_2$是一致可积族.

下面的定理像Ascoli-Arzela定理.

定理 5.3.1 设$\mathcal{H}\subset L^1(\Omega,\mathscr{F},P)$, 则

$$\mathcal{H}\text{是一致可积族}\Leftrightarrow\begin{cases}L^1\text{有界: }\sup\limits_{\xi\in\mathcal{H}}\mathbb{E}|\xi|<\infty,\\ \text{一致绝对连续: }\forall\varepsilon>0,\exists\delta>0,\text{若}P(A)\leq\delta,\text{则}\sup\limits_{\xi\in\mathcal{H}}\displaystyle\int_A|\xi|\mathrm{d}P\leq\varepsilon.\end{cases}$$

证明: “$\Rightarrow$”: (1)$\forall\varepsilon>0$, 存在C_0, 当$C>C_0$时, $\sup\limits_{\xi\in\mathcal{H}}\int_{[|\xi|\geq C]}|\xi|\mathrm{d}P\leq\frac{\varepsilon}{2}$. 所以

$$\sup_{\xi\in\mathcal{H}}\mathbb{E}|\xi|\leq\sup_{\xi\in\mathcal{H}}\int_{[|\xi|\geq C]}|\xi|\mathrm{d}P+\sup_{\xi\in\mathcal{H}}\int_{[|\xi|<C]}|\xi|\mathrm{d}P\leq\frac{\varepsilon}{2}+C.$$

(2)取$\delta=\frac{\varepsilon}{2C}$, 则当$P(A)<\delta$时,

$$\int_A|\xi|\mathrm{d}P\leq\int_{[|\xi|<C]\cap A}|\xi|\mathrm{d}P+\int_{[|\xi|\geq C]\cap A}|\xi|\mathrm{d}P\leq CP(A)+\frac{\varepsilon}{2}<\varepsilon.$$

“$\Leftarrow$”: $\forall\varepsilon>0,\exists\delta>0$, 当$C\geq\frac{1}{\delta}\sup\limits_{\xi\in\mathcal{H}}\mathbb{E}|\xi|$时,

$$P(|\xi|\geq C)\leq\frac{\mathbb{E}|\xi|}{C}\leq\frac{a}{C}\leq\delta.$$

所以$\sup\limits_{\xi\in\mathcal{H}}\int_{[|\xi|\geq C]}|\xi|\mathrm{d}P\leq\varepsilon$, 则$\mathcal{H}$一致可积. □

推论 5.3.2 设$\eta\in L^1(\Omega,\mathscr{F})$是可积随机变量, $\mathcal{H}$是一致可积族, 则$\eta+\mathscr{H}=\{\xi+\eta|\xi\in\mathcal{H}\}$是一致可积的.

推论 5.3.3 若$\mathcal{H}\subset L^1(\Omega,\mathscr{F},P)$是一致可积族, 则$\mathcal{H}$在$L^1(\Omega,\mathscr{F},P)$中的闭凸包也一致可积.

定理 5.3.4 (L^1收敛准则) 设$\{\xi_n\}\subset L^1(\Omega,\mathscr{F},P)$, *TFAE:*
(1)$\xi_n\xrightarrow{L^1}\xi$;
(2)$\xi_n\xrightarrow{P}\xi$, 且$\{\xi_n\}$一致可积;
(3)$\xi_n\xrightarrow{P}\xi$, 且$\mathbb{E}|\xi_n|\to\mathbb{E}|\xi|<\infty$.

证明: “(1) $\Leftrightarrow$ (3)”以前证过了. 下面证明“(1) $\Leftrightarrow$ (2)”.

“$\Rightarrow$”: 设$\xi_n \xrightarrow{L^1} \xi$, 即$\mathbb{E}|\xi_n-\xi| \to 0$. 令$A \in \mathscr{F}$, 则

$$\int_A |\xi_n| \mathrm{d}P \le \int_A |\xi| \mathrm{d}P + \mathbb{E}|\xi_n - \xi|,$$

给定$\varepsilon > 0$, 取正数N, 使得当$n > N$时, $\mathbb{E}|\xi_n-\xi| \le \dfrac{\varepsilon}{2}$ (处理无限个),

再选取$\delta > 0$, 使得对任何满足$P(A) \le \delta$的$A \in \mathscr{F}$, 有

$$\int_A |\xi| \mathrm{d}P \le \frac{\varepsilon}{2}, \int_A |\xi_n| \mathrm{d}P \le \frac{\varepsilon}{2}, n = 1, 2, \cdots, N.$$

(处理有限个). 所以对任何满足$P(A) \le \delta$的$A \in \mathscr{F}$, 有$\sup\limits_n \int_A |\xi_n| \mathrm{d}P \le \varepsilon$. 此外有$\sup\limits_n \mathbb{E}|\xi_n| < \infty$. 所以由定理5.3.1, $\{\xi_n\}$是一致可积族. 最后显然有$\xi_n \xrightarrow{P} \xi$.

“$\Leftarrow$”: 设$\{\xi_n\}$一致可积, 且$\xi_n \xrightarrow{P} \xi$. 由Fatou引理与定理5.3.1,

$$\mathbb{E}|\xi| \le \sup_n \mathbb{E}|\xi_n| < +\infty,$$

所以ξ可积, 所以$\xi_n - \xi$一致可积.

下证$\mathbb{E}|\xi_n - \xi| \to 0$. 对任何$\varepsilon > 0$, 由定理5.3.1, 存在$\delta > 0$, 使得对任意满足$P(A) < \delta$的$A \in \mathscr{F}$, 有

$$\sup_n \int_A |\xi_n - \xi| \mathrm{d}P \le \varepsilon.$$

由于$\xi_n \xrightarrow{P} \xi$, 可以取$N$充分大使得当$n \ge N$时$P([|\xi_n - \xi| \ge \varepsilon]) < \delta$, 故当$n \ge N$时, 有

$$\mathbb{E}|\xi_n - \xi| = \int_{[|\xi_n - \xi| < \varepsilon]} |\xi_n - \xi| \mathrm{d}P + \int_{[|\xi_n - \xi| \ge \varepsilon]} |\xi_n - \xi| \mathrm{d}P \le 2\varepsilon.$$

所以$\xi_n \xrightarrow{L^1} \xi$. □

把前一定理的L^1改为L^p, 会有如下结果:

定理 5.3.5 *设$(\Omega, \mathscr{F}, P)$是概率空间, $0 < p < \infty$. 若$\xi_n \xrightarrow{P} \xi$且$\{|\xi_n|^p\}$一致可积, 则$\xi_n \xrightarrow{L^p} \xi$.*

证明: 只需注意不等式

$$|a-b|^p \le \gamma_p(|a|^p + |b|^p), \text{其中} \gamma_p = \max\{1, 2^{p-1}\}.$$

来推$|\xi_n - \xi|^p$一致可积. 其他步骤与前一定理类似. 【待补充】

定理 5.3.6 *设$\mathcal{H} \subset L^1(\Omega, \mathscr{F}, P)$, TFAE:*

(1)$\mathcal{H}$是一致可积的;

(2)存在函数$\varphi: [0, +\infty) \to [0, +\infty)$, 满足$\lim\limits_{t \to \infty} \dfrac{\varphi(t)}{t} = \infty$, 且$\sup\limits_{\xi \in \mathcal{H}} \mathbb{E}(\varphi \circ |\xi|) < \infty$.

证明: “(1) $\Rightarrow$ (2)”: 设$\mathcal{H}$是一致可积族, 由于对任何$a>0$, 有

$$\int_\Omega (|\xi|-a)^+\mathrm{d}P \le \int_{[|\xi|>a]}(|\xi|-a)^+\mathrm{d}P \le \int_{[|\xi|>a]}|\xi|\mathrm{d}P \to 0(a\to\infty),$$

故存在自然数$n_k \nearrow \infty$, 使得

$$\sup_{\xi\in\mathcal{H}}\int_\Omega (|\xi|-n_k)^+\mathrm{d}P < 2^{-k}, k\ge 1.$$

令

$$\varphi(t)=\sum_{k\ge 1}([t]-n_k)^+, t\in[0,+\infty)$$

则φ非负、单调非降且右连续, 而且由Fatou引理,

$$\lim_{n\to\infty}\frac{\varphi(n)}{n}=\lim_{n\to\infty}\sum_{k\ge1}\left(1-\frac{n_k}{n}\right)^+ \ge \sum_{k\ge1}\inf_{n\to\infty}\left(1-\frac{n_k}{n}\right)^+=\sum_{k\ge1}1=\infty.$$

所以$\lim\limits_{t\to\infty}\dfrac{\varphi(t)}{t}=\infty$. 最后有

$$\begin{aligned}
\mathbb{E}(\varphi\circ|\xi|) &= \int_\Omega \varphi(|\xi|)\mathrm{d}P=\sum_{n=0}^\infty\int_{n\le|\xi|<n+1}\varphi(|\xi|)\mathrm{d}P\\
&=\sum_{n=0}^\infty\sum_{k=1}^\infty (n-n_k)^+P(n\le|\xi|<n+1)\\
&=\sum_{k=1}^\infty\sum_{n=0}^\infty (n-n_k)^+P(n\le|\xi|<n+1) \qquad (\text{非负可换序})\\
&\le\sum_{k=1}^\infty\int_\Omega(|\xi|-n_k)^+\mathrm{d}P<1.
\end{aligned}$$

“(2) $\Rightarrow$ (1)”: 设(2)成立, 对$\varepsilon>0$, 令$a=M/\varepsilon$, 其中$M=\sup\limits_{\xi\in\mathcal{H}}\mathbb{E}(\varphi(|\xi|))$. 取充分大的$C$使得当$t\ge C$时, $\dfrac{\varphi(t)}{t}\ge a$. 则在$[|\xi|\ge C]$上, 有$|\xi|\le\dfrac{\varphi\circ|\xi|}{a}$, 故

$$\int_{[|\xi|\ge C]}|\xi|\mathrm{d}P\le\frac{1}{a}\int_{[|\xi|\ge C]}\varphi\circ|\xi|\mathrm{d}P\le\frac{M}{a}=\varepsilon,\xi\in\mathcal{H}.$$

所以$\mathcal{H}$是一致可积族.

推论 5.3.7 设$\mathcal{H}\subset L^p(\Omega,\mathscr{F},P),p>1$. 如果$\sup\limits_{\xi\in\mathcal{H}}\mathbb{E}|\xi|^p<\infty$, 则$\mathcal{H}$是一致可积族.

注: 注意对比定理5.3.1的条件, 这里少了积分一致绝对连续的条件.

证明: 【方法一】定理5.3.6中让$\varphi(t)=t^p$立得.

【方法二】设$a=\sup\limits_{\xi\in\mathcal{H}}\mathbb{E}|\xi|^p$, 则对任意$C>0$, 有

$$\int_{[|\xi|>C]}|\xi|\mathrm{d}P\le\int_{[|\xi|>C]}\frac{|\xi|^p}{C^{p-1}}\mathrm{d}P\le\frac{1}{C^{p-1}}\mathbb{E}|\xi|^p\le\frac{a}{C^{p-1}}\to0(C\to\infty).$$

所以由定义可知$\mathcal{H}$是一致可积族. □

定理 5.3.8 设$(\Omega,\mathscr{F},P)$是概率空间，ξ是可积随机变量，$\{\mathscr{G}_i\}_{i\in I}$是$\mathscr{F}$的一族子$\sigma$-代数．令$\eta_i=\mathbb{E}(\xi|\mathscr{G}_i)$，则$\{\eta_i\}_{i\in I}$是一致可积族.

证明: 注意用条件期望的定义以及“条件期望的期望等于无条件期望”即可.

对任何$C>0$, 有

$$P(|\eta_i|\geq C)\leq \frac{1}{C}\mathbb{E}|\eta_i|=\frac{1}{C}\mathbb{E}|\mathbb{E}(\xi|\mathscr{G}_i)|\leq\frac{1}{C}\mathbb{E}(\mathbb{E}(|\xi||\mathscr{G}_i))=\frac{1}{C}\mathbb{E}|\xi|.$$

注意$[|\eta_i|\geq C]\in\mathscr{G}_i$, 则

$$\begin{aligned}\int_{[|\eta_i|\geq C]}|\eta_i|\mathrm{d}P\leq\int_{[|\eta_i|\geq C]}|\xi|\mathrm{d}P&=\int_{[|\eta_i|\geq C]\cap[|\xi|<\delta]}|\xi|\mathrm{d}P+\int_{[|\eta_i|\geq C]\cap[|\xi|\geq\delta]}|\xi|\mathrm{d}P\\&\leq\delta P([|\eta_i|\geq C])+\int_{[|\xi|\geq\delta]}|\xi|\mathrm{d}P\\&\leq\frac{\delta}{C}\mathbb{E}|\xi|+\int_{[|\xi|\geq\delta]}|\xi|\mathrm{d}P.\end{aligned}$$

由积分的绝对连续性, 对$\varepsilon>0$, 取$\delta>0$, 使得$\int_{[|\xi|\geq\delta]}|\xi|\mathrm{d}P\leq\frac{\varepsilon}{2}$, 则当$C\geq\frac{2\delta}{\varepsilon}\mathbb{E}|\xi|$时, 有$\int_{[|\eta_i|\geq C]}|\eta_i|\mathrm{d}P\leq\varepsilon, i\in I$. 所以$\{\eta_i\}$是一致可积族. □

§ 5.4 第五章习题

5.4.1 随机变量的独立性

例 5.4.1 设$(X_n,n\geq1)$是独立$r.v.$序列，则$\limsup\limits_{n\to\infty}X_n$与$\liminf\limits_{n\to\infty}X_n$是退化随机变量(即$a.s.$等于某个常数).

证明: 对任意$c>0$, $\sum\limits_{n=1}^{\infty}P(X_n\geq c)=\sum\limits_{n=1}^{\infty}P(X_1\geq c)=\begin{cases}+\infty, & P(X_1\geq c)\neq0,\\0, & P(X_1\geq c)=0.\end{cases}$. 由Borel-Cantelli引理及$(X_n)$是独立r.v.列, 则

$$P\Big(X_n\geq c,i.o.)=\begin{cases}0, & P(X_1\geq c)=0,\\1, & P(X_1\geq c)\neq0.\end{cases}$$

所以$P(\limsup\limits_{n\to\infty}X_n=c)=0$或1. □

例 5.4.2 设$\{X_n\}$是独立随机变量，证明: $X_n\xrightarrow{a.s.}0\Leftrightarrow\forall c>0,\sum\limits_{n=1}^{\infty}P(|X_n|\geq c)<+\infty.$

例 5.4.3 设$\{X_n\}$是$i.i.d.$，则$\frac{X_n}{\log n}\xrightarrow{a.s.}0\Leftrightarrow\forall c>0,\mathbb{E}e^{c|X_1|}<\infty.$

证明: 由a.s.收敛的刻画, $\frac{X_n}{\log n}\xrightarrow{a.s.}0$等价于对任意$c>0$, 有$\lim\limits_{k\to\infty}P\left(\bigcup\limits_{n\geq k}\left[\frac{|X_n|}{\log n}>\frac{1}{c}\right]\right)=0$. 利

用独立性(或者Borel-Cantelli引理),

$$P\left(\bigcup_{n\geq k}\left[\frac{|X_n|}{\log n}>\frac{1}{c}\right]\right)=\sum_{n=k}^{\infty}P\left(\frac{|X_n|}{\log n}>\frac{1}{c}\right)=\sum_{n=k}^{\infty}P\left(\frac{|X_1|}{\log n}>\frac{1}{c}\right)=\sum_{n=k}^{\infty}P\left(e^{c|X_1|}>n\right).$$

运用Nice引理(引理4.5.1)即可得欲证结论. □

例 5.4.4 设X,Y相互独立, X有密度函数, 则$X+Y$也有密度函数.

证明: 设$Z=X+Y$, 则由独立性,

$$F_Z(z)=\int_{[x+y\leq z]}\mathrm{d}F_{(X,Y)}(x,y)=\int_{\mathbb{R}}\int_{[x\leq z-y]}\mathrm{d}F_X(x)\mathrm{d}F_Y(y)=\int_{\mathbb{R}}F_X(z-y)\mathrm{d}F_Y(y).$$

若X有密度函数f_X, 则$F_X(z-y)=\displaystyle\int_0^{z-y}F_X(x)\mathrm{d}x$, 所以

$$\begin{aligned}F_Z(z)&=\int_{\mathbb{R}}\int_{-\infty}^{z-y}f_X(x)\mathrm{d}x\mathrm{d}F_Y(y)\\&\xlongequal{x=u-y}\int_{\mathbb{R}}\int_{-\infty}^{z}f_X(u-y)\mathrm{d}u\mathrm{d}F_Y(y)\\&=\int_{-\infty}^{z}\int_{\mathbb{R}}f_X(u-y)\mathrm{d}F_Y(y)\mathrm{d}u\qquad\text{(Fubini)}\end{aligned}$$

所以$f_Z(z)=\displaystyle\int_{\mathbb{R}}f_X(z-y)\mathrm{d}F_Y(y)$. □

例 5.4.5 设X,Y是相互独立可积随机变量, 且$\mathbb{E}X=0$, 则$\mathbb{E}|X+Y|\geq\mathbb{E}|Y|$.

证明: 注意$|y|=|\mathbb{E}(y+X)|\leq\mathbb{E}|y+X|$, 两边关于$y$积分可得

$$\mathbb{E}|Y|=\int_{\mathbb{R}}|y|\mathrm{d}P_Y(y)\leq\int_{\mathbb{R}}\mathbb{E}|y+X|\mathrm{d}P_Y(y)=\int_{\mathbb{R}}\int_{\mathbb{R}}|y+x|\mathrm{d}P_X(x)\mathrm{d}P_Y(y)=\mathbb{E}|X+Y|.$$

例 5.4.6 设$(X_n,n\geq1)$是$i.i.d.r.v.$列, X_n都服从指数为1的指数分布, 即$P(X_n>x)=e^{-x},x\geq0$.

(1)证明:$P(X_n>\alpha\log n,i.o.)=\begin{cases}0, & 若\alpha>1,\\1, & 若\alpha\leq1.\end{cases}$

(2)令$L=\limsup\limits_{n\to\infty}\dfrac{X_n}{\log n}$, 证明: $P(L=1)=1$. (提示: 证明$P(L\geq1)=1,P(L>1)=0$.)

证明: (1)由Borel-Cantelli引理, 欲证命题等价于$\displaystyle\sum_{n=1}^{\infty}P(X_n>\alpha\log n)\begin{cases}<\infty, & 若\alpha>1,\\+\infty, & 若\alpha\leq1.\end{cases}$ 事实上, 有

$$\sum_{n=1}^{\infty}P(X_n>\alpha\log n)=\sum_{n=1}^{\infty}e^{-\alpha\log n}=\sum_{n=1}^{\infty}n^{-\alpha}\begin{cases}<\infty, & 若\alpha>1,\\+\infty, & 若\alpha\leq1.\end{cases}$$

证明完毕.

(2)回顾: $[\inf\limits_n f_n<x]=\displaystyle\bigcup_{n=1}^{\infty}[f_n<x]$, 而$[\sup\limits_n f_n>x]=\displaystyle\bigcup_{n=1}^{\infty}[f_n>x]$.

注意到

$$\begin{aligned}
P(L>1) = P(\inf_k \sup_{n\ge k} \frac{X_n}{\log n} > 1) &= P\left(\bigcup_{m=1}^{\infty}\left[\inf_k \sup_{n\ge k}\frac{X_n}{\log n}\ge 1+\frac{1}{m}\right]\right)\\
&= P\left(\bigcup_{m=1}^{\infty}\bigcap_{k=1}^{\infty}\left[\sup_{n\ge k}\frac{X_n}{\log n}\ge 1+\frac{1}{m}\right]\right)\\
&= P\left(\bigcup_{m=1}^{\infty}\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left[\frac{X_n}{\log n}\ge 1+\frac{1}{m}\right]\right)\\
&= \lim_{m\to\infty} P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left[\frac{X_n}{\log n}\ge 1+\frac{1}{m}\right]\right) = 0, \qquad \text{(由第(1)小问)}
\end{aligned}$$

$$\begin{aligned}
P(L\ge 1) = P(\inf_k \sup_{n\ge k} \frac{X_n}{\log n} \ge 1) &= P\left(\bigcap_{k=1}^{\infty}\left[\sup_{n\ge k}\frac{X_n}{\log n}\ge 1\right]\right)\\
&= P\left(\bigcap_{k=1}^{\infty}\bigcap_{m=1}^{\infty}\left[\sup_{n\ge k}\frac{X_n}{\log n}> 1-\frac{1}{m}\right]\right)\\
&= P\left(\bigcap_{k=1}^{\infty}\bigcap_{m=1}^{\infty}\bigcup_{n=k}^{\infty}\left[\frac{X_n}{\log n}> 1-\frac{1}{m}\right]\right)\\
&= \lim_{m\to\infty} P\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left[\frac{X_n}{\log n}> 1-\frac{1}{m}\right]\right) = 1.
\end{aligned}$$

例 5.4.7 设$\{X_n : n\ge 1\}$是独立的$N(0,1)$随机变量, 证明:

$$P\left(\limsup_{n\to\infty}\frac{|X_n|}{\sqrt{\log n}}=\sqrt{2}\right)=1.$$

证明: 类似前一题(2), 只需要证明$P(|X_n|>\alpha\log n, i.o.) = \begin{cases}0, & \alpha>\sqrt{2},\\ 1, & \alpha\le\sqrt{2}.\end{cases}$.

引理 5.4.1 (Mills's Ratio) 设$\phi(x)=e^{-\frac{x^2}{2}}, \Phi(x)=\int_{-\infty}^{x}\phi(t)\mathrm{d}t$, 则

$$\frac{1}{x}-\frac{1}{x^3}<\frac{1-\Phi(x)}{\phi(x)}<\frac{1}{x}-\frac{1}{x^3}+\frac{3}{x^5}, x>0.$$

这个估计式我们非常感兴趣, 因为$\Phi(x)$没有封闭形式的表达式.

证明: $\phi(x)$满足$\phi'=-x\phi$. 所以不断用分部积分可得

$$\begin{aligned}
1-\Phi(x) &= \int_x^{\infty}\phi(t)\mathrm{d}t = -\int_x^{\infty}\frac{\phi'(t)}{t}\mathrm{d}t\\
&= \frac{\phi(x)}{x}+\int_x^{\infty}\frac{\phi'(t)}{t^3}\mathrm{d}t\\
&= \frac{\phi(x)}{x}-\frac{\phi(x)}{x^3}-\int_x^{\infty}\frac{\phi'(t)}{t^5}\mathrm{d}t\\
&= \frac{\phi(x)}{x}-\frac{\phi(x)}{x^3}+\frac{3\phi(x)}{x^5}-\int_x^{\infty}\frac{15\phi(t)}{t^6}\mathrm{d}t.
\end{aligned}$$

证明完毕. □

根据引理, 当$\alpha > \sqrt{2}$时,

$$\sum_{n=3}^{\infty} P(|X_n| > \alpha\sqrt{\log n}) < \sum_{n=3}^{\infty} n^{-\frac{\alpha^2}{2}} \cdot \frac{1}{\alpha\sqrt{\log n}} \le \sum_{n=3}^{\infty} n^{-\frac{\alpha^2}{2}} < \infty.$$

当$0 < \alpha \le \sqrt{2}$时,

$$\sum_{n=3}^{\infty} P(|X_n| > \alpha\sqrt{\log n}) > \sum_{n=3}^{\infty} n^{-\frac{\alpha^2}{2}} \cdot \frac{1}{\alpha\sqrt{\log n}} \left(1 - \frac{1}{\alpha^2 \log n}\right) = +\infty.$$

由Borel-Cantelli引理, $P(|X_n| > \alpha\sqrt{\log n}, i.o.) = \begin{cases} 0, & 若\alpha > \sqrt{2}, \\ 1, & 若\alpha \le \sqrt{2}. \end{cases}$. 用类似前一题(2)的说明可知$P\left(\limsup\limits_{n\to\infty} \frac{|X_n|}{\sqrt{\log n}} = \sqrt{2}\right) = 1$. □

例 5.4.8 设(ξ_n)是一列非负实值随机变量, 则存在一正实数序列(c_n), 使得$\sum\limits_{n=1}^{\infty} c_n\xi_n < \infty$, *a.s.*.

提示: 取正实数序列(a_n), 使得$\sum\limits_{n=1}^{\infty} P(\xi_n > a_n) < \infty$, 用*Borel-Cantelli*引理并令$c_n = (2^n a_n)^{-1}$.

例 5.4.9 设$(\Omega, \mathscr{F}, P)$是概率空间.

(1)设$\xi_1, \cdots, \xi_n$是非负随机变量, $\mathbb{E}\xi_i = 1, 1 \le i \le n$. 则$\prod\limits_{i=1}^{n} \frac{1}{n} \sum\limits_{j=1}^{n} \mathbb{E}(\xi_i\xi_j) \ge 1$.

(2)设$A_1, \cdots, A_n \in \mathscr{F}$, 且每个$P(A_i) > 0$. 令$A = \bigcup\limits_{i=1}^{n} A_i$, 则$\prod\limits_{i=1}^{n} \frac{1}{n} \sum\limits_{j=1}^{n} \frac{P(A_iA_j)}{P(A_i)P(A_j)} \ge \left(\frac{1}{P(A)}\right)^n$.

5.4.2 条件期望

例 5.4.10 设$X, Y \in L^2(\Omega, \mathscr{F}, P)$满足$\mathbb{E}(Y|X) = X, \mathbb{E}(X|Y) = Y$, 则$X = Y$, *a.s.*.

证明: 注意到

$$\begin{aligned} 0 \le \mathbb{E}(X-Y)^2 &= \mathbb{E}X^2 - 2\mathbb{E}XY + \mathbb{E}Y^2 \\ &= \mathbb{E}X^2 - 2\mathbb{E}[\mathbb{E}(XY|X)] + \mathbb{E}Y^2 \text{【性质①】} \\ &= \mathbb{E}X^2 - 2\mathbb{E}(X\mathbb{E}(Y|X)) + \mathbb{E}Y^2 \text{【性质⑦】} \\ &= \mathbb{E}Y^2 - \mathbb{E}X^2. \end{aligned}$$

同理, 如果对上面第二行的式子改为作用Y的条件期望, 可得$\mathbb{E}(X-Y)^2 = EX^2 - EY^2$. 一个数同时等于另一个数与它的相反数, 则这个数只能为0, 即$\mathbb{E}(X-Y)^2 = 0$, 则$X = Y$, a.s..

例 5.4.11 设$X, Y \in L^1(\Omega, \mathscr{F}, P)$满足$\mathbb{E}(Y|X) = X, \mathbb{E}(X|Y) = Y$, 则$X = Y$, *a.s.*.

提示: 只需考虑$\mathbb{E}(X-Y)(\arctan X - \arctan Y)$, 这个依然是非负的, 而且$\arctan x$有界, 则$(X - Y)(\arctan X - \arctan Y)$必定可积.

例 5.4.12 *已知X是一可积$r.v.$, $\mathscr{C}$是$\mathscr{F}$的子$sigma$代数. 令$Y=\mathbb{E}(X|\mathscr{C})$, 假定$X$与$Y$同分布, 证明:*
(1)若X平方可积, 则$X=Y$, $a.s.$.
(2)若$\forall a,b\in\mathbb{R},a<b$, 有$(X\vee a)\wedge b=(Y\vee a)\wedge b$, $a.s.$, 则$X=Y$, $a.s.$.

证明: (1)注意到

$$\begin{aligned}\mathbb{E}(X-Y)^2&=\mathbb{E}X^2-2\mathbb{E}XY+\mathbb{E}Y^2=2\mathbb{E}Y^2-2\mathbb{E}XY\text{【同分布】}\\&=2\mathbb{E}Y^2-2\mathbb{E}[\mathbb{E}(XY|\mathscr{C})]\text{【性质①】}\\&=2\mathbb{E}Y^2-2\mathbb{E}(Y\mathbb{E}(X|\mathscr{C}))\text{【性质⑦】}\\&=2\mathbb{E}Y^2-2\mathbb{E}Y^2\text{【题目条件】}\\&=0,\Rightarrow X=Y,a.s..\end{aligned}$$

(2)下证$\mathbb{E}(X\vee a|\mathscr{C})=Y\vee a(=\mathbb{E}(X|\mathscr{C})\vee a),\forall a\in\mathbb{R}$, 则

$$\mathbb{E}[(X\vee a)\wedge b|\mathscr{C}]=(Y\vee a)\wedge b,a.s.,\forall a,b\in\mathbb{R},a<b.$$

(把并的证明推广到交的证明, 结果是一样的, 所以下面只证明并的情况)

首先考虑到函数$f(x)=x\vee a$是凸的, 根据Jensen不等式有

$$\mathbb{E}(X\vee a|\mathscr{C})\geq\mathbb{E}(X|\mathscr{C})\vee a,a.s..$$

只需证$P[\mathbb{E}((X\vee a)|\mathscr{C})>(\mathbb{E}(X|\mathscr{C})\vee a)]=0$. 事实上,

$$\begin{aligned}\mathbb{E}[(X\vee a)|\mathscr{C}-(\mathbb{E}(X|\mathscr{C})\vee a)]&=\mathbb{E}[\mathbb{E}(X\vee a-\mathbb{E}(X|\mathscr{C})\vee a|\mathscr{C})]\text{【性质①】}\\&=\mathbb{E}(X\vee a-\mathbb{E}(X|\mathscr{C})\vee a)\text{【线性性】}\\&=\mathbb{E}X\vee a-\mathbb{E}(\mathbb{E}(X|\mathscr{C})\vee a)\\&=\mathbb{E}X\vee a-\mathbb{E}Y\vee a=0.\text{【同分布】}\end{aligned}$$

证明完毕. □

例 5.4.13 *设$X\in L^2(\Omega,\mathscr{F},P)$, $\mathscr{C}_1\subset\mathscr{C}_2$是$\mathscr{F}$的子$\sigma$-代数, 则*

$$\mathbb{E}(\mathbb{E}(X|\mathscr{C}_1)-X)^2\geq\mathbb{E}(\mathbb{E}(X|\mathscr{C}_2)-X)^2.$$

证明: 仿照条件期望的投影定理来证明.

例 5.4.14 *设ξ,η是$(\Omega,\mathscr{F},P)$中的随机变量, $\mathbb{E}(\xi|\mathscr{C})=\eta$, 且$\mathbb{E}\xi^2=\mathbb{E}\eta^2<+\infty$. 证明: $\xi=\eta$, $a.s.$.*

证明: 注意

$$\begin{aligned}\mathbb{E}(\xi-\eta)^2&=\mathbb{E}\xi^2+\mathbb{E}\eta^2-2\mathbb{E}(\xi\eta)\\&=2\mathbb{E}\eta^2-2\mathbb{E}(\mathbb{E}(\xi\eta|\mathscr{C}))\\&=2\mathbb{E}\eta^2-2\mathbb{E}(\eta\mathbb{E}(\xi|\mathscr{C}))\\&=2\mathbb{E}\eta^2-2\mathbb{E}\eta^2=0.\end{aligned}$$

5.4.3 一致可积

例 5.4.15 设$\{\xi_n\}$是一致可积随机变量序列, 则有$\lim\limits_{n\to\infty}\mathbb{E}\left(\frac{1}{n}\sup\limits_{1\le k\le n}|\xi_k|\right)=0$.

证明: (1)我们有

$$\begin{aligned}
&\mathbb{E}(|X|\vee|Y|I_{[|X|\vee|Y|\ge C]})\\
=&\int_{[|X|\ge|Y|]}|X|\vee|Y|I_{[|X|\vee|Y|\ge C]}\mathrm{d}P+\int_{[|X|\le|Y|]}|X|\vee|Y|I_{[|X|\vee|Y|\ge C]}\mathrm{d}P\\
=&\int_{[|X|\ge|Y|]}|X|I_{[|X|\ge C]}\mathrm{d}P+\int_{[|X|\le|Y|]}|Y|I_{[|Y|\ge C]}\mathrm{d}P\\
\le&\mathbb{E}(|X|I_{[|X|\ge C]})+\mathbb{E}(|Y|I_{[|Y|\ge C]}).
\end{aligned}$$

(2)注意到

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{n}\sup_{1\le k\le n}|\xi_k|\right)&=\mathbb{E}\left(\frac{1}{n}\sup_{1\le k\le n}|\xi_k|I_{[\sup\limits_{1\le k\le n}|\xi_k|\ge C]}\right)+\mathbb{E}\left(\frac{1}{n}\sup_{1\le k\le n}|\xi_k|I_{[\sup\limits_{1\le k\le n}|\xi_k|<C]}\right)\\
&\le\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(|\xi_i|I_{[|\xi_i|\ge C]})+\frac{C}{n}\le\sup_{\xi\in\mathcal{H}}\mathbb{E}(|\xi|I_{[|\xi|\ge C]})+\frac{C}{n}.
\end{aligned}$$

由一致可积性, 对$\varepsilon>0$, 存在C_0, 当$C>C_0$时, $\left|\sup\limits_{\xi\in\mathcal{H}}\mathbb{E}(|\xi|I_{[|\xi|\ge C]})\right|<\varepsilon/2$.

对上述ε, 取$N=\frac{2C}{\varepsilon}$, 则当$n>N$时, $\frac{C}{n}<\frac{\varepsilon}{2}$. 所以$\mathbb{E}\Big(\frac{1}{n}\sup\limits_{1\le k\le n}|\xi_k|\Big)<\varepsilon$. □

例 5.4.16 设$\mathcal{H}\subset L^1(\Omega,\mathscr{F},P)$, 若$\mathcal{H}$满足:

$$A_n\in\mathcal{F},A_n\searrow\varnothing\Rightarrow\lim_{n\to\infty}\sup_{\xi\in\mathcal{H}}\int_{A_n}|\xi|\mathrm{d}P=0,$$

则对任给$\varepsilon>0$, 存在$\delta>0$, 使得

$$A\in\mathscr{F},P(A)<\delta\Rightarrow\sup_{\xi\in\mathcal{H}}\int_A|\xi|\mathrm{d}P\le\varepsilon.$$

证明: (反证)若不然, 存在$\varepsilon>0$, 对$\delta_n=\frac{1}{2^n}$, 都存在$B_n\in\mathscr{F}$, 使得

$$P(B_n)<\delta_n=\frac{1}{2^n}\text{且}\sup_{\xi\in\mathcal{H}}\int_{B_n}|\xi|\mathrm{d}P\ge\varepsilon.$$

由Borel-Cantelli引理, 根据$\sum\limits_{n=1}^{\infty}P(B_n)<\infty$, 可得$P\Big(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}B_k\Big)=0$.

取$C_n=\bigcup\limits_{k=n}^{\infty}B_k$, 则$\sup\limits_{\xi\in\mathcal{H}}\int_{C_n}|\xi|\mathrm{d}P\ge\varepsilon$, 且$P\Big(\bigcap\limits_{m=1}^{\infty}C_m\Big)=0$.

取$A_n=C_n\setminus\bigcap\limits_{m=1}^{\infty}C_m$, 则$A_n$单调下降趋于$\varnothing$, 但对任意$n$, $\sup\limits_{\xi\in\mathcal{H}}\int_{A_n}|\xi|\mathrm{d}P\ge\varepsilon$, 这与条件矛盾. □

第 6 章　鞅论及其应用

§ 6.1　鞅与停时

参考书: R. Durrett, Probability Theory and Examples, 4th edition, 2010.

6.1.1　鞅

定义 6.1.1 设$(\Omega, \mathscr{F}, P)$是概率空间, $\{\mathscr{F}_n\}_{n\geq 1}$是$\mathscr{F}$的一列子$\sigma$-代数. 若$\mathscr{F}_n \subset \mathscr{F}_{n+1}, \forall n \geq 1$, 则称$\{\mathscr{F}_n\}$是***($\sigma$-代数)流(filtration)***.

设$(\Omega, \mathscr{F}, P)$是带流$\{\mathscr{F}_n\}$的概率空间, $\{X_n\}$是一列随机变量, 如果对任意$n \geq 1, X_n$关于$\mathscr{F}_n$可测, 则称$\{X_n\}$关于$\{\mathscr{F}_n\}$**适应**$(\{X_n\}$ *is adapted to* $\{\mathscr{F}_n\}.)$

设随机变量序列$\{X_n\}$满足下面三个条件:

(1)$\mathbb{E}|X_n| < \infty$;

(2)$\{X_n\}$关于$\{\mathscr{F}_n\}$适应;

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

则称$\{(X_n, \mathscr{F}_n)\}_{n\geq 1}$是***鞅(martingale)***. 把(3)中的等号改为"$\leq$"或"$\geq$"就得到***上鞅(supermartingale)***或***下鞅(submartingale)***.

例 6.1.1 设$\{\xi_n\}$是$i.i.d.r.v.$列, 且$\mathbb{E}\xi_1 = 0$, $X_n = \xi_1 + \cdots + \xi_n, \mathscr{F}_n = \sigma(\xi_1, \cdots, \xi_n)$, 则$\{X_n, \mathscr{F}_n\}$是鞅.

证明: 由条件期望的性质,

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathscr{F}_n) &= \mathbb{E}(X_n|\mathscr{F}_n) + \mathbb{E}(\xi_{n+1}|\mathscr{F}_n) \\ &= X_n + \mathbb{E}\xi_{n+1} \qquad (X_n\text{关于}\mathscr{F}_n\text{可积}, \ \xi_{n+1}\text{与}\mathscr{F}_n\text{独立}) \\ &= X_n.\end{aligned}$$

注: 不同的流会得到不同的鞅. 如果没有指定流, 一般$\mathscr{F}_n = \sigma(X_1, \cdots, X_n)$. 这是使得$\{X_n\}$关于$\{\mathscr{F}_n\}$适应的最小$\sigma$-代数, 表示前$n$次观测得到的信息量.

例 6.1.2 设ξ是可积随机变量, $\{\mathscr{F}_n\}$是流, 则$\{\mathbb{E}(\xi|\mathscr{F}_n), \mathscr{F}_n\}_{n\geq 1}$是一致可积鞅.

命题 6.1.1 (1)设X_n是上鞅, $n > m$, 则$\mathbb{E}(X_n|\mathscr{F}_m) \leq X_m$.

(2)设X_n是下鞅, $n > m$, 则$\mathbb{E}(X_n|\mathscr{F}_m) \geq X_m$.

(3)设X_n是鞅, $n > m$, 则$\mathbb{E}(X_n|\mathscr{F}_m) = X_m$.

证明: (1)当$n = m + 1$时结论正确. 设$n = m + k, k \geq 2$, 则由条件期望的平滑性,

$$\mathbb{E}(X_{m+k}|\mathscr{F}_m) = \mathbb{E}(\mathbb{E}(X_{m+k}|\mathscr{F}_{m+k-1})|\mathscr{F}_m) \leq \mathbb{E}(X_{m+k-1}|\mathscr{F}_m).$$

归纳可得欲证结论.

(2)当X_n是下鞅时, $-X_n$是上鞅. 用(1)即可.

(3)只需注意鞅同时为上鞅和下鞅. □

注: 上鞅和下鞅作为对偶的定义, 可以通过取相反数来作替换. 所以我们接下来对下面的部分结果, 都只会陈述上鞅或下鞅的结果.

命题 6.1.2 *鞅的均值是常数, 上鞅的期望递减, 下鞅的期望递增.*

证明: 只需对定义两边取期望即可: 比如如果$\{X_n, \mathscr{F}_n\}$是鞅, 则

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

命题 6.1.3 *设$\{X_n, \mathscr{F}_n\}, \{Y_n, \mathscr{F}_n\}$是鞅(resp. 下鞅), 则$\{aX_n + bY_n, \mathscr{F}_n\}$是鞅(resp. 下鞅).*

注: 一定要关于同一个流!

定理 6.1.4 *设$\{X_n, \mathscr{F}_n\}_{n\geq 1}$是鞅(resp. 下鞅), f是连续(resp. 连续非降)凸函数. 若$\mathbb{E}|f(x_n)| < +\infty, \forall n$, 则$\{f(X_n), \mathscr{F}_n\}_{n\geq 1}$是下鞅.*

证明: 由Jensen不等式,

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

由此定理可以引出下面的推论:

推论 6.1.5 *设$\{X_n, \mathscr{F}_n\}_{n\geq 1}$是鞅(resp. 非负下鞅), $p \geq 1$是常数, 若$\mathbb{E}|X_n|^p < +\infty$, 则$\{|X_n|^p, \mathscr{F}_n\}_{n\geq 1}$是非负下鞅.*

推论 6.1.6 *(1)设$\{X_n, \mathscr{F}_n\}_{n\geq 1}$是下鞅, 则$(X_n - a)^+$是下鞅;*
(2)设$\{X_n, \mathscr{F}_n\}_{n\geq 1}$是上鞅, 则$X_n \wedge a$是上鞅.

6.1.2 停时

定义 6.1.2 *设$T: \Omega \to \overline{\mathbb{N}} \triangleq \mathbb{N} \cup \{+\infty\}$. 如果$\forall n \in \mathbb{N}, [T = n] \in \mathscr{F}_n$, 则称$T$是$\{\mathscr{F}_n\}_{n\geq 1}$的**停时(stopping time)**.*

若T是停时, 定义

$$\mathscr{F}_T \triangleq \{A \in \mathscr{F}_\infty | A \cap [T = n] \in \mathscr{F}_n, \forall n \geq 1\},$$

*其中$\mathscr{F}_\infty = \sigma\Big(\bigcup_{n=1}^{\infty} \mathscr{F}_n\Big)$, 则不难验证$\mathscr{F}_T$是$\sigma$-代数, 叫**$T$前$\sigma$-代数**.*

注: (1)停时不是随机变量, 因为可以取正无穷.

(2)定义中$[T = n] \in \mathscr{F}_n$可以改为$[T \leq n] \in \mathscr{F}_n$. 这是因为$[T = n] = \underbrace{[T \leq n]}_{\in \mathscr{F}_n} \setminus \underbrace{[T \leq n-1]}_{\in \mathscr{F}_{n-1} \subset \mathscr{F}_n}$.

(3)如果把T看作一个赌客停止赌博的时间, 那么上述条件就是说对任意正整数n, 在n时刻停止赌博的决定需要关于此时的信息可测.

(4)确定时间都是停时, 即任意$n \in \overline{\mathbb{N}}$(只取$n$的随机变量), n都是停时. 停时是确定时间的推广.

设T是停时, $n \in \overline{\mathbb{N}}$, 则$T + n$是停时, 但$T - n$不一定是停时. 这是因为$\mathscr{F}_{n+m} \not\subseteq \mathscr{F}_n$.

命题 6.1.7 设S,T是停时, $\{S_m\}$是一列停时, 则:

(1)$S+T$是停时;

(2)$\bigvee\limits_m S_m, \bigwedge\limits_m S_m$是停时;

(3)$A\in\mathscr{F}_S \Rightarrow A\cap[S\le T]\in\mathscr{F}_T, A\cap[S=T]\in\mathscr{F}_T$;

(4)$S\le T\Rightarrow \mathscr{F}_S\subset\mathscr{F}_T$. **(重要)**

(5)$A\in\mathscr{F}_S$, 令$S_A\triangleq SI_A+(+\infty)I_{A^c}$, 则$S_A$是停时, 且$A\in\mathscr{F}_{S_A}=\mathscr{F}_S\cap A$.

证明: (1)$[S+T=n]=\sum\limits_{i=0}^{n}[S=i,T=n-i]=\sum\limits_{i=0}^{n}[S=i][T=n-i]\in\mathscr{F}_n,\forall n\in\mathbb{N}$.

(2)$\left[\bigwedge\limits_m S_m\le n\right]=\bigcup\limits_{m=1}^{\infty}[S_m\le n]\in\mathscr{F}_n, \left[\bigvee\limits_m S_m\le n\right]=\bigcap\limits_{m=1}^{\infty}[S_m\le n]\in\mathscr{F}_n,\forall n\in\mathbb{N}$.

(3)$A[S\le T][T=n]=\underbrace{A[S\le n]}_{\in\mathscr{F}_S\subseteq\mathscr{F}_n}[T=n]\in\mathscr{F}_n,\forall n\in\mathbb{N}$.

(4)由(3), $\forall A\in\mathscr{F}_S, A=A[S\le T]\in\mathscr{F}_T$.

(5)$[S_A=n]=A[S=n]\in\mathscr{F}_n,\forall n\in\mathbb{N}$(只能取$A$, 不然为$+\infty$). 由(4), $\mathscr{F}_S\subset\mathscr{F}_A$; 对任意$B\in\mathscr{F}_{S_A}$, 有$A\cap B[S=n]=[S_A=n]B\in\mathscr{F}_n$, 则$A\cap B\in\mathscr{F}_S$, 从而$A\cap B\in A\cap\mathscr{F}_S$. □

定理 6.1.8 设$\{X_n,\mathscr{F}_n\}$是适应的$r.v.$列, T是停时, 则$X_TI_{[T<+\infty]}$关于$\mathscr{F}_T$可测.

证明: 对任意$B\in\mathscr{B}(\mathbb{R})$, 有

$$[X_TI_{[T<+\infty]}\in B]\cap[T=+\infty]=\begin{cases}\varnothing, & 0\in B^c,\\ [T=\infty], & 0\in B,\end{cases}\in\mathscr{F}_T.$$

$$[X_TI_{[T<+\infty]}\in B]\cap[T=n]=\underbrace{[X_n\in B]}_{X_n\text{适应}}\cap\underbrace{[T=n]}_{T\text{是停时}}\in\mathscr{F}_T.$$

所以$[X_TI_{[T<+\infty]\in B}\in\mathscr{F}_T$. □

§ 6.2 鞅不等式

下面我们把鞅变成停时(把确定时间变成随机时间).

定理 6.2.1 (Doob停止定理) 设$(X_n,\mathscr{F}_n)$是鞅(resp. 上鞅), S,T是**有界**停时, 且$S\le T$, 则$\mathbb{E}(X_T|\mathscr{F}_S)=X_S$(resp. $\le X_S$), a.s..

证明: 只证明上鞅的情形. 由S,T有界, 设$N_0\in\mathbb{N}$, 使得$T(\omega)\vee S(\omega)\le N_0,\forall\omega\in\Omega$. 所以$|X_T|\le\sum\limits_{j=1}^{N_0}|X_j|, |X_S|\le\sum\limits_{j=1}^{N_0}|X_j|$, 从而$X_S,X_T$可积.

(1)先设$T-S\le 1$. 对任意$A\in\mathscr{F}_S$, 下证$\int_A X_S\mathrm{d}P\ge\int_A X_T\mathrm{d}P$. (条件期望的定义!) 事实上,

$$\int_A(X_S-X_T)\mathrm{d}P=\sum_{j=0}^{N_0}\int_{A[S=i][T=i+1]}(X_S-X_T)\mathrm{d}P. \tag{6.1}$$

(注意当$S=T=i$时, $X_S=X_T$.) 由于$A\in\mathscr{F}_S$, 则$A[S=i]\in\mathscr{F}_i$, 由$[T=i+1]=[T>i]-[T>i+1]$, 则$[T=i+1][S=i]=\underbrace{[T>i]}_{\in\mathscr{F}_i}[S=i]\in\mathscr{F}_i$. 所以(6.1)式$\geq 0$.

(2)对一般情况, 由于$S\leq S+1\leq S+2\leq\cdots\leq S+N_0$, 所以$S\leq(S+1)\wedge T\leq\cdots\leq(S+N_0)\wedge T$. 所以由条件期望的平滑性,

$$\mathbb{E}(X_T|\mathscr{F}_S)=\mathbb{E}(X_{(S+N_0)\wedge T}|\mathscr{F}_S)=\mathbb{E}(\mathbb{E}(X_{(S+N_0)\wedge T}|\mathscr{F}_{(S+N_0-1)\wedge T})|\mathscr{F}_S)\geq\mathbb{E}(X_{(S+N_0-1)\wedge T}|\mathscr{F}_S).$$

不断进行下去可得$\mathbb{E}(X_T|\mathscr{F}_S)\geq\mathbb{E}(X_S|\mathscr{F}_S)=X_S$, a.s.. □

定理 6.2.2 设$\{X_n,\mathscr{F}_n\}_{n\leq k}$是上鞅, 对$\lambda>0$, 有

(1)$\lambda P\Big(\sup\limits_{n\leq k}X_n\geq\lambda\Big)\leq\mathbb{E}X_1-\displaystyle\int_{[\sup\limits_{n\leq k}X_n<\lambda]}X_k\mathrm{d}P$.

(2)$\lambda P\Big(\inf\limits_{n\leq k}X_n\leq-\lambda\Big)\leq\displaystyle\int_{[\sup\limits_{n\leq k}X_n\leq-\lambda]}(-X_k)\mathrm{d}P$.

(3)$\lambda P\Big(\sup\limits_{n\leq k}|X_n|\geq\lambda\Big)\leq\mathbb{E}X_1+2\mathbb{E}(X_k^-)$.

证明: (1)令$T=\inf\{n\geq 1|X_n\geq\lambda\}\wedge k$(表示首次$\geq\lambda$的$X_n$), 则$T$是有界停时, 事实上,

$$\begin{aligned}&[T=1]=[X_1\geq\lambda]\in\mathscr{F}_1,\\&[T=m]=[X_m\geq\lambda][X_1<\lambda]\cdots[X_{m-1}<\lambda]\in\mathscr{F}_m, m=2,3,\cdots,k-1,\\&[T=k]=[X_1<\lambda]\cdots[X_{k-1}<\lambda]\in\mathscr{F}_k.\end{aligned}$$

由有界停时的性质可知,

$$\begin{aligned}\mathbb{E}X_1\geq\mathbb{E}X_T&=\int_{[\sup\limits_{n\leq k}X_n\geq\lambda]}X_T\mathrm{d}P+\int_{[\sup\limits_{n>k}X_n\geq\lambda]}X_T\mathrm{d}P\\&\geq\lambda P\Big(\sup_{n\leq k}X_n\geq\lambda\Big)+\int_{[\sup\limits_{n>k}X_n\geq\lambda]}X_T\mathrm{d}P.\end{aligned}$$

(2)与(1)同理. (3)用(1)(2)可以推出. □

定理 6.2.3 (极大值不等式) 设$\{X_n,\mathscr{F}_n\}_{n\leq k}$是鞅(或非负下鞅), 记$X_k^*=\sup\limits_{n\leq k}|X_n|$, 则对任意$\lambda>0,p\geq 1$, 有$P(X_k^*\geq\lambda)\leq\lambda^{-p}\mathbb{E}|X_k|^p$.

证明: 若$\mathbb{E}|X_k|^p=+\infty$, 则结论显然. 若$\mathbb{E}|X_k|^p<+\infty$, 由Jensen不等式与鞅的性质,

$$\mathbb{E}|X_n|^p=\mathbb{E}|\mathbb{E}(X_k|\mathscr{F}_n)|^p\leq\mathbb{E}(\mathbb{E}(|X_k|^p|\mathscr{F}_n))=\mathbb{E}|X_k|^p<+\infty(1\leq n\leq k).$$

所以$\{-|X_n|^p\}$是上鞅. 从而由前一定理,

$$\begin{aligned}P(X_k^*\geq\lambda)=P((X_k^*)^p\geq\lambda^p)&=P\Big(\inf_{n\leq k}(-|X_n|^p)\leq-\lambda^p\Big)\\&\leq\frac{1}{\lambda^p}\int_{[\inf\limits_{n\leq k}(-|X_n|^p)\leq-\lambda^p]}|X_k|^p\mathrm{d}P\\&\leq\frac{1}{\lambda^p}\mathbb{E}|X_k|^p.\end{aligned}$$

定理 6.2.4 (Doob不等式) *设$\{X_n,\mathscr{F}_n\}_{n\le k}$是鞅(或非负下鞅), 记$X_k^*=\sup\limits_{n\le k}|X_n|$, 则对任意$p\ge 1$, 有$\|X_k^*\|_p\le \dfrac{p}{p-1}\|X_k\|_p$.*

注: 此不等式很像Hardy不等式.

证明: 只需要注意到

$$\begin{aligned}
\|X_k^*\|_p=\mathbb{E}(X_k^*)^p&=\int_0^{+\infty}py^{p-1}P(X_k^*\ge y)\mathrm{d}y\\
&=\int_0^{+\infty}py^{p-1}P\Big(\inf_{n\le k}(-|X_n|)\Big)\mathrm{d}y\\
&\le\int_0^{+\infty}py^{p-1}\frac{1}{y}\int_{[\inf\limits_{n\le k}(-|X_n|)\le -y]}|X_k|\mathrm{d}P\mathrm{d}y\\
&=p\int_0^{+\infty}\int_\Omega y^{p-2}|X_k|I_{[X_k^*\ge y]}\mathrm{d}P\mathrm{d}y\\
&=p\int_\Omega\int_0^{+\infty}y^{p-2}|X_k|I_{[X_k^*\ge y]}\mathrm{d}y\mathrm{d}P\\
&=p\int_\Omega\frac{1}{p-1}|X_k^*|^{p-1}|X_k|\mathrm{d}P\\
&\le\frac{p}{p-1}\|X_k\|_p\|X_k^*\|_p^{p-1}.(\text{Hölder不等式}).
\end{aligned}$$

下面来介绍上穿不等式(upcrossing inequality).

设$\{X_n\}_{n\ge1}$是适应r.v.列,$a,b\in\mathbb{R},a<b$. 令

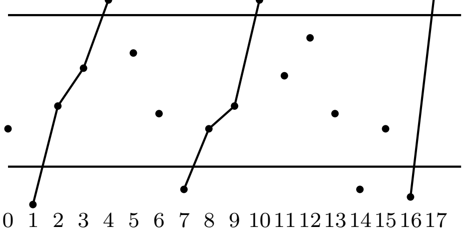
$$\begin{aligned}
T_0=\inf\{n\ge 0|X_n\le a\},T_1=\inf\{n>T_0|X_n\ge b\},\\
T_{2j}=\inf\{n>T_{2j-1}|X_n\le a\},T_{2j+1}=\inf\{n>T_{2j}|X_n\ge b\}.
\end{aligned}$$

则

$$[T_1=n]=\sum_{i=0}^{n-1}[T_0=i][X_{i+1}<b]\cdots[X_{n-1}<b][X_n\ge b]\in\mathscr{F}_n,$$

所以T_1是停时, 以此类推可得T_k是停时.

如下图, 下面对是某个$\omega\in\Omega$的轨道, 则$T_0=1,T_1=4,T_2=7,T_3=10,T_4=16,T_5=17$, 从$T_{2j}$到$T_{2j+1}$叫一次上穿.



令$U_a^b(X,k)$表示$\{X_n\}_{0\le n<k}$上穿$[a,b]$的次数. 则

$$[U_a^b(X,k)=j]=[T_{2j-1}\le k<T_{2j+1}]=[T_{2j-1}\le k]\cap[T_{2j+1}>k]\in\mathscr{F}_k.$$

所以$U_a^b(X,k)$是$\mathscr{F}_k$可测的随机变量.

定理 6.2.5 (上穿不等式) 设$k \geq 1$, $\{X_n\}_{n\leq k}$是上鞅, 则

$$\mathbb{E}(U_a^b(X,k)) \leq \frac{1}{b-a}\mathbb{E}(X_k-a)^-.$$

证明: 由Doob停止定理, $\mathbb{E}(X_{T_{2j+1}\wedge k}|\mathscr{F}_S) \leq X_{T_{2j}\wedge k}$, 两边取期望得$\mathbb{E}_{X_{T_{2j+1}\wedge k}} \geq \mathbb{E}X_{T_{2j}\wedge k}$. 所以

$$\begin{aligned}
0 &\geq \mathbb{E}(X_{T_{2j+1\wedge k}} - X_{T_{2j\wedge K}}) \\
&= \mathbb{E}((X_{T_{2j+1\wedge k}} - X_{T_{2j\wedge K}})(I_{[T_{2j}\leq k<T_{2j+1}]} + I_{[k\geq T_{2j+1}]}) \\
&\geq \mathbb{E}((X_k-a)I_{[T_{2j}\leq k<T_{2j+1}]}) + (b-a)\mathbb{E}I_{[T_{2j}\leq k<T_{2j+1}]}
\end{aligned}$$

注意$[k \geq T_{2j+1}] \supset [U_a^b(X,k) \geq j+1], [T_{2j} \leq k < T_{2j+1}] \supset [U_a^b(X,k) = j+1]$, 则

$$P(U_a^b(X,k) \geq j+1) \leq \frac{1}{b-a}\mathbb{E}[(X_k-a)^- I_{[U_a^b(X,k)=j]}].$$

两边对j求和可得$\mathbb{E}U_a^b(X,k) = \sum\limits_{j=0}^{\infty} P(U_a^b(X,k) \geq j+1) \leq \dfrac{1}{b-a}\mathbb{E}(X_k-a)^-.$ □

§ 6.3 鞅收敛定理

本节介绍Doob收敛定理, 最后再引入Kolmogorov强大数定律. Kolmogorov三级数定理暂时不讲.

定理 6.3.1 (Doob收敛定理) *设$\{X_n\}_{n\geq 0}$是上鞅, 若$\sup\limits_n \mathbb{E}|X_n| < +\infty$ (等价于$\sup\limits_n \mathbb{E}X_n^- < +\infty$), 则存在可积随机变量$X_\infty$使得$X_n \xrightarrow{a.s} X_\infty$.*

进一步, 若$\{X_n\}$非负, 则$\mathbb{E}(X_\infty|\mathscr{F}_n) \leq X_n$, a.s..

证明: 对任意$k \geq 1$, $a,b \in \mathbb{Q}$, $a < b$, 令$U_a^b(X,k)$表示$\{X_n\}_{n\leq k}$上穿$[a,b]$的次数, $U_a^b(X)$表示$\{X_n\}_{n\geq 0}$上穿$[a,b]$的次数, 则

$$\lim_{k\to\infty} U_a^b(X_k) = U_a^b(X), a.s..$$

由

$$\begin{aligned}
\mathbb{E}U_a^b(X) &= \lim_{k\to\infty}\mathbb{E}U_a^b(X,k) \qquad \text{(积分的单调收敛定理)} \\
&\leq \frac{1}{b-a}\lim_{k\to\infty}\mathbb{E}(X_k-a)^- \qquad \text{(上穿不等式)} \\
&\leq \frac{1}{b-a}\sup_n \mathbb{E}((X_n-a)^-) < +\infty.
\end{aligned}$$

所以$U_a^b(X) < \infty$, a.s..

令$W_a^b = \{\omega | \liminf\limits_{n\to\infty} X_n < a, \limsup\limits_{n\to\infty} X_n > b\}$(极限不存在的部分), 则$P(W_a^b) = 0$, 而$W \triangleq \{\omega|\{X_n(\omega)\}\text{极限不存在}\} = \bigcup\limits_{\substack{a,b\in\mathbb{Q}\\ a<b}} W_a^b$, 则$P(W) = 0$. 所以$\{X_n(\omega)\}$极限a.s.存在.

定义$X_\infty(\omega) = \begin{cases} \lim\limits_{n\to\infty} X_n(\omega), & \omega \in W^c, \\ 0, & \omega \in W, \end{cases}$(逐点定义), 则$X_n \to X_\infty$, a.s.. 所以

$$\mathbb{E}(|X_\infty||\mathscr{F}_n) = \mathbb{E}(\lim_{m\to\infty}|X_m||\mathscr{F}_n) \underbrace{\leq}_{\text{Fatou}} \liminf_{n\to\infty}\mathbb{E}(|X_m||\mathscr{F}_n) \leq X_n \qquad \text{(上鞅)} \qquad □$$

推论 6.3.2 设$\{X_n\}$是鞅(上鞅), 且$\{X_n\}$一致可积, 则存在可积随机变量X_∞, 使得$X_n \xrightarrow{L^1} X_\infty$, 且$X_n \xrightarrow{a.s.} X_\infty$且$\mathbb{E}(X_\infty|\mathscr{F}_n) = X_n(\le X_n), a.s.$.

证明: 只证明上鞅的情形. 由于$\{X_n\}$一致可积, 则$\sup\limits_n \mathbb{E}|X_n| < +\infty$. 由前一定理, 存在可积随机变量$X_\infty$使得$X_n \xrightarrow{a.s.} X_\infty$. 又$\{X_n\}$一致可积, 则利用一致可积的$L^1$判别准则可知$X_n \xrightarrow{L^1} X_\infty$. 对任意$A \in \mathscr{F}_n$, 由于$\int_A X_\infty \mathrm{d}P = \lim\limits_{m\to\infty}\int_A X_m \mathrm{d}P = \int_A X_n \mathrm{d}P$. (当$m > n$时$\mathbb{E}(X_m|\mathscr{F}_n) = X_n$). □

注: 如果不加一致可积的条件, 仅用Doob收敛定理的条件不能推出$X_n \xrightarrow{L^1} X_\infty$. 反例如下: 设$\{\xi_n\}$是i.i.d.r.v., $S_0 = 1, S_n = S_{n-1} + \xi_n, P(\xi_1 = 1) = P(\xi_1 = -1) = \dfrac{1}{2}$. 取$N = \inf\{n|S_n = 0\}$, 并令$X_n = S_{N\wedge n}$. 则$X_n$是非负鞅. 由前一定理, $X_n \xrightarrow{a.s.} X_\infty$, 且必有$X_\infty \equiv 0$(不可能收敛到$k > 0$. 如果$X_n = k > 0$, 则$X_{n+1} = k \pm 1$.)由于$\mathbb{E}X_n = \mathbb{E}X_0 = 1(\forall n)$, 且$X_\infty = 0$, 则不满足$X_n \xrightarrow{L^1} X_\infty$..

定理 6.3.3 (Lévy) 设X是可积随机变量.

(1)若$\{\mathscr{F}_n\}$是流, 则$\mathbb{E}(X|\mathscr{F}_n) \xrightarrow{L^1/a.s.} \mathbb{E}(X|\mathscr{F}_\infty)$.

(2)若$\{\mathscr{G}_n\}$是一列单调递减的σ-代数序列, 则$\mathbb{E}(X|\mathscr{G}_n) \xrightarrow{L^1/a.s.} \mathbb{E}\left(X|\bigcap\limits_{n=1}^{\infty}\mathscr{G}_n\right)$.

回顾: $\mathbb{E}(\xi|\mathscr{F}_n)$是一致可积鞅.

证明: (1)由前一推论, 存在可积随机变量X_∞, 使得$\mathbb{E}(X|\mathscr{F}_n) \xrightarrow{L^1/a.s.} X_\infty$. 下证$\mathbb{E}(X|\mathscr{F}_\infty) = X_\infty$, 即证$\int_A X_\infty \mathrm{d}P = \int_A X \mathrm{d}P, \forall A \in \bigcup\limits_{n=1}^{\infty}\mathscr{F}_n$. 存在$n_0$使得$A \in \mathscr{F}_{n_0}$, 所以

$$\begin{aligned}\int_A X_\infty \mathrm{d}P = \lim_{m\to\infty}\int_A \mathbb{E}(X|\mathscr{F}_m)\mathrm{d}P &\overset{m>n_0}{=\!=\!=} \int_A \mathbb{E}(X|\mathscr{F}_{n_0})\mathrm{d}P \\ &\overset{\text{条件期望定义}}{=\!=\!=\!=\!=} \int_A X \mathrm{d}P.\end{aligned}$$

(利用单调类定理可以证明$X_\infty \triangleq \sigma\left(\bigcup\limits_{n=1}^{\infty}\mathscr{F}_n\right) = \bigcup\limits_{n=1}^{\infty}(\mathscr{F}_n)$)

(2)需要用到反向鞅, 略. □

推论 6.3.4 设$p > 1$且$\{X_n\}$是鞅, $\sup\limits_n \mathbb{E}|X_n|^p < +\infty$, 则存在可积$r.v.X_\infty$, 使得$X_n \xrightarrow{a.s./L^p} X_\infty$.

证明: 由Doob收敛定理, 存在可积随机变量X_∞使得$X_n \xrightarrow{a.s.} X_\infty$. 所以

$$\begin{aligned}\mathbb{E}\sup_n |X_n|^p &= \lim_{k\to\infty}\mathbb{E}\Big(\sup_{n\le k}|X_n|^p\Big) \qquad \text{(单调收敛定理)} \\ &\le \lim_{k\to\infty}\frac{p}{p-1}\mathbb{E}|X_k|^p \qquad \text{(Doob不等式)} \\ &\le \frac{p}{p-1}\sup_n \mathbb{E}|X_n|^p < \infty,\end{aligned}$$

再由$|X_n - X_\infty|^p \le 2^{p-1}(|X_n|^p + |X_\infty|^p)$与控制收敛定理可得

$$\lim_{n\to\infty}\mathbb{E}|X_n - X_\infty|^p = \mathbb{E}(\lim_{n\to\infty}|X_n - X_\infty|^p).$$

□

命题 6.3.5 设$\{X_n\}$是鞅, $X_n \xrightarrow{a.s.} X_\infty$, 且存在可积$r.v.X$使得$|X_n| \le |X|$, 则$\mathbb{E}(X_n|\mathscr{F}_n) \xrightarrow{a.s./L^1} \mathbb{E}(X_\infty|\mathscr{F}_\infty)$.

证明: 令$U_n = \inf\limits_{k\ge n} X_k, V_n = \sup\limits_{k\ge n} X_k$, 则$U_n \nearrow X_\infty, V_n \searrow X_\infty, U_n \le X_n \le V_n$, 且$U_n, V_n$均可积. 对任一给定的$m \ge n$, 有

$$\mathbb{E}(U_n|\mathscr{F}_m) \le \mathbb{E}(X_m|\mathscr{F}_m) \le \mathbb{E}(V_n|\mathscr{F}_m).$$

让$m \to \infty$可得

$$\mathbb{E}(U_n|\mathscr{F}_\infty) \le \liminf_{m\to\infty} \mathbb{E}(X_m|\mathscr{F}_m) \le \limsup_{m\to\infty} \mathbb{E}(X_m|\mathscr{F}_m) \le \mathbb{E}(V_n|\mathscr{F}_\infty)(a.s.).$$

再让$n \to \infty$, 由单调收敛定理(或控制收敛定理, $|U_n| \le |X|, |V_n| \le |X|$), 可得

$$\lim_{m\to\infty} \mathbb{E}(X_m|\mathscr{F}_m) = \mathbb{E}(X_\infty|\mathscr{F}_\infty)(a.s.). \qquad (*)$$

由$|\mathbb{E}(X_m|\mathscr{F}_m)| \le \mathbb{E}(|X_m||\mathscr{F}_m) \le \mathbb{E}(|X||\mathscr{F}_m)$, 且$\{\mathbb{E}(|X||\mathscr{F}_m)\}$是一致可积鞅(回顾“一致可积”一节的定理), 所以$(*)$式也是L^1收敛, 且$\{\mathbb{E}(X_m|\mathscr{F}_m)\}$一致可积. □

6.3.1 Kolmogorov强大数定律

在证明之前先作一些准备.

定义 6.3.1 设$\{\xi_n\}$是一列随机变量, 称$\mathscr{D} = \bigcap\limits_{n=1}^{\infty} \sigma(\xi_j : j > n)$ 为$\{\xi_n\}_{n\ge1}$的**尾σ-代数**, $\mathscr{D}$中元叫$\{\xi_n\}_{n\ge1}$的**尾事件*(tail event)***.

定理 6.3.6 (Kolmogorov 0-1律) 设$\{\xi_n\}_{n\ge1}$相互独立, 则$\forall D \in \mathscr{D}, P(D) = 0$或$1$.

证明: 只需证$P^2(D) = P(D) \Leftrightarrow P(D\cap D) = P^2(D) \Leftrightarrow \mathscr{D}$与$\mathscr{D}$独立.

由独立类扩张定理, $\forall n \ge 1, \sigma(\xi_1, \cdots, \xi_n)$与$\sigma(\xi_{n+1}, \xi_{n+2}, \cdots)$独立, 所以$\sigma(\xi_1, \cdots, \xi_n)$与$\mathscr{D}$独立. 令$\mathscr{A} = \bigcup\limits_{n=1}^{\infty} \sigma(\xi_1, \cdots, \xi_n)$, 则$\mathscr{A}$和$\mathscr{D}$独立(定义), 由于$\mathscr{A}$是$\pi$类, 由独立类扩张定理, $\sigma(A)$与$\mathscr{D}$独立. 由$\mathscr{D} \subset \sigma(\mathscr{A})$, 则$\mathscr{D}$与$\mathscr{D}$独立, 则对任意$D \in \mathscr{D}$, 有$P(D) = P(D\cap D) = P(D)^2 \Rightarrow P(D) = 0$或$P(D) = 1$. □

注: 函数$f : (\Omega, \mathscr{F}_\infty) \to (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ 叫尾函数(tail function). 尾函数几乎必然是常数, 考虑尾事件$[f < c]$, 其中$c \in \overline{\mathbb{R}}$, 则$P(f < c) = 0$或1, 如果$k = \sup\{c \in \overline{\mathbb{R}} : P(f < c) = 0\}$, 则$f = k$(a.e.).

命题 6.3.7 设X, Y, Z是随机变量, $(X, Z), (Y, Z)$有相同的联合分布, f是非负Borel可测函数, 则$\mathbb{E}[f(X)|Z] = \mathbb{E}[f(Y)|Z]$, *a.s.*.

证明: 对$A \in \sigma(Z)$, 根据$f(X), f(Y)$同分布, 可得

$$\int_A \mathbb{E}[f(X)|Z]\mathrm{d}P = \int_A f(X)\mathrm{d}P = P\circ[f(X)]^{-1}(A) = P\circ[f(Y)]^{-1}(A) = \int_A f(Y)\mathrm{d}P = \int_A \mathbb{E}[f(Y)|Z]\mathrm{d}P.$$

因此$\mathbb{E}[f(X)|Z] = \mathbb{E}[f(Y)|Z], a.s.$. □

注: 若X, Y是i.i.d.r.v., 由前一命题, $\mathbb{E}(X|X+Y) = \mathbb{E}(Y|X+Y)$, 从而$\mathbb{E}(X|X+Y) = \frac{1}{2}\mathbb{E}(X+Y|X+Y) = \frac{X+Y}{2}$, a.s..

定理 6.3.8 (Kolmogorov强大数定律) 设$\{\xi_n\}$是$i.i.d.r.v.$列, $\mathbb{E}|\xi_1| < \infty$, 则$\frac{1}{n}\sum_{i=1}^{n}\xi_i \xrightarrow{a.s.} \mathbb{E}\xi_1$.

证明: 由前一命题可知

$$\mathbb{E}\Big(\xi_1\Big|\sum_{i=1}^{n}\xi_i\Big) = \mathbb{E}\Big(\xi_j\Big|\sum_{i=1}^{n}\xi_i\Big), j = 1, 2, \cdots, n.$$

所以

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\xi_i &= \mathbb{E}\Big(\xi_1\Big|\sum_{i=1}^{n}\xi_i\Big) \qquad \text{(用前面的“注”的处理方式)} \qquad (*)\\ &= \mathbb{E}\Big(\xi_1\Big|\sum_{i=1}^{n}\xi_i, \xi_{n+1}, \xi_{n+2}, \cdots\Big) \qquad \text{(独立)}\\ &= \mathbb{E}\Big(\xi_1\Big|\sum_{i=1}^{n}\xi_i, \sum_{i=1}^{n+1}\xi_i, \sum_{i=1}^{n+2}\xi_i, \cdots\Big) \qquad \text{(相互可测)}\end{aligned}$$

令$\mathscr{G}_n = \sigma\Big(\sum_{i=1}^{n}\xi_i, \sum_{i=1}^{n+1}\xi_i, \sum_{i=1}^{n+2}\xi_i, \cdots\Big)$, 则$\mathscr{G}_n \searrow$. 由Lévy定理(定理6.3.3), $\frac{1}{n}\sum_{i=1}^{n}\xi_i \xrightarrow{a.s.} \mathbb{E}\Big(\xi_1\Big|\bigcap_{n=1}^{\infty}\mathscr{G}_n\Big)$.

由$\mathbb{E}\Big(\xi_1\Big|\bigcap_{n=1}^{\infty}\mathscr{G}_n\Big)$关于$\mathscr{D}$可测, 则它是个尾函数, 根据Kolmogorov 0-1律, $\mathbb{E}\Big(\xi_1\Big|\bigcap_{n=1}^{\infty}\mathscr{G}_n\Big)$a.s.等于常数, 而且对$(*)$式两边取期望可得这个常数就是$\mathbb{E}\xi_1$. □

6.3.2 上鞅的分解定理

定义 6.3.2 (1)称$r.v.$列$\{X_n\}_{n\geq 0}$关于$\{\mathscr{F}_n\}_{n\geq 0}$**可料*(predictable)***, 若$X_0 \in \mathscr{F}$且$X_n \in \mathscr{F}_{n-1}, \forall n \geq 1$.

(2)称$r.v.$列$\{A_n\}_{n\geq 0}$为**增过程**, 若$0 = A_0 \leq A_1 \leq \cdots \leq A_n \leq \cdots$. 进一步, 若$\lim_{n\to\infty}\mathbb{E}A_n < +\infty$, 则称$\{A_n\}$**可积**.

定理 6.3.9 (Doob分解) 设$\{X_n, \mathscr{F}_n\}_{n\geq 0}$是上鞅, 则它可以唯一分解为

$$X_n = M_n - A_n (\forall n \geq 0), \tag{6.2}$$

其中$\{M_n, \mathscr{F}_n\}$是鞅, $\{A_n\}_{n\geq 0}$是可料增过程.

证明: 设(M_n, A_n)满足(6.2)式的分解, 则

$$A_{n+1} - A_n = \mathbb{E}(A_{n+1} - A_n|\mathscr{F}_n) = \mathbb{E}(M_{n+1} - X_{n+1} - M_n + X_n|\mathscr{F}_n) = X_n - \mathbb{E}(X_{n+1}|\mathscr{F}_n), \forall n \geq 1.$$

所以

$$A_n = \sum_{j=0}^{n-1}\underbrace{(X_j - \mathbb{E}(X_{j+1}|\mathscr{F}_j))}_{\geq 0,\text{上鞅}} + \underbrace{A_0}_{=0}. \tag{6.3}$$

这说明了分解的唯一性, 且A_n是可料增过程.

下设$A_0 = 0$且A_n是由(6.3)式定义的过程, 则$M_0 = X_0$, 且

$$M_n = X_n + A_n = X_n + \sum_{j=0}^{n-1}(X_j - \mathbb{E}(X_{j+1}|\mathscr{F}_j)) = X_0 + \sum_{j=1}^{n}(X_j - \mathbb{E}(X_j|\mathscr{F}_{j-1})),$$

所以$M_n - M_{n-1} = X_n - \mathbb{E}(X_n|\mathscr{F}_{n-1})$, 进一步有

$$\mathbb{E}(M_n|\mathscr{F}_{n-1}) = M_{n-1} + \underbrace{\mathbb{E}[(X_n - \mathbb{E}(X_n|\mathscr{F}_{n-1}))|\mathscr{F}_{n-1}]}_{=0} = M_{n-1}.$$

所以$(M_n, \mathscr{F}_n)$是鞅. □

定义 6.3.3 (1)称非负上鞅$\{X_n\}$为**位势**, 若$\lim\limits_{n\to\infty} \mathbb{E}X_n = 0$.
(2)设$\{X_n\}$是上鞅, 称$X_n = Y_n + Z_n$为$\{X_n\}$的***Riesz*分解**, 其中$\{Y_n\}$为鞅, $\{Z_n\}$是位势.

注: 位势是一致可积上鞅.

Riesz分解是唯一的: 设$X_n = Y_n + Z_n = Y_n' + Z_n'$, 则$Z_n - Z_n' = Y_n' - Y_n$是鞅, 故$\{Z_n - Z_n'\}$是一致可积鞅, 从而$Z_n - Z_n' = \mathbb{E}(0|\mathscr{F}_n) = 0, \forall n \geq 1$.

定理 6.3.10 设$\{X_n\}$是上鞅. 则X_n有*Riesz*分解$\Leftrightarrow \lim\limits_{n\to\infty} \mathbb{E}X_n > -\infty$.

证明: "$\Rightarrow$": 必要性显然.

"$\Leftarrow$": 由Doob分解定理, $X_n = M_n - A_n$, 其中M_n是鞅, A_n是零初值可料增过程. 由$\lim\limits_{n\to\infty} \mathbb{E}X_n > -\infty$可知$\lim\limits_{n\to\infty} \mathbb{E}A_n < +\infty$. 记$A_\infty = \lim\limits_{n\to\infty} A_n$, 则由单调收敛定理, A_∞可积.

最后我们说明

$$X_n = \underbrace{M_n - \mathbb{E}(A_\infty|\mathscr{F}_n)}_{\text{鞅}} + \underbrace{\mathbb{E}(A_\infty|\mathscr{F}_n) - A_n}_{\text{位势}},$$

显然$M_n - \mathbb{E}(A_\infty|\mathscr{F}_n)$是鞅. 由于

$$\begin{aligned}\mathbb{E}(\mathbb{E}(A_\infty|\mathscr{F}_{n+1}) - A_{n+1}|\mathscr{F}_n) &= \mathbb{E}(\mathbb{E}(A_\infty|\mathscr{F}_{n+1})|\mathscr{F}_n) - A_{n+1} \qquad (A_{n+1}\text{关于}\mathscr{F}_n\text{可测, 可料定义}) \\ &= \mathbb{E}(A_\infty|\mathscr{F}_n) - A_{n+1} \leq \mathbb{E}(A_\infty|\mathscr{F}_n) - A_n \qquad (\text{增过程}).\end{aligned}$$

所以$\mathbb{E}(A_\infty|\mathscr{F}_n) - A_n$是上鞅. 又由于$\mathbb{E}(A_\infty|\mathscr{F}_n) - A_n \geq 0$, 故$\mathbb{E}(A_\infty|\mathscr{F}_n) - A_n$是位势. □

§6.4　大数定律

设$\{X_n\}_{n\geq 1}$是r.v.列, $S_n = \sum\limits_{i=1}^n X_i, n \geq 1$. 是否存在$a_n, b_n$使得$\dfrac{S_n - a_n}{b_n} \to 0$? (依概率收敛: 弱大数定律; a.s.收敛: 强大数定律.)

定理 6.4.1 设$\{X_n\}_{n\geq 1}$两两不相关且$\sup\limits_n \mathbb{D}X_n < +\infty$, 则$\dfrac{S_n - \mathbb{E}S_n}{n} \xrightarrow{P} 0$.

证明: $\mathbb{E}\left(\dfrac{S_n - \mathbb{E}S_n}{n}\right)^2 = \dfrac{1}{n^2}\mathbb{E}\left(\sum\limits_{i=1}^n (X_i - \mathbb{E}X_i)\right)^2 = \dfrac{1}{n^2}\sum\limits_{i=1}^n DX_i \leq \dfrac{1}{n}\sup\limits_n \mathbb{D}X_n \to 0 (n\to\infty)$. 所以$P\left(\left|\dfrac{S_n - \mathbb{E}S_n}{n}\right| \geq \varepsilon\right) \leq \dfrac{1}{\varepsilon^2}\displaystyle\int_\Omega \left(\frac{S_n - \mathbb{E}S_n}{n}\right)^2 \mathrm{d}P \to 0 (n \to \infty)$. □

注: 若$\{X_n\}$i.i.d.且$\mathbb{E}X^2 < +\infty$, 则$\dfrac{S_n}{n} \xrightarrow{P} \mathbb{E}X_1$.

例 6.4.1 (多项式逼近) 设$f \in C^0[0,1]$, 令$f_n(x) = \sum\limits_{m=0}^n \dbinom{n}{m} x^m (1-x)^{n-m} f\left(\dfrac{m}{n}\right)$,
则$\lim\limits_{n\to\infty} \sup\limits_{x\in[0,1]} |f_n(x) - f(x)| = 0$.

证明: 对任意$x \in [0,1]$, 设$\{\xi_n\}$是i.i.d.r.v., 满足两点分布$P(\xi_1 = 1) = x, P(\xi_1 = 0) = 1 - x$.
则$\mathbb{E}f\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) = \sum_{m=0}^{n} P\left(\sum_{i=1}^{n}\xi_i = m\right) f\left(\frac{m}{n}\right) = f_n(x)$.
注意$\forall\varepsilon > 0, \exists\delta > 0$, 当$x, y \in [0,1]$满足$|x-y| \le \delta$时, $|f(x) - f(y)| \le \varepsilon$, 从而有

$$\begin{aligned}
|f_n(x) - f(x)| &= \left|\mathbb{E}f\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) - f(x)\right| \\
&\le \left|\mathbb{E}\left[f\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) - f(x)\right]\left(I_{[|\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i - x| \ge \delta]} + I_{[|\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i - x| < \delta]}\right)\right| \\
&\le 2\|f\|_\infty \mathbb{E}I_{[|\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i - x| \ge \delta]} + \varepsilon \le 2\|f\|_\infty \frac{1}{n^2\delta^2} n\mathbb{D}\xi_1 + \varepsilon \qquad \text{(Chebyshev)} \\
&= 2\|f\|_\infty \frac{1}{n\delta^2} x(1-x) + \varepsilon \le 2\|f\|_\infty \frac{1}{n\delta^2}\cdot\frac{1}{4} + \varepsilon
\end{aligned}$$

上式对任意$x \in [0,1]$都成立, 所以$\sup\limits_{x\in[0,1]} |f_n(x) - f(x)| \le \frac{\|f\|_\infty}{2n\delta^2} + \varepsilon$. 让$n \to \infty$可得欲证结论. □

6.4.1 截断

对随机变量X在M处截断指考虑考虑随机变量$\overline{X} = XI_{[|X| \le M]}$. 下面考虑随机变量三角矩阵列

$$\begin{matrix}
X_{11} & & & & \\
X_{21} & X_{22} & & & \\
\vdots & \vdots & \vdots & & \\
X_{n1} & X_{n2} & \cdots & X_{ni} & \cdots & X_{nn}
\end{matrix}$$

对任意$n \ge 1$, $\{X_{ni}\}_{i \le n}$相互独立. 那么我们有

定理 6.4.2 设$0 < b_n \nearrow \infty$, 记$\overline{X_{ni}} = X_{ni}I_{[|X_{ni}| \le b_n]}$. *若满足:* ①$\sum\limits_{i=1}^{n} P(|X_{ni}| > b_n) \to 0 (n \to \infty)$, 且②$\frac{1}{b_n^2}\sum\limits_{i=1}^{n}\mathbb{E}|\overline{X_{ni}}|^2 \to 0 (n \to \infty)$, 则$\frac{1}{b_n}\left(\sum\limits_{i=1}^{n} X_{ni} - \mathbb{E}\sum\limits_{i=1}^{n}\overline{X_{ni}}\right) \xrightarrow{P} 0$.

证明: 记$S_n = \sum\limits_{i=1}^{n} X_{ni}, \overline{S_n} = \sum\limits_{i=1}^{n}\overline{X_{ni}}, a_n = \mathbb{E}\overline{S_n}$. 对任意$\delta > 0$,

$$\begin{aligned}
P\left(\left|\frac{S_n - a_n}{b_n}\right| \ge \delta\right) &= P\left(\left[\left|\frac{S_n - a_n}{b_n}\right| \ge \delta\right] \cap [S_n = \overline{S_n}]\right) + P\left(\left[\left|\frac{S_n - a_n}{b_n}\right| \ge \delta\right] \cap [S_n \ne \overline{S_n}]\right) \\
&\le P\left(\left|\frac{\overline{S_n} - a_n}{b_n}\right| \ge \delta\right) + P(S_n \ne \overline{S_n}) \\
&\le \frac{\mathbb{E}(\overline{S_n} - \mathbb{E}\overline{S_n})^2}{\delta^2 b_n^2} + \sum_{i=1}^{n} P(X_{ni} \ne X_{ni}I_{[|X_{ni} \le b_n]}) \\
&\le \frac{1}{\delta^2 b_n^2}\sum_{i=1}^{n}\mathbb{E}|\overline{X_{ni}}|^2 + \sum_{i=1}^{n} P(|X_{ni}| > |b_n|).
\end{aligned}$$

让$n \to \infty$即可得欲证结论. □

定理 6.4.3 (弱大数定律) 设$\{X_n\}i.i.d.$, 且$\lim\limits_{n\to\infty} xP(|X_1|>x)=0$, 则

$$\frac{1}{n}\sum_{i=1}^{n}X_i-\mathbb{E}(X_1I_{[|X_1|\le n]})\xrightarrow{P}0.$$

证明: 对任意$n\ge 1$, 在前一定理中取$X_{ni}=X_i, 1\le i\le n$, 且$b_n=n$. 只需验证两个条件即可:
(1)由条件可得$\sum\limits_{i=1}^{n}P(|X_{ni}|>b_n)=nP(|X_i|>n)\to 0$.
(2)记$\overline{X_i}=X_iI_{[|X_i|\le n]}$, 则

$$\begin{aligned}\frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}|\overline{X_i}|^2=\frac{1}{n}\mathbb{E}|\overline{X_1}|^2=\frac{2}{n}\int_0^\infty yP(|\overline{X_1}|>y)\mathrm{d}y&=\frac{2}{n}\int_0^n yP(|\overline{X_1}|>y)\mathrm{d}y\\&=\frac{2}{n}\int_0^n y(P(|X_1|>y)-P(|X_1|>n))\mathrm{d}y\\&\le\frac{2}{n}\int_0^n yP(|X_1|>y)\mathrm{d}y\to 0(n\to\infty).\end{aligned}$$

注: $xP(|X_1|>x)\le\mathbb{E}|X_1|I_{[|X_1|\ge x]}$, 所以这个定理的条件很弱.

6.4.2 强大数定律

定理 6.4.4 设$\{X_n\}_{n\ge1}$是独立$r.v.$列, 若$\sum\limits_{n=1}^{\infty}\mathbb{D}X_n<+\infty$, 则$\sum\limits_{n=1}^{\infty}(X_n-\mathbb{E}X_n)$ $a.s.$收敛.

证明: 不妨设$EX_n=0,\forall n\ge 1$. 令$S_n=\sum\limits_{i=1}^{n}X_i,\forall n\ge 1$. 先对欲证命题进行转化:

$$\begin{aligned}S_n\text{a.s.收敛}\Leftrightarrow&S_n-S_m\xrightarrow{a.s.}0(n,m\to\infty)\ (\text{Cauchy准则})\\\Leftrightarrow&\forall\varepsilon>0,P\left(\bigcap_{k=1}^{\infty}\bigcup_{m,n=k}^{\infty}[|S_n-S_m|\ge\varepsilon]\right)=0\ (\text{a.s.收敛刻画})\\\Leftrightarrow&\forall\varepsilon>0,\lim_{n\to\infty}P\left(\bigcup_{m,n=k}^{\infty}[|S_n-S_m|\ge\varepsilon]\right)=0\qquad(\text{单调性})\\\Leftrightarrow&\forall\varepsilon>0,P\left(\bigcup_{l=1}^{\infty}[|S_{l+k}-S_k|\ge\varepsilon]\right)\to0(k\to\infty).\end{aligned}$$

固定上面的k, 注意到

$$\begin{aligned}P\left(\bigcup_{l=1}^{\infty}[|S_{l+k}-S_k|\ge\varepsilon]\right)&=\lim_{m\to\infty}P\left(\bigcup_{l=1}^{m}[|S_{l+k}-S_k|\ge\varepsilon]\right)\\&=\lim_{m\to\infty}P\left(\max_{1\le l\le m}|S_{l+k}-S_k|\ge\varepsilon\right)\\&\le\lim_{m\to\infty}\frac{1}{\varepsilon^2}D|S_{m+k}-S_k|\ (\text{极大值不等式, 定理6.2.3})\\&=\frac{1}{\varepsilon^2}\lim_{m\to\infty}\sum_{i=k+1}^{k+m}DX_i=\frac{1}{\varepsilon^2}\sum_{i=k+1}^{\infty}DX_i\to0(k\to\infty).\end{aligned}$$

这样证明已完成. □

引理 6.4.5 (Kronecker) 设$\sum_{k=1}^{\infty} a_k$存在, 且$p_n \nearrow \infty$, 则$\lim_{n\to\infty} \frac{p_1 a_1 + \cdots + p_n a_n}{p_n} = 0$.

定理 6.4.6 (Kolmogorov强大数定律) 设$\{X_n\}_{n\geq 1}$为独立$r.v.$列, $\{b_n\}$单调递增趋于正无穷, 若

$$\sum_{n=1}^{\infty} \frac{D(X_n)}{b_n^2} < +\infty,$$

则$\frac{1}{b_n}\sum_{i=1}^{n}(X_i - EX_i) \xrightarrow{a.s.} 0$.

证明: 根据定理6.4.4, $\sum_{n=1}^{\infty} \frac{X_n}{b_n}$几乎必然收敛. 根据Kronecker引理,

$$\frac{S_n - ES_n}{b_n} = \frac{1}{b_n}\sum_{k=1}^{n}\left(\frac{X_k - EX_k}{b_k}\right) b_k \xrightarrow{a.s.} 0. \qquad \square$$

推论 6.4.7 设$\{X_n\}_{n\geq 1}$相互独立$r.v.$列, 且有相同的均值与方差, 则$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} EX_1$.

证明: 在Kolmogorov强大数定律中令$b_n = n, D(X_n) = C$即可. □

例 6.4.2 (强大数定律) 设$\{\xi_n\}_{n\geq 1}$是$i.i.d.r.v.$列, 则

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{a.s.} a(< +\infty) \Leftrightarrow \xi_1 \text{可积且} a = \mathbb{E}\xi_1.$$

证明: “⇒”: 如下.

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{a.s.} a \Rightarrow & \frac{\xi_n}{n} = \frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{n-1}{n}\left(\frac{1}{n-1}\sum_{i=1}^{n-1}\xi_i\right) \xrightarrow{a.s.} 0 \\
& \Rightarrow P\left(\left[\left|\frac{\xi_n}{n}\right| \geq 1\right] \text{发生无穷次}\right) = 0 \text{ (a.s.收敛的刻画)} \\
& \Rightarrow \sum_{n=1}^{\infty} P\left(\left|\frac{\xi_n}{n}\right| \geq 1\right) < +\infty \text{ (独立性, Borel-Cantelli)} \\
& \Rightarrow \sum_{n=1}^{\infty} P(|\xi_n| \geq n) < +\infty \text{ (同分布)} \\
& \Rightarrow \mathbb{E}|\xi_1| \leq \sum_{n=1}^{\infty} P(|\xi_n| \geq n) + 1 < +\infty. \text{ (Nice引理)}
\end{aligned}$$

另外,

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \xrightarrow{a.s.} a \Rightarrow \mathbb{E}\frac{1}{n}\sum_{i=1}^{n}\xi_i \xrightarrow{a.s.} \mathbb{E}a = a,$$

根据同分布可得$\mathbb{E}\xi_i \xrightarrow{a.s.} a$, 而$E\xi_i$是常数, 所以$\mathbb{E}\xi_i = a$. □

“$\Leftarrow$”：不妨设$\mathbb{E}\xi_1=0$(否则可以作中心化处理). 令$\eta_i=\xi_i I_{[|\xi_i|<i]},\forall i\geq 1$, 则

$$\begin{aligned}\sum_{n=1}^{\infty}P(\xi_n\neq\eta_n)&\leq\sum_{n=1}^{\infty}P(|\xi_n|\geq n)\ \text{(i.i.d)}\\&\leq\mathbb{E}|\xi_1|<\infty\ \text{(Nice引理)}.\end{aligned}$$

根据Borel-Cantelli引理, 可得$P(\xi\neq\eta_i\text{发生无穷多次})=0$. 也就是说只有有限个$\xi_i\neq\eta_i$, 于是

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i\xrightarrow{a.s.}0\Leftrightarrow\frac{1}{n}\sum_{i=1}^{n}\eta_i\xrightarrow{a.s.}0\Leftrightarrow\frac{1}{n}\sum_{i=1}^{n}(\eta_i-E\eta_i)\xrightarrow{a.s.}0.$$

考虑到

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{D}\eta_n<\infty&\Rightarrow\sum_{n=1}^{\infty}\frac{\eta_n-\mathbb{E}\eta_n}{n}\text{a.s.收敛(定理6.4.4)}\\&\Rightarrow\frac{1}{n}\sum_{i=1}^{n}(\eta_i-\mathbb{E}\eta_i)\xrightarrow{a.s.}0.\ \text{(Kronecker)}\end{aligned}$$

只需证$\sum\limits_{n=1}^{\infty}\frac{1}{n^2}\mathbb{D}\eta_n<\infty$. 事实上,

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{D}\eta_n&\leq\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\xi_n^2 I_{[|\xi_n|<n]}\ \text{(方差定义)}\\&=\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{k=0}^{n-1}\mathbb{E}\xi_1^2 I_{[k\leq|\xi_1|<k+1]}=\sum_{k=0}^{\infty}\sum_{n=k+1}^{\infty}\frac{1}{n^2}\mathbb{E}\xi_1^2 I_{[k\leq|\xi_1|<k+1]}\\&<\sum_{k=1}^{\infty}\sum_{n=k+1}^{\infty}\left(\frac{1}{n(n-1)}\mathbb{E}\xi_1^2 I_{[k\leq|\xi_1|<k+1]}\right)+\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}\xi_1^2 I_{[0\leq|\xi_1|<1]}\ \text{(提取}k=0)\\&\leq\sum_{k=1}^{\infty}\left(\frac{1}{k}\mathbb{E}\xi_1^2 I_{[k\leq|\xi_1|<k+1]}\right)+\sum_{n=1}^{\infty}\frac{1}{n^2}\ \text{(后一项用}|\xi_1|<1\text{放缩)}\\&\leq\sum_{k=1}^{\infty}\left(\frac{(k+1)^2}{k}P(k\leq|\xi_1|<k+1)\right)+\frac{\pi^2}{6}\ \text{(前面项用}|\xi_n|<k+1\text{放缩)}\\&\leq 3\sum_{k=1}^{\infty}kP(k\leq|\xi_1|<k+1)+\frac{\pi^2}{6}\leq 3\sum_{k=1}^{\infty}\mathbb{E}|\xi_1|I_{[k\leq|\xi_1|<k+1]}+\frac{\pi^2}{6}\\&\leq 3\mathbb{E}|\xi_1|+\frac{\pi^2}{6}<+\infty.\square\end{aligned}$$

定理 6.4.8 设$\{X_n\}i.i.d.$, 且$\mathbb{E}X_1=+\infty$, 则$\frac{1}{n}\sum\limits_{i=1}^{n}X_i\xrightarrow{a.s.}+\infty$.

证明： 给定m, 令$Y_i=X_i\wedge m(\forall i\geq 1)$, 则$\{Y_i\}$i.i.d., 且$\mathbb{E}|Y_1|<+\infty$. 由前一定理, $\frac{1}{n}\sum\limits_{i=1}^{n}Y_i\rightarrow\mathbb{E}Y_1=\mathbb{E}X_1\wedge m$, 再由

$$\mathbb{E}X_1\wedge m=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Y_i\leq\liminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}X_i.$$

因此让$m\to\infty$可得$\liminf\limits_{n\to\infty}\frac{1}{n}\sum\limits_{i=1}^{n}X_i=+\infty$. $\square$

定理 6.4.9 (Glivenko-Cantelli) 设$\{X_n\}i.i.d.$, 分布函数为F, 定义**经验分布函数**为

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[X_i\le x]},$$

则$\sup\limits_x |F_n(x) - F(x)| \xrightarrow{a.s.} 0.$

证明: 见Durrett书的定理2.4.7.

定理 6.4.10 (Marcinkiewicz-Zygmund) 设$\{X_n\}i.i.d.$, $\mathbb{E}X_1 = 0$, 对$1 < p < 2$, 有

$$\frac{1}{n^{1/p}}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} 0 \Leftrightarrow \mathbb{E}|X_1|^p < +\infty.$$

证明: 见Durrett书的定理2.5.8.

强大数定律(定理6.4.2)的逆命题不对, 需要加一致有界的条件.

定理 6.4.11 设$\{X_n\}$是独立$r.v.$列, $\mathbb{E}X_n = 0(\forall n > 0)$, 且存在$C > 0$使得

$$|X_n| \le C, a.s., \forall n \ge 1.$$

则:

(1)$\forall \varepsilon > 0, P\Big(\max\limits_{1\le j\le n} |S_j| \le \varepsilon\Big) \le \dfrac{(C+\varepsilon)^2}{\mathbb{E}S_n^2}$.

(2)若$\sum\limits_{n=1}^{\infty} X_n\, a.s.$收敛, 则$\sum\limits_{n=1}^{\infty} \mathbb{D}X_n < +\infty$.

证明: (1)一般出现$\max\limits_{1\le j\le n} |S_j|$这种东西都考虑停时.

令$T = \inf\{k \ge 1 | |S_k| \le \varepsilon\}$, 则$T$为停时, 且

$$[T \ge k] = \bigcap_{j=1}^{k-1} [|S_j| \le \varepsilon] \in \sigma(X_1, X_2, \cdots, X_{k-1})\text{与}\sigma(X_k, X_{k+1}, \cdots)\text{独立}.$$

又$|S_{T\wedge k}| \le \begin{cases} |S_{T-1}| + |X_T|, & T \le n \\ \varepsilon, & T > n \end{cases} \le C + \varepsilon$, 从而

$$\begin{aligned}(C+\varepsilon)^2 \ge \mathbb{E}|S_{T\wedge n}|^2 &= \sum_{i=1}^{n} \mathbb{E}|X_i|^2 I_{[T\ge i]} + 2\sum_{1\le i<j\le n} \mathbb{E}\, \underbrace{X_j}\, \underbrace{I_{[T\ge j]} X_i I_{[T\ge i]}}_{\text{独立}} \\ &\ge \sum_{i=1}^{n} \mathbb{E}X_i^2 P(T\ge i) + 0 \ge \sum_{i=1}^{n} \mathbb{E}X_i^2 P(T \ge n) \\ &= \sum_{i=1}^{n} \mathbb{D}X_i P\Big(\max_{1\le k\le n} |S_k| \le \varepsilon\Big). \qquad \square\end{aligned}$$

注: $\sum\limits_{j=1}^{T\wedge n} X_j = \sum\limits_{j=1}^{n} X_j I_{[T\wedge j]}$这样写会好很多.

§6.5 三级数定理

下面记$X_n^a \triangleq X_n I_{[|X_n|\le a]} < +\infty$.

定理 6.5.1 (三级数定理) 设$\{X_n\}$是独立$r.v.$列, 则

$$\sum_{n=1}^{\infty} X_n\, a.s.\text{收敛} \Leftrightarrow \forall a\in(0,+\infty), \begin{cases} ① \sum\limits_{n=1}^{\infty} P(|X_n|>a)<+\infty; \\ ② \sum\limits_{n=1}^{\infty} \mathbb{E}X_n^a<+\infty; \\ ③ \sum\limits_{n=1}^{\infty} \mathbb{D}X_n^a<+\infty; \end{cases}$$

注: 在必要性中, ①+②+③对所有$a\in(0,+\infty)$均成立. 在充分性中, 若对**某个**$a\in[0,+\infty]$(可以取0或∞), ①+②+③成立, 则$\sum\limits_{n=1}^{\infty} X_n$a.s.收敛.

证明: "$\Leftarrow$": 若$a\in(0,+\infty)$, 由③可得$\sum\limits_{n=1}^{\infty}(X_n^a-\mathbb{E}X_n^a)$a.s.收敛(定理6.4.4). 所以由②可得$\sum\limits_{n=1}^{\infty} X_n^a$a.s.收敛. 由①以及$P(X_n\ne X_n^a, i.o.)=0$, 则$\sum\limits_{n=1}^{\infty} X_n$a.s.收敛.

"$\Rightarrow$": (1)由条件可得$\lim\limits_{n\to\infty} X_n=0(a.s.)$, 等价于$P([|X_n|>a], i.o.)=0$, 由Borel-Cantelli引理, 这也等价于$\sum\limits_{n=1}^{\infty} P(|X_n|>a)<+\infty$.

(2)取$\{Y_n\}_{n\ge1}$为$\{X_n^a\}$的独立复制(即$\{X_n^a\}$与$\{Y_n\}$独立同分布), 由于$P([|X_n|>a], i.o.)=0$, 所以$\sum\limits_{n=1}^{\infty} X_n^a$a.s.收敛. 则$\sum\limits_{n=1}^{\infty} Y_n$a.s.收敛. 则$\sum\limits_{n=1}^{\infty}(X_n^a-Y_n)$a.s.收敛.

又$\mathbb{E}X_n^a=\mathbb{E}Y_n$, 且$|X_n^a-Y_n|\le 2a$, a.s., $(\forall n\ge1)$. 由定理6.4.11, $\sum\limits_{n=1}^{\infty}\mathbb{D}(X_n^a-Y_n)<+\infty$, 从而$\sum\limits_{n=1}^{\infty}\mathbb{D}X_n^a=\frac{1}{2}\sum\limits_{n=1}^{\infty}\mathbb{D}(X_n^a-Y_n)<+\infty$, 故$\sum\limits_{n=1}^{\infty}(X_n^a-\mathbb{E}X_n^a)$a.s.收敛(定理6.4.4).

又知$\sum\limits_{n=1}^{\infty} X_n^a$a.s.收敛, 则$\sum\limits_{n=1}^{\infty}\mathbb{E}X_n^a<+\infty$. □

注: 这里证明的时候采用了独立复制的技巧, 不可以使用中心化, 因为不知道中心化之后的序列如何表现.

例 6.5.1 设$\{X_n\}_{n\ge1}$相互独立, 且对每个n, X_n服从参数为λ_n的$Poisson$分布, 则$\sum\limits_{n=1}^{\infty} X_n\, a.s.$收敛$\Leftrightarrow \sum\limits_{n=1}^{\infty}\lambda_n<+\infty$.

证明: $P(X_n=k)=e^{-\lambda_n}\frac{\lambda_n^k}{k!}, \lambda_n>0$.

"$\Leftarrow$": 易证$\sum\limits_{n=1}^{\infty}\mathbb{D}X_n^a<+\infty$, 且$\sum\limits_{n=1}^{\infty}\mathbb{E}X_n^a<+\infty$.

"$\Rightarrow$": 取$a=\frac{1}{2}$, 则$\sum\limits_{n=1}^{\infty}P(X_n>\frac{1}{2})<+\infty$, 所以$\sum\limits_{n=1}^{\infty}(1-e^{-\lambda_n})<+\infty$, 故$\lambda_n\to0$.

所以$\sum_{n=1}^{\infty}\lambda_n=\sum_{n=1}^{\infty}(1-e^{-\lambda_n})\cdot\underbrace{\frac{\lambda_n}{1-e^{-\lambda_n}}}_{\to 1}<+\infty.$ □

例 6.5.2 设$\{X_n\}_{n\geq 1}$独立, 且对每个n, $X_n\sim N(\mu_n,\sigma_n^2)$, 则$\sum_{n=1}^{\infty}X_n\,a.s.$收敛$\Leftrightarrow\begin{cases}\sum_{n=1}^{\infty}\sigma_n^2<+\infty\\ \sum_{n=1}^{\infty}\mu_n<+\infty\end{cases}.$

例 6.5.3 设$\{X_n\}_{n\geq 1}$独立, 且对每个n, $X_n\sim E(\lambda_n)$, 则$\sum_{n=1}^{\infty}X_n\,a.s.$收敛$\Leftrightarrow\sum_{n=1}^{\infty}\lambda_n^{-1}<+\infty$.

定理 6.5.2 设$\{X_n\}$是独立$r.v.$列, 则$\sum_{n=1}^{\infty}X_n\,a.s.$收敛$\Leftrightarrow\sum_{n=1}^{\infty}X_n$依概率收敛.

补充习题:

例 6.5.4 设$\{X_n\}$是一列独立$r.v.$且均值为0, $\sum_{n=1}^{\infty}n\mathbb{E}|X_n|^2<\infty$, 则$S_n=\sum_{i=1}^{n}X_i\,a.s.$收敛.

例 6.5.5 设$X,Y\ i.i.d.$且$\mathbb{E}X=0$, 则当$1\leq p\leq 2$时有$\frac{1}{2}\mathbb{E}|X-Y|^p\leq\mathbb{E}|X|^p\leq\mathbb{E}|X-Y|^p$.

例 6.5.6 设(X_n)是$r.v.$列, $\lim_{a\to\infty}\sup_{n\geq 1}P(|X_n|>a)=0$, 且$Y_n\xrightarrow{P}0$, 则$X_nY_n\xrightarrow{P}0$.

例 6.5.7 设(X_n)独立, $X_n\geq 0(n=1,2,\cdots)$, 证明: *TFAE*

(1) $\sum_{n=1}^{\infty}X_n<\infty,\ a.s.$.

(2) $\sum_{n=1}^{\infty}[P(|X_n|>1)+\mathbb{E}(X_nI_{[X_n\leq 1]})]<\infty$;

(3) $\sum_{n=1}^{\infty}\mathbb{E}\left(\frac{X_n}{1+X_n}\right)<\infty$.

第 7 章　测度的收敛

§ 7.1　测度的几种收敛

设(E,ρ)是度量空间, $\mathscr{B}(E)$是E上的Borel σ-代数, $B_b(E)$表示E是有界可测函数全体, $C_b(E)$表示E上有界连续函数全体, $C_c(E)$表示E上具有紧支集连续函数全体, $\mathscr{M}$表示$(E,\mathscr{B}(E))$上**有限测度**全体, $\mathscr{P}$表示$(E,\mathscr{B}(E))$上**概率测度**全体,

定义 7.1.1 设$\{\mu_n\}\subset\mathscr{M}$.

(1)称μ_n**一致收敛**于μ, 若$\sup\limits_{A\in\mathscr{B}(E)}|\mu_n(A)-\mu(A)|\to 0$.

(2)称μ_n**强收敛**于μ, 若$\mu_n(A)\to\mu(A),\forall A\in\mathscr{B}(E)$, 等价于$\mu_n(f)\to\mu(f),\forall f\in B_b(E)$.

(3)称μ_n**弱收敛**于μ, 若$\mu_n(f)\to\mu(f),\forall f\in C_b(E)$, 记为$\mu_n\xrightarrow{W}\mu$.

(4)称μ_n**淡收敛**于μ, 若$\mu_n(f)\to\mu(f),\forall f\in C_c(E)$, 记为$\mu_n\xrightarrow{v}\mu$.

定义 7.1.2 设$\mu\in\mathscr{M}$, $A\in\mathscr{B}(E)$. 称A是μ-**连续集**, 若$\mu(A^\circ)=\mu(\overline{A})$, 即$\mu(\partial A)=0$.

注: 全集E是μ-连续集.

例: 在$(\mathbb{R},\mathscr{B}(\mathbb{R}),\mathrm{d}x)$中考虑$A=(a,b)$, 则$A$是$\mathrm{d}x$-连续集.

定理 7.1.1 设$\{\mu_n\}\subset\mathscr{M}$, 则以下叙述等价:

(1)$\mu_n\xrightarrow{W}\mu$;

(2)对任意有界一致连续函数f, 有$\mu_n(f)\to\mu(f)$.

(3)对任意有界*Lipschitz*连续函数f, 有$\mu_n(f)\to\mu(f)$.

(4)对E上的任意闭集F, 有$\limsup\limits_{n\to\infty}\mu_n(F)\le\mu(F)$且$\mu_n(E)\to\mu(E)$.

(5)对E上的任意开集F, 有$\liminf\limits_{n\to\infty}\mu_n(F)\ge\mu(F)$且$\mu_n(E)\to\mu(E)$.

(6)对任意μ-连续集A, 有$\mu_n(A)\to\mu(A)$.

注: 定义与等价条件都要知道, 期末必考叙述.

证明: “$(1)\Rightarrow(2)\Rightarrow(3)$”、“$(4)\Leftrightarrow(5)$”显然, 只需证“$(3)\Rightarrow(4)$”、“$(4)+(5)\Rightarrow(6)$”、“$(6)\Rightarrow(1)$”.

“$(3)\Rightarrow(4)$”: 设F是E中的闭集, 对任意$n\ge 1$, 令$f_n(x)=\dfrac{1}{1+nd(x,F)},x\in E$, 则$f_n\searrow I_F$, 且$f_n$是有界Lipschitz连续函数, 从而

$$\mu(F)=\lim_{n\to\infty}\mu(f_n)=\lim_{n\to\infty}\lim_{m\to\infty}\mu_m(f_n)\ge\limsup_{m\to\infty}\mu_m(F).$$

取$f=\chi_E$可得$\mu_n(E)\to\mu(E)$.

“$(4)+(5)\Rightarrow(6)$”: 设A是μ-连续集, 即$\mu(A^\circ)=\mu(\overline{A})$, 由(4),

$$\mu(A)=\mu(\overline{A})\ge\limsup_{n\to\infty}\mu_n(\overline{A})\ge\limsup_{n\to\infty}\mu_n(A).$$

由(5),

$$\mu(A)=\mu(A^\circ)\le\liminf_{n\to\infty}\mu_n(A^\circ)\le\liminf_{n\to\infty}\mu_n(A).$$

“(6) ⇒ (1)”：设$f \in C_b(E)$, 由$\mu(E) < \infty$, 则$D \triangleq \{a \in \mathbb{R} | \mu([f = a]) > 0\} = \bigcup_{n=1}^{\infty} \underbrace{\{a \in \mathbb{R} | \mu([f = a]) \geq 1/n\}}_{\text{有限个}}$是至多可数集. 令$\{r_i\}_{1 \leq i \leq n}$为$[-(\|f\|_\infty + 1), \|f\|_\infty - 1]$的一个分割, 且$\{r_i\}_{0 \leq i \leq n} \subset D^c$(即$\mu([f = r_i]) = 0$), 则$\delta_n \triangleq \max_{1 \leq i \leq n}(r_i - r_{i-1}) \to 0(n \to \infty)$. 令$f_n = \sum_{i=1}^{n} r_{i-1} I_{[r_{i-1} \leq f \leq r_i]}$, 则$\|f_n - f\|_\infty \leq \delta_n \to 0(n \to \infty)$. 所以

$$\begin{aligned}
|\mu_n(f) - \mu(f)| &\leq |\mu_m(f) - \mu_m(f_n)| + |\mu_m(f_n) - \mu(f_n)| + |\mu(f_n) - \mu(f)| \\
&\leq [\mu_m(E) + \mu(E)]\|f - f_n\|_\infty + \sum_{i=1}^{n} |r_{i-1}||\mu_m(r_{i-1} \leq f < r_i) - \mu(r_{i-1} \leq f < r_i)|.
\end{aligned}$$

(注意$B_i = [r_{i-1} \leq f \leq r_i]$是$\mu$-连续集, 这是由于$f$连续, $\partial B_i = [f = a_{i-1}] \cup [f = a_i]$.)

对上式令$m \to \infty$, 再令$n \to \infty$可得$\lim_{m \to \infty} \mu_m(f) = \mu(f)$. □

§ 7.2 胎紧

定理 7.2.1 设(E, ρ)是紧度量空间, $\{\mu_n\} \subset \mathscr{M}$满足$\sup_n \mu_n(E) < \infty$, 则存在子列$\{\mu_{n_k}\}$与$\mu \in \mathscr{M}$使得$\mu_{n_k} \overset{W}{\to} \mu$. *(即: 紧度量空间中的测度序列有弱收敛子列.)*

我们要想办法把‘紧’的条件减弱, 因为$\mathbb{R}^n$不是紧度量空间. 下面的胎紧的概念很重要, 要牢牢记住.

定义 7.2.1 (胎紧) 设(E, ρ)是度量空间, $\mathscr{M}' \subset \mathscr{M}$. 称$\mathscr{M}'$是**胎紧的*(tight)***, 如果对任意$\varepsilon > 0$, 存在紧集$K_\varepsilon \subset E$, 使得$\sup_{\mu \in \mathscr{M}'} \mu(K_\varepsilon^c) \leq \varepsilon$.

注: 对概率测度空间, 可以把定义中的不等式改写为$\inf_{\mu \in \mathscr{M}'} \mu(K_\varepsilon^c) \geq 1 - \varepsilon$.

例 7.2.1 设$\mathscr{P}'$是$\mathbb{R}^n$上的一族概率测度, 若存在$r \geq 2$使得$\sup_{\mu \in \mathscr{P}'} \int_{\mathbb{R}^n} |x|^r \mu(\mathrm{d}x) < +\infty$, 则$\mathscr{P}'$胎紧.

证明: 对任意$m \geq 1$, 集合$K_m = \{x \in \mathbb{R}^n | |x| \leq m\}$是紧集, 且

$$\mu(K_m^c) = \mu(\{x \in \mathbb{R}^n | |x| > m\}) \leq \frac{1}{m^r} \int_{\mathbb{R}^n} |x|^r \mu(\mathrm{d}x),$$

所以$\sup_{\mu \in \mathscr{P}'} \mu(K_m^c) \leq \frac{1}{m^r} \sup_{\mu \in \mathscr{P}'} \int_{\mathbb{R}^n} |x|^r \mu(\mathrm{d}x) \to 0(m \to \infty)$. □

定义 7.2.2 (Polish空间) 设(X, ρ)是度量空间. 若(X, ρ)是完备可分的, 则称(X, ρ)是***Polish*空间**.

定理 7.2.2 (Prohorov) 设(E, ρ)是度量空间, $\{\mu_n\} \subset \mathscr{M}$, $\mathscr{M}' \subset \mathscr{M}$.
(1)若$\{\mu_n\}$胎紧, 则存在μ使得$\mu_n \xrightarrow{W} W$.
(2)若(E, ρ)是*Polish*空间, 则$\mathscr{M}'$胎紧$\Leftrightarrow \mathscr{M}'$弱相对紧, 即$\overline{\mathscr{M}'}$是$\mathscr{M}$中的弱紧集.

证明: 略.

定义 7.2.3 (依分布收敛) *设ξ_n是$(\Omega_n, \mathscr{F}_n, P_n)$上的$r.v.$列, ξ是$(\Omega, \mathscr{F}, P)$上的$r.v.$列. 称ξ_n依分布收敛于ξ, 记为$\xi_n \xrightarrow{d} \xi$, 若$P \circ \xi_n^{-1} \xrightarrow{W} P \circ \xi^{-1}$.*

注: 定义等价于对任意$f \in C_b(\mathbb{R}^n)$有$\mathbb{E}f(\xi_n) \to \mathbb{E}f(\xi)$.

显然, 由控制收敛定理可得$\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{d} \xi$. 反之不对, 因为r.v.列$\{\xi_n\}$可能不是定义在同一个概率空间.

§ 7.3 测度空间上的距离

回顾: 设(E, ρ)是Polish空间, 则$C_b(E)$在一致范数$(\|\cdot\|_\infty)$下也是Polish空间, 从而存在可数稠密子集$\{f^{(n)}\}_{n\geq 1} \subset C_b(E)$.

令

$$d(u, v) = \sum_{n=1}^{\infty} \frac{|\mu(f^{(n)}) - \nu(f^{(n)})| \wedge 1}{2^n}, \forall u, v \in \mathscr{P}.$$

定理 7.3.1 *设(E, ρ)是Polish空间, 则$(\mathscr{P}(E), d)$是可分度量空间, 且对$\{\mu_m, \mu\} \subset \mathscr{P}$, 有$\mu_m \xrightarrow{W} \mu \Leftrightarrow d(\mu_m, \mu) \to 0$.*

若(E, ρ)是局部紧空间, 则$(\mathscr{P}(E), d)$是完备度量空间.

证明: (1)先证明d为距离. 显然$d(\mu, \nu) \geq 0$. 若$d(\mu, \nu) = 0$, 则对任意$f \in C_b(E)$, $\mu(f) = \nu(f)$. (注意稠密性) 对称性与三角不等式都是显然的.

(2) "$\Leftarrow$": 设$d(\mu_m, \mu) \to 0$, 下证对任意$f \in C_b(E)$, $\mu_m(f) \to \mu(f)$. 对给定$f \in C_b(E)$, 对任意$\varepsilon > 0$, 存在$f^{(n')} \in \{f^{(n)}\}$, 使得$\|f^{n'} - f\|_\infty \leq \varepsilon$, 则$\|f\|_\infty \leq \|f^{(n')}\|_\infty + \varepsilon$, 所以

$$\begin{aligned} |\mu_m(f) - \mu(f)| &\leq |\mu_m(f - f^{(n')}) - \mu(f - f^{(n')})| + |\mu_m(f^{(n')}) - \mu(f^{(n')})| \\ &\leq 2\varepsilon + 2^{n'} d(\mu_m, \mu). \end{aligned}$$

让$m \to \infty$, 再利用m的任意性可得欲证结论.

"$\Rightarrow$": 注意

$$\begin{aligned} \limsup_{m\to\infty} d(\mu_m, d) &= \limsup_{m\to\infty} \sum_{n=1}^{\infty} \frac{|\mu_m(f^{(n)}) - \mu(f^{(n)})| \wedge 1}{2^n} \\ &\leq \sum_{n=1}^{\infty} \limsup_{m\to\infty} \frac{|\mu_m(f^{(n)}) - \mu(f^{(n)})| \wedge 1}{2^n} \qquad \text{(Fatou)} \\ &= 0 \qquad \text{(弱收敛)} \end{aligned}$$

(3)可分性: 只需找$\mathcal{P}(E)$的可数稠密子集.

对$m \geq 1$, 令

$$U_m = \{(\mu(f^{(1)}), \cdots, \mu(f^{(n)}))|\mu \in \mathcal{P}(E)\} \subset \mathbb{R}^m,$$

由于$\mathbb{R}^m$有可数稠密子集(可分), 则U_m也有可数稠密子集, 即存在$\mathcal{P}(E)$的至多可数子集$\mathcal{P}_m(E)$使得

$$\tilde{U}_m \triangleq \{(\mu(f^{(1)}), \cdots, \mu(f^{(n)}))|\mu \in \mathcal{P}_m(E)\}$$

为U_m的可数稠密子集. 令

$$\mathcal{P}_\infty(E)=\bigcup_{m=1}^{\infty}\mathcal{P}_m(E),$$

则$\mathcal{P}_\infty(E)$为$\mathcal{P}(E)$的可数子集.

又对任意$\mu\in\mathcal{P}(E)$, 利用$\tilde{U}_m$定义可知存在$\mu_m\in\mathcal{P}_m(E)$使得

$$|\mu_m(f^{(i)})-\mu(f^{(i)})|\le\frac{1}{m},\forall 1\le i\le m,$$

所以

$$\begin{aligned}d(\mu_m,\mu)&=\sum_{n=1}^{\infty}\frac{|\mu_m(f^{(n)})-\mu(f^{(n)})|\wedge 1}{2^n}\\&=\sum_{n=1}^{m}\frac{|\mu_m(f^{(n)})-\mu(f^{(n)})|\wedge 1}{2^n}+\sum_{n=m+1}^{\infty}\frac{|\mu_m(f^{(n)})-\mu(f^{(n)})|\wedge 1}{2^n}\\&\le\frac{1}{m}+\frac{1}{2^m}\to 0(m\to\infty).\end{aligned}$$

□

7.3.1 最优运输问题与Wasserstein距离

定义 7.3.1 *设E是集合, $\mu,\nu\in\mathscr{P}(E)$. 称$\pi\in\mathscr{P}(E\times E)$是$\mu$和$\nu$的一个**耦合(coupling)**, 若$\forall A\in E,\pi(A\times E)=\mu(A),\pi(E\times A)=\nu(A)$.*

注: 耦合可以不唯一, 考虑二维正态分布.

对于$\Omega=\{x_1,\cdots,x_n\},\mu=(\mu_1,\cdots,\mu_n),\nu=(\nu_1,\cdots,\nu_n)$, 用$\pi_{ij}$表示$x_i$向$x_j$所运输的市场份额. 定义

$$\mu_i=\sum_{j=1}^{n}\pi_{ij},\nu_j=\sum_{i=1}^{n}\pi_{ij},1\le i,j\le n.$$

则$\sum\limits_{i=1}^{n}\mu_i=\sum\limits_{i=1}^{n}\nu_i=1$, 从而$\sum\limits_{i=1}^{n}\mu_i$与$\sum\limits_{i=1}^{n}\nu_i$是$\Omega$的概率测度, $\{\pi_{ij}\}_{1\le i,j\le n}$为$\{x_1,\cdots,x_n\}\times\{x_1,\cdots,x_n\}$上的概率测度, 且$\pi$为$\mu$与$\nu$的耦合. 反之, 任一耦合可以看作一种运输方案.

设ρ_{ij}为从x_i到x_j运输单位产品所需运费, 则π方案的运输成本为

$$\sum_{i,j}\pi_{ij}\rho_{ij}=\int_{\{x_1,\cdots,x_n\}^2}\rho\mathrm{d}\pi.$$

问什么时候成本最小, 即求

$$\inf_{\pi\in(\mu,\nu)}\int\rho\mathrm{d}\pi.$$

其中$\mathcal{C}(\mu,\nu)$表示以μ,ν为边际分布的所有耦合全体.

定义 7.3.2 *设(E,ρ)为度量空间, $p\in[1,+\infty)$, 称*

$$W_p^\rho(\mu,\nu)=\inf_{\pi\in\mathcal{C}(\mu,\nu)}\left(\int_{E\times E}\rho^p(x,y)\pi(\mathrm{d}x,\mathrm{d}y)\right)^{\frac{1}{p}}$$

*为由ρ诱导的L^p-**Wasserstein距离**.*

注: ρ可能无界, 此时积分可能为正无穷. 我们定义

$$\mathcal{P}_p(E) = \{\mu \in \mathcal{P}(E) | \int_E \rho^p(x,0)\mu(\mathrm{d}x) < +\infty\}.$$

定理 7.3.2 设(E,ρ)为$Polish$空间, 则对任意$\mu,\nu \in \mathcal{P}_p(E)$, 存在$\pi_0 \in \mathcal{C}(\mu,\nu)$, 使得

$$W_p^\rho(\mu,\nu) = \left[\int_{E\times E} \rho^p(x,y)\pi_0(\mathrm{d}x,\mathrm{d}y)\right]^{\frac{1}{p}}.$$

即最优运输可以取到.

证明: 由于$\mu,\nu \in \mathcal{P}_p(E)$, 从而$W_p^\rho(\mu,\nu) < \infty$. 由下确界定义, 对任意$n \geq 1$, 存在$\pi_n \in \mathcal{C}(\mu,\nu)$, 使得

$$[\pi_n(\rho^p)] \leq (W_p^\rho(\mu,\nu))^p + \frac{1}{n}.$$

下证存在$\pi_0 \in \rho(E\times E)$使得$\pi_n \xrightarrow{W} \pi_0$. (此时还不能说$\pi_0$是$\mu$与$\nu$的耦合) 由Prohorov定理, 只需证$\{\pi_n\}$是胎紧的.

注意到$\{\mu,\nu\}$弱相对紧, 从而$\{\mu,\nu\}$胎紧, 即对任意$\varepsilon > 0$, 存在紧集$K_\varepsilon \subset E$, 使得$\mu(K_\varepsilon^c)+\nu(K_\varepsilon^c) \leq \varepsilon$. 则

$$\pi_n((K_\varepsilon \times K_\varepsilon)^c) \leq \pi_n(K_\varepsilon^c \times E) + \pi_n(E \times K_\varepsilon^c) = \mu(K_\varepsilon^c) + \nu(K_\varepsilon^c) \leq \varepsilon. \qquad \text{(耦合的定义)}$$

所以$\{\pi_n\}$胎紧, 从而有弱收敛子列, 即存在$\pi_{n_k} \in \mathcal{P}(E\times E)$与$\pi_0 \in \mathcal{P}(E\times E)$, 使得$\pi_{n_k} \xrightarrow{W} \pi_0 (k\to\infty)$, ($\mathcal{P}(E\times E)$中的元素不一定有界) 从而

$$\pi_0(\rho^p \wedge N) = \lim_{k\to\infty} \pi_{n_k}(\rho^p \wedge N) \leq W_p^\rho(\mu,\nu)^p + \lim_{k\to\infty}\frac{1}{n_k}.$$

令$N\to\infty$, 则$\pi_0(\rho^p) \leq W_p^\rho(\mu,\nu)^p$.

对任意$f \in C_b(E)$, $\pi_n(f(x)I_E(y)) \to \pi_0(f(x)I_E(y))$, 其中$f(x) = \mu(f)$, 所以$\pi_0$的一个边际分布为$\mu$. 而对于$\nu$是同理的. 所以$\pi_0 \in \mathcal{C}(\mu,\nu)$, 从而$\pi_0(\rho^p) = W_p^\rho(\mu,\nu)^p$. □