测度与概率 习题参考答案

编者: 陈攀屹 黄奕杰
北京师范大学

2024 年 9 月 22 日

序言

严士健和刘秀芳两位老师编著的《测度与概率》是国内一流的基于测度的概率论教材. 在学习该教材时我们对一些习题感到困扰, 但这部教材并未提供习题答案, 因此我们结合自己的作业和前人整理的文档, 整理出了此份参考答案. 尽管参考答案可能使读者变得怠惰, 但我们依然认为它会对学习主动性强的同学有巨大的帮助. 主动学习是数学学习中不可或缺的一环, 因此我们希望您在每次使用本文档前尽可能地独立思考. 希望本文档能够开拓您的思路, 对您学习基于测度的概率论有所帮助.

教材中一些习题的错误已经在答案中改正并指出. 同时, 我们尽可能地收集了一些优秀的解法, 并以"注"的形式补充了一些与题目有关的思考. 在此一并感谢提出新解法与新思路的助教师兄以及同学.

由于我们水平有限, 加之时间仓促, 编写的参考答案难免存在不严谨、不正确之处. 如果您在使用时发现错误, 或者想要分享新的解法、有关题目的新的思考, 欢迎向我们投稿. 为使您的投稿有效, 投稿前烦请阅读下一节"投稿前须知".

我们会不定期地将新投稿整合到本文档里, 此后更新的版本将在我们的博客[1]发布. 请注意, 只有当文件的哈希值与我们博客上列出的值相同时, 才可以认为是我们发布的.

编者

2024 年 9 月 22 日

[1] 我们的博客地址——地址 1: https://map-abook.gitlab.io/; 地址 2: https://map-abook.pages.dev/

[2] 详见https://creativecommons.org/licenses/by-nc-sa/4.0/deed.zh-hans

投稿前须知

首先, 十分感谢您愿意向我们以及其他读者分享您的想法! 您的投稿将丰富本文档的内容, 使本文档不断得到完善, 对我们和其他读者起到更大的帮助. 不过, 为使我们能理解您的投稿内容, 请您先阅读以下的投稿规范.

我们接受以下三种稿件:

1. 手写稿. 电子手写笔记请导出为 PDF 格式, 纸质手写笔记请扫描为 PDF 电子档, 或者拍照并打包为 ZIP 压缩包 (压缩包内文件的命名需体现先后顺序). 文件内容需体现涉及到的题目编号.

2. LaTeX 稿件. 推荐使用我们的模版[1]. 提交时请提交 .tex 文件. 文件内容需体现涉及到的题目编号. (推荐会使用 LaTeX 的同学使用这种方式, 这能极大减少我们的工作量.)

3. 网页或网页链接 (只能通过电子邮件投稿). 请将网页链接放在邮件正文, 或者将网页导出为 PDF/Markdown/MHTML 格式作为附件发送.

请以涉及的习题编号 + 投稿者署名的形式命名您的文件, 例如习题 3.3.4-张三.tex 或者习题 2.1.6, 2.2.1-匿名.pdf. 我们将在文档里注明投稿者的署名. (必要时我们可能修改您稿件中的部分语句. 同时请文明用语, 避免出现敏感内容, 否则我们可能将您的投稿标记为匿名.)

请发送电子邮件到map_abook@163.com进行投稿. 您需要将要投稿的文件添加到邮件附件里, 并以涉及的习题编号 + 投稿者署名的格式命名邮件主题, 例如习题 3.3.4, 5.1.1-李四.

我们将尽快处理您的投稿, 正确且不同于已有解法的解答将会被采用. 如您认为您的投稿在某些小细节上不完全符合上述规范, 请不必过分担忧, 我们会尽力理解您的投稿内容. 不过, 即使您的解法正确, 我们也不保证采用不符合上述规范的投稿. 如您有疑问、意见或建议, 欢迎发送电子邮件到map_abook@163.com. 我们保留最终的解释权.

感谢您的理解与支持!

编者

2024 年 9 月 22 日

[1]下载地址——地址 1: https://map-abook.gitlab.io/file/template.tex; 地址 2: https://map-abook.pages.dev/file/template.tex

目录

第一章 集合、映射与势 **1**

§ 1.1 集合及其运算 1

§ 1.2 映射与势 7

§ 1.3 可数集 8

§ 1.4 不可数集 9

第二章 距离空间 **13**

§ 2.1 定义及例 13

§ 2.2 开集、闭集 16

§ 2.3 完备性 18

§ 2.4 可分性、列紧性与紧性 20

§ 2.5 距离空间上的映射与函数 23

第三章 测度空间与概率空间 **25**

§ 3.1 集类 25

§ 3.2 单调函数与测度的构造 35

§ 3.3 测度空间的一些性质 48

第四章 可测函数与随机变量 **55**

§ 4.1 可测函数与分布 55

§ 4.2 可测函数的构造性质 59

第五章 积分与数学期望 **61**

§ 5.1 积分的定义 61

§ 5.2 积分的性质 64

§ 5.3 期望的性质及 L-S 积分表示 67

§ 5.4 积分收敛定理 76

第六章 乘积测度与无穷乘积概率空间 **79**

§ 6.1 乘积测度与转移测度 79

§ 6.2 Fubini 定理及其应用 88

§ 6.3 无穷维乘积概率 90

第七章 不定积分与条件期望 **93**

§ 7.1 符号测度的分解 93

§ 7.2 Lebesgue 分解定理与 Radon-Nikodym 定理 95

§ 7.3 条件期望的概念 99

§ 7.4 条件期望的性质 101

§ 7.5 条件概率分布 105

第八章 收敛概念 **109**

§ 8.1 几乎处处收敛 109

§ 8.2 依测度收敛 112

§ 8.3 L^r 收敛 113

§ 8.5 概率测度的收敛 116

第九章 大数定律、随机级数 **119**

§ 9.1 简单的极限定理及其应用 119

§ 9.2 弱大数定律 121

§ 9.3 随机级数的收敛 124

§ 9.4 强大数律 125

第十章 特征函数和中心极限定理 **127**

§ 10.1 特征函数的定义及简单性质 127

§ 10.2 逆转公式及连续性定理 129

§ 10.3 中心极限定理 131

参考文献 **137**

第一章　集合、映射与势

§ 1.1　集合及其运算

1.1.1 证明: $(A\cup B)\backslash B=A\Leftrightarrow A\cap B=\varnothing$.

证明: 我们有

$$\begin{aligned}(A\cup B)\backslash B&=(A\cup B)\cap B^c=(A\cap B^c)\cup(B\cap B^c)=A\cap B^c\\&=(A\cap A^c)\cup(A\cap B^c)=A\cap(A^c\cup B^c)=A\cap(A\cap B)^c.\end{aligned}$$

故 $(A\cup B)\backslash B=A\Leftrightarrow A\cap B=\varnothing$. □

1.1.2 证明: $(A\backslash B)\cup B=A\Leftrightarrow B\subset A$.

证明: 我们有

$$(A\backslash B)\cup B=(A\cap B^c)\cup B=(A\cup B)\cap(B^c\cup B)=A\cup B,$$

故 $(A\backslash B)\cup B=A\Leftrightarrow B\subset A$. □

1.1.3 $(A\backslash B)\cup C=A\backslash(B\backslash C)$ 成立的充分必要条件是什么?

证明: 我们知道

$$(A\backslash B)\cup C=(A\cap B^c)\cup C,\quad A\backslash(B\backslash C)=A\cap(B\cap C^c)^c,$$

因此

$$\begin{aligned}(A\backslash B)\cup C=A\backslash(B\backslash C)&\Leftrightarrow(A\cap B^c)\cup C=A\cap(B\cap C^c)^c\\&\Leftrightarrow\mathbb{1}_A(1-\mathbb{1}_B)+\mathbb{1}_C-\mathbb{1}_A(1-\mathbb{1}_B)\mathbb{1}_C=\mathbb{1}_A(1-\mathbb{1}_B(1-\mathbb{1}_C))\\&\Leftrightarrow\mathbb{1}_C=\mathbb{1}_A\mathbb{1}_C\Leftrightarrow C=A\cap C\Leftrightarrow C\subset A.\end{aligned}$$

□

1.1.4 证明下述等式:

(1) $A\cap B=A\backslash(A\backslash B)$;

(2) $A\backslash(B\backslash C)=(A\backslash B)\cup(A\cap C)$;

(3) $A\backslash(B\cap C)=(A\backslash B)\cup(A\backslash C)$;

(4) $A\backslash(B\cup C)=(A\backslash B)\cap(A\backslash C)$;

(5) $(A\backslash B)\cap(C\backslash D)=(A\cap C)\backslash(B\cup D)$;

(6) $(A\Delta B)\Delta C=A\Delta(B\Delta C)$;

(7) $(A\Delta B)\cap C=(A\cap C)\Delta(B\cap C)$;

(8) $B\cap\left(\bigcup_{\alpha\in I}A_\alpha\right)=\bigcup_{\alpha\in I}(B\cap A_\alpha)$;

(9) $B\cap\left(\bigcap_{\alpha\in I}A_\alpha\right)=\bigcap_{\alpha\in I}(B\cap A_\alpha)$.

证明: (1) $A\backslash(A\backslash B)=A\backslash(A\cap B^c)=A\cap(A^c\cup B)=A\cap B$;

(2) 我们有

$$\begin{aligned}A\backslash(B\backslash C)&=A\backslash(B\cap C^c)=A\cap(B^c\cup C)\\&=(A\cap B^c)\cup(A\backslash C)=(A\backslash B)\cup(A\backslash C);\end{aligned}$$

(3) $A\backslash(B\cap C)=A\cap(B^c\cup C^c)=(A\cap B^c)\cup(A\cap C^c)=(A\backslash B)\cup(A\backslash C)$;

(4) $A\backslash(B\cup C)=A\cap(B^c\cap C^c)=A\cap B^c\cap A\cap C^c=(A\backslash B)\cup(A\backslash C)$;

(5) 我们有

$$\begin{aligned}(A\backslash B)\cap(C\backslash D)&=(A\cap B^c)\cap(C\cap D^c)=(A\cap C)\cap(B^c\cap D^c)\\&=(A\cap C)\backslash(B\cup D);\end{aligned}$$

(6) 我们有

$$\begin{aligned}(A\Delta B)\Delta C&=((A\backslash B)\cup(B\backslash A))\Delta C\\&=(((A\backslash B)\cup(B\backslash A))\backslash C)\cup(C\backslash((A\backslash B)\cup(B\backslash A)))\\&=(((A\cap B^c)\cup(B\cap A^c))\cap C^c)\cup(C\cap(((A\backslash B)\cup(B\backslash A))^c)\\&=((A\cap B^c\cap C^c)\cup(B\cap A^c\cap C^c))\cup((A^c\cap B^c\cap C)\cup(A\cap B\cap C))\\&=((B\cap C^c)\cup(C\cap B^c)\cap A^c)\cup(A\cap((B\cap C^c)\cup(C\cap B^c))^c)\\&=((B\backslash C)\cup(C\backslash B)\backslash A)\cup(A\backslash((B\backslash C)\cup(C\backslash B)))\\&=A\Delta((B\backslash C)\cup(C\backslash B))\\&=A\Delta(B\Delta C);\end{aligned}$$

(7) 我们有

$$\begin{aligned}(A\Delta B)\cap C&=((A\backslash B)\cup(B\backslash A))\cap C\\&=((A\cap B^c)\cup(B\cap A^c))\cap C\\&=((A\cap B^c)\cap C)\cup((B\cap A^c)\cap C)\\&=((A\cap C)\cap(B^c\cup C^c))\cup((B\cap C)\cap(A^c\cup C^c))\\&=((A\cap C)\cap(B\cap C)^c)\cup((B\cap C)\cap(A\cap C)^c)\\&=((A\cap C)\backslash(B\cap C))\cup((B\cap C)\backslash(A\cap C))\\&=(A\cap C)\Delta(B\cap C);\end{aligned}$$

(8) 我们有

$$\begin{aligned}x\in B\cap(\cup_{\alpha\in I}A_\alpha)&\Leftrightarrow x\in B,x\in\cup_{\alpha\in I}A_\alpha\\&\Leftrightarrow\exists\alpha_0\in I,x\in B,x\in A_{\alpha_0}\\&\Leftrightarrow\exists\alpha_0\in I,x\in B\cap A_{\alpha_0}\\&\Leftrightarrow x\in\cup_{\alpha\in I}(B\cap A_\alpha);\end{aligned}$$

(9) 我们有

$$\begin{aligned} x \in B \cap (\cap_{\alpha\in I} A_\alpha) &\Leftrightarrow x \in B, x \in \cap_{\alpha\in I} A_\alpha \\ &\Leftrightarrow \forall \alpha \in I, x \in B, x \in A_\alpha \\ &\Leftrightarrow \forall \alpha \in I, x \in B \cap A_\alpha \\ &\Leftrightarrow x \in \cap_{\alpha\in I}(B \cap A_\alpha). \end{aligned}$$

□

1.1.5 下列等式是否成立? 若不成立, 有怎样的包含关系?

(1) $A \cup (B \backslash C) = (A \cup B) \backslash (A \cup C)$;

(2) $A \cup (B \Delta C) = (A \cup B) \Delta (A \cup C)$;

(3) $A \backslash (B \cup C) = (A \backslash B) \cup (A \backslash C)$;

(4) $(A \backslash B) \cup C = (A \cup C) \backslash B$.

证明: (1) 有

$$\begin{aligned} (A \cup B) \backslash (A \cup C) &= (A \cup B) \cap (A^c \cap C^c) = (A \cap A^c \cap C^c) \cup (B \backslash (A \cup C)) \\ &= B \backslash (A \cup C) \subseteq B \backslash C \subseteq A \cup (B \backslash C), \end{aligned}$$

当 $A = \varnothing$ 时等号成立.

(2) 有

$$(A \cup B) \Delta (A \cup C) = (B \backslash C) \cup (C \backslash B) = B \Delta C \subseteq A \cup (B \Delta C),$$

当 $A = \varnothing$ 时等号成立.

(3) 有

$$A \backslash (B \cup C) = A \cap B^c \cap C^c \subseteq A \cap B^c \subseteq (A \cap B^c) \cup (A \cap C^c) = (A \backslash B) \cup (A \backslash C),$$

当 $(A \cap B^c) \subseteq (A \cap C^c)$ 时成立.

(4) 有

$$(A \cup C) \backslash B = (A \cup C) \cap B^c \subset (A \cap B^c) \cup C = (A \backslash B) \cup C,$$

故等式成立当且仅当 $C \subset B^c$. □

1.1.6 试化简集合 $(A \cup B^c \cup C^c) \cap (A \cup (B \cup C^c))$.

证明: 我们有

$$\begin{aligned} (A \cup B^c \cup C^c) \cap (A \cup (B \cup C^c)) &= (A \cup C^c \cup B^c) \cap (A \cup C^c \cup B) \\ &= A \cup C^c = A \backslash C. \end{aligned}$$

□

1.1.7 设 $\{A_n : n = 1, 2, \cdots\}$ 为单调减集序列, 则有 $A_1 = \left(\bigcap_{n=1}^{\infty} A_n\right) \cup \left(\bigcup_{n=1}^{\infty} (A_n - A_{n+1})\right)$, 且右端各项互不相交.

证明: 我们知道 $A_n - A_{n+1} := \{x \in A_n, x \notin A_{n+1}\}$, 而 $\forall k \in \mathbb{N}^*, A_{n+k} \subset \cdots \subset A_{n+1}$. 故 $(A_n - A_{n+1}) \cap (A_{n+k} - A_{n+k+1}) = (A_n - A_{n+1}) \cap \left(\bigcap_{n=1}^{\infty} A_n\right) = \varnothing$.

与此同时, 对于 $x \in A_1$, 若 $\forall n \in \mathbb{N}$, $x \in A_n$, 则 $x \in \bigcap_{n=1}^{\infty} A_n$;

若 $\exists k \geqslant 2$, s.t. $x \notin A_k$, 则 $x \in A_1 \backslash A_k = \bigcup_{i=1}^{k-1}(A_i \backslash A_{i+1}) \subset \bigcup_{n=1}^{\infty}(A_n - A_{n+1})$.

故 $A_1 \subseteq \left(\bigcap_{n=1}^{\infty} A_n\right) \cup \left(\bigcup_{n=1}^{\infty} (A_n - A_{n+1})\right)$,

同时, $\forall x \in \left(\bigcap_{n=1}^{\infty} A_n\right) \cup \left(\bigcup_{n=1}^{\infty} (A_n - A_{n+1})\right)$, 有 $x \in \bigcap_{n=1}^{\infty} A_n$ 或 $x \in \bigcup_{n=1}^{\infty} (A_n - A_{n+1})$,

若 $x \in \bigcap_{n=1}^{\infty} A_n$, 则 $\forall n \in \mathbb{N}, x \in A_n \subset A_1$;

若 $x \in \bigcup_{n=1}^{\infty} (A_n - A_{n+1})$, 则 $\exists k \in \mathbb{N}^*, x \in A_k - A_{k+1} \subset A_k \subset A_1$.

故 $\left(\bigcap_{n=1}^{\infty} A_n\right) \cup \left(\bigcup_{n=1}^{\infty} (A_n - A_{n+1})\right) \subseteq A_1$. 故它们相等. □

1.1.8 设 R 为 Ω 的一切子集组成的集类, 则 R 对集合的交 (看成乘法)、对称差 (看成加法) 运算作成一个环. Ω 是单位元, $\varnothing$ 是零元.

证明: 回忆环的定义, 我们只需证明:

(i) (R, Δ) 是 Abel 群且有单位元 $\varnothing$;

(ii) $(R, \cap)$ 是半群且有单位元 Ω;

(iii) $\cap$ 对于 Δ 满足左右分配律.

这意味着, 我们需要证明:

(i) (a) $\forall A, B \in R,\ A\Delta B = B\Delta A \in R$;

(b) $\forall A, B, C \in R,\ (A\Delta B)\Delta C = A\Delta(B\Delta C)$;

(c) $\forall A \in R,\ A\Delta\varnothing = \varnothing\Delta A = A$;

(d) $\forall A \in R,\ \exists B \in R,\ A\Delta B = B\Delta A = \varnothing$;

(ii) (a) $\forall A, B \in R,\ A \cap B \in R$;

(b) $\forall A, B, C \in R,\ A \cap (B \cap C) = (A \cap B) \cap C$;

(c) $\forall A \in R,\ A \cap \Omega = \Omega \cap A$;

(iii) $\forall A, B, C \in R, \begin{cases} A \cap (B\Delta C) = (A \cap B)\Delta(A \cap C), \\ (B\Delta C) \cap A = (B \cap A)\Delta(C \cap A). \end{cases}$

其中 (i)(a), (ii)(a), (ii)(c) 是显然的, (i)(b) 为习题 1.1.4(6), (iii) 为习题 1.1.4(7)(由于交运算是可交换的, 所以两个式子等价).

下面证明 (i)(c), (i)(d), (ii)(b):

(i)(c): 我们有 $A\Delta\varnothing=(A\backslash\varnothing)\cup(\varnothing\backslash A)=(A\cap\Omega)\cup(\varnothing\cap A^c)=A$;

(i)(d): 取 $B=A$ 即可;

(ii)(b): 我们有 $x\in A\cap(B\cap C)\Leftrightarrow x\in A,x\in B,x\in C\Leftrightarrow x\in(A\cap B)\cap C$.

$\square$

1.1.9 设 $\{A_n:n=1,2,\cdots\}$ 是一集序列, 令 $B_1=A_1,B_n=A_n\backslash\left(\bigcup\limits_{k=1}^{n-1}A_k\right)$, 则 $B_n,n=1,2,\cdots$ 两两不交, 且 $\bigcup\limits_{n=1}^{\infty}A_n=\bigcup\limits_{n=1}^{\infty}B_n$.

证明: 我们知道 $B_n\subset A_n$, 故 $\bigcup\limits_{n=1}^{\infty}B_n\subset\bigcup\limits_{n=1}^{\infty}A_n$;

同时 $\forall x\in\bigcup\limits_{n=1}^{\infty}A_n$, 考虑最小的 $j\in\mathbb{N}$ 使得 $x\in A_j$, 则 $x\in A_j\backslash\left(\bigcup\limits_{k=1}^{j-1}A_k\right)\subset\bigcup\limits_{n=1}^{\infty}B_n$. 故 $\bigcup\limits_{n=1}^{\infty}A_n\subset\bigcup\limits_{n=1}^{\infty}B_n$, 因此它们相等. $\square$

1.1.10 试证明定理 1.1.7.

1.1.7 定理: 设 $\{A_n:n\in\mathbb{N}\}$ 为任一集序列.

(1) $\limsup\limits_{n\to\infty}A_n=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k,\quad \liminf\limits_{n\to\infty}A_n=\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{k=n}^{\infty}A_k$.

(2) 若 $\{A_n\}$ 单调, 则 $\lim\limits_{n\to\infty}A_n$ 存在, 且

$$\lim_{n\to\infty}A_n=\begin{cases}\bigcup\limits_{n=1}^{\infty}A_n, & \text{若}\ \{A_n\}\ \text{单调增},\\ \bigcap\limits_{n=1}^{\infty}A_n, & \text{若}\ \{A_n\}\ \text{单调减}.\end{cases}$$

证明: 我们知道

$$\limsup_{n\to\infty}A_n=\{x:\forall n\in\mathbb{N},\exists k\geqslant n,\text{s.t.}x\in A_k\},$$
$$\liminf_{n\to\infty}A_n=\{x:\exists n\in\mathbb{N},\text{s.t.}\forall k\geqslant n,x\in A_k\}.$$

于是

$$\begin{aligned}\limsup_{n\to\infty}A_n&=\{x:\forall n\in\mathbb{N},\exists k\geqslant n,\text{s.t.}x\in A_k\}\\&=\{x:\forall n\in\mathbb{N},x\in\bigcup_{k=n}^{\infty}A_k\}\\&=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k,\end{aligned}$$

类似地,

$$\begin{aligned}\liminf_{n\to\infty} A_n &= \{x : \exists n \in \mathbb{N}, \text{s.t.} \forall k \geqslant n, x \in A_k\}\\ &= \{x : \exists n \in \mathbb{N}, x \in \bigcap_{k=n}^{\infty} A_k\}\\ &= \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k.\end{aligned}$$

由于 (2) 中的集合是单调的, 由 (1) 立刻可证. □

1.1.11 试举例说明: $\{A_n : n = 1, 2, \cdots\}$ 不单调, 但 $\lim\limits_{n\to\infty} A_n$ 存在.

证明: 取 $A_n = \begin{cases}\left(\dfrac{1}{2k-1}, 1\right), & n = 2k-1\\ \left(0, 1-\dfrac{1}{2k}\right), & n = 2k\end{cases}$, 则显然 $\{A_n : n = 1, 2, \cdots\}$ 不单调, 但 $\lim\limits_{n\to\infty} A_n = (0, 1)$ (可以验证 $\limsup\limits_{n\to\infty} A_n = (0, 1) = \liminf\limits_{n\to\infty} A_n$). □

注: 一个常见的错误例子: 取 $A_n = \left[\dfrac{(-1)^n}{n}, 1\right)$. 此时, 0 在无穷多个 A_n 里, 但 0 并非不属于有限个 A_n, 即 $0 \in \limsup\limits_{n\to\infty} A_n$ 但是 $0 \notin \liminf\limits_{n\to\infty} A_n$ (见定义 1.1.6), 或者也可由定理 1.1.7(1) 计算得 $\limsup\limits_{n\to\infty} A_n = [0, 1)$ 但 $\liminf\limits_{n\to\infty} A_n = (0, 1)$, 总之得到 $\limsup\limits_{n\to\infty} A_n \neq \liminf\limits_{n\to\infty} A_n$, 即 $\lim\limits_{n\to\infty} A_n$ 不存在. 取 $A_n = \left(\dfrac{(-1)^n}{n}, 1\right)$ 也是同样的道理. 举出这样的例子可能是因为题目要求 $\{A_n\}$ 不单调, 但是不单调并不一定要取 $\left\{\dfrac{(-1)^n}{n}\right\}$ 这种 "上下摆动" 的数列作为区间端点, 如参考答案这样对端点取值进行简单的分类讨论也是可以的.

1.1.12 设 $\forall k = 1, 2, \cdots$, 定义 $A_{2k+1} = \left[0, 2 - \dfrac{1}{2k+1}\right]$, $\quad A_{2k} = \left[0, 1 + \dfrac{1}{2k}\right]$, 试求 $\liminf\limits_{n\to\infty} A_n, \limsup\limits_{n\to\infty} A_n$.

证明: 我们有

$$\liminf_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k = \bigcap_{n=1}^{\infty}[0, 1] = [0, 1];$$

$$\limsup_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k = \bigcup_{n=1}^{\infty}[0, 2) = [0, 2).$$

□

1.1.13 给定非零自然数 m 及 m 个集合 $B_0, B_1, \cdots, B_{m-1}$, 设 $A_n = B_k$, 当 m 整除 $n-k$ 时, 试求 $\liminf\limits_{n\to\infty} A_n, \limsup\limits_{n\to\infty} A_n$.

证明: 我们知道 $A_{Nm+k} = B_k$, 这里 $N \in \mathbb{N}$. 因此

$$\liminf_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} A_j = \bigcap_{k=0}^{m-1} B_k;$$

$$\limsup_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{j=n}^{\infty} A_j = \bigcup_{k=0}^{m-1} B_k.$$

□

1.1.14 试证定理 1.1.10 的 (1)(2)(3)(4).

1.1.10 定理: 给定非空集合 Ω, 下述集合都是 Ω 的子集, 则 $\forall x \in \Omega$,

(1) $A = \Omega \Leftrightarrow \mathbb{1}_A(x) \equiv 1$;

(2) $A = \varnothing \Leftrightarrow \mathbb{1}_A(x) \equiv 0$;

(3) $A \subset B \Leftrightarrow \mathbb{1}_A(x) \leqslant \mathbb{1}_B(x), \forall x \in \Omega$;

(4) $\mathbb{1}_{\cup_{\alpha\in J} A_\alpha}(x) = \max\limits_{\alpha\in J} \mathbb{1}_{A_\alpha}(x), \quad \mathbb{1}_{\cap_{\alpha\in J} A_\alpha}(x) = \min\limits_{\alpha\in J} \mathbb{1}_{A_\alpha}(x)$;

证明: (1)(2) 是显然的, 对 (3), 我们有 $\mathbb{1}_B(x) - \mathbb{1}_A(x) = \mathbb{1}_{B-A}(x) \geqslant 0$. 下面我们证明 (4): 我们有

$$\forall x \in \cup_{\alpha\in J} A_\alpha \Leftrightarrow \exists \alpha_0 \in J, x \in A_{\alpha_0} \Leftrightarrow \mathbb{1}_{\cup_{\alpha\in J} A_\alpha} = \mathbb{1}_{A_{\alpha_0}} = 1 = \max_{\alpha\in J} \mathbb{1}_{A_\alpha},$$

$$\forall x \notin \cup_{\alpha\in J} A_\alpha \Leftrightarrow \forall \alpha \in J, x \notin A_\alpha \Leftrightarrow \mathbb{1}_{\cup_{\alpha\in J} A_\alpha} = \mathbb{1}_{A_\alpha} = 0 = \max_{\alpha\in J} \mathbb{1}_{A_\alpha},$$

$$\forall x \in \cap_{\alpha\in J} A_\alpha \Leftrightarrow \forall \alpha \in J, x \in A_\alpha \Leftrightarrow \mathbb{1}_{\cup_{\alpha\in J} A_\alpha} = \mathbb{1}_{A_\alpha} = 1 = \min_{\alpha\in J} \mathbb{1}_{A_\alpha},$$

$$\forall x \notin \cap_{\alpha\in J} A_\alpha \Leftrightarrow \exists \alpha_0 \in J, x \notin A_{\alpha_0} \Leftrightarrow \mathbb{1}_{\cap_{\alpha\in J} A_\alpha} = \mathbb{1}_{A_{\alpha_0}} = 0 = \min_{\alpha\in J} \mathbb{1}_{A_\alpha}.$$

□

1.1.15 设 $A, B \subset \Omega$, 试将 $\mathbb{1}_{A\setminus B}, \mathbb{1}_{A^c}, \mathbb{1}_{A\Delta B}$ 用 $\mathbb{1}_A, \mathbb{1}_B$ 表示出来.

证明: 我们有

$$\begin{aligned}
\mathbb{1}_{A\setminus B} &= \mathbb{1}_{A\cap B^c} = \mathbb{1}_A \mathbb{1}_{B^c} = \mathbb{1}_A(1 - \mathbb{1}_B); \quad \mathbb{1}_{A^c} = 1 - \mathbb{1}_A; \\
\mathbb{1}_{A\Delta B} &= \mathbb{1}_{(A\setminus B)\cup(B\setminus A)} = \mathbb{1}_{A\setminus B} + \mathbb{1}_{B\setminus A} - \mathbb{1}_{A\setminus B}\mathbb{1}_{B\setminus A} \\
&= \mathbb{1}_A - \mathbb{1}_A\mathbb{1}_B + \mathbb{1}_B - \mathbb{1}_A\mathbb{1}_B - (\mathbb{1}_A - \mathbb{1}_A\mathbb{1}_B)(\mathbb{1}_B - \mathbb{1}_A\mathbb{1}_B) \\
&= \mathbb{1}_A + \mathbb{1}_B - 2\mathbb{1}_A\mathbb{1}_B \\
&= \mathbb{1}_A^2 + \mathbb{1}_B^2 - 2\mathbb{1}_A\mathbb{1}_B \\
&= (\mathbb{1}_A - \mathbb{1}_B)^2 = |\mathbb{1}_A - \mathbb{1}_B|.
\end{aligned}$$

□

§ 1.2 映射与势

1.2.1 证明定理 1.2.3 的证明 (i) 中所提出的事实.

(设 $\{A_\alpha : \alpha \in I\}$, $\{B_\alpha : \alpha \in I\}$ 是两个集类, $\forall \alpha \in I$, $A_\alpha \sim B_\alpha$, 且 $A_\alpha, \alpha \in I$ 两两不交, $B_\alpha, \alpha \in I$ 两两不交, 则 $\bigcup\limits_{\alpha\in I} A_\alpha \sim \bigcup\limits_{\alpha\in I} B_\alpha$.)

证明: 设双射 $f_\alpha : A_\alpha \mapsto B_\alpha$, $x \to f_\alpha(x)$, $\forall x \in A_\alpha$. 则

$$\begin{aligned}
f : \bigcup_{\alpha\in I} A_\alpha &\mapsto \bigcup_{\alpha\in I} B_\alpha, \\
x &\to \sum_{\alpha\in I} \mathbb{1}_{A_\alpha} f_\alpha(x)
\end{aligned}$$

也是双射. 证毕.

□

1.2.2 试作开上半平面与开单位圆间的一一映射.

证明: 回忆复变函数中的分式线性映射:

$$\begin{aligned} f:\{z:\operatorname{Im} z>0\} &\mapsto B(0,1) \\ z &\to \mathrm{e}^{\mathrm{i}\theta}\frac{z-\bar{z}_0}{z-z_0}. \end{aligned}$$

取 $z=x+\mathrm{i}y$, $z_0=x_0+\mathrm{i}y_0$, 可得一一映射:

$$\begin{aligned} g:\mathbb{R}\times\mathbb{R}^* &\mapsto B(0,1) \\ (x,y) &\to \left(\operatorname{Re}\left(\mathrm{e}^{\mathrm{i}\theta}\frac{z-\bar{z}_0}{z-z_0}\right),\operatorname{Im}\left(\mathrm{e}^{\mathrm{i}\theta}\frac{z-\bar{z}_0}{z-z_0}\right)\right) \end{aligned}$$

□

1.2.3 设集合 A 有 n 个元素 $(n=1,2,\cdots)$, 在 A 的子集和它的余集间建立一一对应, 由此证明

$$C_n^k=C_n^{n-k},\quad 0\leqslant k\leqslant n,$$

其中 C_n^k 表示从 n 个元中取出 k 个元的组合数.

证明: 考虑集类

$$\mathscr{A}_1:=\{A_1\subset A,|A_1|=k\},\quad \mathscr{A}_2:=\{A_2\subset A,|A_2|=n-k\},$$

则有一一映射

$$\begin{aligned} f:\mathscr{A}_1 &\mapsto \mathscr{A}_2; \\ A_1 &\to A\backslash A_1. \end{aligned}$$

故 $\overline{\overline{\mathscr{A}_1}}=C_n^k=\overline{\overline{\mathscr{A}_2}}=C_n^{n-k}$. □

§ 1.3 可数集

1.3.1 $\mathbb{R}^d$ 中以有理点为中心, 以正有理数为半径的球的全体是可数集.

证明: 令 $\mathscr{A}:=\{B(x,r),x\in\mathbb{Q}^d,r\in\mathbb{Q}^*\}$, 则 $\mathscr{A}\sim\mathbb{Q}^d\times\mathbb{Q}^*$. 我们知道 $\mathbb{Q}$ 和 $\mathbb{Q}^*$ 是可数集, 而有限个可数集的直积仍是可数集, 故 $\mathscr{A}$ 是可数集. □

1.3.2 直线上一个由长度不为零的互不相交的开区间组成的集至多可数.

证明: 令 $\mathscr{A}:=\{(a_\alpha,b_\alpha),\alpha\in I\}$, 这里 I 是指标集. 则有单射:

$$\begin{aligned} \mathscr{A} &\mapsto \mathbb{Q}, \\ (a_\alpha,b_\alpha) &\to c_\alpha \end{aligned}$$

这里 $c_\alpha\in\mathbb{Q}$ 且 $c_\alpha\in(a_\alpha,b_\alpha)$. 则 $\overline{\overline{\mathscr{A}}}\leqslant\overline{\overline{\mathbb{Q}}}=\aleph_0$ 至多可数. □

1.3.3 证明任一可数集的所有有限子集的全体组成可数集.

证明: 设 A 可数, 其元素为 $a_1, a_2, \cdots, a_n, \cdots$, 以 $\mathscr{A}_n$ 表示 A 中 n 个元素组成的子集的全体. 定义 $\mathcal{F}$ 为 A 中所有有限子集全体, 则 $\mathcal{F} = \bigcup\limits_{n=0}^{\infty} \mathscr{A}_n$.

另设 $\mathscr{D} = \{(k_1, k_2, \cdots, k_n) | k_i \in \mathbb{N}\}$, 则 $\mathscr{D} \sim \mathbb{N}^n \sim \mathbb{N}$ 为可数集. 同时 $\mathscr{A}_n$ 中的每个元素 $\{a_{k_1}, a_{k_2}, \cdots, a_{k_n}\}$ 都可以对应到 $\mathscr{D}$ 中的元素, 由于 $\mathscr{D}$ 是可数集, 所以 $\mathscr{A}_n$ 至多可数. 我们知道 $\mathscr{A}_n$ 是无限集, 故 $\mathscr{A}_n$ 可数. 故 $\mathcal{F} = \bigcup\limits_{n=0}^{\infty} \mathscr{A}_n$ 也可数. □

1.3.4 给定平面上一个集, 若此集中的任意两点间距离大于某个固定正数 α, 则此集至多是可数集.

证明: 令此集合为 A, 则对某个 $a_k \in A$, 在平面上取以 a 为中心而边长为 $\sqrt{2}\alpha$ 的正方形 Ω_k, 则 $\Omega_k \cap A = \{a_k\}$. 我们可将全平面分为可数个, 不交的, 边长为 $\sqrt{2}\alpha$ 的正方形, 而任意一个正方形中至多有两个属于 A 的点. 因此 A 至多可数. □

1.3.5 设 A 是有限集或者可数集, B 是无限集, 则 $A \cup B \sim B$.

证明: 首先考虑 $A \cap B = \varnothing$ 的情形:

(1) 若 A 有限, 设 $A = \{a_1, a_2, \cdots, a_N\}$; 又由定理 1.3.2, 可取 B 的可数子集 $B_1 := \{b_1, b_2, \cdots\}$, $B_2 = B \backslash B_1$, 则可构造这样的一一映射 f:

$$f: A \cup B \to B$$

$$\begin{cases} f(a_i) = b_i, & a_i \in A, \\ f(b_i) = b_{N+i}, & b_i \in B_1, \\ f(x) = x, & x \in B_2, \end{cases}$$

从而 $A \cup B \sim B$.

(2) 若 A 可数, 设 $A = \{a_1, a_2, \cdots\}$; 又由定理 1.3.2, 可取 B 的可数子集 $B_1 := \{b_1, b_2, \cdots\}$, $B_2 = B \backslash B_1$, 则可构造这样的一一映射 f:

$$f: A \cup B \to B$$

$$\begin{cases} f(a_i) = b_{2i-1}, & a_i \in A, \\ f(b_i) = b_{2i}, & b_i \in B_1, \\ f(x) = x, & x \in B_2, \end{cases}$$

从而 $A \cup B \sim B$.

然后考虑 $A \cap B \neq \varnothing$ 的情形: 由于 A 至多可数, 则 $A \backslash B$ 也至多可数, 从而 $A \cup B = (A \backslash B) \cup B \sim B$. □

§ 1.4 不可数集

1.4.1 证明: 定义在 $[a, b]$ 上的连续函数的全体组成的集 $C[a, b]$ 的势为 $\aleph$.

证明: 我们知道常数函数是连续的, 常数函数的全体 K 的势为 $\aleph$. 所以

$$E^{\infty} \sim K \subset C[a, b],$$

考虑 Bernstein 定理, 我们只需证明 E^∞ 的某个子集与 $C[a,b]$ 等势.

将 $[a,b]$ 中的有理数全体排成一列, 记为 $r_1, r_2, \cdots$, 任何一个连续函数 $f(x)$ 在 $[a,b]$ 上的取值都可以由 $f(r_1), f(r_2), \cdots$, 完全确定. 因此存在可逆映射

$$\varphi : C[a,b] \mapsto E^\infty,$$
$$f \to (f(r_1), f(r_2), \cdots)$$

故 $C[a,b] \sim \varphi(C[a,b]) \subset E^\infty$, 故 $C[a,b] \sim E^\infty$, 其势为 $\aleph$. □

1.4.2 (1) 证明定义在 $[a,b]$ 上的右连续的单调函数全体的势为 $\aleph$;

(2) 定义在 $[a,b]$ 上的单调函数全体具有怎么样的势?

证明: (1) 由 §1.3 节例 3 知 f 的间断点至多可数;

类似于习题 1.4.1, 将 $[a,b]$ 中的有理数全体排成一列, 记为 $r_1, r_2, \cdots$, 任一右连续单调函数 $f(x)$ 在 $[a,b]$ 上的取值都可以由 $f(x_1), f(x_2), \cdots, f(r_1), f(r_2), \cdots$, 完全确定. 因此存在可逆映射

$$\varphi : C_r[a,b] \mapsto E^\infty,$$
$$f \to (r_1, r_2, \cdots, f(x_1), f(x_2), \cdots, f(r_1), f(r_2), \cdots).$$

故 $C_r[a,b] \sim \varphi(C_r[a,b]) \subset E^\infty$, 而常数函数全体 K 的势为 $\aleph$, 所以

$$E^\infty \sim K \subset C_r[a,b], \quad C_r[a,b] \sim \varphi(C_r[a,b]) \subset E^\infty.$$

由 Bernstein 定理知道 $C_r[a,b] \sim E^\infty$, 其势为 $\aleph$.

(2) 定义 $M[a,b]$ 为 $[a,b]$ 上的单调函数全体, 对 $f \in M[a,b]$, 由 §1.3 节例 3 知单调函数的间断点至多可数. 与 (1) 同理, $\overline{\overline{M[a,b]}} = \aleph$. □

证法二: 仍记 $M[a,b]$ 为 $[a,b]$ 上的单调函数全体.

首先易证 $\overline{\overline{M[a,b]}} \geqslant \aleph$, 这是因为取这样的单调函数 $f_k : f_k(x) = k(x-a)$, $k \in \mathbb{R}$, 则 $M[a,b] \supset \{f_k : k \in \mathbb{R}\} \sim \mathbb{R}$;

其次, 由 §1.3 节例 3 知任意单调函数 f 都可以分解成 $f = g + h$ 的形式, 其中 $g \in C[a,b]$ 是连续函数, h 是仅在 $[a,b]$ 上至多可数个点取非零值、且这些非零值或者同时为正或者同时为负的函数.

现在分别对上述分解 $f = g + h$ 中的 g 和 h 进行进一步的讨论:

(1) 对于 $g \in C[a,b]$, 习题 1.4.1已经证明 $\overline{\overline{C[a,b]}} = \aleph$;

(2) 而要确定函数 h, 只需给出跳跃点的位置以及在跳跃点处的取值, 即每个 h 可对应某个实数列; 而由定理 1.4.3 知全体实数列的势是 $\aleph$, 从而 h 的全体的势不超过 $\aleph$.

因此, 满足上述分解 $f = g + h$ 的 g 全体的势、h 全体的势均不超过 $\aleph$, 从而 $\overline{\overline{M[a,b]}} \leqslant \aleph^2 = \aleph$ ($\aleph^2$ 表示对于两个势为 $\aleph$ 的集合 A, B 而言 $A \times B$ 的势, 由 $\overline{\overline{\mathbb{R}^2}} = \overline{\overline{\mathbb{R}}} = \aleph$(为什么?) 即得 $\aleph^2 = \aleph$).

综合 $\overline{\overline{M[a,b]}} \geqslant \aleph$ 及 $\overline{\overline{M[a,b]}} \leqslant \aleph$, 由 Berstein 定理 (定理 1.2.3) 即可得 $\overline{\overline{M[a,b]}} = \aleph$. □

注: 两个证法都用到这样的结论: $\mathbb{R}$ 上单调函数的间断点的全体至多可数 (§1.3 节例 3). 现给出这一结论的另一个证明 (by 229):

首先证明闭区间 $[a,b]$ 上不降函数 f 的间断点的全体可数. 不妨假设 $-\infty < f(a) \leqslant f(b) < +\infty$, 否则取 $\alpha = \inf\{x \in [a,b]:\ f(x) > -\infty\}$, $\beta = \sup\{x \in [a,b]:\ f(x) < +\infty\}$, 则可以说明 f 在 $[\alpha,\beta]$ 上间断点可数 (α,β 至多是间断点, 即至多再添加两个间断点), 从而 f 在 $[a,b]$ 上间断点可数.

现假设 f 在 $[a,b]$ 上有无穷多个间断点, 这些间断点的全体记为 $\{x_\alpha:\ \alpha \in \Lambda\}$, 则 $\sum\limits_{\alpha\in\Lambda}\left(f\left(x_\alpha^+\right) - f\left(x_\alpha^-\right)\right) \leqslant f(b) - f(a) < +\infty$, 由此断言 Λ 可数: 事实上, 可以证明这样的结论: 若 $\forall \alpha \in \Lambda$, $x_\alpha > 0$且 $\sum\limits_{\alpha\in\Lambda} x_\alpha < +\infty$, 则 Λ可数. 反设 Λ 不可数, 设 $\Lambda_n = \left\{\alpha \in \Lambda:\ x_\alpha > \dfrac{1}{n}\right\}$, 则 $\bigcup\limits_{n=1}^{\infty} \Lambda_n = \Lambda$, 从而至少存在 $n_0 \in \mathbb{N}$ 使得 Λ_{n_0} 中有无穷个元素, 进而有 $\sum\limits_{\alpha\in\Lambda} x_\alpha \geqslant \sum\limits_{\alpha\in\Lambda_{n_0}} x_\alpha \geqslant \sum\limits_{\alpha\in\Lambda_{n_0}} \dfrac{1}{n_0} \xlongequal{\Lambda_{n_0}\text{中有无限个元素}} +\infty$, 矛盾! 故 Λ 可数, 从而证得在 $[a,b]$ 上单调的函数 f 的间断点的全体必然至多可数.

最后, 由 $\mathbb{R} = \bigcup\limits_{n=-\infty}^{\infty} [n, n+1]$ 以及定理 1.3.4(2) 即可知 $\mathbb{R}$ 上单调函数的间断点的全体至多可数. □

1.4.3 证明自然数列全体的势为 $\aleph$.

证明: 考虑 $[0,1]$ 上的二进制小数集 $B := \left\{\sum\limits_{n=1}^{\infty} \dfrac{x_n}{2^n}, x_n = 0, 1\right\}$, 则 $\overline{\overline{B}} = \aleph$. 定义自然数列全体为:

$$A := \{(a_1, a_2, \cdots), a_k \in \mathbb{N}\},$$

于是有一一映射:

$$\varphi : A \mapsto B,$$
$$(a_1, a_2, \cdots) \to \sum_{n=1}^{\infty}\left(\sum_{k=1}^{n-1} a_k + k + 1\right)\left(1 - 2^{-a_n}\right).$$

这里 $\sum\limits_{n=1}^{\infty}\left(\sum\limits_{k=1}^{n-1} a_k + k + 1\right)\left(1 - 2^{-a_n}\right)$ 是在 B 中取

$$x_1, \cdots, x_{a_1} = 1, \quad x_{a_1+1} = 0, \quad x_{a_1+2}, \cdots, x_{a_1+a_2+1} = 1,$$

以此类推并求和得到的. □

证法二: 由定理 1.4.3 知自然数列全体的势小于等于 $\aleph$; 而将 $[0,1]$ 中的数用二进制小数去表达, 又可知自然数列全体的势大于等于 $\aleph$; 由 Bernstein 定理即得证. □

1.4.4 证明自然数的全体子集组成的集 $\mathscr{P}(\mathbb{N})$ 的势为 $\aleph$.

证明: 我们仍然考虑做一个 $\mathscr{P}(\mathbb{N})$ 到 $(0,1)$ 的映射. 考虑二进制小数组成的集 $B := \left\{\sum\limits_{n=1}^{\infty} \dfrac{x_n}{2^n}, x_n = 0, 1\right\}$, 设 $\mathscr{P}(\mathbb{N})$ 的元为 $A_1, A_2, \cdots$, 则有一一映射

$$\varphi : \mathscr{P}(\mathbb{N}) \mapsto B,$$
$$A_k \to \sum_{n=1}^{\infty} \frac{\mathbb{1}_{\{n \in A_k\}}}{2^n},$$

于是 $\mathscr{P}(\mathbb{N}) \sim B \sim (0,1)$. □

第二章 距离空间

§ 2.1 定义及例

2.1.1 若 $p>1,\ q>1,\ \dfrac{1}{p}+\dfrac{1}{q}=1$, 则 $\forall a,b\in\mathbb{R}_+,\ a^{\frac{1}{p}}b^{\frac{1}{q}}\leqslant\dfrac{a}{p}+\dfrac{b}{q}$.

证明: 注意 $\ln x$ 是凹的, 由 Jensen 不等式可得 $\dfrac{1}{p}\ln a+\dfrac{1}{q}\ln b\leqslant\ln\left(\dfrac{a}{p}+\dfrac{b}{q}\right)$. 也即 $a^{\frac{1}{p}}b^{\frac{1}{q}}\leqslant\dfrac{a}{p}+\dfrac{b}{q}$. □

2.1.2 设 $[a,b]$ 是给定的区间, $C[a,b]$ 是 $[a,b]$ 上全体连续函数的集, 则由

$$\rho(f,g):=\sup_{a\leqslant t\leqslant b}|f(t)-g(t)|,\quad f,g\in C[a,b]$$

定义的 ρ 为 $C[a,b]$ 上的一个距离.

设 $\varphi\in C[a,b]$, 且 $\forall t\in[a,b]$, $\varphi(t)>0$, $p\geqslant 1$, 则由

$$\rho_p(f,g):=\left[\int_a^b|f(t)-g(t)|^p\varphi(t)\mathrm{d}t\right]^{\frac{1}{p}},\quad f\in C[a.b]$$

定义的 ρ_p 也是 $C[a,b]$ 的距离.

证明: (1)

(i) 显然有 $\rho(f,g)\geqslant 0$;

(ii) 实际上 $\rho(f,g)=0\Rightarrow\sup\limits_{a\leqslant t\leqslant b}|f(t)-g(t)|=0\Rightarrow\forall t\in[a,b],f(t)=g(t)$;

(iii) $\rho(f,g)=\rho(g,f)$;

(iv) $\rho(f,g)=\sup\limits_{a\leqslant t\leqslant b}|f(t)-g(t)|\leqslant\sup\limits_{a\leqslant t\leqslant b}(|f(t)-h(t)|+|h(t)-g(t)|)=\rho(f,h)+\rho(h,g)$.

(2)

(i) 由于 $\varphi>0$, 故 $\rho_p(f,g)\geqslant 0$;

(ii) $\rho_p(f,g)=0\Rightarrow|f(t)-g(t)|=0$, a.e.. 又 $f,g\in C[a,b]$, 故 $f=g$;

(iii) 显然 $\rho_p(f,g)=\rho_p(g,f)$;

(iv) 由 Minkowski 不等式, 我们有

$$\rho_p(f,g)=\||f-g|\varphi(t)^{1/p}\|_p\leqslant\||f-h|\varphi(t)^{1/p}\|_p+\||h-g|\varphi(t)^{1/p}\|_p=\rho_p(f,h)+\rho_p(h,g).$$

故 ρ, ρ_p 都是 $C[a,b]$ 上的距离. $\square$

2.1.3 设 $\alpha := (\alpha_1, \alpha_2, \cdots)$ 为正实数序列, 满足 $\sum\limits_{k=1} \alpha_k < \infty, p \geqslant 1, E := \mathbb{R}$ 或 $\mathbb{C}$ (全体复数集). 令

$$E^{\mathbb{N}}(\alpha; p) = \left\{ x : x = (x_1, x_2, \cdots), x_k \in E, k \in \mathbb{N}, \sum_{k=1}^{\infty} \alpha_k |x_k|^p < \infty \right\},$$

$$\rho_p(x, y) = \left[\sum_{k=1}^{\infty} \alpha_k |x_k - y_k|^p \right]^{\frac{1}{p}}, \quad x, y \in E^{\mathbb{N}}(\alpha; p),$$

则 $\left(E^{\mathbb{N}}(\alpha; p), \rho_p\right)$ 为一距离空间. 此外, 设

$$E^{\mathbb{N}} := \left\{ x : x = (x_1, x_2, \cdots), \sup_{x \in \mathbb{N}} |x_k| < \infty \right\},$$

$$\rho(x, y) := \sup_{k \in \mathbb{N}} |x_k - y_k|, x, y \in E^{\mathbb{N}},$$

则 $\left(E^{\mathbb{N}}, \rho\right)$ 也是一距离空间. $\lim\limits_{p \to \infty} \rho_p = \rho$ 是否还成立?

证明: (1)

(i) 显然 $\rho_p \geqslant 0$;

(ii) 显然 $\rho_p = 0 \Rightarrow x = y$;

(iii) 显然 $\rho_p(x, y) = \rho_p(y, x)$;

(iv) 由 Minkowski 不等式, 有

$$\rho_p(x, y) = \|\alpha^{1/p}(x - y)\|_p \leqslant \|\alpha^{1/p}(x - z)\|_p + \|\alpha^{1/p}(z - y)\|_p = \rho_p(x, z) + \rho_p(z, y).$$

因此 ρ_p 是 $E^{\mathbb{N}}(\alpha; p)$ 上的距离. 类似易证 ρ 是 $E^{\mathbb{N}}$ 上的距离.

(2) 我们有

$$\rho_p = \left[\sum_{k=1}^{\infty} \alpha_k |x_k - y_k|^p \right]^{\frac{1}{p}} \leqslant \sup_{k \in \mathbb{N}} |x_k - y_k| \left(\sum_{k=1}^{\infty} \alpha_k \right)^{\frac{1}{p}},$$

又 $\sum\limits_{k=1}^{\infty} \alpha_k < \infty$, 因此 $\limsup\limits_{p \to \infty} \rho_p \leqslant \rho$.

又 $\rho = \sup\limits_{k \in \mathbb{N}} |x_k - y_k|$, $\forall \varepsilon > 0$, $\exists J \subset \mathbb{N}$, s.t. $\forall k \in J$, $\rho < |x_k - y_k| + \varepsilon$, $\rho_p > \left[\sum\limits_{k \in J} \alpha_k |x_k - y_k|^p \right]^{\frac{1}{p}} > p - \varepsilon$.
由 ε 的任意性知道 $\liminf\limits_{p \to \infty} \rho_p \geqslant \rho$. $\square$

2.1.4 设 $E := \mathbb{R}$ 或 $\mathbb{C}, E^{\mathbb{N}} := \{x : x = (x_1, x_2, \cdots), x_n \in E, n \in \mathbb{N}\}$, $\{\alpha_n : n \in \mathbb{N}\}$ 为一可和的正数列, 试证: $\left(E^{\mathbb{N}}, \rho\right)$ 是距离空间, 其中 ρ 有如下定义

$$\rho(x, y) := \sum_{n=1}^{\infty} \alpha_n \frac{|x_n - y_n|}{1 + |x_n - y_n|}, x, y \in E^{\mathbb{N}}.$$

证明:

(i) 显然 $\rho \geqslant 0$;

(ii) $\rho(x,y)=0 \Rightarrow \forall k \in \mathbb{N}, x_k=y_k, x=y$;

(iii) $\rho(x,y)=\rho(y,x)$;

(iv) 我们有

$$\begin{aligned}\rho(x,y) &= \sum_{n=1}^{\infty} \alpha_n \frac{|x_n-y_n|}{1+|x_n-y_n|} \\ &\geqslant \sum_{n=1}^{\infty} \alpha_n \left(\frac{|x_n-z_n|}{1+|x_n-z_n|} = \frac{|z_n-y_n|}{1+|z_n-y_n|} \right) \\ &= \rho(x,z)+\rho(z,y).\end{aligned}$$

因此 ρ 是 $E^{\mathbb{N}}$ 上的距离. □

2.1.5 设 $E := \{x : x=(x_1,x_2,\cdots), x_k \in \{0,1\}, k \in \mathbb{N}\}$, d 为 $\{0,1\}$ 上的离散距离, 试证; 如下定义的 ρ 是 E 上的距离,

$$\rho(x,y) := \sum_{k=1}^{\infty} \frac{1}{2^k} d(x_k,y_k), x,y \in E.$$

证明: 注意到在 $\{0,1\}$ 上, $d(x,y) := \mathbb{1}_{\{x\neq y\}} = |x-y|$, 故只需在习题 2.1.4中取 $\alpha_n = 2^{-n}$ 即可. □

2.1.6 设 p 是一给定素数, 对每一给定的 $n \in \mathbb{N}$, 令 $U_p(n)$ 是能整除 n 的 p 的幂的最大指数 (即 $p^{U_p^{(n)}}$ 能整除 n, 但 $p^{U_p(n)+1}$ 不能整除 n). 规定 $U_p(0)=0$. 设 $x=\pm\frac{r}{s}$ 为有理数, $r,s \in \mathbb{N}$, 定义 $U_p(x) := U_p(r)-U_p(s)$. 试证: (1) $U_p(x), x \in \mathbf{Q}$ 是 $\mathbf{Q}$ 到 $\mathbf{Z}$ 的一个映射. (2) $\forall x,y \in \mathbf{Q}$, 令

$$d(x,y) := \begin{cases} 0, & x=y, \\ p^{-U_p(x-y)}, & x \neq y, \end{cases}$$

则 $d(x,y)$ 是 $\mathbf{Q}$ 上的一个距离 (此距离称为 p-adic 距离). 事实上可以证明

$$d(x,y) \leqslant \max(d(x,z), d(z,y)), \quad \forall x,y,z \in \mathbf{Q}.$$

注: 对给定的素数 p, 由 p-adic 距离定义的有理数基本列 (基本列的一般定义参看定义 2.3.1) 出发, 采用由绝对值定义距离的有理数基本列出发构造实数一样的方法, 也可以构造出一个完备数域 (对 p-adic 距离而言). 这个数域与 p 有关, 称为 p-adic 域. p-adic 域是一种非阿基米德域, 近来发现它在理论物理中有用.

解: (1) 即证: 对于 $x \in \mathbb{Q}$ 而言, $U_p(x)$ 的值仅与 x 自身有关, 而与 x 被表示成的整数之比的形式无关, 即当 $x=\frac{r}{s}=\frac{m}{n}$ $(r,s,m,n \in \mathbb{Z},\ s,n \neq 0,\ r \neq m)$ 时, 有 $U_p(r)-U_p(s)=U_p(m)-U_p(n)$.

显然, 对于任意某个非零整数 x, 可以对其作因数分解, 把因子 p 全部提取出来, 写成 $x=y\cdot p^z$ 的形式. 由此, 联想题目中 U_p 在整数上的定义, 不难想到 (其中"$\nmid$" 表示" 不整除"):

$$\text{对于非零整数 } x, \text{有 } U_p(x)=z \Longleftrightarrow \exists \text{非零整数} y:\ p \nmid y,\ \exists \text{非负整数} z,\ \text{s.t. } x=y\cdot p^z. \qquad (*)$$

例如, $12=3\cdot 2^2\cdot$, 因此 $U_2(12)=2$; $210=30\cdot 7$, 因此 $U_7(210)=1$.

不妨设 $r \neq 0$ 且 $m = kr$, $n = ks$ (k是非零整数). 基于 (*), 我们可设 $U_p(r) = b$, $U_p(s) = d$, $U_p(k) = h$, 同时将 r, s, k 分别作因数分解, 写成 $r = a \cdot p^b$, $s = c \cdot p^d$, $k = g \cdot p^h$ 的形式 (其中 a, c, g 为非零整数, b, d, h 为非负整数), 则 $p \nmid a, c, g$. 由此, $m = kr = ag \cdot p^{b+h}$ 且 $p \nmid ag$, 从而 $U_p(m) = b + h$; $n = ks = cg \cdot p^{d+h}$ 且 $p \nmid cg$, 从而 $U_p(m) = d + h$. 因此 $U_p(m) - U_p(n) = (b + h) - (d + h) = b - d = U_p(r) - U_p(s)$.

(2) 即验证 $d(x, y)$ 满足定义 2.1.2 的 4 个条件. $d(x, y)$ 显然满足定义 2.1.2 的条件 (1)(2) (其中满足条件 (2) 在本题第 1 小问已证明). 根据 U_p 的定义, $U_p(x) = U_p(-x)$, 因此 $d(x, y)$ 显然也满足条件 (3).

至于条件 (4), 事实上可以直接证明更强的结论: $\forall x, y, z \in \mathbb{Q}$, $d(x, y) \leqslant \max\{d(x, z), d(z, y)\}$. 这一结论等价于 $\forall x, y, z \in \mathbb{Q}$, $U_p(x - y) \geqslant \min\{U_p(x - z), U_p(z - y)\}$. 由于 $x - y = (x - z) + (z - y)$, 这又等价于 $\forall x, y \in \mathbb{Q}$, $U_p(x + y) \geqslant \min\{U_p(x), U_p(y)\}$. 下面将证明这一等价结论. 不妨假设 $x, y, x + y \neq 0$.

先考虑 $x, y \in \mathbb{Z}$ 的情形. 设 $x = a \cdot p^b, y = c \cdot p^d, x + y = g \cdot p^h$ 且 $U_p(x) = b$, $U_p(y) = d$, $U_p(x + y) = h$, 则根据 (*) 有 $p \nmid a, c, g$. 不妨设 $b \geqslant d$, 则要证明的是 $h \geqslant d$. 由 $g \cdot p^h = a \cdot p^b + c \cdot p^d$ 知 $g = a \cdot p^{b-h} + c \cdot p^{d-h} = (a \cdot p^{b-d} + c) \cdot p^{d-h}$. 由于 $p \nmid g$, 必然有 $d - h \leqslant 0$, 即 $h \geqslant d$.

再考虑 $x, y \in \mathbb{Q}$ 的情形. 设 $x = \dfrac{m}{n}$, $y = \dfrac{r}{s}$, $m, n, r, s \in \mathbb{Z}$, 则 $x + y = \dfrac{ms + nr}{ns}$, 从而

$$
\begin{aligned}
U_p(x + y) &= U_p(ms + nr) - U_p(ns) \\
&\geqslant \min\{U_p(ms), U_p(nr)\} - U_p(ns) \\
&= \min\{U_p(ms) - U_p(ns), U_p(nr) - U_p(ns)\} \\
&= \min\left\{U_p\left(\frac{ms}{ns}\right), U_p\left(\frac{nr}{ns}\right)\right\} \\
&= \min\{U_p(x), U_p(y)\}.
\end{aligned}
$$

综上所述证毕. □

§ 2.2 开集、闭集

2.2.1 给出 $\overline{B(x, r)} \neq \overline{B}(x, r)$ 的例子.

证明: 考察整数集合 $\mathbb{Z}$, 定义距离 $d(m, n) = |m - n|$, 则 $B(0, 1) = \{0\}$, $\overline{B(0, 1)} = \{0\}$. 但 $\overline{B}(0, 1) = \{-1, 0, 1\}$. □

2.2.2 给定 $a \in E$, 则任何 a 的邻域 A, 一定存在 $n \in \mathbb{N}$ 使 $B\left(a, \dfrac{1}{n}\right) \subset A$.

证明: 我们知道 $\exists r$, s.t. $B(a, r) \subset A$, 因此只需取 $\dfrac{1}{n} < r$ 即可. □

2.2.3 设 $\varnothing \neq A \subset E$, 则 A 的任意有限个邻域的交仍然是 A 的邻域.

证明: A 的任意邻域是开集, 因此存在 $\{O_n\}$ 为开集, $A \subset O_n$, 且 $\bigcap_n O_n$ 仍为开集. 又 $A \subset \bigcap_n O_n$, 故 $\bigcap_n O_n$ 仍为 A 的邻域. □

2.2.4 $A \subset E$ 为开集的充要条件是 $\partial A \cap A = \varnothing$.

证明: 当 A 为开集, 则 $\partial A = \overline{A} \backslash A^o = \overline{A} \backslash A$, 也即 $\partial A \cap A = \varnothing$. 当 $\partial A \cap A = \varnothing$, 则 $\overline{A} \backslash A^o \cap A = \varnothing$, 也即

$A^o \subset A$, $A \subset A^o$. 因此 $A = A^o$, 为开集. □

2.2.5 $A \subset E$ 为闭集的充要条件是 $\partial A \subset A$.

证明: 当 A 为闭集, 则 $\overline{A} = A \cup \partial A = A$, 故 $\partial A \subset A$. 若 $\partial A \subset A$, 则 $\partial A = \overline{A} \backslash A^o \subset A$, 故 $\overline{A} \subset A$, 也即 $A = \overline{A}$ 为闭集. □

2.2.6 $x \in \partial A$ 的充要条件是 x 的任何邻域既包含 A 的点又包含 A^c 的点.

证明: 由于 $x \in \partial A = \overline{A} \cap \overline{A^c}$, 故 $x \notin A^o$, $x \notin (A^c)^o$. 故 $\forall r > 0$, $B(x,r) \not\subset A$, $B(x,r) \not\subset A^c$. 因此 $B(x,r) \cap A \neq \varnothing$, $B(x,r) \cap A^c \neq \varnothing$. 也即 $x \in A \cap A^c$. □

2.2.7 $A \subset E$ 是开集的充要条件是对任意 $B \subset E$, 都有 $A \cap \overline{B} \subset \overline{A \cap B}$.

证明: 必要性 ($\Rightarrow$): $\forall x \in A \cap \overline{B}$, 由本书闭包的定义 (定义 2.2.5) 知 $x \notin (B^c)^o$. 又由于 $x \notin A^c$ 且 $(A^c)^o \subset A^c$, 可知也有 $x \notin (A^c)^o$. 又由定理 2.2.6(4) 知 $(A^c)^o \cup (B^c)^o \subset ((A \cap B)^c)^o$, 从而 $x \notin ((A \cap B)^c)^o$, 也即 $x \in \overline{A \cap B}$.

充分性 ($\Leftarrow$): 取 $B = A^c$, 可知 $A \cap \overline{A^c} = \varnothing$, 由本书闭包的定义 (定义 2.2.5) 知也即 $A \cap (A^o)^c = \varnothing$, 从而 $A \subset A^o$, 也即 $A^o = A$, 由定理 2.2.6(1) 知 A 是开集. □

2.2.8 给定距离空间 (E, d), $A \subset E$, 则

(1) A' 为闭集;

(2) $A \subset B \Rightarrow A' \subset B'$;

(3) $(A \cap B)' \subset A' \cap B'$;

(4) $(A \cup B)' = A' \cup B'$.

证明: (1) 只需证明 $(A')^c$ 是开集. 事实上, 由导集的定义 (定义 2.4.8) 知 $\forall x \in (A')^c$, 存在 x 的一个邻域 $N(x)$, 使得 $A \cap (N(x) \backslash \{x\}) = \varnothing$. 这等价于 $\exists r > 0$, s.t. $A \cap (B(x,r) \backslash \{x\}) = \varnothing$(请读者自行证明), 也就是说 $(A')^c = \{x: \ \exists r_x > 0, \text{ s.t. } A \cap (B(x, r_x) \backslash \{x\}) = \varnothing\}$.

现设 $A \cap (B(x, r_0) \backslash \{x\}) = \varnothing$ $(r_0 > 0)$, 则 $\forall y \in B\left(x, \dfrac{r_0}{2}\right)$, $A \cap \left(B\left(y, \dfrac{r_0}{2}\right) \backslash \{y\}\right) = \varnothing$ (因为易证 $\left(B\left(y, \dfrac{r_0}{2}\right) \backslash \{y\}\right) \subset (B(x, r_0) \backslash \{x\})$, 可以画图来帮助理解这一事实), 也即 $\forall x \in (A')^c$, $\exists r = \dfrac{r_0}{2} > 0$, s.t. $\forall y \in B(x, r)$, $y \in (A')^c$, 说明 $(A')^c$ 是开集, 因此 A' 是闭集.

(2) 由导集的定义容易说明. $\forall x \in A'$, 存在 x 的一个邻域 $N(x)$ 使得 $A \cap (N(x) \backslash \{x\}) = \varnothing$, 又因为 $A \subset B$, 所以 $B \cap (N(x) \backslash \{x\}) = \varnothing$, 从而 $x \in B'$, 也即 $A' \subset B'$.

(3) 由导集的定义显然, 因为 $A \cap B \cap (N(x) \backslash \{x\}) = \varnothing \Longrightarrow A \cap (N(x) \backslash \{x\}), B \cap (N(x) \backslash \{x\}) = \varnothing$.

(4) 由导集的定义显然, 因为 $(A \cup B) \cap (N(x) \backslash \{x\}) = (A \cap (N(x) \backslash \{x\})) \cup (B \cap (N(x) \backslash \{x\}))$. □

2.2.9 作一实数列使其极限点集为空集.

解: 取 $\{a_n: \ n \in \mathbb{N}\} \subset \mathbb{R}$ 使得 $a_n = n$ 即可. □

2.2.10 作一实数列使其极限点集为全体实数集.

解: 取 $\{a_n: \ n \in \mathbb{N}\} \subset \mathbb{R}$ 使得 $a_n = r_n \in \mathbb{Q}$ 为全体有理数的一个排列即可. □

2.2.11 设 (E,d) 为距离空间, 给定 $A\subset E$. 试证: $\overline{A}\backslash A'$ 为 A 的全体孤立点组成的集.

证明: 由本书导集的定义 (见定义 2.2.8) 知 $\forall x\in (A')^c$, $\exists x$的邻域$N(x)$, s.t. $A\cap(N(x)\backslash\{x\})=\varnothing$, 因此由本书孤立点的定义 (见定义 2.2.8) 知只需证明 $\forall x\in\overline{A}\backslash A'$, $x\in A$. 而由定理 2.2.9 知 $A\cup A'=\overline{A}$, 故证毕. □

2.2.12 $\forall x_0\in\overline{A}$, $\exists A$ 中序列 $\{x_n\}$ 使 $d(x_n,x_0)\to 0\ (n\to\infty)$.

证明: 由本书闭包的定义 (见定义 2.2.5) 知 $\forall x\in\overline{A}$, $\forall r>0$, $B(x,r)\cap A\neq\varnothing$. 由此, 对于 $x_0\in\overline{A}$, 取 $r_1=1$, 则存在 $x_1\in B(x_0,1)\cap A$; 取 $r_2=\dfrac{1}{2}$, 则存在 $x_2\in B\left(x_0,\dfrac{1}{2}\right)\cap A$; 取 $r_3=\dfrac{1}{3}$, 则存在 $x_3\in B\left(x_0,\dfrac{1}{3}\right)\cap A$; 以此类推, 取 $r_n=\dfrac{1}{n}$, 则存在 $x_n\in B\left(x_0,\dfrac{1}{n}\right)\cap A$. 由此我们得到了满足题意的数列 $\{x_n\}$. □

§ 2.3 完备性

2.3.1 $\mathbb{R}$ 为实数集. 设

(1) $\rho_1(x,y)=|\arctan x-\arctan y|$;

(2) $\rho_2(x,y)=|\mathrm{e}^x-\mathrm{e}^y|$.

证明: $(\mathbb{R},\rho_1)$, $(\mathbb{R},\rho_2)$ 都不是完备距离空间.

证明: (1) 令 $\{x_n\}=n$, 则 $\arctan x_n\to\dfrac{\pi}{2}$. 因此 $\forall\varepsilon>0$, $\exists N>0$, s.t. $\forall m,n>N$, $\rho(x_m,x_n)=|\arctan x_n-\arctan x_m|<\varepsilon$. 但 x_n 发散, 因此其不完备.

(2) 令 $\{x_n\}=-n$, 与 (1) 同理可推出不完备. □

2.3.2 设 $\rho(m,n)=|m^{-1}-n^{-1}|$, $m,n\in\mathbb{N}$, 证明: $(\mathbb{N},\rho)$ 不完备.

证明: 令 $\{x_n\}=n$, 与习题 2.3.1(1) 同理. □

2.3.3 $\mathbb{Z}$ 为整数集, $\rho(m,n)=|m-n|$, 证明: $(\mathbb{Z},\rho)$ 是完备距离空间.

证明: 假设 $\{x_n\}\subset\mathbb{Z}$ 为 Cauchy 列, 则 $\forall\varepsilon>0$, $\exists N>0$, s.t. $m.n>N$ 时有 $|x_n-x_m|<\varepsilon$. 也即 $x_n=x_m$, 因此 $x_n\to c$ 为一常数且为整数. 容易验证 ρ 是距离, 因此 $(\mathbb{Z},\rho)$ 是完备距离空间. □

2.3.4 考虑三个定义在整个 $\mathbb{R}$ 上的连续函数集:

(1) 有界连续函数集;

(2) 满足条件 $\lim\limits_{x\to\infty}f(x)=0$ 的连续函数集;

(3) 在有限区间外恒等于零的连续函数集.

如果规定 $\rho(f,g)=\sup\limits_{-\infty<x<\infty}|f(x)-g(x)|$, 哪一个是完备空间, 哪一个不是?

证明: (1) 完备

(2) 不完备, 考虑 $f_n(x)=\begin{cases}x^{-\frac{1}{2n+1}}, & |x|\geqslant 1\\ x, & |x|<1\end{cases}$, 则有 $\lim\limits_{x\to\infty}f(x)=0$. 但 $f_n\to f=\begin{cases}1, & |x|\geqslant 1\\ x, & |x|<1\end{cases}$, 此时 $\lim\limits_{x\to\infty}f(x)\neq 0$.

(3) 不完备, 只需考虑 $f_n = \begin{cases} 0, & x \notin [-1,1] \\ 1, & x \in \left(-1+\dfrac{1}{n}, 1-\dfrac{1}{n}\right) \\ n(x+1), & x \in \left[-1, -1+\dfrac{1}{n}\right) \\ -n(x-1), & x \in \left(1-\dfrac{1}{n}, 1\right] \end{cases}$. 此时 $f_n \to \mathbb{1}_{(-1,1)} \notin C(\mathbb{R})$. □

2.3.5 给定 $\Omega := \{\Delta : \Delta = [a,b], a < b, a, b \in \mathbb{R}\}$. 对于 $\Delta_1 = [a,b]$, $\Delta_2 = [c,d]$, 规定 $\rho(\Delta_1, \Delta_2) := |a-c| + |b-d|$, 证明: ρ 是 Ω 上的距离函数, 但 (Ω, ρ) 不完备.

证明: 易证 ρ 是距离, 取 $\Delta_n = \left[-\dfrac{1}{n}, \dfrac{1}{n}\right]$, 则 $\Delta_n \to \{0\} \notin \Omega$, 不完备. □

2.3.6 若 (X_1, ρ_1) 与 (X_2, ρ_2) 等距同构, (X_1, ρ_1) 完备, 则 (X_2, ρ_2) 完备.

证明: 考虑 X_2 中的基本列 $\{x_n\}$, 则 $\exists N > 0$, $\forall n > N$, $\rho_2(x_n, x_m) < \varepsilon$. 令 φ 是 X_2 到 X_1 的同构映射, 由于同构是等距的, 所以 $\forall n > N$, $\rho_1(\varphi(x_n), \varphi(x_m)) = \rho_2(x_n, x_m) < \varepsilon$. 因此 $\{\varphi(x_n)\}$ 是 X_1 中的基本列, 又 X_1 完备, 故 $\exists y \in X_1$, $\varphi(x_n) \to y$. 且必然存在 $x \in X_2$, s.t. $\varphi(x) = y$. 又 $\rho_2(x_n, x) = \rho_1(\varphi(x_n), y) < \varepsilon$, 故 (X_2, ρ_2) 完备. □

2.3.7 证明 $(C[a,b], \rho)$ 完备.

证明: 显然 ρ 是距离, 下面证明其完备.

取 $\{f_n\} \subset C[a,b]$ 为基本列, 也即 $\forall \varepsilon > 0, \exists N \in \mathbb{N}$, s.t. $\forall m, n > N$, $\rho(f_n, f_m) = \max\limits_{a \leqslant t \leqslant b} |f_m(t) - f_n(t)| < \varepsilon$. □

2.3.8 证明引理 2.3.4.

证明: 我们只需证明有收敛子列的基本列收敛. 设 $\{a_n\} \subset E$ 为基本列, $\{a_{nk}\}$ 为其收敛到 a 的子列, 则 $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, s.t. $\forall k > N$, $|a_{nk} - a| < \varnothing$. 又 $|a_n - a| \leqslant |a_n - a_{nk}| + |a_{nk} - a|$, 由于 $\{a_n\}$ 为基本列, 故 $|a_n - a_{nk}|$ 也趋于 0, 因此 $a_n \to a$. □

2.3.9 证明引理 2.3.5.

证明: $\Longrightarrow$: 若 (F, d) 为完备子空间, 则 $\overline{F} = F$, 故 F 是闭集.

$\Longleftarrow$: 若 F 为闭集, 任取基本列 $\{x_n\} \subset F$, 则它也是 E 中的基本列. 故其存在极限点 $x \in E$. 又 F 闭, 因此 $x \in F$, 故 F 完备. □

2.3.10 (Sierpinski 距离空间) 设 $E := \{x_n : n \in \mathbb{N}\}$, 试证: 由

$$d(x_n, x_m) := \begin{cases} 1 + (n+m)^{-1}, & n \neq m, \\ 0, & n = m, \quad n, m \in \mathbb{N} \end{cases}$$

定义的 d 是 E 上的一个距离; (E, d) 的每一单点集都是开集, 因而没一点都是孤立点; (E, d) 是完备距离空间.

再证: $\overline{B}(x_n, 1 + (2n)^{-1}), n \in \mathbb{N}$, 是一递降的闭球列, 但它们的交集为空集.

证明: d 是距离是显然的.

设 $A=\{x_0\}\in E$, 则 $A^o=A$, 因此 A 是开集, 且 x_n 均为孤立点. 对任意 E 中的基本列 $\{x_n\}$, $\forall\varepsilon>$, $\exists N\in\mathbb{N}$, $\forall m,n>N$, $d(x_n,x_m)<\varepsilon$. 则 $x_n=x_m=x_N\in E$, 因此 (E,d) 完备. 同时注意到不同点之间距离至少为 1, 因此 $B(x_n,1+(2n)^{-1})$ 是闭球, 且 $d(x_{n+1},x_n)=1+\dfrac{1}{2n+1}<1+\dfrac{1}{2n}$, 因此是递降闭球列. 又 $B(x_n,1+(2n)^{-1})$ 的元素为 $\{x_m\}_{m\geqslant n}$, 故 $\bigcap\limits_{n=1}^{\infty}B(x_n,1+(2n)^{-1})=\varnothing$. □

§ 2.4 可分性、列紧性与紧性

2.4.1 试证: $(\mathbb{R},d)$ $(d(x,y)=|x-y|)$ 完备, 可分, 但不是紧空间.

证明: 完备性与可分性是显然的. 下面证明 $\mathbb{R}$ 不存在有限开覆盖, 考虑反证法, 否则令 $\mathbb{R}=\bigcup\limits_{m=1}^{n}O_n$, 令 $N=\max\{\delta(O_n):\max|x-y|,x,y\in O_n\}$, 则 $\bigcup\limits_{m=1}^{n}O_n\subset B(0,2N)\subsetneqq\mathbb{R}$, 矛盾! □

2.4.2 试证: $([a,b],d)$ 为完备, 可分, 紧距离空间.

证明: $[a,b]$ 是 $\mathbb{R}$ 的闭子集, 故根据习题 2.4.1可得结论. □

2.4.3 试证: 紧集的闭子集也是紧集, 闭集是否一定是紧集?

证明: 实际上引理 2.4.8(1) 便是本题的第一问. 后者则有反例: $\mathbb{R}$ 是闭集, 但不是紧集. □

2.4.4 若 $\{A_n\subset E,n\in\mathbb{N}\}$ 是递降非空紧集序列, 且 $\delta(A_n)=\sup\limits_{x,y\in A_n}d(x,y)\to 0$, 则存在唯一的 $x_0\in\bigcap\limits_{n=1}^{\infty}A_n$, 将 $\{A_n\}$ 改为闭集序列结论还成立吗?

证明: 先证明 x_0 的存在性, 我们知道在 A_n 中任取一点 a_n, $\{a_n\}$ 都满足 $d(a_m,a_n)<\min(\delta(A_m),\delta(A_n))\to 0$, 则 $\{a_n\}$ 是基本列. 而距离空间中的紧集等价于有界, 因此 $\{a_n\}$ 存在收敛子列. 根据定理 2.3.4 可知 $\{a_n\}$ 收敛, 也即 $\bigcap\limits_{n=1}^{\infty}A_n$ 非空.

再证明 x_0 是唯一的. 若存在 $x_1,x_2\in\bigcap\limits_{n=1}^{\infty}A_n$ 且 $x_1\neq x_2$, 记 $d=d(x_1,x_2)$, 则 $\delta\left(\bigcap\limits_{n=1}^{\infty}A_n\right)>d$, 这和题设矛盾.

若将 $\{A_n\}$ 改为闭集序列, 考虑反例: 在 $\mathbb{N}$ 上装备度量 $d(x,y)=\begin{cases}\dfrac{1}{\min(x,y)}, & x\neq y,\\ 0, & x=y,\end{cases}$ 则 d 是距离, 考虑 $\{A_n:n\in\mathbb{N}\}\subset\mathbb{N}$ 且 $A_n:=\mathbb{N}\cap[n,\infty)$, 则 $A_n\downarrow\varnothing$, $\delta(A_n)\leqslant\dfrac{1}{n}\to 0$. 但 $\bigcup\limits_{n=1}^{\infty}A_n$ 中的点不唯一. □

2.4.5 设 $A\subset B\subset E$, 若 $B\subset\overline{A}$, 则称 A 在 B 中稠, 试证: 设 $A,B,C\in E$ 且 A 在 B 中稠, B 在 C 中稠, 则 A 在 C 中稠.

证明: $A\subset B\subset C$, $B\subset\overline{A}$, $C\subset\overline{B}$, 因此只需证明 $C\subset\overline{A}$.

而由定理 2.3.4, $\forall x\in\overline{B}$, $\exists x_n\in B\subset\overline{A}$, 使得 $x_n\to x\in\overline{A}$. 故 $C\subset\overline{B}\subset\overline{A}$, A 在 C 中稠. □

2.4.6 A 为疏朗集的充要条件是任何非空开集 B 都有一非空开集 $C \subset B$, 使 $C \cap A = \varnothing$.

证明: 当 A 为疏朗集, 则 $\overline{A}$ 中不包含任何开集, 从而对于任何非空开集 B, $\exists B(x_0, r_0) \in B$, 且 $\exists x_1 \in B(x_0, r_0), x_1 \notin \overline{A}$, 而 $\overline{A}$ 为闭集, 故必然存在 $\varepsilon_1 > 0$, s.t. $\overline{B}(x_1, \varepsilon_1) = \varnothing$. (否则该点为聚点, 则与 $x_1 \notin \overline{A}$ 矛盾.). 取 $0 < r_1 < \min(\varepsilon_1, r_0 - d(x_0, x_1))$, 便有 $B(x_1, r_1) \subset B(x_0, r_0)$ 且 $B(x_1, r_1) \cap \overline{A} = \varnothing$.

如果 A 不为疏朗集, 也即 $\overline{A}$ 中有内点, 则 $\exists B(x_0, r_0) \subset \overline{A}$. 由题设条件, 存在非空开集 (不妨设为非空开球), 设 $B(x_1, r_1) \subset B(x_0, r_0)$, s.t. $\overline{B}(x_1, r_1) \cap \overline{A} \neq \varnothing$, 则与 $B(x_1, r_1) \subset B(x_0, r_0) \subset \overline{A}$ 矛盾. 故 $\overline{A}$ 为疏朗集. □

2.4.7 疏朗集的余集是稠集, 并举例说明稠集的余集不一定是疏朗集.

证明: 设 A 是 (E, d) 中的疏朗集, 则根据习题 2.4.6, $\forall x \in E$, $\forall r > 0$, $B(x, r) \cap A = \varnothing$, 则推出 $B(x, r) \cap A^c \neq \varnothing$, 因此 A^c 在 E 中稠. 在实数集中, 无理数集是稠集, 但其余集有理数集仍为稠集. □

2.4.8 给定 (E, d), $A \subset E$, 则下列命题成立.

(1) $a \in \overline{A}$ 当且仅当存在 $\{a_n : n \in \mathbb{N}\} \subset A$ 使 $d(a_n, a) \to 0$, 即 $(a_n \to a)$, $n \to \infty$.

(2) $a \in A'$ 当且仅当存在 $\{a_n : n \in \mathbb{N}\} \subset A$, 且 $a_n, n \in \mathbb{N}$ 两两不同, 使 $d(a_n, a) \to 0$.

证明: (1) $\Leftarrow$: 显然成立, 若存在 $\{a_n : n \in \mathbb{N}\} \subset A$ 使 $d(a_n, a) \to 0$, 则 a 为聚点, 故 $a \in \bar{A}$.

$\Rightarrow$: $a \in A$ 时显然成立, 取 $a_n = a$ 即可. $a \in \bar{A} \backslash A$ 时, 若存在 $\varepsilon_0 > 0$ 使得 A 中任何点列 $\{a_n\}$ 都有 $d(a_n, a) \geqslant \varepsilon_0$, 则 $B(a, \varepsilon_0) \backslash \{a\}$ 不在 A 中, 即 $B(a, \varepsilon_0) \backslash \{a\} \subset A^c$. 而 $a \in \bar{A}$, 故 $a \notin (A^c)^o$, 矛盾.

(2) $\Leftarrow$: 显然成立, 若存在 $\{a_n : n \in \mathbb{N}\} \subset A$, 且 $a_n, n \in \mathbb{N}$ 两两不同, 使 $d(a_n, a) \to 0$, 则 a 为聚点, 故 $a \in A'$.

$\Rightarrow$: 当 $a \in A'$, 则 $\forall N(a), A \cap (N(a) \backslash \{a\}) \neq \varnothing$. 取 a 的邻域 $B\left(a, \dfrac{1}{n}\right), n \in \mathbb{N}$, 则每个 $A \cap (N(a) \backslash \{a\})$ 中都可选出一点 a_n, 与 a_{n-1} 不同. 且 $d(a_n, a) < \dfrac{1}{n}$. 故存在 $\{a_n : n \in \mathbb{N}\} \subset A$, 且 $a_n, n \in \mathbb{N}$ 两两不同, 使 $d(a_n, a) \to 0$. □

2.4.9 设 (E, d) 为一紧距离空间, $\{G_\lambda : \lambda \in A\}$ 是 E 的一个开覆盖, 则 $\exists \alpha > 0$ 使任何半径为 a 的开球至少被包含在一个 G_λ 之中 (Lebesgue 性). 并试举一反例说明: 当 E 为全有界时, 上述结论不真.

证明: 若任何 $\alpha > 0$, 半径为 α 的开球都不被包含在一个 G_λ 之中, 则选取一列递降开球列 B_n, 且第 n 个开球 B_n 的半径小于 $\dfrac{1}{n}$, 同时在每个球中选取一点 a_n.

则在紧距离空间中, $\{a_n\}$ 存在聚点, 设为 a_0, 则存在 $\{a_n\}$ 的子列 $\{a_{nk}\}$ 收敛到 a_0. 因为 $\{G_\lambda : \lambda \in A\}$ 是 E 的一个开覆盖, 故 $\exists \lambda$, s.t. $a_0 \in G_\lambda$, 且 $E \backslash G_\lambda \neq \varnothing$ (否则 E 中的每个开球已经被 G_λ 覆盖), 令 d 为 a_0 到 $E \backslash G_\lambda$ 的距离, 且严格大于 0 , 由于 $\{a_{nk}\}$ 收敛到 a_0, 故存在 $\exists n_k > \dfrac{2}{d}$, s.t. $d(a_{nk}, a_0) < \dfrac{d}{2}$, 而

$$\forall x \in A_{nk}, d(x, a_0) \leqslant d(x, a_k) + d(a_k, a_0) < \frac{1}{n_k} + \frac{d}{2} = \frac{d}{2} + \frac{d}{2} = d.$$

则 $A_{nk} \in G_\lambda$ 与假设矛盾, 故 (E, d) 具有 Lebesgue 性.

设 $E = \left\{\dfrac{1}{n} : n \in \mathbb{N}\right\}, d(x, y) = |x - y|$, 为完全有界, 但不是紧集, $\left(\dfrac{1}{n^2}, 1\right)$ 是 E 的一个开覆盖, 但永远存在其无法包含的开球. □

2.4.10 在紧距离空间中, 若 $\{x_n\}$ 两两不同, 且只有一个聚点 a, 则 $x_n \to a$ $(n \to \infty)$.

证明: 设此空间为 E, 由紧性得 $\{x_n\}$ 有界, 若 $x_n \to a$ 不成立, 则 $\forall r>0$, $(E\backslash B(a,r))\cap\{x_n\}$ 有无限个元素, 则必然存在聚点, 这与题设矛盾. □

2.4.11 相对紧集一定全有界, 而在完备距离空间中, 全有界集一定相对紧.

证明: 见定理 2.4.11. □

2.4.12 $\mathbb{R}^n$ 中的子集为紧集当且仅当它是有界闭集.

证明: $\Rightarrow$: 当 $\mathbb{R}^n$ 中的子集为紧集, 由定理 2.4.10 知道其为闭集且列紧, 又由定理 2.4.9 知道其有界.

$\Leftarrow$: 考虑 $\mathbb{R}^n$ 中的有界闭子集, 我们将证明其为紧集. 根据定理 2.4.11, 只需证明有界闭集为列紧集. 任取 $A\subset\mathbb{R}^n$ 为有界闭集, 则其中任一序列 $\{x_n\}$ 为有界序列, 存在聚点. 又因为 A 为闭集, 故此聚点在 A 中, 而 $\overline{A}=A$, 故由定理 2.4.11 知其为紧集. □

2.4.13 (E,d) 中任何两紧集之并仍为紧集.

证明: 设 A,B 为 E 中的紧集, 则任意 A,B 的开覆盖 A_α,B_β 存在有限开覆盖. 故 $A\cup B$ 的任意开覆盖自然是 A,B 的开覆盖, 故存在有限开覆盖 A_α,B_β, 且 $A_\alpha\cup B_\beta$ 为 $A\cup B$ 的有限开覆盖, 故 $A\cup B$ 为紧集. □

2.4.14 设 $(E_n,d_n),n\in\mathbb{N}$ 为一距离空间列.

$$E^\infty := E_1\times E_2\times\cdots := \{x: x=(x_1,x_2,\cdots),x_n\in E_n,n\in\mathbb{N}\},$$
$$d(x,x') := \sum_{n=1}^{\infty}\frac{1}{2^n}\min\left(d_n\left(x_n,x_n'\right),1\right),$$

试证:

(1) (E^∞,d) 为一距离空间;

(2) 设 $n\in\mathbb{N},U_k$ 是 E_k 的开子集 $(k=1,2,\cdots,n)$, 则 (E^∞,d) 中一切形如

$$\begin{aligned} G &:= U_1\times\cdots\times U_n\times E_{n+1}\times\cdots \\ &:= \left\{x\in E^\infty: \begin{array}{l} x=(x_1,\cdots,x_n,\cdots),x_k\in U_k,k=1,2,\cdots,n, \\ x_l\in E_l,l=n+1,n+2,\cdots \end{array}\right\} \end{aligned}$$

都是 E^∞ 的开集;

(3) (E^∞,d) 的任一开集都是形如 (2) 的集的并, 即一切由 (2) 列出的集族是 (E^∞,d) 的一个拓扑基;

(4) 若每一 (E_n,d_n) 可分, 则 (E^∞,d) 可分;

(5) 若每一 (E_n,d_n) 是紧空间, 则 (E^∞,d) 也是紧空间.

证明: (1) 只需验证 d 是距离.

(i) $\forall n\in\mathbb{N}$, $\min(d_n(x_n,y_n),1)\geqslant 0$, 故 $d(x,y)\geqslant 0$;

(ii) $d(x,y)=0 \Longleftrightarrow \min(d_n(x_n,y_n,1))=0,\forall n\in\mathbb{N} \Longleftrightarrow x_n=y_n,\forall n\in\mathbb{N} \Longrightarrow x=y$;

(iii) 显然 $d(x,y)=d(y,x)$;

(iv) $d(x,y)\leqslant\sum_{n=1}^{\infty}2^{-n}\min(d_n(x_n,z_n)+d_n(z_n,y_n),1)\leqslant\sum_{n=1}^{\infty}2^{-n}\min(d_n(x_n,z_n),1)+\sum_{n=1}^{\infty}2^{-n}\min(d_n(z_n,y_n),1)=d(x,z)+d(z,y)$.

因此 d 是距离, (E,d) 是距离空间.

(2) 我们知道 $\forall k=1,2,\cdots,n$, $\forall x_k\in U_k$, 存在它的邻域 $N(x_k)\subset U_k$. 故 $N(x_1)\times N(x_2)\times\cdots\times N(x_n)\times E_{n+1}\times\cdots$ 是 x 在 G 中的邻域, 故 G 是 E^∞ 中的开集.

(3) 设 U 是 E^∞ 中的开集, 则 $\forall x_0\in U$, $\exists\delta\in(0,1)$, s.t. $d(x,x_0)<\delta$ 时有 $x\in U$. 因此 $d_n(x_{0n},x_n)<\delta$ 时 $x\in U$, 故 $x\in U_1\times U_2\times\cdots\times U_n\times\cdots$, 其中 $U_n=B(x_{0n},\delta)$. 遍历 x_0 即可得到结论.

(4) 若 E_n 可分, 则存在稠子集 $A_n\subset E_n$. 下面证明 $A:=A_1\times A_2\times\cdots\times A_n\times\cdots$ 为 E^∞ 的稠子集. 我们知道存在 x_n, s.t. $d_n(x_{0n},x_n)<\varepsilon$, 故 $x:=(x_1,x_2,\cdots,x_n,\cdots)\in A$, 且 $d(x,x_0)<\varepsilon$. 因此 E^∞ 可分.

(5) 根据定理 2.4.11, E_n 完备且全有界. 因此 E^∞ 也是完备的, 我们只需证明其全有界. 我们知道 $\forall n\in\mathbb{N}$, $\forall\varepsilon>0$, 都存在 $k_n\in\mathbb{N}$, s.t. $\{x_{n1},x_{n2},\cdots,x_{nk_n}\}\subset E_n\subset\bigcup\limits_{n=1}^{k_n}B(x_{k_n},\varepsilon)$. 同时取充分大的整数 N, s.t. $\sum\limits_{n=N+1}^{\infty}2^{-n}<\varepsilon$, 令 $x':=(x_1,x_2,\cdots,x_N)$, $x_i\in\{x_{n1},x_{n2},\cdots,x_{nk_n}\}$, $i=1,2,\cdots,N$. 令 $A:=\bigcup\limits_{n=1}^{k_n}B(x'_n,2\varepsilon)$, 则 $\forall x\in E$, $\exists x'\in A$, s.t. $d(x,x')<\sum\limits_{n=1}^{N}\frac{\varepsilon}{2^n}+\varepsilon<2\varepsilon$. 故 A 是 E^∞ 的有限 ε 网, 故 E^∞ 全有界, 为紧空间. □

§ 2.5 距离空间上的映射与函数

2.5.1 完成引理 2.5.2、引理 2.5.3、引理 2.5.4、引理 2.5.8 和引理 2.5.9 的证明.

证明: 引理 2.5.2:

(2) ⇒ (3): 当 F 为闭集, 显然 F^c 是开集. 而 $f^{-1}(F^c)=(f^{-1}(F))^c$ 为开集, 故 $f^{-1}(F)$ 为闭集.

(3) ⇒ (5): 我们知道任意 S 中的闭集 F, 都有 $f^{-1}(F)$ 为闭集. 又 $A\subset f^{-1}(f(A))$ 且 $\overline{f(A)}$ 为闭集,$f^{-1}(\overline{f(A)})$ 为闭集. 因此

$$\overline{A}\subset\overline{f^{-1}(\overline{f(A)})}=f^{-1}(f(\overline{A})),\quad f^{-1}(\overline{A})\subset f(f^{-1}(\overline{f(A)}))\subset\overline{f(A)}.$$

(5) ⇒ (4): 我们知道 $\forall A\in S$ 都有 $f(f^{-1}(A))\subset A$, 因此

$$f(\overline{f^{-1}(A)})\subset\overline{f(f^{-1}(A))}\subset A.$$

故 $\overline{f^{-1}(A)}\subset f^{-1}(\overline{A})$.

引理 2.5.3: 我们知道 f_2,f_3 都是连续的, 因此根据引理 2.5.2(2), 任意 E_3 中的开集 A 都有 $f_2^{-1}(A)$ 是 E_2 的开集, $f_1^{-1}(f_2^{-1}(A))$ 是 E_1 的开集. 因此 $f_2\circ f_1$ 是连续的.

引理 2.5.4: 我们有 $\forall x_0\in E$, $\forall\varepsilon>0$, $\exists\delta>0$, 当 $d(x,x_0)<\delta$ 时有 $|f(x)-f(x_0)|<\varepsilon$, $|g(x)-g(x_0)|<\varepsilon$.

我们有

$$
\begin{aligned}
||f(x)|-|f(x_0)|| &\leqslant |f(x)-f(x_0)|,\\
|\alpha f(x)-\alpha f(x_0)| &= |\alpha||f(x)-f(x_0)|,\\
|f(x)+g(x)-f(x_0)g(x_0)| &\leqslant |f(x)-f(x_0)|+|g(x)-g(x_0)|,\\
|f(x)g(x)-f(x_0)g(x_0)| &\leqslant |f(x)||g(x)-g(x_0)|+|g(x_0)||f(x)-f(x_0)|\\
&\leqslant (|f(x_0)+\varepsilon|)|g(x)-g(x_0)|+|g(x_0)||f(x)-f(x_0)|,\\
\left|\frac{1}{f(x)}-\frac{1}{f(x_0)}\right| &= \left|\frac{1}{f(x)f(x_0)}\right||f(x)-f(x_0)|,\\
f\vee g &= \frac{|f-g|+f+g}{2},\\
f\wedge g &= \frac{|f-g|-(f+g)}{2}.
\end{aligned}
$$

因此各函数均为实值连续函数.

引理 2.5.8: 我们知道 $\forall x\in E$, $\forall\varepsilon>0$, $\exists\delta>0$, 当 $d(x,x_0)<\delta$ 时 $\rho(f(x)-f(x_0))<\varepsilon$. 而 $\{B(x_0,\delta(x_0,\varepsilon))\}$ 组成了 E 的一组开覆盖, 因此由紧集的性质, 必然存在有限子覆盖

$$B(x_1,\delta(x_1,\varepsilon)),B(x_2,\delta(x_2,\varepsilon)),\cdots,B(x_n,\delta(x_n,\varepsilon)).$$

取 $\delta(\varepsilon):=\min\limits_{1\leqslant k\leqslant n}\delta(x_k,\varepsilon)$, 则 $\forall x_1,x_2\in E$, $\forall\varepsilon>0$, $\exists\delta(\varepsilon)>0$, s.t. 当 $d(x_1,x_2)<\delta(\varepsilon)$ 时有 $\rho(f(x_1),f(x_2))<\varepsilon$. 这便是一致连续.

引理 2.5.9: 设 $\{G_\alpha\}$ 是 $f(E)$ 的一组开覆盖, 我们知道 f 是连续的, 因此 $\{f^{-1}(G_\alpha)\}$ 是 E 的一组开覆盖. 故存在有限子覆盖 $\{f^{-1}(G_\beta)\}$. 此时 $\{G_\beta\}$ 是 $f(E)$ 的有限开覆盖, 因此 $f(E)$ 也是紧集. 又 $f(E)\subset\mathbb{R}$, 因此是有界闭集. 因此存在 $a=\max f(E)$, $b=\min f(E)$. 又因为 a,b 是聚点, 所以存在 $a_n\to a$, $b_n\to b$. 因此存在 $f^{-1}(a_n)$ 的子列 $\{f^{-1}(a_{nk})\}$ 收敛, 设其极限为 a_0, 则 $f(f^{-1}(a_{nk}))=a_{nk}\to f(a_0)=a$. 因此 $f(x)$ 存在最大值, 同理存在最小值. □

第三章 测度空间与概率空间

§ 3.1 集类

3.1.1 试验证 3.1.3 例 2 的 $d=1,3$ 情形.

例 2: 设 $\Omega=\mathbb{R}^d$, $d\geqslant 1$, 则

$$\mathscr{S}:\left\{(a,b]\subset\mathbb{R}^d: a=(a_1,\cdots,a_d), b=(b_1,\cdots,b_d), -\infty\leqslant a_k\leqslant b_k\leqslant\infty, k=1,\cdots,d\right\}$$

是 $\mathbb{R}^d$ 的一个半集代数.

证明: 我们需要证明 $\mathscr{S}$ 对交封闭, 包含全集和空集, 并且 $\forall A, A_1\in\mathscr{S}$, $A_1\subset A$, $\exists\{A_2,A_3,\cdots,A_n\}\subset\mathscr{S}$, $A_1,A_2,\cdots,A_n$ 两两不交, 且 $A=\bigcup\limits_{k=1}^{n}A_k$.

我们只需取 $a=b$, 则 $(a,b]=\varnothing$, 而取 $a_k=-\infty, b_k=\infty$, 则 $(a,b]=\mathbb{R}^d$. 并且其对交封闭是显然的. 下面证明: $\forall A, A_1\in\mathscr{S}$, $A_1\subset A$, $\exists\{A_2,A_3,\cdots,A_n\}\subset\mathscr{S}$, $A_1,A_2,\cdots,A_n$ 两两不交, 且 $A=\bigcup\limits_{k=1}^{n}A_k$.

若 $d=1$, 有

$$(a_1,b_1]=(a_1,a_2]\cup(a_2,b_2]\cup(b_2,b_1].$$

若 $d=3$, 有

$$(a,b]=(a',b']\cup\left(\bigcup_{k=1}^{26}A_k\right).$$

这里 A_k 有 26 个的原因是 a',b' 将 a,b 分为 27 个小立方体. □

3.1.2 设 $\mathscr{S}$ 是 Ω 的半集代数, 试证:

(1) 若 $\{A,A_1,\cdots,A_n\}\subset\mathscr{S}$, 且 $A_k\subset A, k=1,2,\cdots,n$ 两两不交, 则存在 $\{A_{n+1},A_{n+2},\cdots,A_s\}\subset\mathscr{S}$, 使 $A_1,A_2,\cdots,A_s$ 两两不交, 且 $A=\bigcup\limits_{k=1}^{s}A_k$;

(2) 若 $\{A_1,A_2,\cdots,A_n\}\subset\mathscr{S}$, 则在 $\mathscr{S}$ 中存在两两不交的有限个集 $B_1,B_2,\cdots,B_t$, 使每个 A_k 可以表成若干个 B_j 之并.

证明: (1) 考虑归纳法. 当 $n=1$, 结论显然.

假设 $n-1$ 时命题成立, 则存在 $B_1,B_2,\cdots,B_m\in\mathscr{S}$, 使得 $A,A_1,\cdots,A_{n-1},B_1,B_2,\cdots,B_m$ 两两不交, 且它们的并 $\left(\bigcup\limits_{k=1}^{n-1}A_k\right)\cup\left(\bigcup\limits_{i=1}^{m}B_i\right)=A$.

考虑引入新集合 $A_n \in \mathscr{S}$, 它至少要与 $\{A_k, 1 \leqslant k \leqslant n-1, k \in \mathbb{N}\}$ 中的任意一集不交, 注意到

$$A = \left(\bigcup_{k=1}^{n-1} A_k\right) \cup \left(\bigcup_{i=1}^{m} B_i\right) = \left(\bigcup_{k=1}^{n-1} A_k\right) \cup \left(A_n \cup \left(\bigcup_{i=1}^{m} (B_i \backslash (B_i \cap A_n))\right)\right),$$

我们知道半集代数对交封闭, 故 $B_i \cap A_n \subset B_i \in \mathscr{S}$, 因此存在 $C_{i1}, C_{i2}, \cdots, C_{it_i} \in \mathscr{S}$, 它们两两不交, 且 $\bigcup\limits_{j=1}^{t_i} C_{ij} = B_i \backslash (B_i \cap A_n)$, 因此

$$A = \left(\bigcup_{k=1}^{n} A_k\right) \cup \left(\bigcup_{i=1}^{m} \bigcup_{j=1}^{t_i} C_{ij}\right) = \bigcup_{k=1}^{s} A_k.$$

(2) 同样考虑归纳法, 当 $n=2$ 时, 存在两两不交的集列 $\{B_{1i}, 1 \leqslant i \leqslant n_1\}$, $\{B_{2i}, 1 \leqslant i \leqslant n_2\}$, 使得

$$A_1 = (A_1 \cap A_2) \cup \left(\bigcup_{i=1}^{n_1} B_{1i}\right), \quad A_2 = (A_1 \cap A_2) \cup \left(\bigcup_{i=1}^{n_2} B_{2i}\right),$$

因此只需令 $B_1 = A_1 \cap A_2$, $\{B_i, 2 \leqslant i \leqslant 1+n_1\} = \{B_{1i}, 1 \leqslant i \leqslant n_1\}$, $\{B_i, 2+n_1 \leqslant i \leqslant 1+n_1+n_2\} = \{B_{2i}, 1 \leqslant i \leqslant n_2\}$ 即可.

设 $\{A_1, A_2, \cdots, A_{n-1}\} \subset \mathscr{S}$ 时命题成立, 也即存在两两不交的 $\{B_1, B_2 \cdots, B_{t_1}\} \subset \mathscr{S}$ 满足条件.

考虑 $A_n \in \mathscr{S}$, 则 $A_n \cap B_k \subset B_k \in \mathscr{S}$, 且它们两两不交. 因此存在两两不交的 $\{C_i^k, 1 \leqslant i \leqslant m_k\} \subset \mathscr{S}$, 它们两两不交且都不和 $A_n \cap B_k$ 相交, 且

$$B_k = (B_k \cap A_n) \cup \left(\bigcup_{l=1}^{m_k} C_l^k\right),$$

考虑到 (1), 我们有: 存在两两不交且与 $A_n \cap B_k$ 两两不交的 $\{C_1, \cdots, C_m\} \subset \mathscr{S}$, 使得

$$A_n = \left(\bigcup_{k=1}^{t_1} (A_n \cap B_k)\right) \cup \left(\bigcup_{i=1}^{m} C_i\right),$$

只需取 $\{B_1, \cdots, B_t\} = B_k \cap A_n, C_j^k$ 即可, 其中 $j = 1, \cdots, m_k$, $k = 1, \cdots, t_1$. □

3.1.3 称 Ω 的子集类 $\mathscr{S}$ 为半环, 如果它满足定义 3.1.1 中的 (ii) (iii). 试证:

(1) $\varnothing \in \mathscr{S}$;

(2) $\mathscr{S} := \{(a,b] : a, b \in \mathbb{R}^d, a \leqslant b\}$ 是 $\mathbb{R}^d$ 的半环.

证明: (1) 是易见的, 只需取两个不交的集 $A, B \in \mathscr{S}$, 由于半环 $\mathscr{S}$ 对交封闭, 则 $\varnothing \in \mathscr{S}$;

(2) $\mathscr{S}$ 对交封闭是显然的, 考虑 $a \leqslant a_1 \leqslant b_1 \leqslant b$, 则

$$(a,b] = (a_1, b_1] \cup \left(\cup_{k=1}^{3^n-1} A_k\right),$$

其中 A_k 是 a_1, b_1 在区间 $(a,b]$ 中划分出的小区间, 且不等于 $(a_1, b_1]$. □

3.1.4 称 Ω 的子集类 $\mathscr{R}$ 为环 (或布尔环), 如果它满足 $A, B \in \mathscr{R}$, 则有 $A \cup B, A \backslash B \in \mathscr{R}$. 试证: Ω 的子集类 $\mathscr{R}$ 是环的充要条件是对并及真差封闭.

证明: 我们知道

$$A \backslash B = A \cup B - B \in \mathscr{R},$$

故子集类 $\mathscr{R}$ 对真差封闭. □

3.1.5 如果将对称差看作集合间的加法运算“+”, 将交看作集合间的乘法运算“·”, 则

(1) Ω 的任一集代数 $\mathscr{A}$ 对"+" "·" 作成一个具单位元的可换环, 而且每个元都是幂等的 (即 $A\cdot A=A$);

(2) Ω 的任一环 $\mathscr{R}$ 对"+" 及"·" 作成一幂等可换环.

证明: (1) 回忆幂等交换环的定义, 我们只需证明:

(i) $(\mathscr{A},\Delta)$ 是 Abel 群且有单位元 $\varnothing$;

(ii) $(\mathscr{A},\cap)$ 是交换半群且有单位元 Ω;

(iii) $\cap$ 对于 Δ 满足左右分配律;

(iv) $A\cap A=A$.

这意味着, 我们需要证明:

(i) (a) $\forall A,B\in\mathscr{A},\ A\Delta B=B\Delta A\in\mathscr{A}$;

(b) $\forall A,B,C\in\mathscr{A},\ (A\Delta B)\Delta C=A\Delta(B\Delta C)$;

(c) $\forall A\in\mathscr{A},\ A\Delta\varnothing=\varnothing\Delta A=A$;

(d) $\forall A\in\mathscr{A},\ \exists B\in\mathscr{A},\ A\Delta B=B\Delta A=\varnothing$;

(ii) (a) $\forall A,B\in\mathscr{A},\ A\cap B=B\cap A\in\mathscr{A}$;

(b) $\forall A,B,C\in\mathscr{A},\ A\cap(B\cap C)=(A\cap B)\cap C$;

(c) $\forall A\in\mathscr{A},\ A\cap\Omega=\Omega\cap A$;

(iii) $\forall A,B,C\in\mathscr{A},\ \begin{cases}A\cap(B\Delta C)=(A\cap B)\Delta(A\cap C),\\(B\Delta C)\cap A=(B\cap A)\Delta(C\cap A).\end{cases}$

我们知道 $A\Delta B=(A\cap B^c)\cup(B\cap A^c)$, 因此 (i)(a) 是显然的. (i)(b) 已经在习题 1.1.4(6) 中证明, 而 (i)(c),(i)(d),(ii)(a),(ii)(b),(ii)(c) 显然. (iii) 已经在习题 1.1.6中证明.

同时 $A\cap A=A$, 因此 $\mathscr{A}$ 是幂等可换环.

(2) 由习题 3.1.4和 (1) 可知结论成立. □

3.1.6 设 $\mathscr{S}$ 为 Ω 的一个半环, 则

$$\{A\subset\Omega: A=\bigcup_{k=1}^{n}A_k, n\in\mathbb{N},\{A_1,A_2,\cdots,A_n\}\subset\mathscr{S}\text{两两不交}\}$$

为包含 φ 的最小环, 由此说明

$$\left\{A\subset\mathbb{R}^d: A=\bigcup_{k=1}^{n}(a_k,b_k], n\in\mathbb{N}, a_k,b_k\in\mathbb{R}^d, a_k\leqslant b_k, k=1,2,\cdots,n\right\}$$

是一个环 (由此可以看出环在表述上有方便之处).

证明: 置

$$\mathscr{A}=\{A\subset\Omega: A=\bigcup_{k=1}^{n}A_k, n\in\mathbb{N}, \{A_1,A_2,\cdots,A_n\}\subset\mathscr{S}\text{两两不交}\},$$

我们将证明 $\mathscr{A}$ 是环, 也即 $\forall A,B\in\mathscr{A}$, $A\cup B\in\mathscr{A}$, 且 $B\subset A$ 时有 $A\backslash B\in\mathscr{A}$.

考虑两两不交的 $A_1,\cdots,A_n\in\mathscr{S}$, $B_1,\cdots,B_m\in S$, 且 $A=\bigcup\limits_{k=1}^{n}A_k$, $B=\bigcup\limits_{k=1}^{m}B_k$.

考虑 $1\leqslant k\leqslant n$, $1\leqslant l\leqslant m$, 若 $A_k\cap B_l=\varnothing$, 则 $A_k\cup B_l\in A$; 若 $\exists k_1,k_2$, $A_k\cap B_l\neq\varnothing$, 于是 $A_k\cap B_l\subset A_k\in\mathscr{S}$, 故 $\exists A_{k_1},A_{k_2},\cdots,A_{k_n}\in\mathscr{S}$, s.t. $A_k=(A_k\cap B_l)\cup\left(\bigcup\limits_{n=k_1}^{k_n}A_n\right)$. 同理也可以对 B_l 作类似操作. 因此

$$A\cup B=\left(\bigcup_{k=1}^{n}\bigcup_{n=k_1}^{k_n}A_n\right)\cup\left(\bigcup_{k=1}^{m}\bigcup_{m=l_1}^{l_m}B_m\right)\in\mathscr{A}.$$

□

3.1.7 若 $A_k\subset\Omega, k=1,2,\cdots,n$ 两两不交, 且 $\bigcup\limits_{k=1}^{n}A_k=\Omega$. 设 $\mathscr{E}:=\{A_1,A_2,\cdots,A_n\}$, 试将 $\mathscr{A}(\mathscr{E})$ 的全部元用 $\mathscr{E}$ 的元表出, 请读者考查此 $\mathscr{A}(\mathscr{E})$ 与 $\Omega_n=\{1,2,\cdots,n\}$ 的一切子集作成的集代数之间的关系.

证明: 回忆集代数: 包含全集, 且对有限交, 有限并, 补封闭. 因此 $\mathscr{A}(\mathscr{E})$ 是 $\mathscr{E}$ 的元进行有限交, 有限并, 补得到的, 因此

$$\mathscr{A}(\mathscr{E})=\left\{\varnothing,\bigcup_{\alpha\in\Omega_n}A_\alpha\right\},$$

这其中自然包含了任取 k 个 A_i 的并和其的补, 也包含了全集 Ω. 而有限交则是空集. 定义 Ω_n 的一切子集作成的集代数为 $\mathscr{P}(\Omega_n)$, 我们可以找到一一映射:

$$\mathscr{P}(\Omega_n)\mapsto\mathscr{A}(\mathscr{E}),$$
$$\{k_1,k_2,\cdots,k_n\}\to\bigcup_{i=1}A_{k_i}.$$

因此 $\overline{\overline{\mathscr{P}(\Omega_n)}}=\overline{\overline{\mathscr{A}(\mathscr{E})}}$. □

3.1.8 设 $\mathscr{E}$ 是 Ω 的任一子集类, 且

$$\mathscr{E}_1:=\{A:A\in\mathscr{E}\text{ 或 }A^c\in\mathscr{E}\}\cup\{\varnothing,\Omega\},$$
$$\mathscr{E}_2:=\left\{\bigcap_{k=1}^{n}A_k:A_k\in\mathscr{E}_1,k=1,2,\cdots,n,n\in\mathbb{N}\right\},$$

则 $\mathscr{A}:=\left\{\bigcup\limits_{i=1}^{m}B_i:B_i\in\mathscr{E}_2,i=1,2,\cdots,m\text{两两不交},m\in\mathbb{N}\right\}$ 是包含 $\mathscr{E}$ 的最小集代数 (即由 $\mathscr{E}$ 生成的集代数 $\mathscr{A}(\mathscr{E})$).

证明: 首先证明 $\mathscr{A}$ 是集代数. $\Omega\in\mathscr{A}$ 和 $\mathscr{A}$ 对有限并封闭是显然的, 所以我们将证明它对有限交、取余集运算封闭.

考虑

$$\bigcup_{i=1}^{m_1}B_{1i}=C_1\in\mathscr{A},\quad\bigcup_{j=1}^{m_2}B_{2j}=C_1\in\mathscr{A}.$$

其中

$$B_{1i} = \bigcap_{k=1}^{n_{1i}} A_{1k} \in \mathscr{E}_2,\ B_{2j} = \bigcap_{k=1}^{n_{2j}} A_{2k} \in \mathscr{E}_2,$$

$$A_{1k} \in \mathscr{E}_1,\ k = 1, 2, \cdots, n_{1i},\ A_{2k} \in \mathscr{E}_1,\ k = 1, 2, \cdots, n_{2j}$$

则

$$C_1 \cap C_2 = \bigcup_{i=1}^{m_1} \bigcup_{j=1}^{m_2} (B_{1i} \cap B_{2j}),\ B_{1i} \cap B_{2j} = \left(\bigcap_{k=1}^{n_{1i}} A_k\right) \cap \left(\bigcap_{k=1}^{n_{2j}} A_k\right) \in \mathscr{E}_2.$$

我们知道 $B_{1i} \cap B_{2j}$ 在 i, j 取到不同的值时, 它们是两两不交的, 因此 $C_1 \cap C_2 \in \mathscr{A}$, 即 $\mathscr{A}$ 对交封闭.

考虑 $C = \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} A_j \in \mathscr{A}$(其中 $A_j \in \mathscr{E}_1,\ j = 1, 2, \cdots, n_i$), 则 $C^c = \bigcap_{i=1}^{m} \bigcup_{j=1}^{n_i} A_j^c,\ A_j^c \in \mathscr{E}_1,\ j = 1, 2, \cdots, n_i$. 首先可以证明 $\bigcup_{j=1}^{n_i} A_j^c \in \mathscr{A}$. 事实上, 考虑 $A, B \in \mathscr{E}_1$, 有 $A^c, B^c \in \mathscr{E}_1$, 从而

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A \cap B^c) \in \mathscr{A}.$$

而我们已经证明 $\mathscr{A}$ 对于有限交封闭, 因此 $C^c \in \mathscr{A}$, 即 $\mathscr{A}$ 对取余集运算封闭.

故 $\mathscr{A}$ 是集代数. 由 $\mathscr{E}_1, \mathscr{E}_2, \mathscr{A}$ 的定义易知 $\mathscr{E} \subset \mathscr{A}$, 因此 $\mathscr{A}(\mathscr{E}) \subset \mathscr{A}$. 又由 $\mathscr{E}_1, \mathscr{E}_2, \mathscr{A}$ 的定义易知 $\mathscr{A} \subset \mathscr{A}(\mathscr{E})$, 因此 $\mathscr{A} = \mathscr{A}(\mathscr{E})$. □

3.1.9 设 Ω 为不可数集, 试证:

(i) Ω 的一切有限集, 可数集以及它们的余集作成一个 σ 代数;

(ii) Ω 的一切有限集, 可数集作成 Ω 的一个 σ 环 (即 Ω 的对可数并及差封闭的子集类).

证明: (i) 回忆 σ 代数的定义: 包含全集, 对余封闭, 对可数并封闭.

设此集类为 $\mathscr{A}$, 我们知道有限集的可数并可数, 可数集的可数并也可数, 考虑 $A_\alpha (\alpha \in I_1)$ 是可数集或有限集, $A_\gamma (\gamma \in I_2)$ 是可数集或有限集的余集, 且 I_1, I_2 是正整数的至多可数子集, 则

$$\left(\bigcup_{\gamma \in I_2} A_\gamma\right) \cup \left(\bigcup_{\alpha \in I_1} A_\alpha^c\right) = \left(\left(\bigcap_{\alpha \in I_1} A_\alpha\right) \cap \left(\bigcap_{\gamma \in I_2} A_\gamma^c\right)\right)^c,$$

我们知道 $\left(\bigcap_{\alpha \in I_1} A_\alpha\right) \cap \left(\bigcap_{\gamma \in I_2} A_\gamma^c\right)$ 至多可数, 所以它们的可数并是至多可数集的余集. 因此 $\mathscr{A}$ 对可数并封闭.

$\mathscr{A}$ 对余封闭是显然的, 而 $\Omega = \varnothing^c$, 则 $\mathscr{A}$ 包含全集. 故 $\mathscr{A}$ 是 σ 代数.

(ii) 设 Ω 的一切有限集, 可数集构成集类为 $\mathcal{F}$, 我们知道至多可数集的可数并至多可数, 所以 $\mathcal{F}$ 对可数并封闭. 而至多可数集的差至多可数, 所以 $\mathcal{F}$ 对差也封闭. □

3.1.10 设 $\mathscr{E}$ 是 Ω 的任意子集类, $A \in \sigma(\mathscr{E})$, 则有 $\mathscr{E}$ 的一个可列子类 $\mathscr{D}$ 使 $A \in \sigma(\mathscr{D})$.

证明: 考虑集类

$$\Lambda := \{A \in \sigma(\mathscr{E}) : \exists \mathscr{D} \subset \mathscr{E}, \mathscr{D}\text{可列}, A \in \sigma(\mathscr{D})\},$$

我们知道 $\mathscr{E} \subset \Lambda$, 只需证明 Λ 是 σ 代数, 这样便有 $\sigma(\mathscr{E}) \subset \Lambda$, 而 $\Lambda \subset \sigma(\mathscr{E})$, 则 $\Lambda = \sigma(\mathscr{E})$. 下面证明 Λ 是 σ 代数:

(1) 显然 $\Omega \in \Lambda$;

(2) 若 $A_1 \in \Lambda$, 则存在 $\mathscr{D}_1 \subset \mathscr{E}$, 使得 $A_1 \in \sigma(\mathscr{D}_1)$. 因此 $A_1^c \in \mathscr{D}_1$, 故 $A_1^c \in \Lambda$;

(3) 考虑 $A_i \in \sigma(\mathscr{D}_i)$, 则 $A_i \in \sigma(\mathscr{D}_i) \subset \sigma\left(\bigcup_{i=1}^{\infty} \mathscr{D}_i\right)$. 我们知道 $\bigcup_{i=1}^{\infty} \mathscr{D}_i \subset \mathscr{E}$ 且可列, 且 $\sigma\left(\bigcup_{i=1}^{\infty} \mathscr{D}_i\right)$ 对可列并封闭, 因此 $\bigcup_{n=1}^{\infty} A_n \in \sigma\left(\bigcup_{i=1}^{\infty} \mathscr{D}_i\right)$, 故 $\bigcup_{n=1}^{\infty} A_n \in \Lambda$.

因此 Λ 是 σ 代数. 证毕. □

3.1.11 设 $\mathscr{E} := \{A_k, k = 1, 2, \cdots\}$, 其中 $A_k \subset \Omega, k = 1, 2, \cdots$ 两两不交, 试求 $\sigma(\mathscr{E})$.

证明: 考虑集类

$$\mathcal{G} := \left\{\bigcup_{i \in I_1} A_i, I_1 \in \mathscr{P}(\mathbb{N})\right\} \cup \left\{\left(\bigcup_{j \in I_2} A_j\right)^c, I_2 \in \mathscr{P}(\mathbb{N})\right\},$$

我们知道, 将 $\mathscr{E}$ 中集类 $\{A_i, i \in I_1\}$ 中的元素以及集类 $\{A_j, j \in I_2\}$ 中元素的余集进行至多可数并, 可以得到 $\left(\bigcup_{i \in I_1} A_i\right)$, $\left(\bigcup_{j \in I_2} A_j\right)^c$, 这里 $I_1, I_2 \subset \mathscr{P}(\mathbb{N})$, 因此 $\mathcal{G} \subset \sigma(\mathscr{E})$.

下面证明 $\mathcal{G}$ 是 σ 代数:

(1) 考虑 $I_1 = I_2 = \varnothing$, 显然 $\Omega \subset \mathcal{G}$;

(2) $\mathcal{G}$ 对余封闭是显然的;

(3) 考虑 $\mathcal{G}$ 中元素的可数并

$$A = \left(\bigcup_{k=1}^{\infty} \bigcup_{i \in I_{1k}} A_i\right) \cup \left(\bigcup_{k=1}^{\infty} \left(\bigcup_{j \in I_{2k}} A_j\right)^c\right)$$

我们知道 $\{A_k, k = 1, 2, \cdots\}$ 是两两不交的, 令 $I^1 = \bigcup_{k=1}^{\infty} I_{1k}$, $I^2 = \bigcap_{k=1}^{\infty} I_{2k}$, 则

$$A = \left(\bigcup_{i \in I^1} A_i\right) \cup \left(\bigcup_{j \in I^2} A_j\right)^c$$

故

$$\begin{aligned}
\mathbb{1}_A &= \sum_{i \in I^1} \mathbb{1}_{A_i} + 1 - \sum_{j \in I^2} \mathbb{1}_{A_j} - \sum_{i \in I^1} \mathbb{1}_{A_i} \left(1 - \sum_{j \in I^2} \mathbb{1}_{A_j}\right) \\
&= 1 - \sum_{j \in I^2} \mathbb{1}_{A_j} + \sum_{l \in I^1 \cap I^2} \mathbb{1}_{A_l} \\
&= 1 - \sum_{l \in I^2 \backslash I^1} \mathbb{1}_{A_l}
\end{aligned}$$

因此

$$A = \left(\bigcup_{l \in I^2 \backslash I^1} A_l\right)^c,$$

而 $I^2 \backslash I^1 \subset \mathbb{N}$, 且至多可数. 因此 $A \in \mathcal{G}$. 故 $\mathcal{G}$ 是 σ 代数, 因此 $\sigma(\mathscr{E}) \subset \mathcal{G}$. 故 $\sigma(\mathscr{E}) = \mathcal{G}$. □

3.1.12 设 A 是 Ω 的一个子集, $\mathscr{E} := \{B : A \subset B \subset \Omega\}$, 试指出 $\sigma(\mathscr{E})$ 由哪些集组成.

证明: 注意到 $\mathscr{E}$ 是 π 系, 因此 $\Lambda(\mathscr{E}) = \sigma(\mathscr{E})$. 我们知道 Λ 系对真差以及不降序列的并封闭, 考虑 $\{B_n : n \in \mathbb{N}\} \subset \Lambda(\mathscr{E})$, $B_n \uparrow$, 则 $B_2 - B_1 \in \Lambda(\mathscr{E})$, $\bigcup\limits_{k=1}^{\infty} B_k \in \Lambda(\mathscr{E})$.

定义 $\mathcal{C} := \{C \in \Omega : C \cap A = \varnothing\}$, 则对任意的 $B_1, B_2 \in \mathscr{E}$, 且 $B_1 \subset B_2$ 都有 $(B_2 \backslash B_1) \cap A = \varnothing$. 而对任意的 $C_1, C_2 \in \mathcal{C}$, $C_1 \subset C_2$, 都有 $C_2 \backslash C_1 \in \mathcal{C}$. 同时 $\mathscr{E}$ 中的集合和 $\mathcal{C}$ 中的集合不可能互相包含. 因此 $\mathscr{E} \cup \mathcal{C}$ 对真差封闭.

下面证明 $\Lambda(\mathscr{E}) = \mathscr{E} \cup \mathcal{C} = \sigma(\mathscr{E})$:

考虑 $\mathscr{D} := \{B_n : n \in \mathbb{N}\} \subset \mathscr{E} \cup \mathcal{C}$, $B_n \uparrow$, 有 $\mathscr{D} \subset \mathscr{E}$ 或 $\mathscr{D} \subset \mathcal{C}$. 容易知道 $\mathscr{D}, \mathcal{C}$ 都对不降序列的并封闭, 所以 $\mathscr{E} \cup \mathcal{C}$ 是 λ 系, $\Lambda(\mathscr{E}) = \sigma(\mathscr{E}) \subset \mathscr{D}$.

同时考虑 $\forall C \in \mathcal{C}$, 都有 $C^c \in \mathscr{D}$, 故 $\mathscr{E} \cup \mathcal{C} \subset \sigma(\mathscr{E})$.

故 $\mathscr{E} \cup \mathcal{C} = \sigma(\mathscr{E})$. □

3.1.13 设 $\mathscr{E}$ 是 Ω 的一个子集类, $\varnothing \neq A \subset \Omega$, 令 $\mathscr{E} \cap A := \{B \cap A : B \in \mathscr{E}\}$.

试证: $\mathscr{A}(\mathscr{E}) \cap A$ 是 A 的子集类 $\mathscr{E} \cap A$ 生成的 A 的集代数, $\sigma(\mathscr{E}) \cap A$ 是 A 的子集类 $\mathscr{E} \cap A$ 生成的 A 的 σ 代数.

证明: (1) 我们先证明 $\mathscr{A}(\mathscr{E}) \cap A = \mathscr{A}_A(\mathscr{E} \cap A)$. 首先证明 $\mathscr{A}(\mathscr{E}) \cap A$ 是 A 的集代数. 考虑 $B_1, B_2 \in \mathscr{A}(\mathscr{E})$, 则有 $B_1 \cap A, B_2 \cap A \in \mathscr{A}(\mathscr{E}) \cap A$, 我们有

$$
\begin{aligned}
(B_1 \cap A) \cap (B_2 \cap A) &= (B_1 \cap B_2) \cap A \in \mathscr{A}(\mathscr{E}) \cap A; \\
(B_1 \cap A) \cup (B_2 \cap A) &= (B_1 \cup B_2) \cap A \in \mathscr{A}(\mathscr{E}) \cap A; \\
A - (B_1 \cap A) = A \backslash B_1 &= B_1^c \cap A \in \mathscr{A}(\mathscr{E}) \cap A.
\end{aligned}
$$

同时由于 $\Omega, \varnothing \in \mathscr{A}(\mathscr{E})$, 因此 $A, \varnothing \in \mathscr{A}(\mathscr{E}) \cap A$. 因此 $\mathscr{A}(\mathscr{E}) \cap A$ 是 A 的集代数. 我们知道 $\mathscr{E} \cap A \subset \mathscr{A}(\mathscr{E}) \cap A$, 因此 $\mathscr{A}_A(\mathscr{E} \cap A) \subset \mathscr{A}(\mathscr{E}) \cap A$.

考虑集类

$$
\mathcal{C} = \{C \in \mathscr{A}, C \cap A \in \mathscr{A}_A(\mathscr{E} \cap A)\},
$$

则 $\mathcal{C}$ 是 Ω 的集代数, 我们知道 $\mathscr{E} \subset \mathcal{C}$, 因此 $\mathscr{A}(\mathscr{E}) \subset \mathcal{C}$, 因此 $\mathscr{A}(\mathscr{E}) \cap A \subset \mathscr{A}_A(\mathscr{E} \cap A)$. 故 $\mathscr{A}(\mathscr{E}) \cap A = \mathscr{A}_A(\mathscr{E} \cap A)$.

(2) 下面证明 $\sigma_A(\mathscr{E} \cap A) = \sigma(\mathscr{E}) \cap A$. 首先证明 $\sigma(\mathscr{E}) \cap A$ 是 A 的 σ 代数. 与 (1) 同理, 可以得到 $\Omega \in \sigma(\mathscr{E}) \cap A$, 且 $\forall B \in \sigma(\mathscr{E}) \cap A$, $A \backslash B \in \sigma(\mathscr{E}) \cap A$. 于是我们只需证明 $\sigma(\mathscr{E}) \cap A$ 对可数并封闭. 考虑

$$
\{B_k, k \in \mathbb{N}\} \subset \sigma(\mathscr{E}), \quad \{B_k \cap A, k \in \mathbb{N}\} \subset \sigma(\mathscr{E}) \cap A,
$$

我们有

$$
\bigcup_{k=1}^{\infty} (B_k \cap A) = \left(\bigcup_{k=1}^{\infty} B_k \right) \cap A \in \sigma(\mathscr{E}) \cap A,
$$

因此 $\sigma(\mathscr{E}) \cap A$ 是 σ 代数. 因此 $\sigma_A(\mathscr{E} \cap A) \subset \sigma(\mathscr{E}) \cap A$.

考虑集类

$$
\mathcal{G} = \{G \in \sigma(\mathscr{E}), G \cap A \in \sigma_A(\mathscr{E} \cap A)\},
$$

我们将证明 $\mathcal{G}$ 是 Ω 的 σ 代数:

(a) 显然 $\Omega \in \mathcal{G}$;

(b) 考虑 $G \in \mathcal{G}$, 则 $G^c \cap A = A \backslash G \in \sigma_A(\mathscr{E} \cap A)$;

(c) 考虑 $\{G_k, k \in \mathbb{N}\} \subset \mathcal{G}$, 则 $\bigcup\limits_{k=1}^{\infty} G_k \in \sigma(\mathscr{E})$, $\left(\bigcup\limits_{k=1}^{\infty} G_k\right) \cap A = \bigcup\limits_{k=1}^{\infty} G_k \cap A \in \sigma_A(\mathscr{E} \cap A)$.

因此 $\mathcal{G}$ 是 Ω 的 σ 代数, 由 $\mathscr{E} \subset \mathcal{G}$, 则 $\sigma(\mathscr{E}) \subset \mathcal{G}$, 则 $\sigma(\mathscr{E}) \cap A \subset \sigma_A(\mathscr{E} \cap A)$. 故 $\sigma(\mathscr{E}) \cap A = \sigma_A(\mathscr{E} \cap A)$. □

3.1.14 若 $\mathscr{A}$ 是集代数且对一切两两不交的集序列的并封闭, 则 $\mathscr{A}$ 是 σ 代数.

证明: 回忆集代数的定义: 包含全集, 对有限交, 有限并, 补封闭. 因此我们只需证明 $\mathscr{A}$ 对可数并封闭. 考虑 $\{A_n : n \in \mathbb{N}\} \subset \mathscr{A}$ 是一集序列, 根据习题 1.1.9, 我们知道, 令

$$B_1 = A_1, \quad B_n = A_n \backslash \left(\bigcup_{k=1}^{\infty} A_k\right) \in \mathscr{A},$$

则 $\{B_n : n \in \mathbb{N}\}$ 两两不交. 但 $\bigcup\limits_{n=1}^{\infty} A_n = \bigcup\limits_{n=1}^{\infty} B_n \in \mathscr{A}$. □

3.1.15 设 $\mathscr{E}$ 是 Ω 的一个子集类, 它含有 Ω 且对对称差与可列交两种运算封闭, 问它是不是一个 σ 代数?

证明: 我们有

(1) 显然 $\Omega \in \mathscr{E}$;

(2) $\forall A \in \mathscr{E}$, $A^c = \Omega \Delta A \in \mathscr{E}$;

(3) 考虑 $\{A_k, k \in \mathbb{N}\} \subset \mathscr{E}$, 则 $\bigcup\limits_{k=1}^{\infty} A_k = \left(\bigcap\limits_{k=1}^{\infty} A_k^c\right)^c \in \mathscr{E}$.

故 $\mathscr{E}$ 是 σ 代数. □

3.1.16 设 S 是一组集运算. 若 Ω 中的非空子集类 $\mathscr{E}$ 对 S 中每一集运算都封闭, 则称 $\mathscr{E}$ 为一 S 类, 试证:

(1) 任意多个 S 类之交还是 S 类;

(2) 设 $\mathscr{E}$ 是 Ω 的一个子集类, 则存在一个唯一的包含 $\mathscr{E}$ 的最小 S 类.

证明: (1) 考虑一组 S 类 $\{\mathscr{E}_\alpha, \alpha \in I\}$, 考虑 $\{A_n\} \subset \bigcap\limits_{\alpha \in I} \mathscr{E}_\alpha$, 则 $\forall \alpha \in I$, $\{A_n\} \subset \mathscr{E}_\alpha$. 故对 $\{A_n\}$ 进行 S 运算得到的集合仍在 $\mathscr{E}_\alpha$ 中, 自然也在 $\bigcap\limits_{\alpha \in I} \mathscr{E}_\alpha$ 中.

(2) 考虑 $\mathscr{E}_0$ 是所有包含 $\mathscr{E}$ 的 S 类的交, 则它被任意包含 $\mathscr{E}$ 的 $\mathscr{S}$ 类包含, 且它是 S 类, 自然也是最小 S 类. □

3.1.17 称空间 Ω 中满足下述条件的集系 $\mathscr{D}$ 为 d 系:

(1) $\Omega \in \mathscr{D}$;

(2) 若 $A, B \in \mathscr{D}$ 且 $A \cap B = \varnothing$, 则 $B \cup A \in \mathscr{D}$ (对不交并封闭);

(3) 若 $A \subset B, A, B \in \mathscr{D}$, 则 $B \backslash A \in \mathscr{D}$ (对真差封闭).

试证: π 系上的最小 d 系等于 π 系上的最小集代数, 从而包含某一 π 系的 d 系必包含此 π 系生成的集代数.

证明: 记 $\mathcal{C}$ 为一 π 系, $\mathscr{D}(\mathcal{C})$, $\mathscr{A}(\mathcal{C})$ 分别为由 $\mathcal{C}$ 生成的最小 d 系与最小集代数, 则我们要证明的是 $\mathscr{D}(\mathcal{C}) = \mathscr{A}(\mathcal{C})$.

首先, 由 d 系及集代数的定义知, 若一个集类是集代数, 则它也是 d 系, 也就是说 $\mathscr{A}(\mathcal{C})$ 是一个包含 $\mathcal{C}$ 的 d 系, 因此 $\mathscr{D}(\mathcal{C}) \subset \mathscr{A}(\mathcal{C})$, 从而只需再证明 $\mathscr{D}(\mathcal{C}) \supset \mathscr{A}(\mathcal{C})$. 类似地, 若能证明 $\mathscr{D}(\mathcal{C})$ 是一个集代数, 就能完成证明.

回顾集代数的一系列定义: 定义 3.1.4、引理 3.1.5, 以及题目给出的 d 系的定义, 若想证明 $\mathscr{D}(\mathcal{C})$ 是一个集代数, 本质上只需证明 $\mathscr{D}(\mathcal{C})$ 对交运算封闭, 即 $\forall A, B \in \mathscr{D}(\mathcal{C})$, $A \cap B \in \mathscr{D}(\mathcal{C})$. 为此, 对于任一集合 A, 我们设

$$\mathscr{D}_A := \{B \in \mathscr{D}(\mathcal{C}) :\ B \cap A \in \mathscr{D}(\mathcal{C})\}$$

则只需证明

$$\forall A \in \mathscr{D}(\mathcal{C}),\ \mathscr{D}(\mathcal{C}) \subset \mathscr{D}_A$$

(注意: 通过引入 $\mathscr{D}_A$, 我们将原本难以直接下手的问题转化为了证明集类相互包含的问题).
然后, 同样地, 我们只需证明

$$\forall A \in \mathscr{D}(\mathcal{C}),\ (\mathrm{i}):\ \mathcal{C} \subset \mathscr{D}_A,\ 且\ (\mathrm{ii}):\ \mathscr{D}_A\ 是\ d\ 系.$$

首先证明 (i): $\forall A \in \mathscr{D}(\mathcal{C})$, $\mathcal{C} \subset \mathscr{D}_A$, 由 $\mathscr{D}_A$ 的定义知也即证明

$$\forall A \in \mathscr{D}(\mathcal{C}),\ \forall B \in \mathcal{C},\ A \cap B \in \mathscr{D}(\mathcal{C})$$

, 也即

$$\forall B \in \mathcal{C},\ \mathscr{D}(\mathcal{C}) \subset \mathscr{D}_B$$

(很巧妙, 之前引入了 $\mathscr{D}_A$, 现在是 $\mathscr{D}_B$, 不过由 $A \in \mathscr{D}(\mathcal{C})$ 变成了 $B \in \mathcal{C}$. 在继续往下看之前, 请读者先自己写出 $\mathscr{D}_B$ 的定义).
再一次, 同样地, 我们只需证明

$$\forall B \in \mathcal{C},\ \mathcal{C} \subset \mathscr{D}_B,\ 且\ \mathscr{D}_B\ 是\ d\ 系.$$

此处, $\mathcal{C} \subset \mathscr{D}_B$ 是显然的 (注意到题设: $\mathcal{C}$ 是 π 系. 请自行验证); $\mathscr{D}_B$ 是 d 系也容易验证 (读者可以尝试先自己验证):

(1) $\Omega \in \mathscr{D}_B$ 显然;

(2) 若 $A_1, A_2 \in \mathscr{D}_B$ 且 $A_1 \cap A_2 = \varnothing$, 则由 $\mathscr{D}_B$ 的定义知 $A_1, A_2 \in \mathscr{D}(\mathcal{C})$ 且 $A_1 \cap B, A_2 \cap B \in \mathscr{D}(\mathcal{C})$. 因为 $\mathscr{D}(\mathcal{C})$ 是 d 系, 所以 $A_1 \cup A_2 \in \mathscr{D}(\mathcal{C})$, 从而 $(A_1 \cup A_2) \cap B = (A_1 \cap B) \cup (A_2 \cap B) \overset{\mathscr{D}(\mathcal{C})\ 是\ d\ 系}{\in} \mathscr{D}(\mathcal{C})$, 即 $A_1 \cup A_2 \in \mathscr{D}_B$;

(3) 若 $A_1, A_2 \in \mathscr{D}_B$ 且 $A_1 \subset A_2$, 则同样可证 $A_2 \backslash A_1 \in \mathscr{D}_B$ (提示: 注意到定理 1.1.5(1): $(A_2 \backslash A_1) \cap B = (A_2 \cap B) \backslash (A_1 \cap B)$. 请读者自行完成这一证明);

从而我们完成了 (i) 的证明.

此外, 在证明 (i) 的过程中也完成了 (ii) 的证明. 由此我们完成了全部的证明. □

注: 读者可以用同样的思路尝试梳理集合形式的单调类定理 (定理 3.1.15) 的证明思路. 之后还会多次用到这样的思路. 总结如下: 利用"生成即最小"的原理 (例如, 由 $\mathcal{C}$ 生成的 σ 代数 $\sigma(\mathcal{C})$ 是包含 $\mathcal{C}$ 的最小的 σ

代数) 倒推, 当写不出来下一步的时候就引入一个新集类, 转化成证明集类相互包含的问题.

3.1.18 Ω 的子集类. $\mathscr{M}$ 称为 Ω 的单调类, 如果它满足:

(i) 对不降集列的并封闭: 即 $A_n \in \mathscr{M}$ 且 $A_n \uparrow, n \in \mathbb{N}$, 则有 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{M}$;

(ii) 对不升序列的交封闭: 即 $A_n \in \mathscr{M}$ 且 $A_n \downarrow, n \in \mathbb{N}$, 则有 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{M}$.

试证:

(1) 若 $\mathscr{A}$ 为 Ω 的集代数且为单调类, 则 $\mathscr{A}$ 为 σ 代数;

(2) Ω 的任一子集 $\mathcal{C}$ 上的最小单调类是存在的;

(3) 若 $\mathscr{A}$ 为 Ω 的集代数, 则包含 $\mathscr{A}$ 的最小单调类等于 $\sigma(\mathscr{A})$, 因而包含 $\mathscr{A}$ 的单调类必包含 $\sigma(\mathscr{A})$.

证明: (1) 只需证明 $\mathscr{A}$ 对于可数并封闭. 对于一列集合 $\{A_n:\ n \in \mathbb{N}\} \subset \mathscr{A}$, 考虑 $B_n = \bigcup\limits_{k=1}^{n} B_k$, 则 $B_n \uparrow$, 且易证 $\bigcup\limits_{n=1}^{\infty} B_n = \bigcup\limits_{n=1}^{\infty} A_n$. 又由于 $\mathscr{A}$ 是集代数, 有 $\forall n \in \mathbb{N}$, $B_n \in \mathscr{A}$; 且 $\mathscr{A}$ 是单调类, $B_n \uparrow$, 因此 $\bigcup\limits_{n=1}^{\infty} A_n = \bigcup\limits_{n=1}^{\infty} B_n \in \mathscr{A}$, 证毕.

(2) 考虑 $\mathscr{M}_\alpha, \alpha \in I$ 是所有 $\mathcal{C}$ 上的单调类. 令 $\mathscr{M}_0 = \bigcap\limits_{\alpha \in I} \mathscr{M}_\alpha$, 则

$$\{A_n, n \in \mathbb{N}\} \subset \mathscr{M}_0 \Rightarrow \forall \alpha \in I, \{A_n, n \in \mathbb{N}\} \subset \mathscr{M}_\alpha,$$

因此 $\forall \alpha \in I$, 若 $A_n \uparrow$, $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{M}_\alpha$; 若 $A_n \downarrow$, $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{M}_\alpha$. 故 $\mathscr{M}_0 = \bigcap\limits_{\alpha \in I} \mathscr{M}_\alpha$ 是单调类.

因为 $\mathscr{M}_0$ 被所有 $\mathcal{C}$ 上的单调类包含, 则 $\mathscr{M}_0$ 便是 $\mathcal{C}$ 上最小的单调类.

(3) 定义包含 $\mathscr{A}$ 的最小单调类为 $\mathscr{M}(\mathscr{A})$, 我们将证明 $\mathscr{M}(\mathscr{A}) = \sigma(\mathscr{A})$.

显然, 任何 σ 代数都是单调类 (见定义 3.1.8 及引理 3.1.9), 所以 $\mathscr{M}(\mathscr{A}) \subset \sigma(\mathscr{A})$. 因此只需证明 $\mathscr{M}(\mathscr{A}) \supset \sigma(\mathscr{A})$, 由 (1) 知只需证明 $\mathscr{M}(\mathscr{A})$ 是集代数.

要证明 $\mathscr{M}(\mathscr{A})$ 是集代数, 由引理 3.1.5 知只需证明 $\mathscr{M}(\mathscr{A})$ 对交运算及补集运算封闭. 为此, 对 $\forall A \in \mathscr{M}(\mathscr{A})$, 考虑集类

$$\mathscr{M}_A := \{B \in \mathscr{M}(\mathscr{A}):\ B \cap A \in \mathscr{M}(\mathscr{A}),\ B^c \in \mathscr{M}(\mathscr{A})\},$$

则只需证明

$$\forall A \in \mathscr{M}(\mathscr{A}),\ \mathscr{M}(\mathscr{A}) \subset \mathscr{M}_A,$$

进而只需证明

$$\forall A \in \mathscr{M}(\mathscr{A}),\ \text{(i)}\ \mathscr{A} \subset \mathscr{M}_A,\ \text{且 (ii)}\mathscr{M}_A\ \text{是单调类}.$$

首先证明 (ii): $\mathscr{M}_A$ 是单调类, 考虑不降集序列 $\{B_n, n \in \mathbb{N}\} \subset \mathscr{M}_A$, 则 $\{B_n \cap A,\ n \in \mathbb{N}\} \subset \mathscr{M}(\mathscr{A})$ 是不降集序列, $\{B_n^c,\ n \in \mathbb{N}\} \subset \mathscr{M}(\mathscr{A})$ 是不升集序列. 因此

$$\left(\bigcup_{n=1}^{\infty} B_n\right) \cap A = \bigcup_{n=1}^{\infty} (B_n \cap A) \in \mathscr{M}(\mathscr{A}),$$

$$\left(\bigcup_{n=1}^{\infty} B_n\right)^c = \bigcap_{n=1}^{\infty} B_n^c \in \mathscr{M}(\mathscr{A}).$$

故 $\bigcup\limits_{n=1}^{\infty} B_n \in \mathscr{M}_A$. 类似可证明 $\mathscr{M}_A$ 对于不升集序列的并也封闭. 故 $\mathscr{M}_A$ 是单调类.

再证明 (i)$\mathscr{A} \subset \mathscr{M}_A$, 即

$$\forall A \in \mathscr{M}(\mathscr{A}),\ \forall B \in \mathscr{A},\ B \in \mathscr{M}_A,$$

也即证明

$$\forall A \in \mathscr{M}(\mathscr{A}),\ \forall B \in \mathscr{A},\ B \cap A \in \mathscr{M}(\mathscr{A}),\ B^c \in \mathscr{M}(\mathscr{A})$$

. $B^c \in \mathscr{M}(\mathscr{A})$ 是显然的 (因为 $\mathscr{A}$ 是集代数), 剩下要证明

$$\forall A \in \mathscr{M}(\mathscr{A}),\ \forall B \in \mathscr{A},\ B \cap A \in \mathscr{M}(\mathscr{A}).$$

为此, 设

$$\mathcal{C}_B := \{A \in \mathscr{M}(\mathscr{A}):\ A \cap B \in \mathscr{M}(\mathscr{A})\}$$

则只需证明

$$\forall B \in \mathscr{A},\ \mathscr{M}(\mathscr{A}) \subset \mathcal{C}_B$$

进而, 只需证明

$$\forall B \in \mathscr{A},\ \mathscr{A} \subset \mathcal{C}_B,\ \text{且}\ \mathcal{C}_B\ \text{是单调类}.$$

这是容易证明的, 请读者自行完成剩下的证明.

由此我们证明了: 若 $\mathscr{A}$ 是集代数, 则 $\mathscr{M}(\mathscr{A}) = \sigma(\mathscr{A})$. □

注: 本题与习题 3.1.17以及集合形式的单调类定理的证明的思路大致相同: 利用 "生成即最小" 的原理 (例如, 由 $\mathcal{C}$ 生成的 σ 代数 $\sigma(\mathcal{C})$ 是包含 $\mathcal{C}$ 的最小的 σ 代数) 倒推, 当写不出来下一步的时候就引入一个新集类, 转化成证明集类相互包含的问题.

§ 3.2 单调函数与测度的构造

3.2.1 设 $(\Omega, \mathcal{F})$ 是可测空间, μ 是 $\mathcal{F}$ 上的可加测度, 且具有次 σ 可加性, 试证 μ 是测度.

证明: 我们只需证明其 σ 可加. 考虑两两不交的 $A_n \in \mathcal{F}(n \in N)$, 由于 μ 是可加的, 因此

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

因此 " $\leqslant$ " 必须取等. 故 μ 可加, 是测度. □

3.2.2 设 $(\Omega, \mathcal{F}, \mu)$ 是测度空间, 则

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

(2) μ 可加;

(3) μ 下方连续: 即对 $\mathcal{F}$ 中任何不降集列 $\{A_n\}$ (即 $A_n \in \mathcal{F}$, 且 $A_n \uparrow$) 都有

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

(4) μ 上方连续: 即对 $\mathcal{F}$ 中任何不升集列 $\{A_n\}$ (即 $A_n \in \mathcal{F}$, 且 $A_n \downarrow$) 且 $m \in \mathbb{N}$, 使 $\mu(A_m) < \infty$, 有

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

提示: 注意 $A_m \backslash A_n \uparrow$.

证明: (1) 我们知道 $A \cap \varnothing = \varnothing$, 则 $\mu(\varnothing) = \mu(A \cup \varnothing) - \mu(A) = 0$.

(2) 取 $A_k \in \mathcal{F}$, $k = 1, 2, \cdots$, 使得 $A_{n+1} = A_{n+2} = \cdots = \varnothing$, 则

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

(3) 我们有

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

(4) 考虑

$$A_n = \left(\bigcup_{k=n}^{\infty} (A_k \backslash A_{k+1})\right) \cup \left(\bigcap_{k=1}^{\infty} A_k\right),$$

故

$$\mu(A_n) = \sum_{k=n}^{\infty} \mu(A_k \backslash A_{k+1}) + \mu\left(\bigcap_{k=1}^{\infty} A_k\right)$$

取极限便得

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

□

3.2.3 设 $(\Omega, \mathcal{F})$ 为可测空间, μ 为 $\mathcal{F}$ 上可加测度, 且满足下述两条件之一:

(1) μ 下方连续;

(2) $\mu(\Omega) < \infty$ 且对 $\mathcal{F}$ 中任何下降到 $\varnothing$ 的集列 $\{A_n\}$ (即 $A_n \in \mathcal{F}$, 且 $A_n \supset A_{n+1}, n \in \mathbb{N}$, $\bigcap\limits_{n=1}^{\infty} A_n = \varnothing$) 都有

$$\lim_{n\to\infty} \mu(A_n) = 0 = \mu(\varnothing)$$

则 μ 为 $\mathcal{F}$ 上的测度.

证明: 我们只需证明 μ 是 σ 可加的. 我们知道 $\mathcal{F}$ 是 σ 代数, 因此 μ 有限可加.

若 (1) 成立, 令 $A_n := \bigcup\limits_{k=1}^{n} B_k \in \mathcal{F}$, 则 $A_n \uparrow \bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$, 因而

$$\mu\left(\bigcup_{n=1}^{\infty} B_k\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) = \sum_{k=1}^{\infty} \mu(B_k),$$

故 μ 是 σ 可加的. 若 (2) 成立, 令 $A_n=\bigcup\limits_{k=n+1}^{\infty}B_k$, 则 $A_n\in\mathcal{F}$ 且 $A_n\downarrow\varnothing$, 因此

$$\mu\left(\bigcup_{k=1}^{\infty}B_k\right)=\mu\left(\bigcup_{k=1}^{n}B_k\right)+\mu\left(\bigcup_{k=n+1}^{\infty}B_k\right)=\sum_{k=1}^{n}\mu(B_k)+\mu\left(\bigcup_{k=n+1}^{\infty}B_k\right),$$

我们知道

$$\lim_{n\to\infty}\mu\left(\bigcup_{k=n+1}^{\infty}B_k\right)=\lim_{n\to\infty}\mu(A_n)=0,$$

因此令 $n\to\infty$ 便有 $\mu\left(\bigcup\limits_{k=1}^{\infty}B_k\right)=\sum\limits_{k=1}^{\infty}\mu(B_k)$, 故 σ 可加. □

3.2.4 设 $(\Omega,\mathcal{F},\mu)$ 为测度空间, $\{A_n\}$ 为 $\mathcal{F}$ 中的集序列, 记

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

试证: $\mu\left(\liminf\limits_{n\to\infty}A_n\right)\leqslant\liminf\limits_{n\to\infty}\mu\left(A_n\right)$.

若 $\exists m\in\mathbb{N}$, 使 $\mu\left(\bigcup\limits_{k=m}^{\infty}A_k\right)<\infty$, 则

$$\mu\left(\limsup_{n\to\infty}A_n\right)\geqslant\limsup_{n\to\infty}\mu\left(A_n\right).$$

若 μ 为有限测度, 且 $\liminf\limits_{n\to\infty}A_n=\limsup\limits_{n\to\infty}A_n=A$, 则

$$\mu(A)=\lim_{n\to\infty}\mu\left(A_n\right).$$

证明: (1) 我们考虑 $\bigcap\limits_{k=n}^{\infty}A_n\uparrow\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{k=n}^{\infty}A_k=\liminf\limits_{n\to\infty}A_k$, 因此

$$\forall m>n,\mu\left(\bigcap_{k=n}^{\infty}A_k\right)\leqslant\mu\left(\bigcap_{k=m}^{\infty}A_k\right)\leqslant\mu(A_m),$$

故考虑 $m,n\to\infty$, 便有

$$\lim_{n\to\infty}\mu\left(\bigcap_{k=n}^{\infty}A_k\right)=\mu\left(\lim_{n\to\infty}\bigcap_{k=n}^{\infty}A_k\right)=\mu\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k\right)=\mu\left(\liminf_{n\to\infty}A_k\right)\leqslant\liminf_{m\to\infty}\mu(A_m).$$

(2) 考虑 $\bigcup\limits_{k=n}^{\infty}A_k\downarrow\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_n=\limsup\limits_{n\to\infty}A_k$, 我们知道 $\exists m\in\mathbb{N}$, 使得 $\mu\left(\bigcup\limits_{k=m}^{\infty}A_k\right)<\infty$, 故

$$\forall n>m,\left(\bigcup_{k=m}^{\infty}A_k\right)\setminus\left(\bigcup_{k=n}^{\infty}A_k\right)\uparrow\left(\bigcup_{k=m}^{\infty}A_k\right)\setminus\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\right)=\liminf_{n\to\infty}\left(\bigcup_{k=m}^{\infty}A_k\backslash A_n\right)$$

因此

$$\lim_{n\to\infty}\mu\left(\left(\bigcup_{k=m}^{\infty}A_k\right)\setminus\left(\bigcup_{k=n}^{\infty}A_k\right)\right)=\mu\left(\left(\bigcup_{k=m}^{\infty}A_k\right)\setminus\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\right)\right),$$

故

$$\limsup_{n\to\infty}\mu(A_n)\leqslant\lim_{n\to\infty}\mu\left(\bigcup_{k=n}^{\infty}A_k\right)=\mu\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\right)=\mu\left(\limsup_{n\to\infty}A_n\right).$$

□

3.2.5 证明引理 3.2.13.

引理 3.2.13: 设 $\forall A\subset\mathbb{R}^d$, 令

$$\lambda^*(A):=\inf\left\{\sum_{n=1}^{\infty}|I_n|:A\subset\bigcup_{n=1}^{\infty}I_n,I_n\subset\mathbb{R}^d\text{为开区间}\right\},$$

其中 $|I_n|$ 表示 I_n 的体积, 则 λ^* 是 $\mathbb{R}^d$ 的一个外测度.

证明: 我们需要证明 λ^* 满足:

(1) $\lambda^*(\varnothing)=0$;

(2) $\forall A\subset B\subset\mathbb{R}^d$, $\lambda^*(A)\leqslant\lambda^*(B)$;

(3) $\forall A_n\subset\mathbb{R}^d$, $n\in\mathbb{N}$, 有 $\lambda^*\left(\bigcup_{n=1}^{\infty}A_n\right)\leqslant\sum_{n=1}^{\infty}\lambda^*(A_n)$.

对于 (1), 任取开区间 I 使得 $|i|<\varepsilon$, 则有 $0\leqslant\lambda^*(\varnothing)<|I|<\varepsilon$.

对于 (2), 考虑 $A\subset B\subset\mathbb{R}^d$, 我们知道 B 的任意开覆盖必是 A 的开覆盖, 因此 $\lambda^*(A)\leqslant\lambda^*(B)$;

对于 (3), 若 $\sum_{n=1}^{\infty}\lambda^*(A_n)=\infty$, 则结论显然. 我们只需考虑 $\sum_{n=1}^{\infty}\lambda^*(A_n)<\infty$ 的情况.

我们知道, $\forall\varepsilon>0$, $n\in\mathbb{N}$, 存在 $\mathbb{R}^d$ 中的开区间族 $\{I_{n,k}:,k\in\mathbb{N}\}$, 使得 $A_n\subset\bigcup_{k=1}^{\infty}I_{n,k}$, 且 $\sum_{k=1}^{\infty}|I_{n,k}|\leqslant\lambda^*(A_n)+\frac{\varepsilon}{2^n}$. 因此

$$\lambda^*\left(\bigcup_{n=1}^{\infty}A_n\right)\leqslant\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\lambda^*(I_{n,k})\leqslant\sum_{n=1}^{\infty}\lambda^*(A_n)+\sum_{n=1}^{\infty}\frac{\varepsilon}{2^n}=\sum_{n=1}^{\infty}\lambda^*(A_n)+\varepsilon.$$

考虑到 ε 的任意性, 便有 $\lambda^*\left(\bigcup_{n=1}^{\infty}A_n\right)\leqslant\sum_{n=1}^{\infty}\lambda^*(A_n)$. □

3.2.6 设 $A_n,n\in\mathbb{N}$ 两两不交, 且 $\bigcup_{n=1}^{\infty}A_n=\Omega$ (称 $\{A_n\}$ 为 Ω 的一个可数划分), 试证: $\mathscr{S}=\{\varnothing\}\cup\left\{A_n,\bigcup_{k=n}^{\infty}A_k:n\in\mathbb{N}\right\}$ 为 Ω 的半集代数.

再设 $q_n\geqslant 0,n\in\mathbb{N}$, 在 $\mathscr{S}$ 上定义 $\mu(A_n)=q_n,\mu\left(\bigcup_{k=n}^{\infty}A_k\right):=\sum_{k=n}^{\infty}q_k,n\in\mathbb{N},\mu(\varnothing)=0$, 试具体写出 $\sigma(\mathscr{S})$ 的元及 μ 在 $\sigma(\mathscr{S})$ 上的扩张.

上述测度空间与下列测度空间 $(\Omega,\mathscr{A},\tilde{\mu})$ 是否相同? 此时, $\mathscr{A}$ 为 $A_n,n\in\mathbb{N}$ 的任意并作成的类,

$$\tilde{\mu}\left(\bigcup_{n\in J}A_n\right)=\sum_{n\in J}q_n,\quad\forall J\subset\mathbb{N}.$$

证明: 首先证明 $\mathscr{S}$ 是半集代数:

(1) 显然 $\varnothing \in \mathscr{S}, \Omega = \bigcup_{n=1}^{\infty} A_n \in \mathscr{S}$;

(2) 设 $A, B \in \mathscr{S}$, 则

i) $A = \varnothing$, 则 $A \cap B = \varnothing \in \mathscr{S}$;

ii) $A = A_i,\ B = A_j,\ i \neq j$ ($i = j$ 时显然), 则 $A \cap B = \varnothing \in \mathscr{S}$;

iii) $A = A_i,\ B = \bigcup_{k=n}^{\infty} A_k$, 则 $A \cap B = \begin{cases} \varnothing \in \mathscr{S},\ i \leqslant n \\ A_i \in \mathscr{S},\ i \geqslant n \end{cases}$

iv) $A = \bigcup_{k=i}^{\infty} A_k,\ B = \bigcup_{k=j}^{\infty} A_k$, 则 $A \cap B = \bigcup_{k=\max\{i,j\}}^{\infty} A_k \in \mathscr{S}$.

说明 $\forall A, B \in \mathscr{S},\ A \cap B \in \mathscr{S}$;

(3) 若 $B_1, B \in \mathscr{S},\ B_1 \subset B$, 不妨设 B_1 是 B 的真子集且 $B_1 \neq \varnothing$, 则

i) $B_1 = A_i \subset \bigcup_{k=n}^{\infty} A_k = B$ 时, 取 $B_2 = A_n,\ B_3 = A_{n+1}, \cdots, B_{i-n+1} = A_{i-1}, A_{i-n+2} = \bigcup_{k=i+1}^{\infty} A_k$, 则此时 $B_j \in \mathscr{S},\ j = 1, 2, 3, \cdots, i-n+2$ 两两不交使得 $B = \bigcup_{j=1}^{i-n+2} B_j$;

ii) $B_1 = \bigcup_{k=i}^{\infty} A_k \subset \bigcup_{k=j}^{\infty} A_k = B$ 时, 取 $B_2 = A_j,\ B_3 = A_{j+1}, \cdots, B_{i-j+1} = A_{i-1}$, 则此时 $B_l \in \mathscr{S},\ l = 1, 2, 3, \cdots, i-j+1$ 两两不交使得 $B = \bigcup_{l=1}^{i-j+1} B_l$.

说明 $\forall B_1, B \in \mathscr{S}:\ B_1 \subset B,\ \exists \{B_2, B_3, \cdots, B_n\} \subset \mathscr{S}$ 且两两不交, $s.t. B = \bigcup_{k=1}^{n} B_k$.

故 $\mathscr{S}$ 是半集代数.

$\sigma(\mathscr{S})$ 的元是 $\left\{ \bigcup_{i \in I} A_i : I \in \mathscr{P}(\mathbb{N}) \right\}$, 其中 $\mathscr{P}(\mathbb{N})$ 是自然数集 $\mathbb{N}$ 的幂集; μ 在 $\sigma(\mathscr{S})$ 上的扩张是 $\mu\left(\bigcup_{i \in I} A_i \right) = \sum_{i \in I} q_i,\ I \in \mathscr{P}(\mathbb{N})$.

由上述 $\sigma(\mathscr{S})$ 中元的形式易知 $(\Omega, \sigma(\mathscr{S}), \mu) = (\Omega, \mathscr{A}, \widetilde{\mu})$. 事实上:

(1) $\sigma(\mathscr{S}) = \mathscr{A}$. 首先显然 $\mathscr{A} \subset \sigma(\mathscr{S})$; 又易证 $\mathscr{A}$ 是 σ 代数且 $\mathscr{S} \in \mathscr{A}$, 从而 $\sigma(\mathscr{S}) \subset \mathscr{A}$, 即证得 $\sigma(\mathscr{S}) = \mathscr{A}$;

(2) $\mu = \widetilde{\mu}$. 首先显然有 $\forall A \in \mathscr{S}$, $\widetilde{\mu}(A) = \mu(A)$; 而由 μ 的定义知 μ 是 σ 有限的, 注意到上面已经证明 $\mathscr{S}$ 是半集代数, 由测度扩张定理 (3.2.7) 知扩张后的 μ 是唯一的, 从而 $\mu = \widetilde{\mu}$.

$\square$

3.2.7 设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是概率场, 若 $B \in \mathcal{F}, \mathbb{P}(B) > 0$, 则 $\forall A \in \mathcal{F}$, 定义

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

称为 A 在 B 之下的条件概率, 试证: $\mathbb{P}(\cdot|B)$ 是 $\mathcal{F}$ 上的概率测度, 因而 $(\Omega, \mathcal{F}, \mathbb{P}(\cdot|B))$ 是概率空间.

证明: 我们知道 $\mathbb{P}(\Omega|B) = \dfrac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = 1$, 故我们只需证明其是测度. 我们知道

(a) $\mathbb{P}(\varnothing|B) = \dfrac{\mathbb{P}(\varnothing)}{\mathbb{P}(B)} = 0$;

(b) 考虑 $\{A_n, n \in \mathbb{N}\} \subset \mathcal{F}$, 且它们两两不交, 则 $\{A_n \cap B, n \in \mathbb{N}\} \subset \mathcal{F}$ 且两两不交. 故

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n \middle| B\right) = \frac{\mathbb{P}((\cup_{n=1}^{\infty} A_n) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\cup_{n=1}^{\infty}(A_n \cap B))}{\mathbb{P}(B)} = \sum_{n=1}^{\infty} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n=1}^{\infty} \mathbb{P}(A_n|B).$$

因此 $\mathbb{P}(\cdot|B)$ 是测度, 因而其为概率测度. □

3.2.8 设 $\{A_k : k \in \mathbb{N}\}$ 是概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 中随机事件系, 试证:

(1) 若 $\mathbb{P}(A_n) = 1, n = 1, 2, \cdots$, 则 $\mathbb{P}\left(\bigcap_{k=1}^{\infty} A_k\right) = 1$;

(2) 若 $\mathbb{P}(A_n) = 0, n = 1, 2, \cdots$, 则 $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 0$.

证法一: 首先证明 (2). 设 $B_n = A_n \backslash \left(\bigcup_{k=1}^{n-1} A_k\right)$, 则 $B_n,\ n = 1, 2, \cdots$ 两两不交, 且由习题 1.1.9知 $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$; 又由 $\varnothing \subset B_n \subset A_n$ 且 $\mathbb{P}(A_n) = 0$ 知 $0 = \mathbb{P}(\varnothing) \leqslant \mathbb{P}(B_n) \leqslant \mathbb{P}(A_n) \leqslant 0$ 即 $\mathbb{P}(B_n) = 0,\ n = 1, 2, \cdots$, 从而 $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=0}^{\infty} \mathbb{P}(B_n) = 0$.

然后, 由 (2) 易证 (1). 由于 $1 = \mathbb{P}(\Omega) = \mathbb{P}(A_n \cup A_n^c) = \mathbb{P}(A_n) + \mathbb{P}(A_n^c)$, 从而 $\mathbb{P}(A_n^c) = 0,\ n = 1, 2, \cdots$, 由 (2) 得 $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) = 0$; 而 $\bigcup_{n=1}^{\infty} A_n^c = \left(\bigcap_{n=1}^{\infty} A_n\right)^c$, 同理可得 $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) + \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1$, 故 $\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1$. □

证法二: ((2) 的第二种解法) 由定义 3.3.1(7) 知测度具有次可加性, 从而 $0 \leqslant \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mathbb{P}(A_n) = 0$, 从而 $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 0$. □

3.2.9 若 $\mathbb{P}_0, \mathbb{P}_1$ 是定义在 $(\Omega, \mathcal{F})$ 上的两个概率测度, 且 $\forall \varepsilon > 0, \exists A_\varepsilon \in \mathcal{F}$, 使得 $\mathbb{P}_1(A_\varepsilon) > 1 - \varepsilon$, 而 $\mathbb{P}_0(A_\varepsilon) < \varepsilon$, 试证: 存在 $A \in \mathcal{F}$, 使得 $\mathbb{P}_1(A) = 1, \mathbb{P}_0(A) = 0$.

提示: 取 $A = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_{2^{-m}}$, 具有上述性质的两个测度是相互奇异的.

证明: 取 $\varepsilon=2^{-m}$, $A=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{m=n}^{\infty}A_{2^{-m}}$, 则

$$1\geqslant \mathbb{P}_1(A)\overset{(1)}{=}\lim_{n\to\infty}\mathbb{P}_1\left(\bigcup_{m=n}^{\infty}A_{2^{-m}}\right)\geqslant\lim_{n\to\infty}\mathbb{P}_1(A_{2^{-n}})\geqslant\lim_{n\to\infty}\left(1-2^{-n}\right)=1$$

$$0\leqslant \mathbb{P}_0(A)\overset{(1)}{=}\lim_{n\to\infty}\mathbb{P}_0\left(\bigcup_{m=n}^{\infty}A_{2^{-m}}\right)\overset{(2)}{\leqslant}\lim_{n\to\infty}\mathbb{P}_0\left(\sum_{m=n}^{\infty}\mathbb{P}_0(A_{2^{-m}})\right)=\lim_{n\to\infty}\mathbb{P}_0\left(\sum_{m=n}^{\infty}2^{-m}\right)=0$$

从而 $\mathbb{P}_1(A)=1$, $\mathbb{P}_0(A)=0$.

(注: 其中 (1) 这一步由习题 3.2.2(4) 可得; (2) 这一步是因为任意测度 μ 都具有次 σ 可加性: 设 $B_n=A_n\backslash\left(\bigcup\limits_{k=1}^{n-1}A_k\right)$, 则 B_n, $n=1,2,\cdots$ 两两不交且 $\mu(B_n)\leqslant\mu(A_n)$, $n=1,2,\cdots$, 且由习题 1.1 的 9 题知 $\bigcup\limits_{n=1}^{\infty}A_n=\bigcup\limits_{n=1}^{\infty}B_n$, 从而 $\mu\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=\mu\left(\bigcup\limits_{n=1}^{\infty}B_n\right)=\sum\limits_{n=1}^{\infty}\mu(B_n)\leqslant\sum\limits_{n=1}^{\infty}\mu(A_n)$.) □

3.2.10 设 $\mathbb{P}',\mathbb{P}$ 是定义在 $(\Omega,\mathcal{F})$ 上的两个概率测度, 若对任何使 $\mathbb{P}(A)=0$ 的 $A\in\mathcal{F}$, 都有 $\mathbb{P}'(A)=0$, 则称 $\mathbb{P}'$ 对 $\mathbb{P}$ 连续, 记作 $\mathbb{P}'\ll\mathbb{P}$. 试证:

(1) 设 $\mathbb{P}_1,\mathbb{P}_2$ 是定义在 $(\Omega,\mathcal{F})$ 上的两个概率测度, 则有 $(\Omega,\mathcal{F})$ 上的一个概率测度 $\mathbb{P}$ 使得 $\mathbb{P}_1\ll\mathbb{P}$ 且 $\mathbb{P}_2\ll\mathbb{P}$;

(2) 将 (1) 推广至无穷多个概率测度 $\mathbb{P}_1,\mathbb{P}_2,\cdots$ 的情形.

证明: (1) 考虑 $\mathbb{P}=\dfrac{\mathbb{P}_1+\mathbb{P}_2}{2}$, 我们首先证明这是概率测度. 显然 $\mathbb{P}(\varnothing)=0$, $\mathbb{P}(\Omega)=1$, 因此我们只需验证 $\mathbb{P}$ 是 σ 可加的. 考虑两两不交的 $\{A_n,n\in\mathbb{N}\}\subset\Omega$, 我们有

$$2\mathbb{P}\left(\bigcup_{n=1}^{\infty}A_n\right)=\mathbb{P}_1\left(\bigcup_{n=1}^{\infty}A_n\right)+\mathbb{P}_2\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mathbb{P}_1(A_n)+\sum_{n=1}^{\infty}\mathbb{P}_2(A_n)=2\sum_{n=1}^{\infty}\mathbb{P}(A_n),$$

因此 $\mathbb{P}$ 是概率测度. 由于 $\mathbb{P},\mathbb{P}_1,\mathbb{P}_2\geqslant0$, 所以 $\mathbb{P}=0\Rightarrow\mathbb{P}_1=\mathbb{P}_2=0$. 因此 $\mathbb{P}_1\ll\mathbb{P}$ 且 $\mathbb{P}_2\ll\mathbb{P}$.

(2) 考虑 $\mathbb{P}=\sum\limits_{n=1}^{\infty}\dfrac{\mathbb{P}_n}{2^n}$, 类似地有 $\mathbb{P}(\varnothing)=0$, $\mathbb{P}(\Omega)=1$. 我们需要证明其 σ 可加. 考虑两两不交的 $\{A_n,n\in\mathbb{N}\}\subset\Omega$, 我们有

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{k=1}^{\infty}\frac{1}{2^k}\mathbb{P}_k\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{k=1}^{\infty}\frac{1}{2^k}\sum_{n=1}^{\infty}\mathbb{P}_k(A_n)=\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\mathbb{P}(A_n)}{2^k}=\sum_{n=1}^{\infty}\mathbb{P}(A_n)$$

因此 $\mathbb{P}$ 是概率测度. 且 $\mathbb{P}=0\Rightarrow\mathbb{P}_n=0,n\in\mathbb{N}$, 故 $\mathbb{P}_n\ll\mathbb{P}$, $\forall n\in\mathbb{N}$. □

3.2.11 设 f 是增函数且存在实数 A 与 B 使得 $\forall x:A\leqslant f(x)\leqslant B$. 证明: $\forall\varepsilon>0$, 大小超过 ε 的跳跃数最多为 $(B-A)\varepsilon^{-1}$. 由此证明: 任何不降函数 f 的不连续点集合最多可数.

提示: 首先就 f 有界的情形来证明, 然后考虑一般情况.

证明: 对于增函数 f: $\forall x:A\leqslant f(x)\leqslant B$, 假设大小超过 ε 的跳跃数大于 $(B-A)\varepsilon^{-1}$, 则由于 $f(x)$ 是增函数, 有

$$\sup_x f(x)>A+(B-A)\varepsilon^{-1}\cdot\varepsilon>B,$$

矛盾! 由此, 我们证明了有界增函数的跳跃点个数有限.

对于一般的不降函数 f, 注意到 $\mathbb{R} = \bigcup\limits_{n=-\infty}^{\infty}(n, n+1]$, 有 $f = \sum\limits_{n=-\infty}^{\infty} f\mathbb{1}_{\{n<f\leqslant n+1\}}$. 由于每个 $f\mathbb{1}_{\{n<f\leqslant n+1\}}$ 都是有界且不降的函数, 具有有限个跳跃点, 因此 f 的跳跃点个数至多可数. □

3.2.12 设 f 是 $(-\infty, +\infty)$ 上的一个任意函数, L 是所有这种 x 的集: f 在 x 处右连续但不左连续, 证明 L 是一有限集或可数集.

提示: 考虑 $L \cap M_n$, 其中

$$M_n = \left\{x | O(f;x) > \frac{1}{n}\right\}, O(f;x) = \inf_{\varepsilon>0} \sup_{t\in(x-\varepsilon,x+\varepsilon)} |f(t) - f(x)|.$$

证明: 我们只需证明 $\forall n$, $L \cap M_n$ 是至多可数的. 定义

$$A_n := \{a \in L \cap M_n : \exists \varepsilon > 0, (a, a+\varepsilon) \cap (L \cap M_n) = \varnothing\},$$

则 A_n 中每个 a 所确定的区间中都可以选取一个有理数, 因此 A_n 至多可数.

考虑反证法, 假设 $\exists n$, s.t. $L \cap M_n$ 不可数, 则 $\exists p \in L \cap M_n$, $p \notin A_n$. 因此

$$\forall \varepsilon > 0, (p, p+\varepsilon) \cap (L \cap M_n) \neq \varnothing,$$

此时若 $x \to p$ 且 $x, y \in (p, p+\varepsilon) \cap M_n$, 则 $|f(x) - f(y)| > \dfrac{1}{n}$. (这是 M_n 的性质) 则 x 处不右连续, 因此 $((p, p+\varepsilon) \cap M_n) \cap L = \varnothing$, 与前文矛盾! 故 L 至多可数. □

3.2.13 设 f 是 D 上增函数, D 在 $(-\infty, +\infty)$ 中稠密, 在 $\mathbb{R}$ 上如下定义 $\tilde{f} : \tilde{f}(x) = \inf\limits_{x<t\in D} f(t), \forall x \in \mathbb{R}$. 证明: f 在 D 上的一致连续性一定蕴涵 $\tilde{f}$ 在 $\mathbb{R}$ 上的一致连续性; 并举一反例说明 f 在 D 上的连续性并不蕴涵 $\bar{f}$ 在 $\mathbb{R}$ 上的连续性.

证明: 由 f 在 D 上一致连续知

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon),\ s.t.\ \forall x_1, x_2 \in D : |x_1 - x_2| < \delta,\ \ |f(x_1) - (x_2)| < \varepsilon$$

现将上述命题中 ε 替换为 $\varepsilon' = \dfrac{\varepsilon}{2} > 0$, 对应 $\delta' = \delta'(\varepsilon') = \delta'(\varepsilon)$, 则可得下述命题:

$$\forall \varepsilon > 0,\ \varepsilon' = \frac{\varepsilon}{2} > 0,\ \exists \delta' = \delta'(\varepsilon') = \delta'(\varepsilon),\ s.t.\ \forall x_1, x_2 \in D : |x_1 - x_2| < \delta',\ \ |f(x_1) - (x_2)| < \varepsilon' \qquad (*)$$

固定上述 δ', 取 $x_1', x_2' \in \mathbb{R}$: $|x_1' - x_2'| < \dfrac{\delta'}{2}$, 不妨设 $x_1' < x_2'$, 则由 D 在 $\mathbb{R}$ 中稠密及 f 是 D 上的增函数知

$$\begin{aligned}
&\exists \widetilde{x_1}, \widetilde{x_2} \in D_{x_1'}^{\delta'} \overset{\text{def}}{=} D \cap (x_1', x_1' + \delta') : \\
&\qquad\qquad \widetilde{f}(x_1') = \inf_{x_1'<t\in D} f(x) = \inf_{x_1'<t\in D_{x_1'}^{\delta'}} f(x) > f(\widetilde{x_1}) - \varepsilon', \\
&\qquad\qquad \widetilde{f}(x_2') = \inf_{x_2'<t\in D} f(x) = \inf_{x_2'<t\in D_{x_1'}^{\delta'}} f(x) \leqslant f(\widetilde{x_2})
\end{aligned}$$

从而

$$\begin{aligned}
\left|\widetilde{f}(x_1') - \widetilde{f}(x_2')\right| = \left|\inf_{x_1'<t\in D} f(x) - \inf_{x_2'<t\in D} f(x)\right| &= \inf_{x_2'<t\in D} f(x) - \inf_{x_1'<t\in D} f(x) \\
&< f(\widetilde{x_2}) - (f(\widetilde{x_1}) - \varepsilon') \\
&< |f(\widetilde{x_1}) - f(\widetilde{x_2})| + \varepsilon' \\
&\overset{(**)}{<} 2\varepsilon' = \varepsilon
\end{aligned}$$

（其中 (**) 是因为 $|\widetilde{x_1}-\widetilde{x_2}|<\delta'$, 从而由 (*) 知 $|f(\widetilde{x_1})-f(\widetilde{x_2})|<\varepsilon'$）这就证明了

$$\forall\varepsilon>0,\ \exists\delta''=\frac{\delta'}{2}=\delta''(\varepsilon),\ s.t.\ \forall x_1',x_2'\in\mathbb{R}:\ |x_1'-x_2'|<\delta'',\ \left|\widetilde{f}(x_1')-\widetilde{f}(x_2')\right|<\varepsilon$$

即 $\widetilde{f}$ 在 $\mathbb{R}$ 上一致连续.

f 在 D 上的连续性并不蕴含 $\widetilde{f}$ 在 $\mathbb{R}$ 上的连续性, 反例如下：令

$$f(x)=\begin{cases}x, & x<\sqrt{2}\\ x+1, & x\geqslant\sqrt{2}\end{cases}$$

则 f 在 $\mathbb{Q}$ 上单增且连续, $\mathbb{Q}$ 在 $\mathbb{R}$ 中稠密, 但是此时对应的 $\widetilde{f}$ 在 $\mathbb{R}$ 上并不连续（$\sqrt{2}$ 是 $\widetilde{f}$ 的间断点）.

□

3.2.14 任给 $\mathbb{R}$ 上的实值函数 f, 存在可数集 D 具有如下性质. $\forall t,\exists t_n\in D,n\in\mathbb{N},t_n\to t(n\to\infty)$, 使得 $f(t)=\lim\limits_{n\to\infty}f(t_n)$. 如果 “$t_n\to t$ ”, 用 “$t_n\downarrow t$ ” 或 “$t_n\uparrow t$ ” 来代替, 上述论断仍成立.

证明: 我们知道 $(\mathbb{R}^2,|x-y|)$ 是可分距离空间, 因此其子空间 $(x,f(x)),x\in\mathbb{R}$ 也是可分的. 考虑距离

$$d((x,f(x)),(y,f(y))):=|x-y|+\left|\frac{f(x)}{1+|f(x)|}-\frac{f(y)}{1+|f(y)|}\right|,$$

容易验证 d 与 $\mathbb{R}^2$ 上的距离 $|x-y|$ 等价, 故 $(x,f(x)),x\in\mathbb{R}$ 在距离 d 下是可分距离空间. 则存在一个稠子集 D, s.t. $\forall t\in\mathbb{R}$, $(t,f(t))$ 中都存在 D 中的元素 $(t_n,f(t_n))$ 使得 $d((t_n,f(t_n)),(t,f(t)))\to 0$, 即证. □

3.2.15 计算下列各 Borel 集的 Lebesgue 测度:

(1) $[0,1]$ 中的无理点集;

(2) Cantor 集;

(3) Sierpinski 海绵;

(4) 圆周;

(5) 开圆 $B(0,1)$;

(6) $[a,b]$ 上的连续函数 $f(x),x\in[a,b]$ 的图 $\{(x,f(x)):x\in[a,b]\}$.

证明:

(1) 由 $\mathbb{Q}$ 是可数集及 Lebesgue 测度的定义易知 $\lambda(\mathbb{Q})=0$, 从而 $0\leqslant\lambda(\mathbb{Q}\cap[0,1])\leqslant\lambda(\mathbb{Q})=0$ 即 $\lambda(\mathbb{Q}\cap[0,1])=0$, 从而 $\lambda([0,1]\setminus\mathbb{Q})=\lambda([0,1])-\lambda(\mathbb{Q}\cap[0,1])=1-0=1$.

下证 $\lambda(\mathbb{Q})=0$. 若设 $\mathbb{Q}=\{r_n:\ n\in\mathbb{N}\}$, 则

$$\forall\varepsilon>0,\ \mathbb{Q}\subset\bigcup_{n=1}^{\infty}\left(r_n-\frac{\varepsilon}{2^{n+1}},r_n+\frac{\varepsilon}{2^{n+1}}\right)$$

从而

$$\lambda^*(\mathbb{Q})\leqslant\lambda\left(\bigcup_{n=1}^{\infty}\left(r_n-\frac{\varepsilon}{2^{n+1}},r_n+\frac{\varepsilon}{2^{n+1}}\right)\right)\leqslant\sum_{n=1}^{\infty}\lambda\left(\left(r_n-\frac{\varepsilon}{2^{n+1}},r_n+\frac{\varepsilon}{2^{n+1}}\right)\right)=\sum_{n=1}^{\infty}\frac{\varepsilon}{2^n}=\varepsilon$$

因此 $\lambda^*(\mathbb{Q})=0$, 从而（注：$\Omega=\mathbb{R}$, 从而 $\mathbb{Q}^c=\mathbb{R}\setminus\mathbb{Q}$）

$$\forall D\subset\mathbb{R},\ \lambda^*(\mathbb{Q}\cap D)+\lambda^*(\mathbb{Q}^c\cap D)\leqslant\lambda^*(\mathbb{Q})+\lambda^*(D)\xlongequal{\lambda^*(\mathbb{Q})=0}\lambda^*(D)$$

由引理 3.2.15 知 $\mathbb{Q}$ 是 Lebesgue 可测的. 由定理 3.2.16(3) 知 Lebesgue 外测度 λ^* 限制在 Lebesgue 可测集上是测度（可看做从半集代数 $\mathscr{S}$ 上扩张的 Lebesgue 测度 λ）, 从而 $\lambda(\mathbb{Q}) = \lambda^*(\mathbb{Q}) = 0$.

(2) 设 Cantor 集为 C, 则 $1-\lambda(C) = \lambda([0,1]\setminus C) = \sum_{n=0}^{\infty}\frac{1}{3}\left(\frac{2}{3}\right)^n = \frac{1}{3}\cdot\frac{1}{1-\frac{2}{3}} = 1$, 从而 $\lambda(C)=0$.

(3) 考虑 $E_0\backslash Q_0$, 这里 E_0 是边长为 $a<\infty$ 的等边三角形, Q_0 为 Sierpinski 海绵. 则

$$\lambda(Q_0) = \lambda(E_0) - \lambda(E_0\backslash Q_0) = \frac{\sqrt{3}}{4}a^2\left(1-\sum_{n=0}^{\infty}\frac{3}{4}\left(\frac{1}{2}\right)^{2n}\right) = \frac{\sqrt{3}}{4}a^2(1-1) = 0.$$

(4) 考虑半径为 $a<\infty$ 的圆周 $\partial B(0,a)$, 则

$$0 \leqslant \lambda(\partial B(0,a)) < \pi\left(\left(a+\frac{\varepsilon}{2a\pi}\right)^2 - \left(a-\frac{\varepsilon}{2a\pi}\right)^2\right) = \varepsilon.$$

考虑到 ε 的任意性, $\lambda(\partial B(0,a)) = 0$.

(5) 有 $\lambda(B(0,1)) = \pi - \lambda(\partial B(0,1)) = \pi$.

(6) 取 $\{r_n:\ n\in\mathbb{N}\} = [a,b]\cap\mathbb{Q}$, 则由 $\mathbb{Q}$ 在 $\mathbb{R}$ 中稠以及 f 是 $[a,b]$ 上的连续函数知

$$\forall\varepsilon>0,\ \ \{(x,f(x)):\ x\in[a,b]\} \subset \bigcup_{n=1}^{\infty} B\left(x_n, \sqrt{\frac{\varepsilon}{2^n\pi}}\right).$$

其中 $B\left(x_n,\sqrt{\frac{\varepsilon}{2^n\pi}}\right)$ 是以点 $x_n := (r_n, f(r_n))$ 为圆心, 半径为 $\sqrt{\frac{\varepsilon}{2^n\pi}}$ 的圆, 注意到 $\lambda\left(B\left(x_n,\sqrt{\frac{\varepsilon}{2^n\pi}}\right)\right) = \frac{\varepsilon}{2^n}$, 类似于第 (1) 小问可证 $\lambda(\{(x,f(x)):\ x\in[a,b]\}) = 0$.

□

3.2.16 若 $\mu^*(A)=0$, 则 $\mu^*(A\cup B) = \mu^*(B)$.

证明: 因为 $A\cup B\supset B$, 由外测度的不降性可知 $\mu^*(A\cup B)\geqslant\mu^*(B)$; 又由外测度的次 σ 可加性可知 $\mu^*(A\cup B)\leqslant\mu^*(A)+\mu^*(B)\overset{\mu^*(A)=0}{=}\mu^*(B)$, 从而 $\mu^*(A\cup B)=\mu^*(B)$. □

3.2.17 若 $\mathbb{R}^n$ 的有界闭区间 $[a,b]$ 至少有一边长为 0 (即至少有一 $k, 1\leqslant k\leqslant n$, 使 $a_k=b_k$), 则

$$\lambda^*([a,b]) = 0.$$

证明: 首先, $\lambda^*([a,b]) = \lambda^*((a,b])$. 事实上, $[a,b] = \{a\}\cup(a,b]$, 而由外测度定义显然 $\lambda^*(\{a\})=0$, 从而由习题 3.2 的 16 题可知 $\lambda^*([a,b]) = \lambda^*((a,b])$.

其次, 由推论 3.2.18 可知 $\lambda^*((a,b]) = \prod_{k=1}^{n}(b_k-a_k) \xlongequal{\exists k,\ s.t.\ a_k=b_k} 0$, 故 $\lambda^*([a,b])=0$. □

3.2.18 设 μ,ν 是可测空间 $(\Omega,\mathcal{F})$ 上的两个有限测度, $\mathcal{C}$ 为 π 系, $\Omega\in\mathcal{C}$ 且 $\sigma(\mathcal{C})=\mathcal{F}$. 若 μ,ν 在 $\mathcal{C}$ 上一致, 则 μ,ν 在 $\mathcal{F}$ 上一致.

证明: 设

$$\Lambda = \{A\in\mathcal{F}:\ \mu(A)=\nu(A)\}$$

则只需证明 $\Lambda\supset\mathcal{F}=\sigma(\mathcal{C})$; 由单调类定理（3.1.15）知只需证明 $\mathcal{C}\subset\Lambda$ 且 Λ 是 λ 系.

由 μ,ν 在 $\mathcal{C}$ 上一致知 $\mathcal{C}\subset\Lambda$; 而

(1) $\Omega\in\mathcal{C}$ 且 $\mathcal{C}\subset\Lambda$, 说明 $\Omega\in\Lambda$;

(2) 若 $A,B\in\Lambda$, $A\subset B$, 则 $\mu(A)=\nu(A)$, $\mu(B)=\nu(B)$, 从而 $\mu(B\setminus A)=\mu(B)-\mu(A)=\nu(B)-\nu(A)=\nu(B\setminus A)$（由 μ,ν 在 $\mathcal{F}$ 上都是有限测度知此处的减法有定义）, 即 $B\setminus A\in\Lambda$;

(3) 对 $\{A_n:\ n\in\mathbb{N}\}\subset\Lambda$, $A_n\uparrow$, 由测度的下方连续性（定理 3.3.2(1)）知 $\mu\left(\bigcup_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}\mu(A_n)$

$\xlongequal{A_n\in\Lambda\Rightarrow\mu(A_n)=\nu(A_n)}\lim_{n\to\infty}\nu(A_n)=\nu\left(\bigcup_{n=1}^{\infty}A_n\right)$, 即 $\bigcup_{n=1}^{\infty}A_n\in\Lambda$;

以上三条说明 Λ 的确是 λ 系. 由单调类定理（3.1.15）知 $\Lambda\supset\sigma(\mathcal{C})=\mathcal{F}$, 说明 μ,ν 在 $\mathcal{F}$ 上一致. □

3.2.19 设 $\Omega=\{a,b,c,d\},\mathcal{C}=\{\{a,b\},\{c,d\},\{a,c\},\{b,d\},\Omega,\varnothing\}$, 令

$$\begin{aligned}
&\mu_1(\{a\})=\mu_1(\{d\})=\mu_2(\{b\})=\mu_2(\{c\})=1,\\
&\mu_1(\{b\})=\mu_1(\{c\})=\mu_2(\{a\})=\mu_2(\{d\})=2,\\
&\mu_i(\{x,y\})=\mu_i(\{x\})+\mu_i(\{y\}),\quad i=1,2,\{x,y\}\in\mathcal{C},\\
&\mu_i(\varnothing)=0,\quad \mu_i(\Omega)=6,\quad i=1,2.
\end{aligned}$$

试证: $\mathcal{C}$ 不是半集代数, 在 $\mathcal{C}$ 上 $\mu_1=\mu_2$ 且都 σ 可加, 但在 $\sigma(\mathcal{C})$ 上 $\mu_1\neq\mu_2$. 这个例子说明了什么问题?

证明: 首先, 显然 $\mathcal{C}$ 不是半集代数: $\{a,b\}\cap\{a,c\}=\{a\}\notin\mathcal{C}$, 说明 $\mathcal{C}$ 对集合的交不封闭;
其次容易验证 σ 可加性: 显然, 若 $\{A_n\}\subset\mathcal{C}$ 两两不相交, 则 $\{A_n\}$ 中至多有两个集合不等于 $\varnothing$; 而 $\mu_i(\varnothing)=0$, $i=1,2$, 故只需考虑两个集合的情形: 对 $A,B\in\mathcal{C}$: $A\cap B=\varnothing$,

i) 若 $A=\varnothing$ 则由 $\mu_i(\varnothing)=0$ 可知 $\mu_i(A\cap B)=\mu_i(A)+\mu_i(B)$, $i=1,2$;

ii) 若 $A=\{x,y\}$, $B=\{z,w\}$, 则 $\mu_i(A\cup B)=\mu_i(\Omega)=6=1+1+2+2=\mu_i(\{x\})+\mu_i(\{y\})+\mu_i(\{z\})+\mu_i(\{w\})=\mu_i(\{x,y\})+\mu_i(\{z,w\})=\mu_i(A)+\mu_i(B)$, $i=1,2$

而由 $\mu_1(\{a\})\neq\mu_2(\{a\})$ 知在 $\sigma(\mathcal{C})$ 上 $\mu_1\neq\mu_2$, 这说明测度扩张定理中去掉 $\mathcal{C}$ 是半集代数的条件后, 扩张到 $\sigma(\mathcal{C})$ 上所得的测度未必唯一. （另, μ_i, $i=1,2$ 可以扩张到 $\sigma(\mathcal{C})$ 上, 按 $\forall A\in\sigma(\mathcal{C})$, $\mu_i(A)=\sum_{x\in A}\mu_i(\{x\})$, $i=1,2$ 定义即可） □

3.2.20 如果测度扩张定理中的“σ 有限”条件去掉, 则扩张的唯一性未必成立. 例如 $\Omega=\mathbb{R}$, 在 $\mathscr{S}$ 上定义 $\mu_1(A)=|A|,\mu_2=2\mu_1$, 试证: μ_1,μ_2 都是 $\mathscr{B}$ 上的测度, 不是 σ 有限的, 但在 $\mathscr{S}$ 上 $\mu_1=\mu_2$, 而在 $\sigma(\mathscr{S})=\mathscr{B}$ 上, $\mu_1\neq\mu_2$.

证明:

(1) μ_1,μ_2 都是 $\mathscr{B}$ 上的测度. 只需证明 μ_1 是 $\mathscr{B}$ 上的测度. 对 A_n, $n=1,2,\cdots$ 两两不交有 $\mu_1\left(\bigcup_{n=1}^{\infty}A_n\right)=\left|\bigcup_{n=1}^{\infty}A_n\right|=\sum_{n=1}^{\infty}|A_n|=\sum_{n=1}^{\infty}\mu_1(A_n)$.

(2) μ_1,μ_2 不是 σ 有限的. 取闭区间 $[0,1]$, 则由定理 1.4.1 知 $[0,1]$ 是不可数集. 若 μ_1 是 σ 可加的, 则 $\exists\{A_n:\ n\in\mathbb{N},\ \mu_1(A_n)=|A_n|<\infty\}\in\mathcal{C},\ s.t.\,[0,1]=\bigcup\limits_{n=1}^{\infty}A_n$, 由 $A_n,\ n=1,2,\cdots$ 是有限集知 $[0,1]=\bigcup\limits_{n=1}^{\infty}A_n$ 是可数集, 矛盾! 说明 μ_1 不是 σ 可加的. 同理可以说明 μ_2 不是 σ 可加的.

(3) 在 $\mathscr{S}$ 上 $\mu_1=\mu_2$, 因为 $\forall\,(a,b]\in\mathscr{S},\ \mu_1((a,b])=\infty$（$(a,b]$ 里有无穷多个数）, 从而 $\mu_2((a,b])=2\mu_1((a,b])=\infty$, 即 $\mu_1((a,b])=\infty=\mu_2((a,b])$, 说明在 $\mathscr{S}$ 上 $\mu_1=\mu_2$.

(4) 然而在 $\mathscr{B}$ 上 $\mu_1\neq\mu_2$. 事实上, 由 σ 代数的定义容易证明 σ 代数对可列交封闭, 从而 $\{b\}=\bigcap\limits_{n=1}^{\infty}\left(b-\dfrac{1}{n},b\right]\in\sigma(\mathscr{S})$, 但是 $\mu_1(\{b\})=1$ 而 $\mu_2(\{b\})=2$, 即 μ_1 和 μ_2 在 $\sigma(\mathscr{S})$ 上并不相等.

□

3.2.21 设 $(\Omega,\mathcal{F},\mu)$ 为一测度空间, Δ 为 Ω 的任一子集, 令

$$\mu_\Delta(A):=\mu^*(\Delta\cap A),$$

试证: $(\Delta,\Delta\cap\mathcal{F},\mu_\Delta)$ 为一以 Δ 为空间的测度空间. 若 $0<\mu^*(\Delta)<\infty$, 则

$$\left(\Delta,\Delta\cap\mathcal{F},\frac{\mu_\Delta}{\mu^*(\Delta)}\right)$$

为一以 Δ 为样本空间的概率空间.

证明: 首先证明 $(\Delta,\Delta\cap\mathcal{F},\mu_\Delta)$ 是一个以 Δ 为空间的测度空间, 为此, 首先说明 $\Delta\cap\mathcal{F}$ 是 Δ 上的 σ 代数:

(1) $\Delta=\Delta\cap\Omega\in\Delta\cap\mathcal{F}$;

(2) 若 $A\subset\Delta$ 且 $A\in\Delta\cap\mathcal{F}$, 即 $\exists B\in\mathcal{F},\ s.t.\ A=\Delta\cap B$, 则 $\Delta\setminus A=\Delta\setminus B=\Delta\cap B^c$, 其中 $B\in\mathcal{F}\Rightarrow B^c\in\mathcal{F}$, 说明 $\Delta\setminus A\in\Delta\cap\mathcal{F}$, 即相对 Δ 取补集封闭;

(3) 对 $\{A_n\in\Delta\cap\mathcal{F}:n\in\mathbb{N}$ 且 A_n 两两不交$\}$, $\forall n\in\mathbb{N},\ \exists B_n\in\mathcal{F},\ s.t.\ A_n=\Delta\cap B_n$, 从而 $\bigcup\limits_{n=1}^{\infty}A_n=\Delta\cap\left(\bigcup\limits_{n=1}^{\infty}B_n\right)$, 其中 $\forall n\in\mathbb{N},\ B_n\in\mathcal{F}\Rightarrow\bigcup\limits_{n=1}^{\infty}B_n\in\mathcal{F}$, 说明 $\bigcup\limits_{n=1}^{\infty}A_n\in\Delta\cap\mathcal{F}$;

以上三条说明 $\Delta\cap\mathcal{F}$ 是 Δ 上的 σ 代数;

然后说明 μ_Δ 是 $\Delta\cap\mathcal{F}$ 上的测度, 只需证明其 σ 可加性, 即对 $\{A_n\in\Delta\cap\mathcal{F}:n\in\mathbb{N}$ 且 A_n 两两不交$\}$ 有 $\mu_\Delta\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=\sum\limits_{n=1}^{\infty}\mu_\Delta(A_n)$. 设 $\forall n\in\mathbb{N},\ \exists B_n\in\mathcal{F},\ s.t.\ A_n=\Delta\cap B_n$, 而 $\mu_\Delta\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=\mu^*\left(\Delta\cap\left(\bigcup\limits_{n=1}^{\infty}A_n\right)\right)=\mu^*\left(\Delta\cap\left(\bigcup\limits_{n=1}^{\infty}B_n\right)\right)$, 由外测度的定义知

$$\forall\varepsilon>0,\ \exists\{D_i\in\mathcal{F}:\ i\in\mathbb{N}\},\ s.t.\ \Delta\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\subset\bigcup_{i=1}^{\infty}D_i\ \text{且}\ \mu^*\left(\Delta\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\right)+\varepsilon>\mu\left(\bigcup_{i=1}^{\infty}D_i\right)$$

不妨设 D_i 两两不交, $i=1,2,\cdots$（由习题 1.1.9知可以做到这一点）, 则

$$\begin{aligned}\mu_{\Delta}\left(\bigcup_{n=1}^{\infty}A_n\right)+\varepsilon=&\mu^*\left(\Delta\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\right)+\varepsilon>\mu\left(\bigcup_{i=1}^{\infty}D_i\right)\geqslant\mu\left(\left(\bigcup_{i=1}^{\infty}D_i\right)\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\right)\\ =&\mu\left(\bigcup_{i=1}^{\infty}\left(D_i\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\right)\right)\xlongequal{D_i\text{两两不交}}\sum_{i=1}^{\infty}\mu\left(D_i\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\right)\\ =&\sum_{i=1}^{\infty}\mu\left(\bigcup_{n=1}^{\infty}(D_i\cap B_n)\right)\xlongequal{\text{由}A_n\text{两两不交知}B_n\text{两两不交}}\sum_{i=1}^{\infty}\sum_{n=1}^{\infty}\mu\left(D_i\cap B_n\right)\\ =&\sum_{n=1}^{\infty}\sum_{i=1}^{\infty}\mu\left(D_i\cap B_n\right)\overset{\text{次}\sigma\text{-可加性（定义 3.3.1(7)）}}{\geqslant}\sum_{n=1}^{\infty}\mu\left(\left(\bigcup_{i=1}^{\infty}D_i\right)\cap B_n\right)\\ \overset{\Delta\cap\left(\bigcup_{n=1}^{\infty}B_n\right)\subset\bigcup_{i=1}^{\infty}D_i}{\geqslant}&\sum_{n=1}^{\infty}\mu^*\left(\Delta\cap B_n\right)=\sum_{n=1}^{\infty}\mu_{\Delta}\left(A_n\right)\end{aligned}$$

由 ε 的任意性, 令 $\varepsilon\to 0^+$, 可得 $\mu_{\Delta}\left(\bigcup_{n=1}^{\infty}A_n\right)\geqslant\sum_{n=1}^{\infty}\mu_{\Delta}\left(A_n\right)$;

而 μ_{Δ} 是用外测度定义的, 由外测度的次 σ-可加性易证 $\mu_{\Delta}\left(\bigcup_{n=1}^{\infty}A_n\right)\leqslant\sum_{n=1}^{\infty}\mu_{\Delta}\left(A_n\right)$, 故 $\mu_{\Delta}\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mu_{\Delta}\left(A_n\right)$, 说明 μ_{Δ} 具有 σ-可加性, 是 $\Delta\cap\mathcal{F}$ 上的测度. 这说明 $(\Delta,\Delta\cap\mathcal{F},\mu_{\Delta})$ 是一个以 Δ 为空间的测度空间.

最后, 当 $0<\mu^*\left(\Delta\right)<\infty$ 时, 我们将证明 $\left(\Delta,\Delta\cap\mathcal{F},\dfrac{\mu_{\Delta}}{\mu^*\left(\Delta\right)}\right)$ 是一个以 Δ 为样本空间的概率空间:

(i) 已经证明 $\Delta\cap\mathcal{F}$ 是 Δ 上的 σ-代数;

(ii) 我们知道 μ_{Δ} 具有 σ-可加性, 因此 $\dfrac{\mu_{\Delta}}{\mu^*\left(\Delta\right)}$ 具有 σ-可加性, 故其是一个 $\Delta\cap\mathcal{F}$ 上的测度;

(iii) 由 μ_{Δ} 的定义知 $\mu_{\Delta}\left(\Delta\right)=\mu^*\left(\Delta\cap\Delta\right)=\mu^*\left(\Delta\right)$, 即 $\dfrac{\mu_{\Delta}\left(\Delta\right)}{\mu^*\left(\Delta\right)}=1$, 说明 $\dfrac{\mu_{\Delta}}{\mu^*\left(\Delta\right)}$ 是一个概率测度.

□

3.2.22 设 (E,ρ) 为一完全可分距离空间, $K_n\subset E,n\in\mathbb{N}$ 为 E 的一个上升的紧集列, $\{S_n:n\in\mathbb{N}\}$ 是作为 E 的可数拓扑基的开球列, $\mathcal{F}$ 为 $\bar{S}_m\cap K_n,m,n\in\mathbb{N}$ 的一切有限并作成的集类. $\nu:\mathcal{F}\longmapsto\mathbb{R}_+$, 具有性质:

(i) 若 $F_i\in\mathcal{F},i=1,2,F_1\subset F_2$, 则 $\nu\left(F_1\right)\leqslant\nu\left(F_2\right)$;

(ii) 若 $F_1,F_2\in\mathcal{F}$, 则 $\nu\left(F_1\cup F_2\right)\leqslant\nu\left(F_1\right)+\nu\left(F_2\right)$;

(iii) 若 $F_1,F_2\in\mathcal{F},F_1\cap F_2=\varnothing$, 则 $\nu\left(F_1\cup F_2\right)=\nu\left(F_1\right)+\nu\left(F_2\right)$.

对于 E 的开集 O 及任何 $A\subset E$, 定义

$$\lambda(O):=\sup\{\nu(F):F\subset O,F\in\mathcal{F}\},$$
$$\lambda^*(A):=\inf\{\lambda(G):A\subset G,G\text{ 为开集 }|,$$

则 λ^* 为 E 上的一个外测度.

证明:

(i) 由 ν 的性质, 可以得到 $\nu(F\cup\varnothing)=\nu(F)+\nu(\varnothing)$, 故 $\nu(\varnothing)=0$, 从而有 $\lambda(\varnothing)=0$, $\lambda^*(\varnothing)=0$;

(ii) 设 O_1,O_2 为 E 中的开集, 且 $O_1\subset O_2$, 则 $\forall F\in\mathcal{F}$, $F\subset O_1\subset O_2$ 有 $\lambda(O_1)\leqslant\lambda(O_2)$, 因此 λ 具有单调性. 设 $A\subset B\subset E$, 则 B 的开覆盖自然是 A 的开覆盖. 又由 λ 在开集上单调, 因此 $\lambda^*(A)\leqslant\lambda^*(B)$;

(iii) 设 $A_N\subset E$, 我们知道 $\exists G_n$ 为 A_n 的开覆盖, 使得

$$\lambda(G_n)-\frac{\varepsilon}{2^n}<\lambda^*(A_n)<\lambda(G_n),\forall\varepsilon>0.$$

则 $\bigcup\limits_{n=1}^{\infty}G_n$ 为 $\bigcup\limits_{n=1}^{\infty}A_n$ 的开覆盖. 我们知道 $\forall m$, $\forall\varepsilon>0$, $\exists F_n\in\mathcal{F}$, s.t. $F_n\subset G_n$, $\nu(F_n)>\lambda(G_n)-\dfrac{\varepsilon}{2^n}$. 故

$$\bigcup_{n=1}^{m}F_n\subset\bigcup_{n=1}^{m}G_n,\quad\sum_{n=1}^{m}\nu(F_n)\geqslant\nu\left(\bigcup_{n=1}^{m}F_n\right)>\lambda\left(\bigcup_{n=1}^{m}G_n\right)-\varepsilon.$$

因此 $\forall m$ 有

$$\lambda^*\left(\bigcup_{n=1}^{m}A_n\right)\leqslant\lambda\left(\bigcup_{n=1}^{m}G_n\right)\leqslant\sum_{n=1}^{m}\lambda(G_n)\leqslant\sum_{n=1}^{m}\lambda^*(A_n)+\varepsilon\leqslant\sum_{n=1}^{\infty}\lambda^*(A_n)+\varepsilon.$$

令 $m\to\infty$, $\varepsilon\to\infty$ 即可.

因此 λ^* 为 E 上的外测度. □

§ 3.3 测度空间的一些性质

3.3.1 设 $(\Omega,\mathcal{F},\mathbb{P})$ 是概率空间, 若 $\{A_n:n\in\mathbb{N}\}\subset\mathcal{F}$ 且 $\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$, 则 $\mathbb{P}(\{A_n\ \text{i.o.}\})=0$, 其中 $\{A_n\ \text{i.o.}\}$ 表示有无穷个 A_n 发生的事件.

提示: 可证 $\{A_n\ \text{i.o.}\}=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_n$. 此结论是著名的 Borel-Cantelli 引理.

证明: 首先可证 $\{A_n\ \text{i.o.}\}=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k$. 事实上, 首先有

$\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k=\{\forall n\in\mathbb{N},\ \exists k\geqslant n,\ s.t.\ A_k\ \text{发生}\}\subset\{A_n\ \text{i.o.}\}$; 又设 $\{A_{n_i}:\ i\in\mathbb{N}\}\subset\{A_n:\ n\in\mathbb{N}\}$ 为某次有无穷个 A_n 发生时 $\{A_n:\ n\in\mathbb{N}\}$ 中发生了的事件, 其中 A_{n_i} 表示 $\{A_{n_i}:\ i\in\mathbb{N}\}$ 中第 i 个发生的事件, 则由 $n_i\geqslant i$ 知 $A_{n_i}\subset\bigcup\limits_{k=i}^{\infty}A_k$; 而这次无穷多个事件的发生可表示为 $\bigcap\limits_{i=1}^{\infty}A_{n_i}$, 且 $\bigcap\limits_{i=1}^{\infty}A_{n_i}\subset\bigcap\limits_{i=1}^{\infty}\bigcup\limits_{k=i}^{\infty}A_k=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k$, 由此可知 $\{A_n\ \text{i.o.}\}\subset\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k$, 从而 $\{A_n\ i.o.\}=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k$.

其次可证 $\mathbb{P}(\{A_n\ \text{i.o.}\})=\mathbb{P}\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k\right)=0$. 由测度的上连续性知 $\mathbb{P}\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k\right)=\lim\limits_{n\to\infty}\mathbb{P}\left(\bigcup\limits_{k=n}^{\infty}A_k\right)$, 而 $\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$, 则 $0\leqslant\lim\limits_{n\to\infty}\mathbb{P}\left(\bigcup\limits_{k=n}^{\infty}A_k\right)\leqslant\lim\limits_{n\to\infty}\sum\limits_{k=n}^{\infty}\mathbb{P}(A_k)=0$, 即 $\mathbb{P}(\{A_n\ i.o.\})=\mathbb{P}\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k\right)=$

$\lim\limits_{n\to\infty}\mathbb{P}\left(\bigcup\limits_{k=n}^{\infty}A_k\right)=0.$ □

3.3.2 设 $\mathscr{A}$ 为 Ω 上的集代数.

(1) 设 μ_1,μ_2 是 $\sigma(\mathscr{A})$ 上的测度且在 $\mathscr{A}$ 上 σ 有限, 若 $A\in\sigma(\mathscr{A}),\mu_i(A)<\infty,i=1,2$, 则 $\forall\varepsilon>0,\exists A_\varepsilon\in\mathscr{A}$ 使 $\mu_i\left(A\Delta A_\varepsilon\right)<\varepsilon,i=1,2$.

(2) 试将 (1) 推广至可数个测度的情形.

证明:

(1) 设 $\mu=\mu_1+\mu_2$, 即 $\forall A\in\sigma\left(\mathscr{A}\right)$, $\mu\left(A\right):=\mu_1\left(A\right)+\mu_2\left(A\right)$, 则由命题 3.3.7 知 $\forall\varepsilon>0$, $\exists A_\varepsilon\in\mathscr{A}$, $s.t.\ \mu\left(A\Delta A_\varepsilon\right)<\varepsilon$, 从而 $\mu_i\left(A\Delta A_\varepsilon\right)\leqslant\mu\left(A\Delta A_\varepsilon\right)<\varepsilon$, $i=1,2$.

(2) 设 $\{\mu_n:n\in\mathbb{N}\}$ 是 $\sigma(\mathscr{A})$ 上的测度列且都在 $\mathscr{A}$ 上 σ 有限. 只需取 $\mu=\sum\limits_{k=1}^{\infty}\dfrac{\mu_k}{2^k}$, 则 μ 在 A 上是 σ 有限的. $\forall A\in\sigma(A)$, $\forall j\in\{k:\mu_k(A)\geqslant\mu_n(A),\forall n\in\mathbb{N}\}$, 我们有 $\mu(A)\leqslant\mu_j(A)$, 因此 $\forall\varepsilon>0$, $\forall j\in\{k:\mu_k(A)\geqslant\mu_n(A),\forall n\in\mathbb{N}\}$, $\exists A_\varepsilon\in\mathscr{A}$, s.t. $\mu_j(A\Delta A_\varepsilon)<\varepsilon$, 故 $\mu_k(A\Delta A_\varepsilon)<\varepsilon$, $\forall k\in\mathbb{N}$.

□

3.3.3 证明 $A\subset\mathbb{R}^n$ Lebesgue 可测的充要条件是对任何 $I\subset\mathbb{R}^n$ 开区间, 有

$$\lambda^*(I)=\lambda^*(A\cap I)+\lambda^*\left(A^c\cap I\right).$$

证明: 必要性（“$\Rightarrow$”）显然; 下证充分性（“$\Leftarrow$”）. 由引理 3.2.15 知只需证明 $\forall D\subset\Omega$, $\lambda^*\left(D\right)\geqslant\lambda^*\left(A\cap D\right)+\lambda^*\left(A^c\cap D\right)$. 首先由 Lebesgue 外测度的定义知

$$\forall D\subset\Omega,\ \forall\varepsilon>0,\ \exists\text{开区间列}\left\{I_j:\ j\in\mathbb{N}\right\},\ s.t.\ D\subset\bigcup_{j=1}^{\infty}I_j\ \text{且}\ \lambda^*\left(D\right)+\varepsilon>\sum_{j=1}^{\infty}\lambda\left(I_j\right)$$

从而

$$\begin{aligned}\lambda^*\left(D\right)+\varepsilon>\sum_{j=1}^{\infty}\lambda\left(I_j\right)&\overset{\text{题设}}{=\!=\!=}\sum_{j=1}^{\infty}\left(\lambda^*\left(A\cap I_j\right)+\lambda^*\left(A^c\cap I_j\right)\right)\\&\overset{\text{外测度的次}\sigma\text{-可加性}}{\geqslant}\lambda^*\left(\bigcup_{j=1}^{\infty}\left(A\cap I_j\right)\right)+\lambda^*\left(\bigcup_{j=1}^{\infty}\left(A^c\cap I_j\right)\right)\\&=\lambda^*\left(A\cap\left(\bigcup_{j=1}^{\infty}I_j\right)\right)+\lambda^*\left(A^c\cap\left(\bigcup_{j=1}^{\infty}I_j\right)\right)\\&\overset{D\subset\bigcup_{j=1}^{\infty}I_j}{\geqslant}\lambda^*\left(A\cap D\right)+\lambda^*\left(A^c\cap D\right)\end{aligned}$$

由 ε 的任意性, 令 $\varepsilon\to0^+$, 即得 $\forall D\subset\Omega$, $\lambda^*\left(D\right)\geqslant\lambda^*\left(A\cap D\right)+\lambda^*\left(A^c\cap D\right)$, 由引理 3.2.15 可知 A 是 Lebesgue 可测的. □

3.3.4 若 I 为 $\mathbb{R}^n$ 的一个有界开区间, 试证: $B\subset I$ 为 Lebesgue 可测集的充要条件是

$$\lambda^*(I)=\lambda^*(B)+\lambda^*\left(I\cap B^c\right).$$

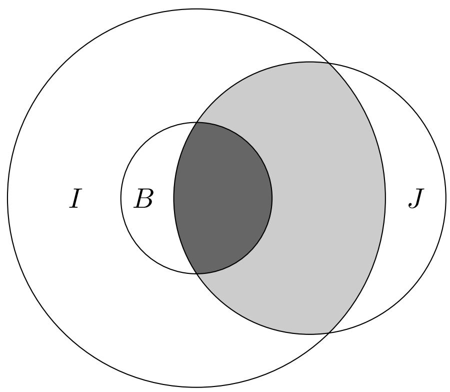
如果对于闭集 F 令 $\lambda(F)$ 为体积测度, 定义 $\lambda_*(B)=\sup\{\lambda(F):F\subset B, F\text{为闭集}\}$ (称为 B 的内测度), 试证: 上述条件等价于 $\lambda^*(B)=\lambda_*(B)$.

证明: 在正式开始证明之前, 需注意 I,F 均是 Lebesgue 可测集, 因为由定理 2.2.11, 所有的开集均 Lebesgue 可测, 而由闭集的定义 (开集的余集, 见定义 2.2.1) 及定理 3.2.16(1)(可测集的全体构成 σ 代数) 知所有的闭集也 Lebesgue 可测. 同时, 因为题目涉及 "有界" 的概念, 所以把 $\mathbb{R}^n$ 当作度量空间看待.

首先证明第一个条件 $(\lambda^*(I)=\lambda^*(B)+\lambda^*(I\cap B^c))$ 的充要性.

必要性 ($\Rightarrow$): 根据可测集的定义显然.

充分性 ($\Leftarrow$): 根据习题 3.3.3, 只需证明对于任意开区间 J, 有 $\lambda^*(J)=\lambda^*(J\cap B)+\lambda^*(J\cap B^c)$. 根据外测度的次可加性 (定义 3.2.11(3)), 只需证明 $\lambda^*(J)\geqslant\lambda^*(J\cap B)+\lambda^*(J\cap B^c)$.



如图所示, B 和 J 将 I "分割" 成四个区域, 但由于我们还没有证明 B 可测, 我们无法直接将 I 的 "面积" (即 Lebesgue 外测度) 拆分成这四个子区域的 "面积" 之和. 不过, 由于 $\lambda^*(I)=\lambda^*(B)+\lambda^*(I\cap B^c)$, 利用 I,J 的可测性, 我们可以做到这一点, 即证明

$$\lambda^*(I)=\lambda^*(B\cap J)+\lambda^*(B\cap J^c)+\lambda^*(I\cap B^c\cap J)+\lambda^*(I\cap B^c\cap J^c). \tag{*}$$

事实上, 由于 I,J 均可测, 由定理 3.2.16(1) 知 $I\cap J$ 也可测. 利用可测集的定义, 取 B 替代定义 3.2.14(又称 Carathéodory 条件) 中的 D(又称试验集), $I\cap J$ 替代定义 3.2.14 中的 A, 可得

$$\lambda^*(B)=\lambda^*(B\cap J)+\lambda^*(B\cap J^c);$$

再取 $I\cap B^c$ 替代定义 3.2.14 中的 D, $I\cap J$ 替代定义 3.2.14 中的 A, 可得

$$\lambda^*(I\cap B^c)=\lambda^*(I\cap B^c\cap J))+\lambda^*(I\cap B^c\cap J^c);$$

又因为 $\lambda^*(I)=\lambda^*(B)+\lambda^*(I\cap B^c)$, 所以 (*) 成立. 于是

$$\begin{aligned}
\lambda^*(J\cap B)+\lambda^*(J\cap B^c) &\leqslant \lambda^*(J\cap B)+\lambda^*(I\cap B^c\cap J)+\lambda^*(J\cap I^c)\ (\text{外测度的次可加性})\\
((*)\ \text{式})\ &= \lambda^*(I)-\lambda^*(B\cap J^c)-\lambda^*(I\cap B^c\cap J^c)+\lambda^*(J\cap I^c)\\
(\text{外测度的次可加性})\ &\leqslant \lambda^*(I)-\lambda^*(I\cap J^c)+\lambda^*(J\cap I^c)\\
&= \lambda^*(I\cup J)-\lambda^*(I\cap J^c)=\lambda^*(J),
\end{aligned}$$

证毕. (注意: 因为 I 有界, 所以以上各式中带减号的各项 $<\infty$, 从而不会出现 $\infty-\infty$ 的情形, 可以相减.)

然后证明第二个条件 $\lambda^*(B)=\lambda_*(B)$ 的充要性.

必要性 (⇒): 由 Lebesgue 测度的定义 (见推论 3.2.18) 知当 B 是 Lebesgue 可测集时 $\lambda(B)=\lambda^*(B)$. 由题目所给的内测度的定义及测度的不降性 (定义 3.3.1(6)) 知 $\lambda_*(B)\leqslant\lambda(B)$; 又由命题 3.3.8 及引理 2.4.8(2)(度量空间下, 紧集是闭集) 知当 B 是 Lebesgue 可测集时有 $\lambda(B)=\sup\{\lambda(F):\ B\supset F,\ F\text{是紧集}\}\leqslant\sup\{\lambda(F):\ B\supset F,\ F\text{是闭集}\}=\lambda_*(B)$, 因此 $\lambda(B)=\lambda_*(B)$, 从而 $\lambda^*(B)=\lambda_*(B)$.

充分性 (⇐): 只需证明 $\lambda^*(I)=\lambda^*(B)+\lambda^*(I\cap B^c)$; 由外测度的次可加性 (定义 3.2.11(3)), 只需证明 $\lambda^*(I)\geqslant\lambda^*(B)+\lambda^*(I\cap B^c)$. 而由于 $\lambda^*(B)=\lambda_*(B)=\sup\{\lambda(F):\ B\supset F,\ F\text{是闭集}\}$, 故 $\forall\varepsilon>0$, $\exists$闭集$F\subset B$, s.t.$\lambda(F)\geqslant\lambda^*(B)-\varepsilon$, 也即 $\lambda^*(B)\leqslant\lambda(F)+\varepsilon$. 另一方面, $F\subset B$ 意味着 $I\cap B^c\subset I\cap F^c$, 从而由外测度的不降性 (定义 3.2.11(2)) 知 $\lambda^*(I\cap B^c)\leqslant\lambda(I\cap F^c)$. 因此, $\lambda^*(B)+\lambda^*(I\cap B^c)\leqslant\lambda(F)+\lambda(I\cap F^c)+\varepsilon=\lambda(I)+\varepsilon$, 由 ε 的任意性知 $\lambda^*(B)+\lambda^*(I\cap B^c)\leqslant\lambda(I)$, 证毕. □

3.3.5 证明对任一集 $E\subset\mathbb{R}$, 有 G_δ 型集 (可表示为可数个开集的交) A, 使 $\lambda(A)=\lambda^*(E)$.

证明: 由外测度的定义知 $\forall\varepsilon>0$, $\exists$开集$G=\bigcup\limits_{j=1}^{\infty}I_j$为开区间之并, $s.t.$ $E\subset G$ 且 $\lambda^*(E)+\varepsilon>\lambda(G)$ 即 $\lambda^*(E)>\lambda(G)-\varepsilon$, 于是

$$
\begin{gathered}
\text{取}\varepsilon=1,\ \text{则}\exists\text{开集}G_1,\ \text{s.t. } E\subset G_1\text{ 且 }\lambda^*(E)>\lambda(G_1)-1\\
\text{取}\varepsilon=\frac{1}{2},\ \text{则}\exists\text{开集}G_2,\ \text{s.t. } E\subset G_2\text{ 且 }\lambda^*(E)>\lambda(G_2)-\frac{1}{2}\\
\text{取}\varepsilon=\frac{1}{3},\ \text{则}\exists\text{开集}G_3,\ \text{s.t. } E\subset G_3\text{ 且 }\lambda^*(E)>\lambda(G_3)-\frac{1}{3}\\
\vdots\\
\text{取}\varepsilon=\frac{1}{n},\ \text{则}\exists\text{开集}G_n,\ \text{s.t. } E\subset G_n\text{ 且 }\lambda^*(E)>\lambda(G_n)-\frac{1}{n}\\
\vdots
\end{gathered}
$$

如此一直进行下去, 可取得一个开集列 $\{G_n:\ n\in\mathbb{N}\}$, 使得对 $\bigcap\limits_{n=1}^{\infty}G_n$ 有 $\forall n\in\mathbb{N}$, $\lambda^*(E)>\lambda(G_n)-\frac{1}{n}\geqslant\lambda\left(\bigcap\limits_{n=1}^{\infty}G_n\right)-\frac{1}{n}$; 且由 $\forall n\in\mathbb{N}$, $E\subset G_n$ 可知 $E\subset\bigcap\limits_{n=1}^{\infty}G_n$, 从而 $\forall n\in\mathbb{N}$, $\lambda\left(\bigcap\limits_{n=1}^{\infty}G_n\right)\geqslant\lambda^*(E)>\lambda\left(\bigcap\limits_{n=1}^{\infty}G_n\right)-\frac{1}{n}$, 由 n 的任意性, 令 $n\to\infty$, 可得 $\lambda\left(\bigcap\limits_{n=1}^{\infty}G_n\right)=\lambda^*(E)$, 取 $A=\bigcap\limits_{n=1}^{\infty}G_n$ 即得证. □

3.3.6 $E\subset\mathbb{R}$ 是 Lebesgue 可测集 $\Leftrightarrow E=A\backslash H$, 其中 A 为 G_δ 型集, $\lambda(H)=0$.

证明: 由习题 3.3.5, $\forall E\subset\mathbb{R}$, 有 G_δ 型集 A, s.t. $\lambda(A)=\lambda^*(E)$. 当 E 为 Lebesgue 可测时, 不妨设 E 有界 (否则可用可数个区间段分割), 则 $\lambda(A\backslash E)=0$. 反之, 若 $\lambda(A\backslash E)=0$, 则 $A\backslash E$ 可测, $E=A\backslash(A\backslash E)$ 也可测. □

3.3.7 $E\subset\mathbb{R}$ 是 Lebesgue 可测集 $\Leftrightarrow E=B\cup H$, 其中 B 为 F_σ 型集 (可表示为可数个闭集的并), 而 $\lambda(H)=0$.

证明: ⇒: 先构造 F_σ 型集, 即闭集列 $\{F_n\}$ 使得 $F_n\subset\mathbb{R}^n$, $B=\bigcup\limits_{n=1}^{\infty}F_n\subset E$, 且 $\lambda^*(E\backslash B)=0$. 由于 E 可测, 且

$$\lambda_*(B)=\sup\{\lambda(F):F\subset B,F\text{为闭集}\}=\lambda(B).$$

由上确界的性质可知, 对于 $\varepsilon_n = \dfrac{1}{n} > 0$, 存在闭集列 F_n 使得

$$\lambda(E \backslash F_n) = \lambda(E) - \lambda(F_n) < \frac{1}{n},$$

不妨设闭集列 $\{F_n\}$ 是递增的, 故 $\{E\backslash F_n\}$ 是递降的, 从而根据测度得上连续性可得

$$\lambda(\lim_{n\to\infty} E\backslash F_n) = \lim_{n\to\infty} \lambda(E\backslash F_n) \leqslant \lim_{n\to\infty} \frac{1}{n} = 0.$$

令 $B = \bigcup\limits_{n=1}^{\infty} F_n$, 则 B 为 F_σ 型集, 且 $B \subset E$. 并且由上式 $\lambda(E\backslash B) = \lambda(\lim\limits_{n\to\infty} E\backslash F_n) = 0$, 也即 $\lambda(E\backslash B) = 0$.

$\Leftarrow$: 设 $\lambda^*(E\backslash B) = 0$, 其中 B 为 F_σ 型集, 则 B 可测. 又

$$0 \leqslant \lambda_*(E\backslash B) \leqslant \lambda^*(E\backslash B) = 0,$$

因此 $\lambda_*(E\backslash B) = \lambda^*(E\backslash B)$. 由习题 3.3.4知 $E\backslash B$ 可测, 故 $B \cup (E\backslash B)$ 可测. □

3.3.8 设 (E, d) 为距离空间, 定义:

(i) 对一切 E 的子集 A, B, 称 $d(A,B) := \inf\{d(x,y) : x \in A, y \in B\}$, 为集合 A, B 的距离, 如果 $d(A,B) > 0$, 称 A, B 是正分离的;

(ii) 对任何 E 的子集 A, 称 $\delta(A) := \sup\{d(x,y) : x, y \in A\}$ 为集合 A 的直径;

(iii) 如果 μ 是 E 上的外测度, 对任何正分离的集合 A, B 有 $\mu(A\cup B) = \mu(A) + \mu(B)$, 则称 μ 为距离外测度;

(iv) 若对 E 的任何子集 $A, s > 0, \varepsilon > 0$, 令

$$\mathscr{H}_\varepsilon^s(A) := \inf\left\{\sum_{i=1}^{\infty} \delta^s(U_i) : A \subset \bigcup_{i=1}^{\infty} U_i, 0 < \delta(U_i) \leqslant \varepsilon\right\},$$
$$\mathscr{H}^s(A) := \lim_{\varepsilon\to 0} \mathscr{H}_s^\infty(A) = \sup_{\varepsilon>0} \mathscr{H}_s^\infty(A).$$

称 $\mathscr{H}^s(A)$ 为集合 A 的 Hausdorff 测度.

试证:

(1)(引理) 若 μ 为 (E,d) 上的距离外测度, $\{A_n\} \subset E$ 为不降集列, $A = \lim\limits_{n\to\infty} A_n$, 且 $\forall n \in \mathbb{N}$, $d(A_n, A\backslash A_{n+1}) > 0$, 则 $\mu(A) = \lim\limits_{n\to\infty} \mu(A_n)$;

(2) (定理) 若 μ 为 (E,d) 上的距离外测度, 则 E 的一切 Borel 子集是 μ 可测集;

(3)(定理) $\mathscr{H}^s$ 是 (E,d) 上的距离外测度.

证明: (1) 我们知道 $\forall k$, 有 $\mu(A_k) \leqslant \mu(A_{k+1}) \leqslant \mu(A)$, 因此 $\lim\limits_{n\to\infty} \mu(A_n) \leqslant \mu(A)$.

若 $\lim\limits_{n\to\infty} \mu(A_n) = \infty$, 则 $\mu(A) = \infty$, 此时结论成立.

下设 $\lim\limits_{n\to\infty} \mu(A_n) < \infty$, 我们有 $\mu(A_n) \leqslant \lim\limits_{n\to\infty} \mu(A_n) < \infty$. 令 $B_n = A_n \backslash A_{n-1}$, 则 $\{B_n : n \in \mathbb{N}\}$ 两两不交, 且 $A = A_n \cup \left(\bigcup\limits_{j=n+1}^{\infty} B_j\right)$. 故 $\forall n$ 有 $\mu(A) \leqslant \mu(A_n) + \sum\limits_{j=n+1}^{\infty} \mu(B_j)$.

下面证明 $\sum\limits_{j=1}^{\infty} \mu(B_j) < \infty$. 注意到 $\forall |i-j| \geqslant 2$ 且 $i, j \in \mathbb{N}$, 有

$$d(B_i, B_j) = d(A_i\backslash A_{i-1}, A_j\backslash A_{j-1}) \geqslant d(A_i, A\backslash A_{i+1}) > 0,$$

因此 $d(B_{2i-1}, B_{2j-1}) > 0$, $d(B_{2i}, B_{2j}) > 0$, 当 $i \neq j$. 故

$$\sum_{k=1}^{2n} \mu(B_k) = \sum_{k=1}^{n} \mu(B_{2k}) + \sum_{k=1}^{n} \mu(B_{2k-1}) \leqslant \mu(A_{2n}) + \mu(A_{2n}) < \infty.$$

因此 $\sum\limits_{j=1}^{\infty} \mu(B_j) < \infty$, 这蕴含着 $\lim\limits_{n\to\infty} \sum\limits_{j=n+1}^{\infty} \mu(B_j) = 0$, 这样

$$\mu(A) \leqslant \lim_{n\to\infty} A_n + \lim_{n\to\infty} \mu(B_j) = \lim_{n\to\infty} \mu(A_n).$$

因此 $\mu(A) = \lim\limits_{n\to\infty} \mu(A_n)$.

(2) 先证明任意 E 中的闭集都是 μ 可测的. 设 $A \subset E$, 且 $\mu(A) < \infty$, 记

$$A_n := \left\{ x \in A \cap F^c, d(x, F) > \frac{1}{n} \right\}, n \in \mathbb{N},$$

则 $\{A_n : n \in \mathbb{N}\}$ 是不降集列, 并且 $d(F \cap A, A_n) \geqslant \dfrac{1}{n}$. 再根据外测度的单调性, 有

$$\mu(A) \geqslant \mu((F \cap A) \cup A_n) = \mu(F \cap A) + \mu(A_n).$$

注意到 $\bigcup\limits_{n=1}^{\infty} A_n = A \cap F^c$, 由 (1), $\lim\limits_{n\to\infty} \mu(A_n) = \mu(A \cap F^c)$, 因此

$$\mu(A) \geqslant \mu(F \cap A) + \mu(F^c \cap A).$$

故任意闭集 $F \subset E$ 均为 μ 可测集, 其余集 F^c 是可测开集. 因此任意开集都是 μ 可测的, 故 E 的一切 Borel 子集都是 μ 可测集.

(3) 先证明 $\mathscr{H}^s$ 是外测度. 显然有 $\mathscr{H}^s(\varnothing) = 0$, 考虑 $A_1 \subset A_2 \subset E$, 则 A_2 的覆盖也是 A_1 的覆盖, 则 $\mathscr{H}^s(A_1) \leqslant \mathscr{H}^s(A_2)$. 再设 $\{A_n : n \in \mathbb{N}\}$ 是 E 的子集列, 则 $\forall \varepsilon > 0$, $\forall n$, 存在直径小于 ε 的覆盖 $\{U_{nk}\}$, 使得 $\sum\limits_{k=1}^{\infty} \delta^s(U_{nk}) \leqslant \mathscr{H}^s_\varepsilon(A_n) + \dfrac{\eta}{2^n}$, $\forall \eta > 0$. 再注意到 $\bigcup\limits_{n=1}^{\infty} \bigcup\limits_{k=1}^{\infty} U_{nk}$ 是 $\bigcup\limits_{n=1}^{\infty} A_n$ 的一个直径不超过 ε 的覆盖, 故

$$\mathscr{H}^s_\varepsilon \left(\bigcup_{n=1}^{\infty} A_n \right) \leqslant \sum_{n=1}^{\infty} \mathscr{H}^s_\varepsilon(A_n) + \eta \leqslant \sum_{n=1}^{\infty} \mathscr{H}^s(A_n) + \eta.$$

我们知道 η 是任意的, 因此 $\mathscr{H}^s \left(\bigcup\limits_{n=1}^{\infty} A_n \right) \leqslant \sum\limits_{n=1}^{\infty} \mathscr{H}^s(A_n)$. 令 $\varepsilon \to 0$ 便有 $\mathscr{H}^s \left(\bigcup\limits_{n=1}^{\infty} A_n \right) \leqslant \sum\limits_{n=1}^{\infty} \mathscr{H}^s(A_n)$. 故 $\mathscr{H}^s$ 是 (E, d) 上的外测度.

下面证明它是距离外测度. 设 $A_1, A_2 \subset E$, $d(A_1, A_2) > 0$, 根据其次可加性我们只需证明 $\mathscr{H}^s(A_1) + \mathscr{H}^s(A_2) \leqslant \mathscr{H}^s(A_1 \cup A_2)$. 不妨设 $\{U_n\}$ 是 $A_1 \cup A_2$ 的一个直径不超过 ε 的覆盖, 则 $\{A_1 \cap U_n : n \in \mathbb{N}\}$ 和 $\{A_2 \cap U_n : n \in \mathbb{N}\}$ 分别是 A_1, A_2 的直径不超过 ε 的覆盖. 我们有

$$\sum_{n=1}^{\infty} \delta^s(A_1 \cap U_n) + \sum_{n=1}^{\infty} \delta^s(A_2 \cap U_n) \leqslant \sum_{n=1}^{\infty} \delta^s(U_n).$$

令 $\varepsilon \to 0$ 即得 $\mathscr{H}^s(A_1) + \mathscr{H}^s(A_2) \leqslant \mathscr{H}^s(A_1 \cup A_2)$.

因此 $\mathscr{H}^s$ 是 (E, d) 上的距离外测度. □

第四章 可测函数与随机变量

§ 4.1 可测函数与分布

4.1.1 试证示性函数有下列性质:

(1) $\mathbb{1}_{A\cap B} = \mathbb{1}_A \cdot \mathbb{1}_B$, $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B}$, $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$, $\mathbb{1}_{A\Delta B} = |\mathbb{1}_A - \mathbb{1}_B|$;

(2)$\mathbb{1}_{\cup_n A_n} = \sup\limits_n \mathbb{1}_{A_n}$, $\mathbb{1}_{\cap_n A_n} = \inf\limits_n \mathbb{1}_{A_n}$, $\mathbb{1}_{\limsup_{n\to\infty} A_n} = \limsup\limits_{n\to\infty} \mathbb{1}_{A_n}$, $\mathbb{1}_{\liminf_{n\to\infty} A_n} = \liminf\limits_{n\to\infty} \mathbb{1}_{A_n}$.

证明: (1) 略.

(2) 我们知道 $x \in \bigcup\limits_n A_n \Leftrightarrow \exists k, x \in A_k$, $x \notin \bigcup\limits_n A_n \Leftrightarrow \forall k, x \notin A_k$, 故 $\mathbb{1}_{\cup_n A_n} = 1 \Leftrightarrow \exists k, \mathbb{1}_{A_k} = 1$, $\mathbb{1}_{\cup_n A_n} = 0 \Leftrightarrow \forall k, \mathbb{1}_{A_k} = 0$, 故 $\mathbb{1}_{\cup_n A_n} = \sup\limits_n \mathbb{1}_{A_n}$;

而 $\bigcap\limits_n A_n = \left(\bigcup\limits_n A_n^c\right)^c$, 故 $\mathbb{1}_{\cap_n A_n} = 1 - \mathbb{1}_{\cup_n A_n^c} = 1 - \sup\limits_n(1 - \mathbb{1}_{A_n}) = \inf\limits_n \mathbb{1}_{A_n}$;

我们知道 $\limsup\limits_{n\to\infty} A_n = \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_n$, 故 $\mathbb{1}_{\limsup_{n\to\infty} A_n} = \inf\limits_{n\in\mathbb{N}} \sup\limits_{k\geqslant n} \mathbb{1}_{A_k} = \limsup\limits_{n\to\infty} \mathbb{1}_{A_n}$;

而 $\liminf\limits_{n\to\infty} A_n = \left(\limsup\limits_{n\to\infty} A_n^c\right)^c$, 故 $\mathbb{1}_{\liminf_{n\to\infty} A_n} = 1 - \limsup\limits_{n\to\infty}(1 - \mathbb{1}_{A_n}) = \liminf\limits_{n\to\infty} \mathbb{1}_{A_n}$. □

4.1.2 设 f, g 是 $(\Omega, \mathcal{F})$ 上的 (广义) 实 (或复) 可测函数, 问下列函数是否 $\mathcal{F}$ 可测函数? 并说明理由.

(1) $f_1(\omega) := \begin{cases} f(\omega), & \omega \in \{|f| < \infty\}, \\ 0, & \omega \in \{|f| = \infty\}, \end{cases}$ 即 $f_1 = f \cdot \mathbb{1}_{\{|f|<\infty\}}$;

(2) $f_2 := f \cdot \mathbb{1}_{\{\omega\}^c} + (f+1)\mathbb{1}_{\{\omega\}}, \omega$ 为 Ω 的给定元;

(3) $f_3 := f \cdot \mathbb{1}_A + g \cdot \mathbb{1}_{A^c}, A \in \mathcal{F}$.

证明:

(1) 我们知道 $f_1(\mathcal{F}) \in \overline{\mathscr{B}}$, 考虑 $\forall B \in \overline{\mathscr{B}}$. 若 $0 \notin B$, 则有 $f_1^{-1}(B) = f^{-1}(B) \in \mathcal{F}$; 若 $0 \in B$, 则有 $f_1^{-1}(B) = f^{-1}(B) \cap f^{-1}(\{|f| = \infty\}) \in \mathcal{F}$. 故 f_1 是 $\mathcal{F}$ 可测的.

(2) 若 $\{\omega\} \notin \mathcal{F}$, 只需取 $B \in \mathscr{B}$, 使得 $f(\omega) + 1 \in B$, 则有 $\{\omega\} \subset f_2^{-1}(B) \notin \mathcal{F}$, 此时 f_2 并非 $\mathcal{F}$ 可测.

若 $\{\omega\} \in \mathcal{F}$, 我们知道 $f_2 = f(1 + \mathbb{1}_{\{\omega\}})$, 容易知道 $1 + \mathbb{1}_{\{\omega\}}$ 是 $\mathcal{F}$ 可测的, 此时 f_2 是 $\mathcal{F}$ 可测的.

(3) 我们知道 $\mathbb{1}_A$, $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$ 都是 $\mathcal{F}$ 可测的, 故 $f_3 = f \cdot \mathbb{1}_A + g \cdot \mathbb{1}_{A^c}$ 是 $\mathcal{F}$ 可测的.

□

4.1.3 f 是实 $(\Omega, \mathcal{F})$ 可测函数, 则 $|f|$ 是 $\mathcal{F}$ 可测的, 逆命题是否成立?

证明: 考虑 $A \notin \mathcal{F}$, 令 $f = 1 - 2\mathbb{1}_A$, 则 $|f| = 1$ 是 $\mathcal{F}$ 可测的, 但 f 本身不是 $\mathcal{F}$ 可测的. □

4.1.4 设 $\mathcal{F} := \{\varnothing, A, A^c, \Omega\}, A \subset \Omega$ 给定, 试写出 $(\Omega, \mathcal{F})$ 上的全部 $\mathcal{F}$ 可测函数.

证法一: 我们证明, f 在 A 上取常数, 在 A^c 上也取常数 (但这两个常数不一定相等).

实际上, 若 f 的值域 D 满足 $\overline{\overline{D}} \geqslant 3$, 取 $r_1, r_2, r_3 \in D$, 则 $f^{-1}(\{r_1\}), f^{-1}(\{r_2\}), f^{-1}(\{r_3\})$ 非空且两两不交. 这代表着 $\Omega = A \cup A^c = f^{-1}(\{r_1\}) \cup f^{-1}(\{r_2\}) \cup f^{-1}(\{r_3\})$, 则至少有一个 $i \in \{1,2,3\}$ 使得 $f^{-1}(\{r_i\}) \notin \mathcal{F}$. 因此 $f^{-1}(\{r_1, r_2, r_3\}) \notin \mathcal{F}$. 这时 f 不是 $\mathcal{F}$ 可测的.

因此 f 的值域至多是两个常数, 即可以写成 $f = c_1 \mathbb{1}_B + c_2 \mathbb{1}_{B^c}$ 的形式. 因为 f 是 $\mathcal{F}$ 可测的, 所以 $B = f^{-1}(\{c_1\}) \in \mathcal{F}$, 因此 $B = A$ 或 $B = \Omega$. □

证法二: 首先, 显然 $\mathcal{F}$ 可测的简单函数至多只能取两个值, 并且分别限定在 A, A^c 上都是常值函数. 而由定理 4.2.4, 非负可测函数可以写成简单函数的极限, 因此非负 $\mathcal{F}$ 可测函数也至多只能取两个值且分别限定在 A, A^c 上都是常值函数, 进而所有 $\mathcal{F}$ 可测函数都至多只能取两个值且分别限定在 A, A^c 上都是常值函数. □

4.1.5 如果两个随机变量几乎处处相等, 则它们具有相同的概率分布测度.

证明: 设 X, Y 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 到 $(\mathbb{R}, \mathscr{B})$ 的随机变量, 则 $\mathbb{P}(\{\omega : X(\omega) = Y(\omega)\}) = 1$. 则 $\forall A \in \mathscr{B}$, 有

$$\begin{aligned}
\mathbb{P}_X(A) &= \mathbb{P}(\{\omega : X(\omega) \in A\}) \\
&= \mathbb{P}(\{\omega : X(\omega) \in A\} \cap \{\omega : X(\omega) = Y(\omega)\}) + \mathbb{P}(\{\omega : X(\omega) \in A\} \cap \{\omega : X(\omega) \neq Y(\omega)\}) \\
&= \mathbb{P}(\{\omega : X(\omega) \in A\} \cap \{\omega : X(\omega) = Y(\omega)\}) \\
&= \mathbb{P}(\{\omega : Y(\omega) \in A\}) = \mathbb{P}_Y(A).
\end{aligned}$$

□

4.1.6 对于给定的 $(\mathbb{R}, \mathscr{B})$ 上的任何概率测度 μ, 定义一个概率分布测度为 μ 的随机变量.

证明: 考虑 $\xi \sim U(0,1)$, $\eta = \inf\{x : \mu((-\infty, x]) \geqslant \xi\} =: \mu^{-1}(\xi)$, 则 $\mathbb{P}(\eta \leqslant t) = \mu((-\infty, t])$.

这是因为 $\mu((-\infty, x]) = \mathbb{P}(\xi \leqslant \mu((-\infty, x])) = \mathbb{P}(\mu^{-1}(\xi) \leqslant \mu^{-1}(\mu((-\infty, x])))$. 我们知道 μ 有下连续性, 所以

$$\mu^{-1}(\mu((-\infty, x])) = \inf\{y : \mu((-\infty, y]) \geqslant \mu((-\infty, x])\} = x,$$

故 $\mathbb{P}(\mu^{-1}(\xi) \leqslant t) = \mathbb{P}(\eta \leqslant t) = \mu((-\infty, t])$. □

4.1.7 设 θ 为 $[0,1]$ 中均匀分布的随机变量, 对于每个概率分布函数 F, 定义 $G(y) = \sup\{x : F(x) < y\}$, 则 $G(\theta)$ 具有概率分布函数 F.

证明: 我们有: $F_{G(\theta)}(t) = \mathbb{P}(G(\theta) \leqslant t) \overset{(*)}{=\!=} \mathbb{P}(\theta \leqslant F(t)) = F(t)$,

其中 (*) 是因为可以证明 $G(y) \leqslant x_0 \Longleftrightarrow y \leqslant F(x_0)$.

事实上, (1) 若 $y \leqslant F(x_0)$, 则 $\forall x:\ F(x) < y$, $F(x_0) \geqslant y > F(x) \overset{F\text{具不降性}}{\Longrightarrow} x_0 \geqslant x$, 则 $x_0 \geqslant \sup\{x:\ F(x) < y\} = G(y)$;

(2) 若 $y > F(x_0)$, 则由 F 的右连续性知存在 $\delta > 0$ 使得 $y > F(x_0 + \delta)$, 从而

$$x_0 + \delta \in \{x:\ F(x) < y\} \Longrightarrow G(y) = \sup\{x:\ F(x) < y\} \geqslant x_0 + \delta > x_0.$$

综合 (1)(2) 即得证. (此证明原型是 [Du] 定理 1.2.2 的证明.) □

4.1.8 设 X 具有连续分布函数 F, 则 $F(X)$ 具有 $[0,1]$ 上的均匀分布, 如果 $F(x)$ 不连续, 情况又怎样?

证明: 对于不降函数 F，记 $G(y) := \sup\{x:\ F(x) \leqslant y\}$，则可证如下结论：

(1) 当 F 具有左连续性时，有 $F(x_0) \leqslant y \Longleftrightarrow x_0 \leqslant G(y)$

(2) 进一步，当 F 连续且是某随机变量的概率分布函数时，有 $F(G(y)) = y,\ \forall y \in (0,1)$

从而 $0 < y < 1$ 时 $\mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant y\}) = \mathbb{P}(\{\omega:\ X(\omega) \leqslant G(y)\}) = F(G(y)) = y$，即 $F(X)$ 服从 $[0,1]$ 上的均匀分布。

（注，$y = 0, 1$ 的情形：

$y = 1$ 时，由分布函数的定义显然有 $\mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant 1\}) = \mathbb{P}(\Omega) = 1$；

$y = 0$ 时，由于 $\forall \epsilon \in (0,1)$，$\mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant 0\}) \leqslant \mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant \epsilon\}) = \epsilon$，令 $\epsilon \to 0^+$ 即得 $\mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant 0\}) = 0$.）

现在证明上述结论：

(1) 证明与习题 4.1.7的证明基本相同：

(a) 当 $F(x_0) \leqslant y$ 时，有 $x_0 \in \{x:\ F(x) \leqslant y\}$，从而 $x_0 \leqslant \sup\{x:\ F(x) \leqslant y\} = G(y)$；

(b) 当 $F(x_0) > y$ 时，由 F 具有左连续性知 $\exists \delta > 0,\ s.t.\ F(x_0 - \delta) > y$，从而 $\forall x:\ F(x) \leqslant y$，$F(x_0 - \delta) > y \geqslant F(x)$，由 F 的不降性知 $\forall x:\ F(x) \leqslant y$，$x_0 - \delta \geqslant x$，从而 $x_0 > x_0 - \delta \geqslant \sup\{x:\ F(x) \leqslant y\} = G(y)$.

综合 (a)(b) 知得证。

(2) 设 $x_0 = G(y)$，$y \in (0,1)$，现在要证 $F(x_0) = y$. 首先，由 (1) 可知 $F(x_0) \leqslant y$；其次，若 $F(x_0) < y$，由 F 连续且 $\lim\limits_{x\to+\infty} F(x) = 1$，$\lim\limits_{x\to-\infty} F(x) = 0$ 知 $\exists \xi \in (-\infty, +\infty)$，$s.t.\ F(\xi) = y$，则由 F 的不降性知只能有 $\xi > x_0$；但 $x_0 = G(y) = \sup\{x:\ F(x) \leqslant y\}$，而此时 $\xi \in \{x:\ F(x) \leqslant y\}$ 且 $\xi > x_0$，矛盾！故只能有 $F(x_0) = y$.

F 不连续时结论未必成立：例如当 $F(x) = \begin{cases} 1, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$（即 $\mathbb{P}(X = 0) = 1$，$\mathbb{P}(X \neq 0) = 0$）时，$\mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant y\}) = \begin{cases} 1, & y \geqslant 1 \\ 0, & y < 1 \end{cases}$（$y \geqslant 1$ 或 $y < 0$ 时由分布函数的性质知显然；$0 \leqslant y < 1$ 时 $\{\omega:\ F(X(\omega)) \leqslant y\} = \{\omega:\ F(X(\omega)) = 0\} = \{\omega:\ X(\omega) < 0\}$，从而 $\mathbb{P}(\{\omega:\ F(X(\omega)) \leqslant y\}) = \mathbb{P}(X < 0) = 0$.），即此时 $F(X)$ 并不服从 $[0,1]$ 上的均匀分布。 □

注: 习题 4.1.8也可用习题 4.1.7中定义的 $G(y)$ 去做, 两种定义的 $G(y)$ 均称为广义逆（generalized inverse）或分位函数（quantile function）, 对于不降函数均存在. 这里给出来自维基百科的一张示意图片的链接. 关于广义逆的一系列性质与结论请参考 [GI].

4.1.9 如果 f 为 Borel 可测, X 与 Y 同分布, 则 $f(X)$ 与 $f(Y)$ 也同分布.

证明: 我们有

$$\mathbb{P}_{f(X)}(A) = \mathbb{P}(f(X) \in A) = \mathbb{P}_X(f^{-1}(A)) = \mathbb{P}_Y(f^{-1}(A)) = \mathbb{P}(f(Y) \in A) = \mathbb{P}_{f(Y)}(A).$$

□

4.1.10 (1) 设 $\sigma(X)$ 是由随机变量 X 所产生的 σ 域, 则 $\Lambda \in \sigma(X)$ 的充要条件是对于某个 $B \in \mathscr{B}$, $\Lambda = X^{-1}(B)$. 问此 B 是否唯一? 能否存在一个集 $A \notin \mathscr{B}$, 使得 $\Lambda = X^{-1}(A)$?

(2) 将上题中的论断推广到有限个随机变量的情况. (推广到任意个随机变量的情况也是可能的)

证明: (1) 实际上此 B 并不一定唯一, 我们可以轻易地举出反例: 考虑 X 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 到 $(\mathbb{R}, \mathscr{B})$ 的随机变量, 且 $\forall \omega$, $|X(\omega)| \leqslant M$, 那么考虑 Borel 集 $B_1 = [-M, M]$, $B_2 = [-M, M+1]$, 便有 $f^{-1}(B_1) = f^{-1}(B_2)$.

对于 (1) 中的第二个问题, 我们只需要考虑常 r.v., 也即 $X(\omega) \equiv C \in \mathbb{R}$, 那么对于任意的 Borel 不可测集 A, $A \cup \{C\} \notin \mathscr{B}$, 但是 $f^{-1}(A \cup \{C\}) = \Omega$. □

4.1.11 设 X 是 n 维实随机变量, F_X, μ_X 分别是 X 的分布函数和概率分布测度, 试用 F_X, μ_X 表示 $f(X)$ 的分布函数或概率分布测度, 其中 $f: \mathbb{R}^n \to \mathbb{R}$ 是下列 Borel 可测函数:

(1) 当 $n = 1$ 时, $f(x) := x^2, x \in \mathbb{R}$;

(2) 当 $n = 1$ 时, $f(x) := (x+4)(x-1)(x-3), x \in \mathbb{R}$;

(3) 当 $n = 1$ 时, $f(x) := \cos kx, x \in \mathbb{R}, k$ 为常数;

(4) $f(x) := \sqrt{\sum_{k=1}^{n} x_k^2}, x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$;

(5) $f(x) := \sum_{k=1}^{n} x_k, x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$;

(6) $f(x) := \max_{1 \leqslant k \leqslant n} x_k, x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$;

(7) $f(x) := \min_{1 \leqslant k \leqslant n} x_k, x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$.

证明:

(1) 我们有

$$F_{f(X)}(y) = \mathbb{P}(x^2 \leqslant y) = \begin{cases} F_X(\sqrt{y}) - F_X(-\sqrt{y}), & y > 0; \\ 0, & y \leqslant 0. \end{cases} = \begin{cases} \mu_X([-\sqrt{y}, \sqrt{y}]), & y > 0; \\ 0, & y \leqslant 0. \end{cases}$$

(2) 十分复杂, 省去计算过程 (可自行搜索三次方程求根公式): 若定义

$$A := \frac{2(54+13\sqrt{39})}{9}, \quad B := 12 - \frac{26}{3}\sqrt{\frac{13}{3}}, \quad u(y) := \sqrt[3]{\frac{\sqrt{27(y-12)^2 - 8788} + 3y - 27}{6}},$$

$$x_1(y) := u(y) + \frac{13}{3u(y)}, \quad x_2(y) := \omega_1 u(y) + \frac{13\omega_2}{3u(y)}, \quad x_3(y) := \omega_2 u(y) + \frac{13\omega_1}{3u(y)},$$

则

$$F_{f(X)}(y) = \mu_X((-\infty, x_1(y)]) + \mu_X([x_2(y), x_3(y)])\, \mathbb{1}_{\{B \leqslant y \leqslant A\}}.$$

其中 $\omega_1 = \dfrac{-1+\sqrt{3}\mathrm{i}}{2}$, $\omega_2 = \dfrac{-1-\sqrt{3}\mathrm{i}}{2}$ 为三次单位根.

(3) 当 $k=0$, 显然有 $F_{f(X)}(y)=\mathbb{1}_{\{y\geqslant 1\}}$. 当 $k>0$, 我们有

$$
\begin{aligned}
F_{f(X)}(y)=\mathbb{P}(\cos kx\leqslant y)&=\begin{cases}1, & y\geqslant 1;\\ \mathbb{P}\left(x\in\bigcup\limits_{n\in\mathbb{Z}}\left[\dfrac{2n\pi+\arccos y}{k},\dfrac{2n\pi+2\pi-\arccos y}{k}\right]\right), & y\in(0,1);\\ 0, & y\leqslant 0.\end{cases}\\
&=\begin{cases}1, & y\geqslant 1;\\ \sum\limits_{n\in\mathbb{Z}}\left(F_X\left(\dfrac{2n\pi+2\pi-\arccos y}{k}\right)-F_X\left(\dfrac{2n\pi+\arccos y}{k}\right)\right), & y\in(0,1);\\ 0, & y\leqslant 0.\end{cases}\\
&=\begin{cases}1, & y\geqslant 1;\\ \sum\limits_{n\in\mathbb{Z}}\mu_X\left[\dfrac{2n\pi+\arccos y}{k},\dfrac{2n\pi+2\pi-\arccos y}{k}\right], & y\in(0,1);\\ 0, & y\leqslant 0.\end{cases}
\end{aligned}
$$

类似地, 当 $k<0$, 有

$$
F_{f(X)}(y)=\begin{cases}1, & y\geqslant 1;\\ \sum\limits_{n\in\mathbb{Z}}\mu_X\left[\dfrac{2n\pi+2\pi-\arccos y}{k},\dfrac{2n\pi+\arccos y}{k}\right], & y\in(0,1);\\ 0, & y\leqslant 0.\end{cases}
$$

(4) 显然有 $F_{f(X)}(y)=\mu_X\left(\sum\limits_{k=1}^n x_k^2\leqslant y^2\right)\mathbb{1}_{\{y\geqslant 0\}}$.

(5) 类似 (4), 有 $F_{f(X)}(y)=\mu_X\left(\sum\limits_{k=1}^n x_k\leqslant y\right)$.

(6) 我们有 $F_{f(X)}(y)=\mathbb{P}\left(\bigcap\limits_{k=1}^n\{x_k\leqslant y\}\right)=\prod\limits_{k=1}^n\mathbb{P}(x_k\leqslant y)=(\mu_X((-\infty,y]))^n$.

(7) 我们有 $F_{f(X)}(y)=1-\mathbb{P}\left(\min\limits_{1\leqslant k\leqslant n}x_k>y\right)=1-\mathbb{P}\left(\bigcap\limits_{k=1}^n\{x_k>y\}\right)=1-(\mu_X((y,\infty)))^n$.

□

§ 4.2 可测函数的构造性质

4.2.1 设 $\mathcal{C}$ 为 Ω 的一个 π 系, $\mathscr{H}$ 为 Ω 上的函数类, 且满足下列条件:

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

(2) $\mathscr{H}$ 对非负线性组合封闭, 且若 $f,g\in\mathscr{H}$, 有界, $f\geqslant g$, 则 $f-g\in\mathscr{H}$;

(3) 若 $f_n\in\mathscr{H},n\in\mathbb{N}$, 且 $0\leqslant f_n\uparrow f$, 则 $f\in\mathscr{H}$;

(4) $\mathscr{H} \supset \{I_A : A \in \mathcal{C}\}$, 则 $\mathscr{H}$ 包含一切非负 $\sigma(\mathcal{C})$ 可测函数.

证明: 考虑集类

$$\Lambda := \{A \subset \Omega, \mathbb{1}_A \in \mathscr{H}\},$$

显然 $\Omega \in \Lambda$; 考虑 $A_1, A_2 \subset \Lambda$, 且 $A_2 \subset A_1$, 则 $\mathbb{1}_{A_1 \setminus A_2} = \mathbb{1}_{A_1} - \mathbb{1}_{A_2} \in \mathscr{H}$.

考虑 $\{A_n, n \in \mathbb{N}\} \subset \Lambda$, 且 $A_n \uparrow$. 我们有: $0 \leqslant \mathbb{1}_{A_n} \uparrow \mathbb{1}_{\bigcup_{n=1}^{\infty} A_n}$, 因此 $\bigcup\limits_{n=1}^{\infty} A_n \in \Lambda$.

故 Λ 是 λ 系, 又 $\mathcal{C}$ 是 π 系, 因此 $\sigma(\mathcal{C}) = \Lambda(\mathcal{C}) \subset \Lambda$. 故 $\forall A \in \sigma(\mathcal{C})$, 都有 $\mathbb{1}_A \in \mathscr{H}$. 我们知道 $\mathscr{H}$ 对非负线性组合封闭, 因此 $\mathscr{H}$ 包含一切非负 $\sigma(\mathcal{C})$ 简单可测函数.

根据定理 4.2.4, 我们知道任意非负 $\sigma(\mathcal{C})$ 可测函数都是一个不降 $\sigma(\mathcal{C})$ 简单函数序列的极限. 考虑 f 是 $\sigma(\mathcal{C})$ 可测的, 则存在非负可测简单函数列 $\{f_n\} \subset \mathscr{H}$, 使得 $0 \leqslant f_n \uparrow f$, 故 $f \in \mathscr{H}$. 因此 $\mathscr{H}$ 包含一切非负 $\sigma(\mathcal{C})$ 可测函数. □

4.2.2 设 $f : \mathbb{R} \longmapsto \mathbb{R}$ 在 $\mathbb{R}$ 的每一点可求导, 试证其导函数 Borel 可测.

证明: 我们知道 f 连续, 因此 Borel 可测. 考虑函数列 $f_n(x) := n\left(f\left(x + \dfrac{1}{n}\right) - f(x)\right)$, 则 $f_n(x)$ 是 Borel 可测的. 我们知道 $\lim\limits_{n\to\infty} f_n(x) = f'(x)$, 故 $f'(x)$ 也是 Borel 可测的. □

4.2.3 Cantor 集 $P_0 = [0,1] \backslash G_0$,

$$G_0 := \bigcup_{n-1}^{\infty} \bigcup_{\substack{a_k=0,2 \\ k=1,2,\cdots n-1}} \left(\frac{1}{3^n}\left[\sum_{k=1}^{n-1} a_{n-k}3^k + 1\right], \frac{1}{3^n}\left[\sum_{k=1}^{n-1} a_{n-k}3^k + 2\right]\right),$$

今定义函数 $f : [0,1] \mapsto [0,1]$ 如下: 当

$$x \in \left(\frac{1}{3^n}\left[\sum_{k=1}^{n-1} a_{n-k}3^k + 1\right], \frac{1}{3^n}\left[\sum_{k=1}^{n-1} a_{n-k}3^k + 2\right]\right)$$

时,

$$f(x) := \frac{1}{2^n}\left(\sum_{k=1}^{n-1} \frac{a_{n-k}}{2} \cdot 2^k + 1\right),$$

于是 f 在 G_0 上有定义, 且在 G_0 上不降. 其次, 令 $f(0) = 0$, 而 $\forall x \in P_0 \backslash \{0\}$, 定义 $f(x) := \sup\{f(t) : t \in G_0, t < x\}$, 试证: f 为 $[0,1]$ 上的不降连续函数, 因而 f 为 $\mathscr{B}[0,1]$ 可测.

4.2.4 设 (E,d) 是一距离空间, 试证:

(1) $\forall G \subset E$ 为开集, 令 $f_n := \dfrac{nd\,(x, G^c)}{1 + nd\,(x, G^c)}, x \in E$, 则 $0 \leqslant f_n \uparrow I_G$;

(2) 设 $\mathscr{L}$ 为 (E,d) 上的全体实值函数类, L 为 $\mathscr{L}$ 系, 且 $L \supset C_b(E,\mathbb{R})$ (即 E 上的有界实连续函数类), 则 L 包含 $\mathscr{B}(E)$ 可测函数类.

证明: (1) 我们知道 G 是开集, 则 $\forall x \in G$, 有 $d(x, G^c) > 0$. 考虑 $x \in G$, 则 f_n 递增且 $\lim\limits_{n\to\infty} \dfrac{nd(x,G^c)}{1 + nd(x,G^c)} = 1$; 若 $x \notin G$, 则 $d(x, G^c) = 0$, $f_n = 0$. 故 $0 \leqslant f_n \uparrow \mathbb{1}_G$.

(2) 考虑 π 系 $\mathcal{C} := \{G \subset E, G\text{是开集}\}$, 则 $\sigma(\mathcal{C}) = \mathscr{B}(E)$. 我们知道实函数 f_n 连续, 且 $|f_n| \leqslant 1$, 则 $f_n \in L$. 又因为 L 是 $\mathscr{L}$ 系, 且 $f_n \uparrow \mathbb{1}_G$, 则 $\mathbb{1}_G \in L$. 由单调类定理便知 L 包含一切属于 $\mathscr{L}$ 的 $\sigma(\mathcal{C})$ 可测函数, 也即 L 包含 $\mathscr{B}(E)$ 可测函数类. □

第五章 积分与数学期望

§ 5.1 积分的定义

5.1.1 给出非负可测函数积分的另一种定义:

(1) 按照定义 5.1.2(1) 定义非负简单函数的积分, 证明定义的合理性;

(2) 证明非负简单函数的积分具有性质: 若 $f \leqslant g$, 则 $\displaystyle\int f \leqslant \int g$;

(3) 如下定义非负可测函数的积分: 若 f 非负可测, $\{f_n\}$ 为简单函数列, 满足 $0 \leqslant f_n \uparrow f$, 则令 $\displaystyle\int f = \lim_{n\to\infty} \int f_n$;

(4) 证明 (3) 所定义的非负可测函数积分的合理性;

(5) 证明单调收敛定理.

证明:

(1) 考虑非负 $\mathcal{F}$ 简单函数 $f = \displaystyle\sum_{k=1}^{m} x_k \mathbb{1}_{A_k} = \sum_{l=1}^{n} y_l \mathbb{1}_{B_l}$, 其中 $A_1, A_2, \cdots, A_m$ 两两不交, 且 $\displaystyle\bigcup_{k=1}^{m} A_k = \Omega$. $B_1, B_2, \cdots, B_n$ 两两不交, $\displaystyle\bigcup_{l=1}^{n} B_l = \Omega$. 则

$$\int_\Omega f \mathrm{d}\mu = \sum_{k=1}^{m} x_k \mu(A_k) = \sum_{k=1}^{m}\sum_{l=1}^{n} x_k \mu(A_k \cap B_l) = \sum_{k=1}^{m}\sum_{l=1}^{n} y_l \mu(A_k \cap B_l) = \sum_{l=1}^{n} y_l \mu(B_l) = \int_\Omega f \mathrm{d}\mu.$$

因此定义合理.

(2) 考虑测度空间 $(\Omega, \mathcal{F}, \mu)$ 上的非负 $\mathcal{F}$ 简单函数 $f = \displaystyle\sum_{k=1}^{m} x_k \mathbb{1}_{A_k}$, 其中 $A_k, k = 1, 2, \cdots, m$ 两两不交, 且 $\displaystyle\bigcup_{k=1}^{m} A_k = \Omega$. 有以及非负 $\mathcal{F}$ 简单函数 $g = \displaystyle\sum_{j=1}^{n} y_j \mathbb{1}_{B_j}$, 其中 $B_j, j = 1, 2, \cdots, n$ 两两不交, 且 $\displaystyle\bigcup_{k=1}^{n} B_j = \Omega$. 若 $g \leqslant f$, 则必然有 $y_j \leqslant x_k$(当 $A_k \cap B_j \neq \varnothing$ 时). 因此我们有

$$\int_\Omega f \mathrm{d}\mu = \sum_{k=1}^{m} x_k \mu(A_k) = \sum_{k=1}^{m}\sum_{j=1}^{n} x_k \mu(A_k \cap B_j) \geqslant \sum_{k=1}^{m}\sum_{j=1}^{n} y_j \mu(A_k \cap B_j) = \sum_{j=1}^{n} y_j \mu(B_j) = \int_\Omega g \mathrm{d}\mu.$$

(4) 考虑非负 $\mathcal{F}$ 可测函数 f, 我们只需证明: 任取两个非负 $\mathcal{F}$ 简单函数列 $\{f_n\}$, $\{g_n\}$, 如果 $f_n \uparrow f$, $g_n \uparrow f$, 那么 $\displaystyle\lim_{n\to\infty} \int_\Omega f_n \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega g_n \mathrm{d}\mu$. 为了证明这个命题, 我们实际上只需要证明 $\forall m \in \mathbb{N}$, 有 $\displaystyle\int_\Omega f_m \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_\Omega g_n \mathrm{d}\mu$. 在这个命题的基础上令 $m \to \infty$ 便有 $\displaystyle\lim_{m\to\infty} \int_\Omega f_m \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_\Omega g_n \mathrm{d}\mu$, 由对

称性可以再得到 $\lim\limits_{n\to\infty}\int_\Omega g_n\,\mathrm{d}\mu\leqslant\lim\limits_{m\to\infty}\int_\Omega f_m\,\mathrm{d}\mu$, 这便证明了 $\lim\limits_{n\to\infty}\int_\Omega f_n\,\mathrm{d}\mu=\lim\limits_{n\to\infty}\int_\Omega g_n\,\mathrm{d}\mu$. 下面我们证明 $\forall m\in\mathbb{N}$, 有 $\int_\Omega f_m\,\mathrm{d}\mu\leqslant\lim\limits_{n\to\infty}\int_\Omega g_n\,\mathrm{d}\mu$.

考虑 $f_n=\sum\limits_{k=1}^{r_n}a_k^{(n)}\mathbb{1}_{A_k^{(n)}}+\infty\mathbb{1}_{A_{r_n+1}^{(n)}}$, 其中 $a_k^{(n)}\geqslant 0$, $A_1^{(n)},A_2^{(n)},\cdots,A_{r_n+1}^{(n)}$ 两两不交, 且 $\bigcup\limits_{k=1}^{r_n+1}A_k^{(n)}=\Omega$. $g_n=\sum\limits_{k=1}^{s_n}b_k^{(n)}\mathbb{1}_{B_k^{(n)}}+\infty\mathbb{1}_{B_{s_n+1}^{(n)}}$, 其中 $b_k^{(n)}\geqslant 0$, $B_1^{(n)},B_2^{(n)},\cdots,B_{s_n+1}^{(n)}$ 两两不交, 且 $\Omega=\bigcup\limits_{k=1}^{s_n+1}B_k^{(n)}$. $\forall m\in\mathbb{N}$, 考虑 $f_m=\sum\limits_{k=1}^{r_m}a_k^{(m)}\mathbb{1}_{A_k^{(m)}}+\infty\mathbb{1}_{A_{r_m+1}^{(m)}}$, 再对 $c\in(0,1)$, $l\in\mathbb{N}$, 考虑 $f_{c,l}^{(m)}=\sum\limits_{k=1}^{r_m}ca_k^{(m)}\mathbb{1}_{A_k^{(m)}}+l\mathbb{1}_{A_{r_m+1}^{(m)}}$, 则当 $c\to 1^-$, $l\to\infty$ 时 $\int_\Omega f_{c,l}^{(m)}\mathrm{d}\mu\to\int_\Omega f_m\,\mathrm{d}\mu$. 现考虑集合

$$\Omega_n=\left\{\omega\in\Omega:\ f_{c,l}^{(m)}(\omega)\leqslant g_n(\omega)\right\},$$

则易证 $\Omega_n\uparrow\Omega$. (事实上, 由 $\{g_n\}$ 是单增的函数列知 $\Omega_n\subset\Omega_{n+1}(\forall n\in\mathbb{N})$; 若 $\bigcup\limits_{n=1}^{\infty}\Omega_n\neq\Omega$, 则 $\bigcap\limits_{n=1}^{\infty}\Omega_n^c\neq\varnothing$, 从而 $\exists\omega_0\in\Omega$, s.t. $\forall n\in\mathbb{N}$, $f_{c,l}^{(m)}(\omega_0)>g_n(\omega_0)$, 令 $n\to\infty$ 即得 $f_{c,l}^{(m)}(\omega_0)\geqslant f(\omega)$, 与 $f_{c,l}^{(m)}(\omega_0)<f_m(\omega_0)\leqslant f(\omega_0)$ 矛盾!)

由引理 5.1.3(3), 对于任意非负可测函数 f, 任意 $A\in\mathcal{F}$, 若定义 $\varphi(A):=\int_A f\,\mathrm{d}\mu$, 则 φ 是 $\mathcal{F}$ 上的一种测度. 而测度具有下连续性, 因此

$$\int_\Omega f_{c,l}^{(m)}\mathrm{d}\mu\xlongequal{\text{测度的下连续性}}\lim_{n\to\infty}\int_{\Omega_n}f_{c,l}^{(m)}\mathrm{d}\mu\leqslant\lim_{n\to\infty}\int_{\Omega_n}g_n\,\mathrm{d}\mu\xlongequal{\text{测度的下连续性}}\int_\Omega g_n\,\mathrm{d}\mu,$$

令 $c\to 1^-$, $l\to\infty$, 便有 $\int_\Omega f_m\mathrm{d}\mu\leqslant\lim\limits_{n\to\infty}\int_\Omega g_n\,\mathrm{d}\mu$. 证毕.

(5) 我们先证明这个定义下, 非负可测函数积分的单调性. 由定理 4.2.4, 对非负可测函数 f,g, 能找到非负简单函数列 $\{f_n,n\in\mathbb{N}\},\{g_n,n\in\mathbb{N}\}$, 使得 $f_n\uparrow f$, $g_n\uparrow g$. 不妨设 $f\leqslant g$, 则我们要证明 $\int_\Omega f\,\mathrm{d}\mu\leqslant\int_\Omega g\,\mathrm{d}\mu$, 按 (3) 的定义, 只需证明 $\lim\limits_{n\to\infty}\int_\Omega f_n\,\mathrm{d}\mu\leqslant\lim\limits_{n\to\infty}\int_\Omega g_n\,\mathrm{d}\mu$, 和 (4) 同样地, 只需证明 $\forall m\in\mathbb{N}$, $\int_\Omega f_m\mathrm{d}\mu\leqslant\lim\limits_{n\to\infty}\int_\Omega g_n\mathrm{d}\mu$. 令 $h_n^{(m)}=\min\{f_m,g_n\}$, 则 $h_n^{(m)}$ 是非负简单函数且 $h_n^{(m)}\uparrow f_m$, 于是 $\lim\limits_{n\to\infty}\int_\Omega h_n^{(m)}\mathrm{d}\mu=\int_\Omega f_m\mathrm{d}\mu$. 又 $\int_\Omega h_n^{(m)}\mathrm{d}\mu\leqslant\int_\Omega g_n\mathrm{d}\mu$ $(\forall n\in\mathbb{N})$, 故 $\forall m\in\mathbb{N}$, 有 $\int_\Omega f_m\leqslant\lim\limits_{n\to\infty}\int_\Omega g_n\mathrm{d}\mu$. 于是

$$\int_\Omega f\,\mathrm{d}\mu=\lim_{m\to\infty}\int_\Omega f_m\,\mathrm{d}\mu\leqslant\lim_{n\to\infty}\int_\Omega g_n\,\mathrm{d}\mu=\int_\Omega g\,\mathrm{d}\mu.$$

下面我们证明单调收敛定理. 考虑非负可测函数列 $\{f_n,n\in\mathbb{N}\}$, 且 $f_n\uparrow f$, 则 $\int_\Omega f_n\mathrm{d}\mu\leqslant\int_\Omega f\mathrm{d}\mu$ $(\forall n\in\mathbb{N})$, 因此 $\lim\limits_{n\to\infty}\int_\Omega f_n\,\mathrm{d}\mu\leqslant\int_\Omega f\,\mathrm{d}\mu$, 接下来只需证明 $\lim\limits_{n\to\infty}\int_\Omega f_n\,\mathrm{d}\mu\geqslant\int_\Omega f\,\mathrm{d}\mu$.

取非负简单函数列 $\{u_n,n\in\mathbb{N}\}$, 使得 $u_n\uparrow f$, 则类似于第 (4) 问只需证明 $\int_\Omega u_m\mathrm{d}\mu\leqslant\lim\limits_{n\to\infty}\int_\Omega f_n\mathrm{d}\mu$ $(\forall m\in\mathbb{N})$. 设 $u_n=\sum\limits_{k=1}^{t_n}d_k^{(n)}\mathbb{1}_{D_k^{(n)}}+\infty\mathbb{1}_{D_{t_n+1}^{(n)}}$, 其中 $d_k^{(n)}\geqslant 0$, $D_1^{(n)},D_2^{(n)},\cdots,D_{t_n+1}^{(n)}$ 两两不交, 且 $\bigcup\limits_{k=1}^{t_n+1}D_k^{(n)}=\Omega$.

取定 $m \in \mathbb{N}$, 考虑 $u_m = \sum_{k=1}^{t_m} d_k^{(m)} \mathbb{1}_{D_k^{(m)}} + \infty \mathbb{1}_{D_{t_{m+1}}^{(m)}}$, 类似于第 (4) 问, 对任意的 $c \in (0,1)$, $l \in \mathbb{N}$, 考虑 $u_{c,l}^{(m)} = \sum_{k=1}^{t_m} c d_k^{(m)} \mathbb{1}_{D_k^{(m)}} + l \mathbb{1}_{D_{t_{m+1}}^{(m)}}$, 再考虑集合

$$\Lambda_n = \left\{\omega \in \Omega:\ u_{c,l}^{(m)}(\omega) \leqslant f_n(\omega)\right\},$$

则同理于 (4) 可证 $\Lambda_n \uparrow \Omega$. 同样地, 由于积分可看作一种测度, 根据测度的下连续性, 有

$$\int_\Omega u_{c,l}^{(m)} \mathrm{d}\mu = \lim_{n\to\infty} \int_{\Lambda_n} u_{c,l}^{(m)} \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_{\Lambda_n} f_n \mathrm{d}\mu = \int_\Omega g_n \mathrm{d}\mu$$

令 $c \to 1^-$, $l \to \infty$, 则得到 $\int_\Omega u_m \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_\Omega f_n \mathrm{d}\mu$ $(\forall m \in \mathbb{N})$, 因此

$$\int_\Omega f \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega u_m \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_\Omega f_n \mathrm{d}\mu,$$

证毕.

$\square$

注: 第 (4) 小问 $\Omega_n \uparrow \Omega$ 要详细说明, 因为 $\Omega_n = \{\omega \in \Omega:\ T_n(\omega)\ \text{成立}\} \not\Longrightarrow \lim_{n\to\infty} \Omega_n = \left\{\omega \in \Omega:\ \lim_{n\to\infty} T_n(\omega)\ \text{成立}\right\}$.
举例: $f(x) \equiv 1,\ f_n(x) \equiv 1 - \dfrac{1}{n}$, 设 $\Omega_n = \{x:\ f_n(x) \geqslant f(x)\}$, 则
$\left\{x:\ \lim_{n\to\infty} f_n(x) \geqslant f(x)\right\} = \Omega$, 但是 $\lim_{n\to\infty} \Omega_n = \varnothing$.

5.1.2 证明注 5.1.10.

实际上, 在测度空间 $(\Omega, \mathcal{F}, \mu)$ 上, 我们需要证明的结论 (或命题) 有如下几个:

(1) 设 f, g 为 $(\Omega, \mathcal{F}, \mu)$ 上的非负 a.e. 可测函数, 且 $g \leqslant f$, 则 $\int_\Omega g \mathrm{d}\mu \leqslant \int_\Omega f \mathrm{d}\mu$.

(2) 若 $\{f_n, n \in \mathbb{N}\}$ 是非负 a.e. 可测函数列, 且 $f_n \uparrow f$, 则 $\lim_{n\to\infty} \int_\Omega f_n \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu$.

(3) 任一以 a.e. 可测函数 f 为极限的非负不降简单函数列 $\{f_n, n \in \mathbb{N}\}$, 有 $\lim_{n\to\infty} \int_\Omega f_n \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu$.

(4) 若 f, g 都是 a.e. 可测函数, 且 $f = g$, μ a.e., 则 $\int_\Omega f \mathrm{d}\mu = \int_\Omega g \mathrm{d}\mu$.

证明:

(1) 我们知道存在可测函数 f', g', 使得 $N_1 = \{\omega \in \Omega : f(\omega) \neq f'(\omega)\}$, $N_2 = \{\omega \in \Omega : g(\omega) \neq g'(\omega)\}$ 是 μ 零集. 这意味着存在 $B_i \supset N_i$, 使得 $\mu(B_i) = 0$. 我们知道, $\forall \omega \in B_1^c \cap B_2^c = (B_1 \cup B_2)^c$, 有 $g'(\omega) = g(\omega) \leqslant f(\omega) = f'(\omega)$. 且 $0 \leqslant \mu(B_1 \cup B_2) \leqslant \mu(B_1) + \mu(B_2) = 0$, 故 $\int_{B_1 \cup B_2} g' \mathrm{d}\mu = \int_{B_1 \cup B_2} f' \mathrm{d}\mu$, 因此

$$\begin{aligned}
\int_\Omega g \mathrm{d}\mu = \int_\Omega g' \mathrm{d}\mu &= \int_{B_1 \cup B_2} g' \mathrm{d}\mu + \int_{\omega \setminus (B_1 \cap B_2)} g' \mathrm{d}\mu \\
&\leqslant \int_{B_1 \cup B_2} g' \mathrm{d}\mu + \int_{\omega \setminus (B_1 \cap B_2)} f' \mathrm{d}\mu \\
&= \int_{B_1 \cup B_2} f' \mathrm{d}\mu + \int_{\omega \setminus (B_1 \cap B_2)} f' \mathrm{d}\mu = \int_\Omega f' \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu.
\end{aligned}$$

(2) 因为 f_n a.e. 可测, 所以存在 μ 零集 $\{A_n, n\in\mathbb{N}\}\subset\mathcal{F}$, 可测集 $\{B_n, n\in\mathbb{N}\}\subset\mathcal{F}$, 可测函数列 $\{g_n, n\in\mathbb{N}\}$. 使得 $\forall\omega\in A_n^c$, $g_n(\omega)=f_n(\omega)$, 且 $\mu(B_n)=0$. 考虑集合 $A=\bigcup\limits_{n=1}^{\infty}B_n$, 则 $g_n'=g_n(1-\mathbb{1}_A)$ 是非负可测函数, 且 $g_n'\uparrow$. 定义 $\lim\limits_{n\to\infty}g_n'=g$, 则根据单调收敛定理, 有 $\displaystyle\int_\Omega g\,\mathrm{d}\mu=\int_\Omega g_n'\,\mathrm{d}\mu$.

考虑集合 $\Omega_n:=\{\omega\in\Omega, g_n'(\omega)\neq g_n(\omega)\}\subset A$, 则 $\mu(\Omega_n)\leqslant\mu(A)\leqslant\sum\limits_{n=1}^{\infty}\mu(B_n)=0$, 因此 $g_n'=g_n$, a.e. 于是

$$\int_\Omega g_n'\,\mathrm{d}\mu=\int_\Omega g_n\,\mathrm{d}\mu=\int_\Omega f_n\,\mathrm{d}\mu,$$

下面我们证明 $f=g$, a.e., 实际上, 我们有

$$\forall\omega\in A^c, f(\omega)=\lim_{n\to\infty}f_n(\omega)=\lim_{n\to\infty}g_n'(\omega)=g(\omega),$$

因此 $N:=\{\omega\in\Omega: f(\omega)\neq g(\omega)\}\subset A$, 故 N 是 μ 零集, 因此 $f=g$, a.e.. 故

$$\int_\Omega f\,\mathrm{d}\mu=\int_\Omega g\,\mathrm{d}\mu=\int_\Omega g_n'\,\mathrm{d}\mu=\int_\Omega f_n\,\mathrm{d}\mu.$$

(3) 我们知道存在可测函数 f', 使得 $\displaystyle\int_\Omega f\,\mathrm{d}\mu=\int_\Omega f'\,\mathrm{d}\mu$, 故由推论 5.1.6 立得.

(4) 我们知道, 存在可测函数 f', g', 使得 $f'=f$, $g'=g$, a.e.. 故

$$N_1=\{\omega\in\Omega: f'(\omega)\neq f(\omega)\},\quad N_2=\{\omega\in\Omega: g'(\omega)\neq g(\omega)\},\quad N_3=\{\omega\in\Omega: f(\omega)=g(\omega)\},$$

均为 μ 零集. 因此存在 $B_i\supset N_i$, $(i=1,2,3)$, 使得 $\mu(B_i)=0$. 令 $N_0=\{\omega\in\Omega: f'(\omega)\neq g'(\omega)\}$, 则

$$N_0=\{\omega\in\Omega: f'(\omega)\neq g'(\omega)\}\subset N_1\cup N_2\cup N_3,$$

故 $N_0\subset B_1\cup B_2\cup B_3$, 而 $\mu(B_1\cup B_2\cup B_3)\leqslant\mu(B_1)+\mu(B_2)+\mu(B_3)=0$, 故 N_0 是 μ 零集. 因此 $f'=g'$, a.e.. 由引理 5.1.8 立得要证结论.

□

§ 5.2 积分的性质

5.2.1 设 μ_1,μ_2 是可测空间 $(\Omega,\mathcal{F})$ 上的测度, a_1,a_2 是两个非负有限数, $\mu=a_1\mu_1+a_2\mu_2$, 试证: 若 f 对 μ_1,μ_2 的积分存在且 $a_1\displaystyle\int f\,\mathrm{d}\mu_1+a_2\int f\,\mathrm{d}\mu_2$ 有意义, 则 f 对 μ 的积分也存在, 且

$$\int f\,\mathrm{d}\mu=a_1\int f\,\mathrm{d}\mu_1+a_2\int f\,\mathrm{d}\mu_2.$$

证明: 按照课本上的方法, 从非负简单函数开始.

若 f 是非负简单函数, $f=\sum\limits_{k=1}^{m}x_k\mathbb{1}_{A_k}$, $A_1,A_2,\cdots,A_m$ 两两不交, 且 $\bigcup\limits_{k=1}^{m}A_k=\Omega$, 则

$$\int_\Omega f\,\mathrm{d}\mu=\sum_{k=1}^{m}x_k\mu(A_k)=a_1\sum_{k=1}^{m}x_k\mu_1(A_k)+a_2\sum_{k=1}^{m}x_k\mu_2(A_k)=a_1\int_\Omega f\,\mathrm{d}\mu_1+a_2\int_\Omega f\,\mathrm{d}\mu_2.$$

若 f 是非负可测函数, 则根据定理 4.2.4, 存在非负简单函数列 $\{f_n, n \in \mathbb{N}\}$ 使得 $0 \leqslant f_n \uparrow f$. 且 $\lim\limits_{n\to\infty}\int_\Omega f_n \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu$. 因此

$$\begin{aligned}\int_\Omega f \mathrm{d}\mu = \lim_{n\to\infty}\int_\Omega f_n \mathrm{d}\mu &= \lim_{n\to\infty}\left(a_1\int_\Omega f_n \mathrm{d}\mu_1 + a_2\int_\Omega f_n \mathrm{d}\mu_2\right)\\ &= a_1\lim_{n\to\infty}\int_\Omega f_n \mathrm{d}\mu_1 + a_2\lim_{n\to\infty}\int_\Omega f_n \mathrm{d}\mu_2\\ &= a_1\int_\Omega f \mathrm{d}\mu_1 + a_2\int_\Omega f \mathrm{d}\mu_2.\end{aligned}$$

若 f 是一般的可测函数, 考虑 $f = f^+ - f^-$, 则 f^+ 和 f^- 是非负可测函数. 故

$$\begin{aligned}\int_\Omega f \mathrm{d}\mu &= \int_\Omega f^+ \mathrm{d}\mu - \int_\Omega f^- \mathrm{d}\mu\\ &= a_1\int_\Omega f^+ \mathrm{d}\mu_1 + a_2\int_\Omega f^- \mathrm{d}\mu_2 - a_1\int_\Omega f^- \mathrm{d}\mu_1 - a_2\int_\Omega f^- \mathrm{d}\mu_2\\ &= a_1\left(\int_\Omega f^+ \mathrm{d}\mu_1 - \int_\Omega f^- \mathrm{d}\mu_1\right) + a_2\left(\int_\Omega f^+ \mathrm{d}\mu_1 - \int_\Omega f^- \mathrm{d}\mu_1\right)\\ &= a_1\int_\Omega f \mathrm{d}\mu_1 + a_2\int_\Omega f \mathrm{d}\mu_2.\end{aligned}$$

(注: 证明过程中, 由于 $a_1\int_\Omega f\mathrm{d}\mu_1 + a_2\int_\Omega f\mathrm{d}\mu_2$ 有意义且 $\forall n \in \mathbb{N}$, $0 \leqslant f_n \leqslant f$, 所以 $a_1\int_\Omega f\mathrm{d}\mu_1 + a_2\int_\Omega f\mathrm{d}\mu_2$ 也有意义. “有意义”是指避免了 $\infty - \infty$ 的情形.) □

5.2.2 (积分中值定理) 设 f, g 是测度空间 $(\Omega, \mathcal{F}, \mu)$ 上的可测函数, g 对 μ 可积, $-\infty < a \leqslant f \leqslant b < \infty$, a. e. , 则存在一个常数 $c \in [a, b]$, 使 $\int_\Omega f|g|\mathrm{d}\mu = c\int_\Omega |g|\mathrm{d}\mu$. 特别, 若 μ 有限, 则 $\int f\mathrm{d}\mu = c\mu(\Omega)$.

证明: 我们知道 g 是可积的, 因此 $\int_\Omega |g|\mathrm{d}\mu < \infty$, 故 $|g|$ 可积. 我们知道 f 是有界的, 因此 $\int_\Omega |f|g||\mathrm{d}\mu = \int_\Omega |f||g|\mathrm{d}\mu < \infty$, 故 $f|g|$ 亦可积. 考虑连续函数

$$F(x) = x\int_\Omega |g|\mathrm{d}\mu - \int_\Omega f|g|\mathrm{d}\mu,$$

则 $F(a) \leqslant 0 \leqslant F(b)$, 故 $\exists c \in [a, b]$, 使得 $F(c) = c\int_\Omega |g|\mathrm{d}\mu - \int_\Omega f|g|\mathrm{d}\mu = 0$.

若 μ 有限, 取 $g = 1$ 便得 $\int_\Omega f\mathrm{d}\mu = c\mu(\Omega)$. □

5.2.3 设 f, g 是 $(\Omega, \mathcal{F}, \mu)$ 上取有限值的 $\mathcal{F}$ 简单函数, 若 f, g 之一对 μ 可积, 则 fg 也可积.

证明: 我们知道 f, g 都是有限简单函数, 设 $f = \sum\limits_{k=1}^m x_k \mathbb{1}_{A_k}$, $A_1, A_2, \cdots, A_m$ 两两不交且 $\bigcup\limits_{k=1}^m A_k = \Omega$. 同时设 $g = \sum\limits_{j=1}^n y_j \mathbb{1}_{B_j}$, $B_1, B_2, \cdots, B_n$ 两两不交且 $\bigcup\limits_{j=1}^n B_j = \Omega$. 不妨设 f 是对 μ 可积的, 则由引理 5.2.4 知

$$\int_\Omega |f|\mathrm{d}\mu = \sum_{k=1}^m |x_k|\mu(A_k) < \infty.$$

我们知道 $fg=\sum_{k=1}^{m}\sum_{j=1}^{n}x_ky_j\mathbb{1}_{A_k\cap B_j}$, 故

$$\begin{aligned}\int_\Omega |fg|\mathrm{d}\mu &= \sum_{k=1}^{m}\sum_{j=1}^{n}|x_ky_j|\mu(A_k\cap B_j)\leqslant \sum_{k=1}^{m}\sum_{j=1}^{n}|x_ky_j|\mu(A_k)\\ &\leqslant \max_j|y_j|\sum_{k=1}^{m}|x_k|\mu(A_k)=\max_j|y_j|\int_\Omega|f|\,\mathrm{d}\mu<\infty\end{aligned}$$

(其中 $\max_j|y_j|<\infty$, 因为 g 取有限值). 因此 fg 也可积. □

5.2.4 设 f 是 $(\Omega,\mathcal{F},\mu)$ 上的可积函数, 则对每一个正数 $\varepsilon,\mu(\{\omega:|f(\omega)|\geqslant\varepsilon\})<\infty$.

证明: 我们考虑反证法, 假设存在某个 $\varepsilon_0>0$, 使得 $\mu(\{\omega:|f(\omega)|\geqslant\varepsilon_0\})=\infty$, 则

$$\int_\Omega f\mathrm{d}\mu\geqslant\int_\Omega f\mathbb{1}_{\{|f|\geqslant\varepsilon_0\}}\geqslant\int_\Omega\varepsilon_0\mathbb{1}_{\{|f|\geqslant\varepsilon_0\}}=\infty.$$

这与可积矛盾! □

5.2.5 设 Ω 是全体正整数组成的空间, $\mathcal{F}$ 是 Ω 的一切子集作成的 σ 代数, 对于 $A\in\mathcal{F},\mu(A)=|A|$($A$ 中元素的个数), 则 $(\Omega,\mathcal{F},\mu)$ 是一个测度空间. 称此 μ 为计数测度. 讨论此空间上的函数的积分存在的充分与必要条件.

证明: 若 f 是非负简单函数, $f=\sum_{k=1}^{m}x_k\mathbb{1}_{A_k}$, 且 $\bigcup_{k=1}^{m}A_k=\Omega$, $A_1,A_2,\cdots,A_m$ 两两不交. 我们有

$$\int_\Omega f\mathrm{d}\mu=\sum_{k=1}^{m}x_k\mu(A_k)=\sum_{l=1}^{\infty}f(l).$$

若 f 是非负可测函数, 考虑非负简单函数 $f_n=f\mathbb{1}_{\Omega_n}$, 这里 $\Omega_n=\{1,2,\cdots,n\}$. 则 $f_n\uparrow f$, 故

$$\int_\Omega f\mathrm{d}\mu=\lim_{n\to\infty}\int_\Omega f_n\mathrm{d}\mu=\lim_{n\to\infty}\sum_{l=1}^{n}f(l)=\sum_{l=1}^{\infty}f(l).$$

若 f 是一般的可测函数, 积分存在等价于 $\int_\Omega f^+\mathrm{d}\mu$ 和 $\int_\Omega f^-\mathrm{d}\mu$ 中有至少一个有限, 即 $\sum_{n=1}^{\infty}\max\{f(n),0\}=\sum_{n=1}^{\infty}\frac{|f(n)|+f(n)}{2}$, $\sum_{n=1}^{\infty}\max\{-f(n),0\}=\sum_{n=1}^{\infty}\frac{|f(n)|-f(n)}{2}$ 中至少有一个有限, 也即 $\sum_{n=1}^{\infty}(|f(n)|+f(n))$ 和 $\sum_{n=1}^{\infty}(|f(n)|-f(n))$ 中有至少一个有限. □

5.2.6 设 f,g 是 $(\Omega,\mathcal{F})$ 上的可测函数, $\mathscr{P}(\mathcal{F})$ 表示 $\mathcal{F}$ 上概率测度的全体, 若 $\forall\nu\in\mathscr{P}(\mathcal{F}),\int f\mathrm{d}\nu=\int g\mathrm{d}\nu$, 则 $f(\omega)=g(\omega),\forall\omega\in\Omega$ 成立.

证明: 采用反证法. 假设 $\exists\omega_0\in\Omega$, 使得 $f(\omega_0)\neq g(\omega_0)$. 考虑单点概率测度 $\mathbb{P}_0(A)=\begin{cases}1,\omega_0\in A\\0,\omega_0\notin A\end{cases}$ (请读者自行证明这的确是一个测度), 则 $\int_\Omega f(\omega)\mathrm{d}\mathbb{P}_0=f(\omega_0)$(因为 $\int_\Omega f(\omega)\mathrm{d}\mathbb{P}_0=\int_{\{\omega_0\}}f(\omega)\mathrm{d}\mathbb{P}_0+\int_{\{\omega_0\}^c}f(\omega)\mathrm{d}\mathbb{P}_0=$

$f(\omega_0)\mathbb{P}_0(\{\omega_0\}) + 0 = f(\omega_0))$. 同理 $\int_\Omega g(\omega)\,\mathrm{d}\mathbb{P}_0 = g(\omega_0)$. 而由题设 $\int_\Omega g(\omega)\,\mathrm{d}\mathbb{P}_0 = \int_\Omega f(\omega)\,\mathrm{d}\mathbb{P}_0$, 故 $f(\omega_0) = g(\omega_0)$, 矛盾! □

§ 5.3 期望的性质及 L-S 积分表示

5.3.1 设 ξ, η 为实 r.v., $\mathbb{E}\xi^2, \mathbb{E}\eta^2$ 有限, 试证: $\mathbb{D}(\xi+\eta) = \mathbb{D}\xi + \mathbb{D}\eta$ 的充要条件是 $\mathbb{E}(\xi\eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$.

证明: 我们知道 $\mathbb{E}\xi^2$ 有限, 因此 $\mathbb{E}\xi\mathbb{E}\xi \leqslant \mathbb{E}\xi^2 < \infty$, 故 $\mathbb{E}\xi$ 有限. 同理 $\mathbb{E}\eta, \mathbb{E}\xi \cdot \mathbb{E}\eta$ 有限. 故

$$\begin{aligned}
\mathbb{D}(\xi+\eta) &= \mathbb{E}|\xi+\eta|^2 - |\mathbb{E}(\xi+\eta)|^2 \\
&= \mathbb{E}|\xi|^2 + \mathbb{E}|\eta|^2 + 2\mathbb{E}(\xi\eta) - (|\mathbb{E}\xi|^2 + 2\mathbb{E}\xi \cdot \mathbb{E}\eta + |\mathbb{E}\eta|^2) \\
&= \mathbb{D}\xi + \mathbb{D}\eta + 2(\mathbb{E}(\xi\eta) - \mathbb{E}\xi \cdot \mathbb{E}\eta).
\end{aligned}$$

故 $\mathbb{D}(\xi+\eta) = \mathbb{D}\xi + \mathbb{D}\eta \iff \mathbb{E}(\xi\eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$. □

5.3.2 若 $X_1, X_2, \cdots, X_n$ 是 n 个实 r.v., 其分布函数分别为 $F_1(x), F_2(x), \cdots, F_n(x), \eta$ 是一 r.v., 其分布函数为 $\frac{1}{n}(F_1(x) + F_2(x) + \cdots + F_n(x))$. 设 r 为一正数, 证明

$$\mathbb{E}|\eta|^r = \frac{1}{n}\sum_{k=1}^n \mathbb{E}\,|X_n|^r\,.$$

证明: 根据习题 5.2.1, 我们有:

$$\mathbb{E}|\eta|^r = \int_\Omega |x|^r \mathrm{d}\mathbb{P}_\eta = \int_\Omega |x|^r \mathrm{d}\left(\frac{1}{n}\sum_{k=1}^n F_n\right) = \frac{1}{n}\sum_{k=1}^n \int_\Omega |x|^r \mathrm{d}F_k = \frac{1}{n}\sum_{k=1}^n \mathbb{E}\,|X_n|^r\,.$$

□

5.3.3 设 X 是一实 r.v., 它的分布函数是 $F(x)$, c 为一正数, 令

$$X^c := \begin{cases} X, & \text{当}|X| < c, \\ c, & \text{当}X \geqslant c, \\ -c, & \text{当}X \leqslant -c, \end{cases}$$

试将 $\mathbb{E}X^c, \mathbb{D}X^c$ 用对于 F 的 L-S 积分表出.

证明: 我们有:

$$\begin{aligned}
\mathbb{E}X^c = \int_{\mathbb{R}} X\,\mathrm{d}F_X &= \int_{-c}^{c} X\,\mathrm{d}F_X + \int_c^\infty c\mathrm{d}F_X + \int_{-\infty}^{-c} (-c)\mathrm{d}F_X \\
&= \int_{-c}^{c} X\,\mathrm{d}F_X + c(1 - F(c)) - cF(-c).
\end{aligned}$$

类似地,

$$\begin{aligned}\mathbb{D}X^c &= \mathbb{E}(X^c)^2 - (\mathbb{E}X^c)^2 \\ &= \int_{-c}^{c} X^2 \mathrm{d}F_X - \left(\int_{-c}^{c} X \mathrm{d}F_X + c(1-F(c)) - cF(-c)\right)^2 \\ &= \int_{-c}^{c} X^2 \mathrm{d}F_X + c^2(1-F(c)) - c^2F(-c) - \left(\int_{-c}^{c} X \mathrm{d}F_X + c(1-F(c)) - cF(-c)\right)^2.\end{aligned}$$

□

5.3.4 设 $F(x)$ 是一分布函数, 按定义

$$\int_a^b f(x)\mathrm{d}F(x) = \int_{(a,b]} f(x)\mathrm{d}F(x),$$

问它是否等于 $\int_{[a,b]} f(x)\mathrm{d}F(x)$? 在什么情况下它们不相等?

证明: 不一定相等. 实际上,

$$\begin{aligned}\left|\int_{[a,b]} f(x)\mathrm{d}F(x) - \int_a^b f(x)\mathrm{d}F(x)\right| &= \left|\lim_{\varepsilon\to 0}\int_{[a,a+\varepsilon)} f(x)\mathrm{d}F(x)\right| \\ &\geqslant \lim_{\varepsilon\to 0}\min_{a\leqslant x<a+\varepsilon}|f(x)|(F(a+\varepsilon-)-F(a-)| \\ &= |f(a)|(F(a)-F(a-)).\end{aligned}$$

故若 F 在 a 处不左连续时, 两者不相等.

反例: 考虑 $F(x) = \operatorname{sgn} x$, $f(0) \neq 0$, 则 $\int_0^1 f(x)\mathrm{d}F(x) = 0 \neq f(0) = \int_{[0,1]} f(x)\mathrm{d}F(x)$. □

5.3.5 设 X 是一实 r.v., m 是一实数, 它满足:

$$\mathbb{P}(\{X \geqslant m\}) \geqslant \frac{1}{2}, \mathbb{P}(\{X \leqslant m\}) \geqslant \frac{1}{2},$$

称满足上述条件的 m 为 X 的中数, 试证

(1) $\forall a \in \mathbb{R}$, $\mathbb{E}|X-m| \leqslant \mathbb{E}|X-a|$;

(2) X 的中数, 数学期望和方差之间有如下关系:

$$\mathbb{E}X - \sqrt{\mathbb{D}X} \leqslant m \leqslant \mathbb{E}X + \sqrt{\mathbb{D}X}.$$

证明: (1) 我们只需证明 $\forall a \in \mathbb{R}$, 有 $\mathbb{E}(|X-a|-|X-m|) \geqslant 0$. 不妨设 $a > m$, 则

$$|X-a|-|X-m| = \begin{cases} m-a, & X > a \\ m+a-2X, & m < X \leqslant a \\ a-m, & X \leqslant m \end{cases}.$$

故

$$\begin{aligned}\mathbb{E}(|X-a|-|X-m|)&=(m-a)\mathbb{P}(\{X>a\})+(a-m)\mathbb{P}(\{X\leqslant m\})+\int_{(m,a]}(m+a-2X)\mathrm{d}\mathbb{P}\\&\geqslant(m-a)\mathbb{P}(\{X>a\})+(a-m)\mathbb{P}(\{X\leqslant m\})+(a-m)\mathbb{P}(\{m<X\leqslant a\})\\&=(a-m)(\mathbb{P}(\{X<a\})-\mathbb{P}(\{X>a\}))\\&\geqslant(a-m)(\mathbb{P}(\{X\leqslant m\})-\mathbb{P}(\{X>m\}))\\&=(a-m)(2\mathbb{P}(\{X\leqslant m\})-1)=0.\end{aligned}$$

而 $a=m$ 时结论显然, $a<m$ 时与上文同理即证.

(2) 即证 $|\mathbb{E}X-m|\leqslant\sqrt{\mathbb{D}X}$.

在 (1) 中取 $a=\mathbb{E}X$, 由 Cauchy-Schwarz 不等式, 便有

$$|\mathbb{E}X-m|\overset{\text{引理 5.2.3(2)}}{\leqslant}\mathbb{E}|X-m|\overset{\text{(1) 小问结论}}{\leqslant}\mathbb{E}|X-\mathbb{E}X|\overset{\text{Cauchy-Schwarz 不等式}}{\leqslant}\sqrt{\mathbb{E}|X-\mathbb{E}X|^2}=\sqrt{\mathbb{D}X}.$$

故 $\mathbb{E}X-\sqrt{\mathbb{D}X}\leqslant m\leqslant\mathbb{E}X+\sqrt{\mathbb{D}X}$. □

5.3.6 设 $X=\sum_{k\in I}a_k\mathbb{1}_{\{X=a_k\}}$, $a_k\in\mathbb{R}$, $I\subset\mathbb{N}$, 试证:

$$\mathbb{P}(X\in B)=\sum_{k\in I:a_k\in B}\mathbb{P}_X(\{a_k\}),\quad \mathbb{E}\left[g(X)\right]=\sum_{k\in I}g(a_k)\mathbb{P}_X(\{a_k\}).$$

证明: 根据定理 5.3.12, 我们有

$$\mathbb{P}(X\in B)=\int_B\mathrm{d}\mathbb{P}=\sum_{k\in I:a_k\in B}\mathbb{P}(\{X=a_k\})=\sum_{k\in I:a_k\in B}\mathbb{P}_X(\{a_k\}).$$

同时, 根据定理 5.3.13, 有

$$\mathbb{E}\left[g(X)\right]=\int_{\mathbb{R}}g\mathrm{d}\mathbb{P}=\sum_{k\in I}g(a_k)\mathbb{P}_X(\{a_k\}).$$

□

5.3.7 设 X_1,X_2 是在 $[a,b]$ 上均匀分布的独立 r.v., 求 $Y=X_1+X_2$ 的分布函数与分布密度 (可以由公式用数学分析计算, 也可用几何求面积方法计算).

证明: 我们知道 (X_1,X_2) 的分布密度为 $p_{(X_1,X_2)}(X_1,X_2)=p_1(X_1)p_2(X_2)=\dfrac{1}{(b-a)^2}$, 故有:

$$\begin{aligned}F_Y(y)=\mathbb{P}(Y\leqslant y)&=\iint_{[a,b]^2}\mathbb{1}_{X_1+X_2\leqslant y}\frac{1}{(b-a)^2}\mathrm{d}X_1\mathrm{d}X_2\\&=\mathbb{1}_{\{2a\leqslant y<a+b\}}\int_a^{y-a}\int_a^{y-x_1}\frac{1}{(b-a)^2}\mathrm{d}X_2\mathrm{d}X_1\\&\quad+\mathbb{1}_{\{a+b\leqslant y<2b\}}\left(1-\int_{y-b}^b\int_{y-x_1}^b\frac{1}{(b-a)^2}\mathrm{d}X_2\mathrm{d}X_1\right)+\mathbb{1}_{\{y\geqslant 2b\}}\\&=\frac{(y-a)^2}{2(b-a)^2}\mathbb{1}_{\{2a\leqslant y<a+b\}}+\left(1-\frac{(2b-y)^2}{2(b-a)^2}\right)\mathbb{1}_{\{a+b\leqslant y<2b\}}+\mathbb{1}_{\{y\geqslant 2b\}}.\end{aligned}$$

□

5.3.8 设 X,Y 为独立 r.v., X 在 $[0,1]$ 上均匀分布, Y 按二项分布律 $B(n,p)$ 分布, 试证 $X+Y$ 是连续型随机变量, 并求其分布密度.

证明: 考虑 $z \geqslant n+1$, 则 $\mathbb{P}(X+Y \leqslant z)=1$;

若 $0 \leqslant z < n+1$, 则

$$\begin{aligned}\mathbb{P}(X+Y \leqslant z) &= \sum_{k=0}^{[z]-1} \mathbb{P}(Y=k)+\mathbb{P}(Y=[z])\mathbb{P}(X<z-[z]) \\ &= \sum_{k=0}^{[z]-1} \binom{n}{k} p^k(1-p)^{n-k} + \binom{n}{[z]} p^{[z]}(1-p)^{n-[z]}(z-[z]).\end{aligned}$$

这里 $[z]$ 是 Gauss 取整函数. 故其分布密度为

$$p_Z(z)=\begin{cases}\dbinom{n}{[z]} p^{[z]}(1-p)^{n-[z]}, & 0 \leqslant z < n+1, \\ 0, & \text{others}.\end{cases}$$

因此 $z=X+Y$ 是连续型随机变量. □

5.3.9 (1) 设 $X=(X_1,X_2)$ 的分布密度为 $p_X(x_1,x_2)$,

$$Y_1 := \sqrt{X_1^2+X_2^2}, \quad Y_2 := \frac{X_1}{X_2},$$

试求 $Y=(Y_1,Y_2)$ 的分布密度 $p_Y(y_1,y_2)$.

(2) 设 $X=(X_1,X_2)$ 的分布密度 $p_X(x_1,x_2) := \dfrac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x_1^2+x_2^2}{2\sigma^2}}$, 试求 (1) 中定义的 $Y=(Y_1,\ Y_2)$ 的分布密度, 并证明 Y_1,Y_2 独立.

证明: (1) 解方程 $y_1=\sqrt{x_1^2+x_2^2}, y_2=\dfrac{x_1}{x_2}$, 便有 $\begin{cases} x_1^{(1)} = \dfrac{y_1y_2}{\sqrt{1+y_1^2}} \\ x_2^{(1)} = \dfrac{y_1}{\sqrt{1+y_1^2}}\end{cases}$, $\begin{cases} x_1^{(2)} = -\dfrac{y_1y_2}{\sqrt{1+y_1^2}} \\ x_2^{(2)} = -\dfrac{y_1}{\sqrt{1+y_1^2}}\end{cases}$. 因此

$$\begin{aligned} p_Y(y_1,y_2) &= p_x\left(\frac{y_1y_2}{\sqrt{1+y_1^2}}, \frac{y_1}{\sqrt{1+y_1^2}}\right)\left|\frac{\partial(x_1^{(1)},x_2^{(1)})}{\partial(y_1,y_2)}\right| + p_x\left(-\frac{y_1y_2}{\sqrt{1+y_1^2}}, -\frac{y_1}{\sqrt{1+y_1^2}}\right)\left|\frac{\partial(x_1^{(2)},x_2^{(2)})}{\partial(y_1,y_2)}\right| \\ &= \begin{cases} \dfrac{y_1}{1+y_2^2}\left(p_x\left(\dfrac{y_1y_2}{\sqrt{1+y_1^2}}, \dfrac{y_1}{\sqrt{1+y_1^2}}\right) + p_x\left(-\dfrac{y_1y_2}{\sqrt{1+y_1^2}}, -\dfrac{y_1}{\sqrt{1+y_1^2}}\right)\right), & y_1>0, \\ 0, & y_1 \leqslant 0. \end{cases}\end{aligned}$$

(2) 代入便有

$$p_Y(y_1,y_2)=\frac{1}{\pi\sigma^2}\frac{y_1}{1+y_2^2}\exp\left(-\frac{y_1^2}{2\sigma^2}\right),$$

我们有

$$\mathbb{P}(y_1 \leqslant t) = \iint_{\mathbb{R}^2} \mathbb{1}_{\{x_1^2+x_2^2 \leqslant t^2\}} p_X(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 = \frac{1}{2\pi\sigma^2}\iint_{x_1^2+x_2^2\leqslant t^2}\exp\left(-\frac{x_1^2+x_2^2}{2\sigma^2}\right)\mathrm{d}x_1\mathrm{d}x_2,$$

考虑极坐标换元 $x_1 = r\cos\theta$, $x_2 = r\sin\theta$, $0 \leqslant r \leqslant t$, $\theta \in [0, 2\pi)$. 则 $\left|\dfrac{\partial(x_1, x_2)}{\partial(r, \theta)}\right| = r$, 故

$$\mathbb{P}(Y_1 \leqslant t) = \frac{1}{2\pi\sigma^2}\int_0^{2\pi}\int_0^t r\exp\left(-\frac{r^2}{2\sigma^2}\right)\mathrm{d}r\,\mathrm{d}\theta = 1 - \exp\left(-\frac{t^2}{2\sigma^2}\right),$$

因此 $p_{Y_1}(y_1) = \dfrac{\mathrm{d}}{\mathrm{d}y_1}\mathbb{P}(Y_1 \leqslant y_1) = \dfrac{y_1}{\sigma^2}\exp\left(-\dfrac{y_1^2}{2\sigma^2}\right)$. 同时,

$$\begin{aligned}
\mathbb{P}(Y_2 \leqslant t) &= \mathbb{P}\left(X_1 \geqslant 0, \frac{X_1}{X_2} \leqslant t\right) + \mathbb{P}\left(X_1 < 0, \frac{X_1}{X_2} \leqslant t\right) \\
&= \int_0^\infty\int_{\mathbb{R}} \mathbb{1}_{\{\frac{X_1}{X_2} \leqslant t\}} p_X(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 + \int_{-\infty}^0\int_{\mathbb{R}} \mathbb{1}_{\{\frac{X_1}{X_2} \leqslant t\}} p_X(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= 2\cdot\frac{1}{2\pi\sigma^2}\int_0^\infty\int_{\mathbb{R}} \mathbb{1}_{\{\frac{X_1}{X_2} \leqslant t\}}\exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \frac{1}{\pi\sigma^2}\int_0^\infty\int_{\mathbb{R}} \mathbb{1}_{\{u \leqslant t\}}\exp\left(-\frac{(1+u^2)x_2^2}{2\sigma^2}\right)\mathrm{d}(ux_2)\mathrm{d}x_2 \\
&= \frac{1}{\pi\sigma^2}\int_0^\infty\int_{-\infty}^t x_2\exp\left(-\frac{(1+u^2)x_2^2}{2\sigma^2}\right)\mathrm{d}u\mathrm{d}x_2 \\
&= \frac{1}{\pi\sigma^2}\int_{-\infty}^t\int_0^\infty x_2\exp\left(-\frac{(1+u^2)x_2^2}{2\sigma^2}\right)\mathrm{d}x_2\mathrm{d}u \\
&= \frac{1}{\pi}\int_{-\infty}^t \frac{1}{1+u^2}\mathrm{d}u.
\end{aligned}$$

故 $p_{Y_2}(y_2) = \dfrac{\mathrm{d}}{\mathrm{d}y_2}\mathbb{P}(Y_2 \leqslant y_2) = \dfrac{1}{\pi}\cdot\dfrac{1}{1+y_2^2}$. 故 $P_Y(y_1, y_2) = p_{Y_1}(y_1)p_{Y_2}(y_2)$, 故 Y_1, Y_2 独立. □

5.3.10 设 $X_1, X_2, \cdots, X_n$ 是具有相同分布密度 $p(x)$ 的独立 r.v., 且当 $x < 0$ 时 $p(x) = 0$, 当 $x \geqslant 0$ 时 $p(x)$ 连续, 其次设 ξ_k^* 为 $X_1, X_2, \cdots, X_n$ 按不降顺序排列的第 k 个值, 试证: $(\xi_1^*,\ \xi_2^*, \cdots, \xi_r^*)$, $1 \leqslant r \leqslant n$ 的分布密度为

$$p(y_1, y_2, \cdots, y_r) = \begin{cases} \dfrac{n!}{(n-r)!}p(y_1)\,p(y_2)\cdots p(y_r)\left(\displaystyle\int_{y_r}^\infty p(x)\mathrm{d}x\right)^{n-r}, & 0 \leqslant y_1 \leqslant y_2 \leqslant \cdots \leqslant y_r, \\ 0, & \text{其他}. \end{cases}$$

证明: 我们有:

$$\begin{aligned}
\mathbb{P}(\{\xi_k^* \leqslant y_k, k \in \mathbb{N}, k \leqslant r\}) &= \mathbb{P}\left(\bigcap_{k=1}^r\{\xi_k^* \leqslant y_k\}\right) = 1 - \mathbb{P}\left(\bigcup_{k=1}^r\{\xi_k^* \geqslant y_k\}\right) \\
&\xlongequal{\text{抽屉原理（定义 3.3.1(5)）}} 1 - \sum_{j=1}^r(-1)^{j-1}\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant r}\mathbb{P}\left(\bigcap_{l=1}^j\{\xi_{i_l}^* \geqslant y_{i_l}\}\right).
\end{aligned}$$

记上述概率为 $\mathbb{P}_0$, 我们知道,

$$\mathbb{P}\left(\bigcap_{l=1}^j\{\xi_{i_l}^* \geqslant y_{i_l}\}\right) = \frac{n!}{(n-j)!}\mathbb{1}_{y_{i_1}\leqslant\cdots\leqslant y_{i_j}}\prod_{k=1}^j\left[\mathbb{P}(y_{i_{k-1}} < X_1 \leqslant \xi_{i_k}^*)\right]^{i_k - i_{k-1}}\left[\mathbb{P}(X_1 > y_{i_j})\right]^{n-i_j},$$

故

$$\mathbb{P}_0 = 1 - \sum_{j=1}^r(-1)^{j-1}\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant r}\mathbb{1}_{y_{i_1}\leqslant\cdots\leqslant y_{i_j}}\frac{n!}{(n-j)!}\prod_{k=1}^j\left[\mathbb{P}(y_{i_{k-1}} < X_1 \leqslant \xi_{i_k}^*)\right]^{i_k - i_{k-1}}\left[\mathbb{P}(X_1 > y_{i_j})\right]^{n-i_j},$$

我们只需考虑含有所有 y_k, $1 \leqslant k \leqslant r$ 的项. 有

$$\begin{aligned}
p(y_1, y_2, \cdots, y_r) &= \frac{\partial^r \mathbb{P}_0}{\partial y_1 \partial y_2 \cdots \partial y_r} \\
&= \frac{\partial^r}{\partial y_1 \partial y_2 \cdots \partial y_r} \mathbb{1}_{y_{i_1} \leqslant \cdots \leqslant y_{i_j}} \frac{n!(-1)^r}{(n-r)!} \left[\mathbb{P}(X_1 > y_r)]^{n-r} \prod_{k=1}^{r} [\mathbb{P}(y_{k-1} < X_1 \leqslant \xi_k^*) \right] \\
&= \frac{n!(-1)^r}{(n-r)!} \frac{\partial^r}{\partial y_1 \partial y_2 \cdots \partial y_r} F(y_1) \left(\prod_{k=1}^{r-1} \left[F(\xi_{k+1}^*) - F(y_k) \right] \right) [1 - F(y_r)]^{n-r} \mathbb{1}_{y_1 \leqslant \cdots \leqslant y_r} \\
&= \frac{n!}{(n-r)!} p(y_1) p(y_2) \cdots p(y_r) \left(\int_{y_r}^{\infty} p(x) \mathrm{d}x \right)^{n-r} \mathbb{1}_{y_1 \leqslant \cdots \leqslant y_r}.
\end{aligned}$$

□

5.3.11 设 $F(x)$ 和 $G(x)$ 是两个有界分布函数 (不一定是概率分布函数), $G(-\infty) = 0, f(x)$ 是连续函数, 且 $0 < c_1 \leqslant f(x) < c_2 < \infty, \forall x \in \mathbb{R}$, 试应用定理 5.3.10 证明: 若

$$F(x) = \int_{-\infty}^{x} \frac{1}{f(x)} \mathrm{d}G(x), \quad \forall x \in \mathbb{R},$$

则

$$G(x) = \int_{-\infty}^{x} f(x) \mathrm{d}F(x), \quad \forall x \in \mathbb{R}.$$

证明: 考虑 $F(x) = \mu((-\infty, x])$, $G(x) = \nu((-\infty, x])$. 则 $F(x) = \displaystyle\int_{(-\infty,x]} \frac{1}{f(x)} \mathrm{d}\nu$. 故 $\mu(A) = \displaystyle\int_A \frac{1}{f} \mathrm{d}\nu$. 故

$$\int_{(-\infty,x]} f(x) \mathrm{d}F(x) = \int_{(-\infty,x]} f(x) \mathrm{d}\mu = \int_{(-\infty,x]} \mathrm{d}\nu = \nu((-\infty, x]) = G(x).$$

□

5.3.12 设 $X_1, X_2, \cdots$ 是无穷个独立 r.v.(即其中任意有限个都独立), 且它们的分布都是以 $\lambda(\lambda > 0)$ 为参数的指数分布, 即 $\mathbb{P}(X_k > t) = \mathrm{e}^{-\lambda t}$, $t \geqslant 0$, 令

$$N(t) := \sup \left\{ n \in \mathbb{N} : \sum_{k=1}^{n} X_k \leqslant t \right\},$$

试证:

(1) 若 $0 < s < t < \infty$, 则 $N(s)$ 与 $N(t) - N(s)$ 独立, 且分别服从参数为 λs 及 $\lambda(t-s)$ 的 Poisson 分布;

(2) $\forall m$ 及任何 $0 = t_0 < t_1 < \cdots < t_n < \infty$, $N(t_k) - N(t_{k-1})$, $k = 1, 2, \cdots, m$ 独立, 且 $N(t_k) - N(t_{k-1})$ 服从以 $\lambda(t_k - t_{k-1})$ 为参数的 Poisson 分布, $k = 1, 2, \cdots, m$.

证明: (1) 我们知道 $X_1, X_2, \cdots \overset{\text{i.i.d}}{\sim} \Gamma(\lambda, 1)$, 则 $\sum\limits_{k=1}^{n} X_k \sim \Gamma(\lambda, n)$. 实际上, 我们有

$$\begin{aligned}
\mathbb{P}(N(s)=n) &= \mathbb{P}\left(\sum_{k=1}^{n} X_k \leqslant s\right) - \mathbb{P}\left(\sum_{k=1}^{n+1} X_k \leqslant s\right) \\
&= \int_0^s \frac{\lambda^n}{\Gamma(n)} x^{n-1} \mathrm{e}^{-\lambda x} \mathrm{d}x - \int_0^t \frac{\lambda^{n+1}}{\Gamma(n+1)} x^n \mathrm{e}^{-\lambda x} \mathrm{d}x \\
&= \frac{\lambda^n}{n!} \int_0^s (n-\lambda x) x^{n-1} \mathrm{e}^{-\lambda x} \mathrm{d}x \\
&= \frac{\lambda^n}{n!} \left. x^n \mathrm{e}^{-\lambda x} \right|_0^s = \frac{\lambda^n s^n \mathrm{e}^{-\lambda s}}{n!}.
\end{aligned}$$

因此 $N(s)$ 服从参数为 λs 的 Poisson 分布. 置 $S_n = \sum\limits_{k=1}^{n} X_k \sim \Gamma(\lambda, n)$, 我们有

$$\begin{aligned}
\mathbb{P}(N(s)=n, N(t)=m+n) &= \mathbb{P}\left(S_n \leqslant s, s-S_n < X_{n+1} < t-S_n, \sum_{k=n+2}^{m+n} X_k \leqslant t-S_n-X_{n+1} < \sum_{k=n+2}^{m+n+1} X_k\right) \\
&= \mathbb{P}\left(S_n \leqslant s, s-S_n < X_{n+1} < t-S_n, S_{m-1} \leqslant t-S_n-X_{n+1} < S_m\right) \\
&= \int_0^s \int_{s-x}^{y-s} \mathbb{P}(N(t-x-y)=m-1) \mathrm{d}\mathbb{P}(X_{n+1} \leqslant y) \mathrm{d}\mathbb{P}(S_n \leqslant x) \\
&= \int_0^s \int_{s-x}^{y-s} \frac{\lambda^{m-1}(t-x-y)^{m-1}\mathrm{e}^{-\lambda(t-x-y)}}{(m-1)!} \lambda \mathrm{e}^{-\lambda y} \frac{\lambda^{n+1}}{\Gamma(n)} x^{n-1} \mathrm{e}^{-\lambda x} \mathrm{d}y \mathrm{d}x \\
&= \frac{\lambda^{m+1}\mathrm{e}^{-\lambda t}}{(m-1)!(n-1)!} \int_0^s \int_{s-x}^{y-s} (t-x-y)^{m-1} x^{n-1} \mathrm{d}y \mathrm{d}x \\
&= \frac{\lambda^{m+1}\mathrm{e}^{-\lambda t}}{(m-1)!(n-1)!} \frac{(t-s)^m s^n}{mn} \\
&= \frac{\lambda^n s^n \mathrm{e}^{-\lambda s}}{n!} \cdot \frac{\lambda^m (t-s)^m \mathrm{e}^{-\lambda(t-s)}}{m!}.
\end{aligned}$$

同时, 我们有

$$\begin{aligned}
\mathbb{P}(N(t)-N(s)=m) &= \sum_{j=0}^{\infty} \mathbb{P}(N(s)=j, N(t)=m+j) \\
&= \sum_{j=1}^{\infty} \frac{\lambda^j s^j \mathrm{e}^{-\lambda s}}{j!} \cdot \frac{\lambda^m (t-s)^m \mathrm{e}^{-\lambda(t-s)}}{m!} \\
&= \frac{\lambda^m (t-s)^m \mathrm{e}^{-\lambda(t-s)}}{m!}.
\end{aligned}$$

因此 $\mathbb{P}(N(s)=n, N(t)-N(s)=m) = \mathbb{P}(N(s)=n)\mathbb{P}(N(t)-N(s)=m)$, 因此它们相互独立, 且分别服从参数为 λs 及 $\lambda(t-s)$ 的 Poisson 分布.

(2) 用数学归纳法证明. 由 (1) 知 $m=1$ 时结论成立. 不妨设 $m=n$ 时结论成立, 令

$$A = \{N(t_1)-N(t_0)=k_1, \cdots, N(t_n)-N(t_{n-1})=k_n\},$$

则有

$$\mathbb{P}(A \cap \{N(t_{n+1})-N(t_n)=k\}) = \mathbb{P}(N(t_{n+1})-N(t_n)=k | A)\mathbb{P}(A).$$

又根据指数分布的无记忆性, 容易得到

$$\mathbb{P}(N(t_{n+1})-N(t_n)=k|A)=\frac{[\lambda(t_{n+1}-t_n)]^k}{k!}\mathrm{e}^{-\lambda(t_{n+1}-t_n)}.$$

□

5.3.13 设 $X_1,X_2,\cdots,Y_1,Y_2,\cdots$ 是无穷个独立的非负 r.v., X_k 都服从以 $\lambda(\lambda>0)$ 为参数的指数分布, Y_k 服从集中在 $(0,\infty)$ 上的分布 μ, 令 $N(t)$ 为一 r.v., 它满足:

$$\begin{aligned}\{N(t)=n\}=\{&X_1+X_2+\cdots+X_n+Y_1+Y_2+\cdots+Y_n\leqslant t,\\ &X_1+X_2+\cdots+X_{n+1}+Y_1+Y_2+\cdots+Y_n>t\}.\end{aligned}$$

试证:

$$\mathbb{P}(\{N(t)=n\})=\int_{(0,t]}\frac{[\lambda(t-x)]^n}{n!}\mathrm{e}^{-\lambda(t-x)}\mu^{*n}(\mathrm{d}x),$$

其中

$$\mu^{*n}((0,x]):=\mathbb{P}\left(\{Y_1+Y_2+\cdots+Y_n\leqslant x\}\right),$$

即 μ 的 n 重卷积.

证明: 我们知道 $\sum\limits_{k=1}^n Y_k\sim\mu^{*n}$, 定义 $S_n:=\sum\limits_{k=1}^n X_k\sim\Gamma(\lambda,n)$. 则

$$\begin{aligned}\mathbb{P}(N(t)=n)&=\mathbb{P}\left(\sum_{k=1}^n X_k+\sum_{k=1}^n Y_k\leqslant t<\sum_{k=1}^{n+1}X_k+\sum_{k=1}^n Y_k\right)\\ &=\int_0^t\mathbb{P}(S_n\leqslant t-x<S_n+X_{n+1})\mu^{*n}(\mathrm{d}x)\\ &=\int_0^t\mathbb{P}(S_n\leqslant t-x,X_{n+1}>t-x-S_n)\mu^{*n}(\mathrm{d}x)\\ &=\int_0^t\int_0^{t-x}\mathbb{P}(X_{n+1}>t-x-y)\frac{\lambda(\lambda y)^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda y}\mathrm{d}y\\ &=\int_0^t\int_0^{t-x}\mathrm{e}^{-\lambda(t-x-y)}\frac{\lambda(\lambda y)^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda y}\mathrm{d}y\\ &=\int_{(0,t]}\frac{[\lambda(t-x)]^n}{n!}\mathrm{e}^{-\lambda(t-x)}\mu^{*}(\mathrm{d}x).\end{aligned}$$

□

5.3.14 设 $X_1,X_2,\cdots,X_n$ 是独立 r.v., 且都服从 $(0,1)$ 上的均匀分布, 试应用广义加法公式 (或称逐步淘汰原则) 及公式

$$\int\cdots\int_{\sum_{k=1}^n x_k\leqslant x,x_l\geqslant 0}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=\frac{x^n}{n!},$$

证明当 $x\in(0,n)$ 时,

$$\mathbb{P}\left(X_1+X_2+\cdots+X_n\leqslant x\right)=\sum_{k=0}^{n-1}(-1)^k\mathrm{C}_n^k\frac{[(x-k)\vee 0]^n}{n!}.$$

证明： 我们知道 $(X_1,\cdots,X_n)$ 的联合密度为

$$p(x_1,x_2,\cdots,x_n)=\begin{cases}1, & 0\leqslant x_1\leqslant 1, i=1,2,\cdots,n,\\ 0, & \text{otherwise}.\end{cases}$$

因此当 $x\in(0,n)$, 有

$$\begin{aligned}\mathbb{P}(X_1+\cdots+X_n\leqslant x)&=\mathbb{P}(X_1+\cdots+X_n\leqslant x, 0\leqslant X_i\leqslant 1, i=1,\cdots,n)\\&=P(X_1+\cdots+X_n\leqslant x, X_i\geqslant 0, i=1,\cdots,n)\\&\quad-\mathbb{P}(X_1+\cdots+X_n\leqslant x, X_i\geqslant 0, i=1,\cdots,n\exists j, X_j\geqslant 1).\end{aligned}$$

令

$$A_j=\{X_1+\cdots+X_n\leqslant x, X_i\geqslant 0, i=1,\cdots,j-1,j+1,\cdots,n, X_j\geqslant 1\},$$

则 A_j 两两不交, 且 $\bigcup\limits_{j=1}^{n}A_j=\{X_1+\cdots+X_n\leqslant x, X_i\geqslant 0, i=1,\cdots,n\exists j, X_j\geqslant 1\}=:A$, 则

$$\mathbb{P}(A)=\mathbb{P}\left(\bigcup_{j=1}^{n}A_j\right)=\sum_{k=1}^{n}(-1)^{k-1}C_n^k\mathbb{P}(A_1\cap\cdots\cap A_k).$$

实际上, 我们有

$$\begin{aligned}\mathbb{P}(A_1)&=\int\cdots\int_{A_1}p(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n\\&=\int\cdots\int_{\substack{\sum_{k=1}^n x_k\leqslant x\\ x_j\geqslant 0, j\geqslant 2\\ x_1\geqslant 1}}p(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n\\&=\int\cdots\int_{\substack{x_1'+\sum_{k=2}^n x_k\leqslant x-1\\ x_j\geqslant 0, j\geqslant 2\\ x_1'=x_1-x\geqslant 1}}p(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n\\&=\frac{[(x-1)\vee 0]^n}{n!}.\end{aligned}$$

类似地, $\mathbb{P}(A_1\cap\cdots\cap A_k)=\dfrac{[(x-k)\vee 0]^n}{n!}$. 又注意到 $x\in(0,n)$ 时, $\mathbb{P}(\bigcup\limits_{k=1}^{n}A_n)=0$, 以及

$$\mathbb{P}(X_1+\cdots+X_n\leqslant x, X_i\geqslant 0, i=1,2,\cdots,n)=\int\cdots\int_{\substack{\sum_{k=1}^n x_k\leqslant x\\ x_i\geqslant 0,\forall i}}p(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n=\frac{(x\vee 0)^n}{n!},$$

故

$$\mathbb{P}\left(X_1+X_2+\cdots+X_n\leqslant x\right)=\sum_{k=0}^{n-1}(-1)^k\mathrm{C}_n^k\frac{[(x-k)\vee 0]^n}{n!}.$$

□

§ 5.4　积分收敛定理

5.4.1 (Fatou 引理的推广) 设 U_n, V_n 可积, 且 $U_n \to U$, a.e. $V_n \to V$, a.e. $\int U_n \to \int U$ 有限, $\int V_n \to \int V$ 有限.

(1) 若 $U_n \leqslant X_n, \forall n \in \mathbb{N}$, 则 $\int \liminf\limits_{n\to\infty} X_n \leqslant \liminf\limits_{n\to\infty} \int X_n$.

(2) 若 $X_n \leqslant V_n, \forall n \in \mathbb{N}$, 则 $\limsup\limits_{n\to\infty} \int X_n \leqslant \int \limsup\limits_{n\to\infty} X_n$.

证明: (1) 实际上, 我们有:

$$0 \leqslant \inf_{k\leqslant n}(X_k - U_k) \uparrow \sup_n \inf_{k\leqslant n}(X_k - U_k) = \liminf_{n\to\infty}(X_n - U_n),$$

故

$$\begin{aligned}\int \liminf_{n\to\infty} X_n - \int U &= \int \left(\liminf_{n\to\infty} X_n - U\right) = \int \liminf_{n\to\infty}(X_n - U_n) = \lim_{n\to\infty}\int \inf_{k\leqslant n}(X_k - U_k)\\ &\leqslant \liminf_{n\to\infty}\int (X_n - U_n) = \liminf_{n\to\infty}\int X_n - U = \liminf_{n\to\infty}\int X_n - \int U.\end{aligned}$$

故 $\int \liminf\limits_{n\to\infty} X_n \leqslant \liminf\limits_{n\to\infty}\int X_n$.

(2) 注意到 $-V_n \leqslant -X_n$, 且 $-V_n, -X_n$ 可积. 故

$$\limsup_{n\to\infty}\int X_n = -\liminf_{n\to\infty}\int(-X_n) \leqslant -\int \liminf_{n\to\infty}(-X_n) = \int \limsup_{n\to\infty} X_n.$$

□

5.4.2 (控制收敛定理推广) 设 $|X_n| \leqslant U_n$ 而 U_n 可积, $U_n \to U$, a.e. 且 $\int U_n \to \int U$ 有限, 则当 $X_n \to X$, a.e. 时, 有 $\int |X_n - X| \to 0$, 因而 $\int X_n \to \int X$.

证明: 我们知道 $-U_n \leqslant X_n \leqslant U_n$, 因此 X_n 可积. 同时

$$|X_n - X| \leqslant |X_n| + |X| \leqslant U_n + \left|\lim_{n\to\infty} X_n\right| \leqslant U_n + \lim_{n\to\infty}|X_n| \leqslant U_n + \lim_{n\to\infty} U_n = U_n + U,$$

故 $|X_n - X|$ 可积. 且 $|X_n - X| \to 0$, a.e., 同时

$$\liminf_{n\to\infty}\int X_n \leqslant \limsup_{n\to\infty}\int X_n \leqslant \int \limsup_{n\to\infty} X_n = \int \liminf_{n\to\infty} X_n \leqslant \liminf_{n\to\infty}\int X_n.$$

故 $\lim\limits_{n\to\infty}\int X_n = \liminf\limits_{n\to\infty}\int X_n = \limsup\limits_{n\to\infty}\int X_n = \int \liminf\limits_{n\to\infty} X_n = \int X$. 也即 $\int X_n \to \int X$. 故 $\int |X_n - X| \to 0$. □

5.4.3 试证: 给定具有有限期望的随机变量 X 及 $\varepsilon > 0$, 存在一个简单函数 X_ε, 使得 $\mathbb{E}|X - X_\varepsilon| < \varepsilon$, $|X_\varepsilon| \leqslant |X|$, 因而, 存在一个简单函数序列 $\{X_m\}$, 使得 $\forall m \in \mathbb{N}$, $|X_m| \leqslant |X|$, 且 $\lim\limits_{m\to\infty}\mathbb{E}|X - X_m| = 0$.

证明: 我们知道 X 有有限期望, 故在测度空间 $(\Omega, \mathcal{F}, \mu)$ 上, 令 $N := \{\omega : |X(\omega)| = \infty\}$, 有 $\mu(N) = 0$. 令

$$X_{(n)}(\omega) = \sum_{k=-n2^n}^{-1}\frac{k+1}{2^n}\mathbb{1}_{\{\frac{k}{2^n}\leqslant X(\omega)\leqslant\frac{k+1}{2^n}\}} + \sum_{k=0}^{n2^n-1}\frac{k}{2^n}\mathbb{1}_{\{\frac{k}{2^n}\leqslant X(\omega)\leqslant\frac{k+1}{2^n}\}},$$

则 $|X_n| \leqslant |X|$. 且 $|X_{(n)} - X|\mathbb{1}_{N^c} \leqslant \frac{1}{2^n}$. 故 $|X - X_{(n)}| \leqslant \frac{1}{2^n}$, a.e., 因此 $\mathbb{E}|X - X_{(n)}| < \frac{1}{2^n}$. 因此, 只需取 $X_\varepsilon = X_{([-\log_2 \varepsilon]+1)}$, 便有 $\mathbb{E}|X - X_\varepsilon| < \frac{1}{2^{[-\log_2 \varepsilon]+1}} \leqslant \varepsilon$. □

5.4.4 对于 $\mathcal{F}$ 中任何两个集 Λ_1 与 Λ_2 定义 $\rho(\Lambda_1,\Lambda_2) = \mathbb{P}(\Lambda_1\Delta\Lambda_2)$, 则 ρ 是 $\mathcal{F}$ 中的集的空间中的伪度量 (即除 $\rho(\Lambda_1,\Lambda_2) = 0 \to \Lambda_1 = \Lambda_2$ 外, 距离的其他假设都满足); 称引入 ρ 后的空间 $\mathcal{F}$ 为度量空间 $M(\mathcal{F},\rho)$, 证明: 对于每个可积的随机变量 X, 由 $\Lambda \to \int_\Lambda X\,\mathrm{d}\mathbb{P}$ 给出的由 $M(\mathcal{F},\rho)$ 到 $\mathbb{R}$ 的映射是连续的. 类似地, 由 $(\Lambda_1,\Lambda_2) \to \Lambda_1\cup\Lambda_2$, $\Lambda_1\cap\Lambda_2$, $\Lambda_1\backslash\Lambda_2$, $\Lambda_1\Delta\Lambda_2$ 给出的由 $M(\mathcal{F},\rho)\times M(\mathcal{F},\rho)$ 到 $M(\mathcal{F},\rho)$ 的映射都是连续的. 如果去掉一个零概率集后, $\limsup\limits_n \Lambda_n = \liminf\limits_n \Lambda_n$（注: "去掉一个零概率集后 $A = B$" 的意思是 $\mathbb{P}(A\Delta B) = 0$）, 则我们用 $\lim\limits_n \Lambda_n$ 表示这两个集共同的等价类. 证明在这种情况下 $\{\Lambda_n\}$ 按度量 ρ 收敛于 $\lim\limits_n \Lambda_n$. 作为一个特殊情况, 试推出:

如果 $\mathbb{E}|X| < \infty$, 且 $\lim\limits_n \mathbb{P}(\Lambda_n) = 0$, 则 $\lim\limits_{n\to\infty}\int_{\Lambda_n} X\,\mathrm{d}\mathbb{P} = 0$, 特别有 $\lim\limits_{n\to\infty}\int_{\{|X|>n\}} X\,\mathrm{d}\mathbb{P} = 0$.

证明: 先证明 ρ 是 $\mathcal{F}$ 上的一个伪度量. 设 $\Lambda_1,\Lambda_2,\Lambda_3 \in \mathcal{F}$, 我们有 $\rho(\Lambda_1,\Lambda_2) = \mathbb{P}(\Lambda_1\Delta\Lambda_2) \geqslant 0$, $\rho(\Lambda_1,\Lambda_2) = \rho(\Lambda_2,\Lambda_1)$. 因此我们只需验证三角不等式即可. 我们有

$$\begin{aligned}
\rho(\Lambda_1,\Lambda_2) =& \mathbb{P}(\Lambda_1\Delta\Lambda_2) \\
&= \mathbb{P}(\Lambda_1\cap\Lambda_2^c) + \mathbb{P}(\Lambda_1^c\cap\Lambda_2) \\
&= \mathbb{P}((\Lambda_1\cap\Lambda_3^c)\cup(\Lambda_2^c\cap\Lambda_3)) + \mathbb{P}((\Lambda_1^c\cap\Lambda_3)\cup(\Lambda_2\cap\Lambda_3^c)) \\
&\leqslant \mathbb{P}(\Lambda_1\cap\Lambda_3^c) + \mathbb{P}(\Lambda_2^c\cap\Lambda_3) + \mathbb{P}(\Lambda_1^c\cap\Lambda_3) + \mathbb{P}(\Lambda_2\cap\Lambda_3^c) \\
&\leqslant \mathbb{P}((\Lambda_1\cap\Lambda_3^c)\cup(\Lambda_1^c\cap\Lambda_3)) + \mathbb{P}((\Lambda_2\cap\Lambda_3^c)\cup(\Lambda_2^c\cap\Lambda_3)) \\
&= \mathbb{P}(\Lambda_1\Delta\Lambda_3) + \mathbb{P}(\Lambda_2\Delta\Lambda_3).
\end{aligned}$$

因此 ρ 是伪度量.

下面证明连续性. 我们有

$$\left|\int_\Lambda X\,\mathrm{d}\mathbb{P} - \int_{\Lambda_0} X\,\mathrm{d}\mathbb{P}\right| = \left|\int_\Omega X(\mathbb{1}_\Lambda - \mathbb{1}_{\Lambda_0})\right| \leqslant \int_\Omega |X||\mathbb{1}_\Lambda - \mathbb{1}_{\Lambda_0}|\mathrm{d}\mathbb{P} = \int_\Omega |X|\mathbb{1}_{\Lambda\Delta\Lambda_0} = \int_{\Lambda\Delta\Lambda_0}|X|\mathrm{d}\mathbb{P},$$

根据推论 5.4.6, 便有: $\rho(\Lambda,\Lambda_0) = \mathbb{P}(\Lambda\Delta\Lambda_0) \to 0$ 时, $\left|\int_\Lambda X\,\mathrm{d}\mathbb{P} - \int_{\Lambda_0} X\,\mathrm{d}\mathbb{P}\right| \to 0$.

下面证明: 由 $(\Lambda_1,\Lambda_2) \to \Lambda_1\cup\Lambda_2, \Lambda_1\cap\Lambda_2, \Lambda_1\backslash\Lambda_2, \Lambda_1\Delta\Lambda_2$ 给出的由 $M(\mathcal{F},\rho)\times M(\mathcal{F},\rho)$ 的映射都是连续的. 这等价于证明 $\forall \Lambda_1,\Lambda_2,\Lambda_1',\Lambda_2'$, 若 $\rho(\Lambda_1,\Lambda_1'),\rho(\Lambda_2,\Lambda_2') \to 0$, 则 $\rho(\Lambda_1\cup\Lambda_2,\Lambda_1'\cup\Lambda_2')$, $\rho(\Lambda_1\cap\Lambda_2,\Lambda_1'\cap\Lambda_2')$, $\rho(\Lambda_1\backslash\Lambda_2,\Lambda_1'\backslash\Lambda_2')$, $\rho(\Lambda_1\Delta\Lambda_2,\Lambda_1'\Delta\Lambda_2')$ 都趋于零.

先证明 $\rho(\Lambda_1\cup\Lambda_2,\Lambda_1'\cup\Lambda_2') \to 0$. 实际上:

$$\begin{aligned}
\rho(\Lambda_1\cup\Lambda_2,\Lambda_1'\cup\Lambda_2') &= \mathbb{P}((\Lambda_1\cup\Lambda_2)\Delta(\Lambda_1'\cup\Lambda_2')) \\
&= \mathbb{P}((\Lambda_1\cup\Lambda_2)\cap(\Lambda_1'\cup\Lambda_2')^c) + \mathbb{P}((\Lambda_1\cup\Lambda_2)^c\cap(\Lambda_1'\cup\Lambda_2')) \\
&= \mathbb{P}((\Lambda_1\cap\Lambda_1'^c)\cup(\Lambda_2\cap\Lambda_2'^c)) + \mathbb{P}((\Lambda_1^c\cap\Lambda_1')\cup(\Lambda_2^c\cap\Lambda_2')) \\
&\leqslant \mathbb{P}(\Lambda_1\cap\Lambda_1'^c) + \mathbb{P}(\Lambda_2\cap\Lambda_2'^c) + \mathbb{P}(\Lambda_1^c\cap\Lambda_1') + \mathbb{P}(\Lambda_2^c\cap\Lambda_2') \\
&= \mathbb{P}(\Lambda_1\Delta\Lambda_1') + \mathbb{P}(\Lambda_2\Delta\Lambda_2') \\
&= \rho(\Lambda_1,\Lambda_1') + \rho(\Lambda_2,\Lambda_2') \to 0.
\end{aligned}$$

注意到

$$\begin{aligned}\rho(A,B)&=\mathbb{P}(A\Delta B)=\mathbb{P}((A\cap B^c)\cup(A^c\cup B))\\&=\mathbb{P}(((A^c)^c\cap B^c)\cup(A^c\cap(B^c)^c))=\mathbb{P}(A^c\Delta B^c)=\rho(A^c,B^c),\end{aligned}$$

故

$$\begin{aligned}\rho(\Lambda_1\cap\Lambda_2,\Lambda_1'\cap\Lambda_2')&=\rho(\Lambda_1^c\cup\Lambda_2^c,\Lambda_1'^c\cup\Lambda_2'^c)\\&\leqslant\rho(\Lambda_1^c,\Lambda_1'^c)+\rho(\Lambda_2^c,\Lambda_2'^c)\\&=\rho(\Lambda_1,\Lambda_1')+\rho(\Lambda_2,\Lambda_2')\to 0.\end{aligned}$$

同理,

$$\rho(\Lambda_1\backslash\Lambda_2,\Lambda_1'\backslash\Lambda_2')=\rho(\Lambda_1\cap\Lambda_2^c,\Lambda_1^c\cap\Lambda_2)\to 0,$$

$$\rho(\Lambda_1\Delta\Lambda_2,\Lambda_1'\Delta\Lambda_2')=\rho((\Lambda_1\cap\Lambda_2^c)\cup(\Lambda_1^c\cap\Lambda_2),(\Lambda_1'\cap\Lambda_2'^c)\cup(\Lambda_1'^c\cap\Lambda_2'))\to 0.$$

下面证明 $\lim\limits_{k\to\infty}\rho(\Lambda_k,\lim\limits_{n\to\infty}\Lambda_n)=0$. 记 $\Lambda=\lim\limits_{n\to\infty}\Lambda_n$, 注意到 $\Lambda_k\backslash\Lambda\subset\bigcup\limits_{n=k}^{\infty}\Lambda_n\backslash\Lambda$, $\Lambda\backslash\Lambda_k\subset\Lambda\backslash\bigcap\limits_{n=k}^{\infty}\Lambda_n$, 我们有

$$\begin{aligned}\rho(\Lambda_k,\Lambda)&=\mathbb{P}(\Lambda_k\backslash\Lambda)+\mathbb{P}(\Lambda\backslash\Lambda_k)\\&\leqslant\mathbb{P}\left(\bigcup_{n=k}^{\infty}\Lambda_n\backslash\Lambda\right)+\mathbb{P}\left(\Lambda\backslash\bigcap_{n=k}^{\infty}\Lambda_n\right)\\&=\mathbb{P}\left(\bigcup_{n=k}^{\infty}\Lambda_n\right)-\mathbb{P}(\Lambda)+\mathbb{P}(\Lambda)-\mathbb{P}\left(\bigcap_{n=k}^{\infty}\Lambda_n\right)\\&=\mathbb{P}\left(\bigcup_{n=k}^{\infty}\Lambda_n\right)-\mathbb{P}\left(\bigcap_{n=k}^{\infty}\Lambda_n\right).\end{aligned}$$

我们知道 $\bigcup\limits_{n=k}^{\infty}\Lambda_n\downarrow\bigcap\limits_{k=1}^{\infty}\bigcup\limits_{n=k}^{\infty}\Lambda_n=\Lambda$, $\bigcap\limits_{n=k}^{\infty}\Lambda_n\uparrow\bigcup\limits_{k=1}^{\infty}\bigcap\limits_{n=k}^{\infty}\Lambda_n=\Lambda$. 因此

$$\lim_{k\to\infty}\rho(\Lambda_k,\Lambda)=\lim_{k\to\infty}\mathbb{P}\left(\bigcup_{n=k}^{\infty}\Lambda_n\right)-\lim_{k\to\infty}\mathbb{P}\left(\bigcap_{n=k}^{\infty}\Lambda_n\right)=\mathbb{P}(\Lambda)-\mathbb{P}(\Lambda)=0.$$

下证: 如果 $\mathbb{E}|X|<\infty$, 且 $\lim\limits_{n}\mathbb{P}(\Lambda_n)=0$, 则 $\lim\limits_{n\to\infty}\int_{\Lambda_n}X\,\mathrm{d}\mathbb{P}=0$, 特别有 $\lim\limits_{n\to\infty}\int_{\{|X|>n\}}X\,\mathrm{d}\mathbb{P}=0$.

实际上, $\mathbb{P}(\Lambda_n)=\rho(\Lambda_n,\varnothing)\to 0$, 因此 $\lim\limits_{n\to\infty}\int_{\Lambda_n}X\,\mathrm{d}\mathbb{P}=\int_{\varnothing}X\,\mathrm{d}\mathbb{P}=0$. 取 $\Lambda_n=\{|X|>n\}$, 则 $\forall n$ 有

$$\infty>\mathbb{E}|X|=\int_{\Lambda_n^c}|X|\mathrm{d}\mathbb{P}+\int_{\Lambda_n}|X|\mathrm{d}\mathbb{P}>\int_{\Lambda_n^c}|X|\mathrm{d}\mathbb{P}+\int_{\Lambda_n}n\,\mathrm{d}\mathbb{P}=\int_{\Lambda_n^c}|X|\mathrm{d}\mathbb{P}+n\mathbb{P}(\Lambda_n),$$

因此 $\lim\limits_{n\to\infty}\mathbb{P}(\Lambda_n)=0$. 故 $\lim\limits_{n\to\infty}\int_{\{|X|>n\}}X\,\mathrm{d}\mathbb{P}=\lim\limits_{n\to\infty}\int_{\Lambda_n}X\,\mathrm{d}\mathbb{P}=0$. □

第六章 乘积测度与无穷乘积概率空间

§ 6.1 乘积测度与转移测度

6.1.1 设 Ω 是一不可数集, $\mathcal{F}$ 是包含 Ω 中一切单点集的最小 σ 代数, 则 $\Omega\times\Omega$ 的对角线 $\Delta := \{(\omega,\omega) : \omega \in \Omega\} \notin \mathcal{F}\times\mathcal{F}$, 但 $\forall \omega_i \in \Omega$, $i = 1, 2$, 有

$$\Delta_{\omega_1} := \{\omega_2 : (\omega_1,\omega_2) \in \Delta\} \in \mathcal{F},$$
$$\Delta_{\omega_2} := \{\omega_1 : (\omega_1,\omega_2) \in \Delta\} \in \mathcal{F}.$$

这个例子说明了什么?

证法一: 根据习题 3.1.9, 我们知道 Ω 的一切有限集, 可数集以及它们的余集作成一个 σ 代数. 设此 σ 代数为 $\mathcal{G}$, 令 $\mathscr{E} = \{\{\omega\} \subset \Omega\}$, 则 $\mathscr{E} \subset \mathcal{G}$. 因此 $\sigma(\mathscr{E}) \subset \mathcal{G}$. 我们将证明 $\mathcal{F} = \sigma(\mathscr{E}) = \mathcal{G}$. 我们知道, 任取一个包含 $\mathscr{E}$ 的 σ 代数 $\mathcal{G}_0$, 只需将 $\mathscr{E}$ 中的元素进行至多可数并便可以得到任意至多可数集. 进而取至多可数集的余集在 $\mathcal{G}_0$ 中. 因此 $\mathcal{G} \subset \mathcal{G}_0$. 因此 $\mathcal{G} = \sigma(\mathscr{E}) = \mathcal{F}$.

考虑集类

$$\mathscr{D}_1 := \left\{A \in \mathcal{F}\times\mathcal{F} : A = \left(\bigcup_{i=1}^{\infty}(A_i \times \{y_i\})\right) \cup \left(\bigcup_{j=1}^{\infty}(\{x_j\} \times B_j)\right)\right\}$$
$$\mathscr{D}_2 := \left\{A \in \mathcal{F}\times\mathcal{F} : A = \left(\bigcup_{i=1}^{\infty}(A_i \times \{y_i\})\right) \cup \left(\bigcup_{j=1}^{\infty}(\{x_j\} \times B_j)\right) \cup \left(\bigcup_{k=1}^{\infty}(C_k \times D_k)\right)\right\}.$$

其中 $A_i, B_j \in \mathcal{F}$, $(C_k)^c$ 和 $(D_k)^c$ 至多可数, $x_j, y_i \in \Omega$ 且两两不相等. 我们知道 $\mathcal{C} = \{A\times B : A, B \in \mathcal{F}\} \subset \mathscr{D}_1 \cup \mathscr{D}_2$, 下面证明 $\mathscr{D}_1 \cup \mathscr{D}_2$ 是 σ 代数:

我们知道 $\Omega\times\Omega \in \mathscr{D}_1 \cup \mathscr{D}_2$, $\varnothing \in \mathscr{D}_1 \cup \mathscr{D}_2$, 且 $\mathscr{D}_1 \cup \mathscr{D}_2$ 对可列并封闭是显然的. 因此我们只需证明 $\mathscr{D}_1 \cup \mathscr{D}_2$ 对余封闭.

考虑 $A = \left(\bigcup_{i=1}^{\infty}(A_i \times \{y_i\})\right) \cup \left(\bigcup_{n=1}^{\infty}(\{x_n\} \times B_n)\right) \in \mathscr{D}_1 \cup \mathscr{D}_2$, 对于 $(x, y) \in A^c$, 我们知道:

(a) 若 $y = y_i$ 且 $y \notin \bigcup_{n=1}^{\infty} B_n$, 则此情况下所有 (x, y) 组成的集合为 $A_i^c \times \{y_i\}$;

(b) 若 $y = y_i$ 且 $y \in \bigcup_{n=1}^{\infty} B_n$, 则 $x \in A_n^c \backslash \{x_{n_1}, x_{n_2}, \cdots, x_{n_n}, \cdots\} \in \mathcal{F}$. 其中 $\{n_k : k \in \mathbb{N}\}$ 是满足 $y_i \in B_{n_k}$ 的所有 n_k. 则此情况下所有 (x, y) 组成的集合为 $(A_i^c \backslash \{x_{n_1}, x_{n_2}, \cdots, x_{n_n}, \cdots\}) \times \{y_i\}$;

(c) 若 $y \neq y_i$ 且 $y \notin \bigcup_{n=1}^{\infty} B_n$, 显然 $\forall x \in \Omega$, $(x,y) \in A^c$. 则此情况下所有 (x,y) 组成的集合为 $\Omega \times \left(\{y_i, i \in \mathbb{N}\} \cup \left(\bigcup_{n=1}^{\infty} B_n\right)\right)^c$;

(d) 若 $y \neq y_i$ 且 $y \in \bigcup_{n=1}^{\infty} B_n$, 这时所有 (x,y) 组成的集合为 $\{x_{n_k}, k \in \mathbb{N}\}^c \times \left(\bigcup_{n=1}^{\infty} B_n \backslash \{y_n, n \in \mathbb{N}\}\right)$. 若 $\bigcup_{n=1}^{\infty} B_n$ 至多可数, 则 $\bigcup_{n=1}^{\infty} B_n \backslash \{y_n, n \in \mathbb{N}\}$ 至多可数. 反之, $\left(\bigcup_{n=1}^{\infty} B_n \backslash \{y_n, n \in \mathbb{N}\}\right)^c$ 至多可数.

因此 A^c 可以写成 $\left(\bigcup_{i=1}^{\infty}(A_i \times \{y_i\})\right) \cup \left(\bigcup_{j=1}^{\infty}(\{x_j\} \times B_j)\right) \cup \left(\bigcup_{k=1}^{\infty}(C_k \times D_k)\right)$ 的形式. 故 $A^c \in \mathscr{D}_2$. 故 $\mathscr{D}_1 \cup \mathscr{D}_2$ 对余封闭, 是 σ 代数.

我们知道 $\mathcal{C} \subset \mathscr{D}_1 \cup \mathscr{D}_2$, 因此 $\sigma(\mathcal{C}) = \mathcal{F} \times \mathcal{F} \subset \mathscr{D}_1 \cup \mathscr{D}_2$. 而 $\Delta \notin \mathscr{D}_1 \cup \mathscr{D}_2$, 故 $\Delta \notin \mathcal{F} \times \mathcal{F}$. 而 $\Delta_{\omega_1} = \{\omega_1\} \in \mathcal{F}$, $\Delta_{\omega_2} = \{\omega_2\} \in \mathcal{F}$,

这个例子说明: 尽管可测集的任意截集皆可测 (定理 6.1.6) , 反之未必成立, 即存在任意截集皆可测但本身不可测的集合. □

证法二: (该证明方法由助教师兄给出) 根据习题 3.1.9, $\mathcal{F}$ 是一个 σ 代数, 且易证 $\mathcal{F} = \{A \subset \Omega:\ A \text{ 可数或 } A^c \text{ 可数}\}$ (证法一已经证明了这一点).

在 $\mathcal{F}$ 上定义测度 $\mu:\ \forall A \in \mathcal{F},\ \mu(A) = \begin{cases} 0, & A \text{ 可数;} \\ 1, & A^c \text{ 可数} \end{cases}$, 易证明 μ 的确是一个测度, 满足可列可加性: 对于一列两两不交的集合 $\{A_n:\ n \in \mathbb{N}\}$, 若 $\forall n \in \mathbb{N}$, A_n 可数, 则 $\bigcup_{n=1}^{\infty} A_n$ 也可数, 从而 $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = 0 = \sum_{n=1}^{\infty} \mu(A_n)$; 若 $\exists n_0 \in \mathbb{N}:\ A_{n_0}^c$ 可数, 则由 A_n 两两不交知 $\forall n \in \mathbb{N}:\ n \neq n_0,\ A_n \subset A_{n_0}^c$, 从而 $\forall n \neq n_0,\ A_n$ 可数, $\mu(A_n) = 0$, 从而 $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = 1 = \sum_{n=1}^{\infty} \mu(A_n)$. 显然, μ 也是 σ 有限测度 (因为 $\mu(\Omega) = 1$).

在可测矩形 $\mathcal{C} := \{A \times B:\ A, B \in \mathcal{F}\}$ 上定义测度 $\nu:\ \nu(A \times B) = \mu(A)\mu(B)$(请自行验证 ν 满足可列可加性), 则根据引理 6.1.2 知 $\mathcal{C}$ 是一个半集代数. 显然, ν 也是 σ 有限测度. 仿照测度扩张定理 (定理 3.2.7) 的证明, 在 $\Omega \times \Omega$ 的全体子集上定义外测度 ν^*:

$$\forall C \subset \Omega \times \Omega,\ \nu^*(C) := \inf\left\{\sum_{k=1}^{\infty} \nu(A_k \times B_k):\ \bigcup_{k=1}^{\infty}(A_k \times B_k) \supset C,\ A_k \times B_k \in \mathcal{C},\ k \in \mathbb{N}\right\}$$

根据测度扩张定理的证明过程, ν^* 在 $\mathcal{F} \times \mathcal{F} = \sigma(\mathcal{C})$ 上是一个测度; 且因为 ν 在 $\mathcal{C}$ 上是 σ 有限的, 由测度扩张定理知这样的 ν^* 在 $\mathcal{F} \times \mathcal{F}$ 上是唯一的.

现说明为什么 $\Delta \notin \mathcal{F} \times \mathcal{F}$. 假设 $\Delta \in \mathcal{F} \times \mathcal{F}$, 则由于 Ω 不可数, 易知 $\nu^*(\Delta^c), \nu^*(\Delta) = 1$(用反证法证明: 假设外测度是 0, 则由以上 ν^* 的定义说明 Δ, Δ^c 可数, 与 Ω 不可数矛盾), 从而 $\nu^*(\Omega \times \Omega) = \nu^*(\Delta) + \nu^*(\Delta^c) = 2$, 与 $\nu^*(\Omega \times \Omega) = \mu(\Omega)\mu(\Omega) = 1$ 矛盾! 因此, 只可能 $\Delta \notin \mathcal{F} \times \mathcal{F}$.

这个例子说明: 尽管可测集的任意截集皆可测 (定理 6.1.6) , 反之未必成立, 即存在任意截集皆可测但本身不可测的集合. □

6.1.2 试问: $\overline{\mathcal{F}}_1 \times \overline{\mathcal{F}}_2 = \overline{\mathcal{F}_1 \times \mathcal{F}_2}$ 吗? 其中 $\overline{\mathcal{F}}_i, i = 1,2$ 表示 $\mathcal{F}_i$ 对 μ_i 的完全化, $\overline{\mathcal{F}_1 \times \mathcal{F}_2}$ 表示 $\mathcal{F}_1 \times \mathcal{F}_2$ 对 $\mu_1 \times \mu_2$ 的完全化, 这个问题对 Lebesgue 可测集说明了什么?

证明: 题目不够严谨, 需要补充的是 μ_1, μ_2 都是 σ 有限的.

我们先证明 $\overline{\mathcal{F}}_1 \times \overline{\mathcal{F}}_2 \subset \overline{\mathcal{F}_1 \times \mathcal{F}_2}$. 我们知道 $\overline{\mathcal{F}}_1 \times \overline{\mathcal{F}}_2 = \sigma(\mathcal{C})$, 其中 $\mathcal{C} = \{A_1 \times A_2 : A_i \in \overline{\mathcal{F}}_i,\ i = 1,2\}$, 而由定理 3.3.5 知 $\overline{\mathcal{F}_1 \times \mathcal{F}_2}$ 是一个 σ 代数, 故只需证明 $\mathcal{C} \subset \overline{\mathcal{F}_1 \times \mathcal{F}_2}$. 对于 $A_1 \times A_2 \in \mathcal{C} : A_i \in \overline{\mathcal{F}}_i,\ i = 1,2$, 由定理 3.3.5 知存在 $A_i' \in \mathcal{F}_i$, 以及 μ_i 零集 N_i, 使得 $A_i = A_i' \cup N_i$. 故

$$A_1 \times A_2 = (A_1' \cup N_1) \times (A_2' \cup N_2) = (A_1' \times A_2') \cup (A_1' \times N_2) \cup (A_2' \times N_1) \cup (N_1 \times N_2),$$

而 $0 \cdot \infty = 0$, 因此 $A_1' \times N_2$, $A_2' \times N_1$, $N_1 \times N_2$ 都是 $\mu_1 \times \mu_2$ 零集, 且 $A_i' \in \mathcal{F}_i$, 因此 $A_1' \times A_2' \in \mathcal{F}_1 \times \mathcal{F}_2$. 因此 $A_1 \times A_2 \in \overline{\mathcal{F}_1 \times \mathcal{F}_2}$, 于是 $\mathcal{C} \subset \overline{\mathcal{F}_1 \times \mathcal{F}_2}$. 证毕.

而 $\overline{\mathcal{F}_1 \times \mathcal{F}_2} \subset \overline{\mathcal{F}}_1 \times \overline{\mathcal{F}}_2$ 不一定成立. 反例: 考虑 $\mathcal{F}_1 = \mathcal{F}_2 = \mathscr{B}$, $\mu_1 = \mu_2 = \lambda$, $\Omega_1 = \Omega_2 = \mathbb{R}$. 考虑 $A \subset \mathbb{R}$ 是 Lebesgue 不可测集, 则任取单点集 $\{\omega_0\} \in \mathbb{R}$, 有 $A \times \{\omega_0\} \subset \Omega \times \{\omega_0\}$. 我们知道 $(\mu_1 \times \mu_2)(A \times \{\omega_0\}) = 0$, 因此 $A \times \{\omega_0\} \in \overline{\mathcal{F}_1 \times \mathcal{F}_2}$. 而 $(A \times \{\omega_0\})_{\omega_0} = A \notin \overline{\mathcal{F}}_1$, 因此这时 $\overline{\mathcal{F}_1 \times \mathcal{F}_2} \subset \overline{\mathcal{F}}_1 \times \overline{\mathcal{F}}_2$ 不成立. □

6.1.3 设 X_1, X_2 是 n 维独立 r.v., $\mathbb{P}_i, F_i$ 分别是 X_i 的概率分布测度和分布函数, $i = 1,2$.

(1) 试用乘积概率定理证明 $X_1 + X_2$ 的概率分布测度和分布函数分别为由

$$\mathbb{P}_1 * \mathbb{P}_2(B) := \int_{\mathbb{R}^n} \mathbb{P}_1(B - y)\mathbb{P}_2(\mathrm{d}y), \quad B \in \mathscr{B}^n,$$
$$F_1 * F_2(x) := \int_{\mathbb{R}^n} F_1(x - y)\mathrm{d}F_2(y), \quad x \in \mathbb{R}^n,$$

定义的 $\mathbb{P}_1 * \mathbb{P}_2, F_1 * F_2$. 它们分别称为 $\mathbb{P}_1, \mathbb{P}_2$ 及 F_1, F_2 的卷积.

(2) 若 $X_i, i = 1,2$ 还具有分布密度 p_i, 则由

$$p_1 * p_2(x) := \int_{\mathbb{R}^n} p_1(x - y)p_2(y)\mathrm{d}y, \quad x \in \mathbb{R}^n$$

定义的 $p_1 * p_2$ 是 $X_1 + X_2$ 的分布密度. $p_1 * p_2$ 也称为 p_1, p_2 的卷积.

(3) 试证: 一切概率分布测度 (相应地: 分布函数) 对卷积运算作成一个可交换半群.

证明: (1) 我们知道

$$\mathbb{P}(X_1 + X_2 \in B) = \int_{\mathbb{R}^n} \mathbb{P}(X_1 \in B - y, X_2 \in \mathrm{d}y) = \int_{\mathbb{R}^n} \mathbb{P}_1(B - y)\mathbb{P}_2(\mathrm{d}y) = \mathbb{P}_1 * \mathbb{P}_2(B),$$

以及

$$F_{X_1+X_2}(x) = \mathbb{P}(X_1 + X_2 \leqslant x) = \int_{\mathbb{R}^n} \mathbb{P}_1((-\infty, x - y])\mathbb{P}_2(\mathrm{d}y) = \int_{\mathbb{R}^n} F_1(x - y)\mathrm{d}F_2(y) = F_1 * F_2(x).$$

(2) 我们有

$$\begin{aligned} p_{X_1+X_2}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}(F_1 * F_2) = \frac{\mathrm{d}}{\mathrm{d}x}\int_{\mathbb{R}^n} F_1(x - y)\mathrm{d}F_2(y) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \mathbb{1}_{\{z \leqslant x - y\}} p_1(z)\mathrm{d}z\right) p_2(y)\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \left(\frac{\mathrm{d}}{\mathrm{d}x}\int_{\mathbb{R}^n} \mathbb{1}_{\{z \leqslant x - y\}} p_1(z)\mathrm{d}z\right) p_2(y)\mathrm{d}y \\ &= \int_{\mathbb{R}^n} p_1(x - y)p_2(y)\mathrm{d}y. \end{aligned}$$

(3) 我们只需验证卷积运算可交换, 且满足结合律. 可交换是显然的, 这是因为 X_1+X_2 和 X_2+X_1 具有相同的概率分布测度. 下面验证结合律, 我们有:

$$\begin{aligned}\mathbb{P}_1*(\mathbb{P}_2*\mathbb{P}_3)(B)&=\int_{\mathbb{R}^n}\mathbb{P}_1(B-y)\mathbb{P}_2*\mathbb{P}_3(\mathrm{d}y)\\&=\int_{\mathbb{R}^n}\mathbb{P}_1(B-y)\int_{\mathbb{R}^n}\mathbb{P}_2(\mathrm{d}y-z)\mathbb{P}_3(\mathrm{d}z)\\&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\mathbb{P}_1(B-y)\mathbb{P}_2(\mathrm{d}y-z)\mathbb{P}_3(\mathrm{d}z)\\&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\mathbb{P}_1(B-t-z)\mathbb{P}_2(\mathrm{d}t)\mathbb{P}_3(\mathrm{d}z)\\&=\int_{\mathbb{R}^n}\mathbb{P}_1*\mathbb{P}_2(B-z)\mathbb{P}_3(\mathrm{d}z)\\&=(\mathbb{P}_1*\mathbb{P}_2)*\mathbb{P}_3(B).\end{aligned}$$

因此一切概率分布测度对卷积运算作成一个可交换半群. □

6.1.4 设 $(\Omega,\mathcal{F},\mathbb{P})$ 是概率空间, $f(t,\omega)$ 作为 $(t,\omega)\in\mathbb{R}\times\Omega$ 的函数是 $\mathscr{B}\times\mathcal{F}$ 可测的, 若 $\forall t\in\mathbb{R},\mathbb{P}(\{\omega\in\Omega: f(t,\omega)=\infty\})=0$, 试证 $\lambda(\{t\in\mathbb{R}:f(t,\omega)=\infty\})=0$, a.e. $\omega(\mathbb{P})$, 其中 λ 表示 $\mathbb{R}$ 上的 Lebesgue 测度.

提示: 令 $A:=\{(t,\omega):f(t,\omega)=\infty\}$, 考虑 $(\lambda\times\mathbb{P})(A)$.

证明: 考虑 $A=\{(t,\omega):f(t,\omega)=\infty\}$, 则 $\mathbb{P}(A_t)=\mathbb{P}(\{\omega:f(t,\omega)=\infty\})=0,\forall t\in\mathbb{R}$. 因此

$$\int_\Omega\lambda(A_\omega)\mathrm{d}\mathbb{P}=(\lambda\times\mathbb{P})(A)=\int_{\mathbb{R}}\mathbb{P}(A_t)\mathrm{d}\lambda=0.$$

又 $\lambda(A_\omega)\geqslant 0$, 由引理 5.2.3(3) 可知 $\lambda(A_\omega)=0$, a.e. $\omega(P)$. 这便是 $\lambda(\{t\in\mathbb{R}:f(t,\omega)=\infty\})=0$, a.e..

□

6.1.5 设 f 是 $(\Omega_1\times\Omega_2\times\cdots\times\Omega_n,\mathcal{F}_1\times\mathcal{F}_2\times\cdots\times\mathcal{F}_n)$ 上的 (实或复) 数值可测函数, 设 $(i_1,i_2,\cdots,i_k,j_1,j_2,\cdots,j_{n-k})$ 是 $(1,2,\cdots,n)$ 的一个置换, $j_1<j_2<\cdots<j_{n-k}$, 试证: $\forall(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})\in\Omega_{i_1}\times\Omega_{i_2}\times\cdots\times\Omega_{i_k}$, f 在 $(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})$ 的截函数 $f_{(\omega_{i_1},\omega_{i_2},\cdots,w_{i_k})}$ 是 $\mathcal{F}_{j_1}\times\mathcal{F}_{j_2}\times\cdots\times\mathcal{F}_{j_{n-k}}$ 可测的.

证明: 先证明 $\forall A\in\prod\limits_{k=1}^n\mathcal{F}_k$, 其在 $(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})$ 处的截集是 $\prod\limits_{\ell=1}^{n-k}\mathcal{F}_{j_\ell}$ 可测的.

令 $\mathcal{C}=\{\prod\limits_{k=1}^n A_k:A_k\in\mathcal{F}_k\}$, $\Lambda=\left\{A\in\prod\limits_{k=1}^n\mathcal{F}_n:A_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})}\in\prod\limits_{\ell=1}^{n-k}\mathcal{F}_{j_\ell}\right\}$.

容易证明 $\mathcal{C}$ 是一个 π 系且 $\mathcal{C}\subset\Lambda\subset\sigma(\mathcal{C})$, 因此我们只需证明 Λ 是一个 λ 系. (实际上接下来直接证明了它是 σ 代数)

显然 $\varnothing,\Omega\in\Lambda$, 同时 $\forall A\in\Lambda$, 有 $A_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})}\in\prod\limits_{\ell=1}^{n-k}\mathcal{F}_{j_\ell}$, 则 $A^c_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})}=(A_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})})^c\in\prod\limits_{\ell=1}^{n-k}\mathcal{F}_{j_\ell}$, 故 $A^c\in\Lambda$.

再考虑集合列 $\{A_n : n \in \mathbb{N}\} \in \Lambda$, 则 $(A_n)_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})} \in \prod_{\ell=1}^{n-k} \mathcal{F}_{j_\ell}$, $\forall n \geqslant 1$. 于是

$$\left(\bigcup_{n=1}^{\infty} A_n\right)_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})} = \bigcup_{n=1}^{\infty} (A_n)_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})} \in \prod_{\ell=1}^{n-k} \mathcal{F}_{j_\ell}.$$

因此 $\bigcup_{n=1}^{\infty} A_n \in \Lambda$, 故 $\Lambda = \sigma(\mathcal{C}) = \prod_{k=1}^{n} \mathcal{F}_k$.

考虑 $\prod_{k=1}^{n} \mathcal{F}_k$ 可测函数 f, 再任取 $B \in \mathscr{B}$, 我们有

$$\begin{aligned}&\{(\omega_{j_1},\omega_{j_2},\cdots,\omega_{j_{n-k}}) : f_{(\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k})}(\omega_{j_1},\omega_{j_2},\cdots,\omega_{j_{n-k}}) \in B\}\\ &\qquad =(\{(\omega_1,\omega_2,\cdots,\omega_n) : f(\omega_1,\omega_2,\cdots,\omega_n) \in B\})_{\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k}} \in \prod_{\ell=1}^{n-k} \mathcal{F}_{j_\ell}.\end{aligned}$$

□

6.1.6 试证定理 6.1.16.

证明: 先证明 $\forall A \in \{A_1 \times A_2, A_i \in \mathcal{F}_i, i = 1,2\} =: \mathcal{C}$, $g(\omega_1) := \int_{\Omega_2} \mathbb{1}_A(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2)$ 是非负的 $\mathcal{F}_1$ 可测函数. 我们知道 λ 是 σ 有限的, 所以对 $i = 1,2$ 存在两两不交的 $\mathcal{F}_i$ 可测集序列 $\{B_{i,n} : n \in \mathbb{N}\}$ 使得 $\Omega_i = \bigcup_{n=1}^{\infty} B_{i,n}$, 且 $\forall m,n \in \mathbb{N}$ 有 $\sup_{\omega_1 \in B_{1,m}} \lambda(\omega_1, B_{2,n}) < \infty$. 我们有

$$g(\omega_1) = \sum_{n=1}^{\infty} \int_{B_{2,n}} \mathbb{1}_A(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) = \sum_{n=1}^{\infty} \lambda(\omega_1, A \cap B_{2,n}).$$

我们知道 $\lambda(\omega_1, A \cap B_{2,n})$ 是非负有限且 $\mathcal{F}_1$ 可测的, 故 g 也是非负 $\mathcal{F}_1$ 可测函数.

再令

$$\Lambda := \{A \in \mathcal{F}_1 \times \mathcal{F}_2, \forall n \geqslant 1, \int_{B_{2,n}} \mathbb{1}_A(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) \text{是} \mathcal{F}_1 \text{可测的}\},$$

我们有 $\mathcal{C} \in \Lambda \in \sigma(\mathcal{C}) = \mathcal{F}_1 \times \mathcal{F}_2$. 又 $\mathcal{C}$ 是 π 系, 所以我们只需证明 Λ 是 λ 系便可. 易得 $\Omega_1 \times \Omega_2 \in \mathcal{C} \subset \Lambda$, 下面证明其对真差封闭.

考虑 $A, B \in \Lambda$ 且 $B \subset A$, 则

$$\int_{B_{2,n}} \mathbb{1}_{A \setminus B}(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) = \int_{B_{2,n}} \mathbb{1}_A(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) - \int_{B_{2,n}} \mathbb{1}_B(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) \in \mathcal{F}_1.$$

因此 $A \setminus B \in \Lambda$. 又设 $\{A_n : n \in \mathbb{N}\} \subset \Lambda$ 且 $A_n \uparrow$, 由单调收敛定理可知

$$\int_{B_{2,n}} \mathbb{1}_{\cup_{n=1}^{\infty} A_n}(\omega_1,\omega_2)\lambda(\omega_1,\mathrm{d}\omega_2) = \int_{B_{2,n}} \lim_{n\to\infty} \mathbb{1}_{A_n}(\omega_1,\omega_2)\lambda(\omega_1,\mathrm{d}\omega_2) = \lim_{n\to\infty} \int_{B_{2,n}} \mathbb{1}_{A_n}(\omega_1,\omega_2)\lambda(\omega_1,\mathrm{d}\omega_2) \in \mathcal{F}_1.$$

因此 $\bigcup_{n=1}^{\infty} A_n \in \Lambda$. 故 Λ 是 λ 系. 因此 $\forall A \in \mathcal{F}_1 \times \mathcal{F}_2$, 有 $\int_{B_{2,n}} \mathbb{1}_A(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) \in \mathcal{F}_1$. 所以对于任意的简单非负函数 $f \in \mathcal{F}_1 \times \mathcal{F}_2$, 有 $\int_{\Omega_2} f(\omega_1,\omega_2)\lambda(\omega_1, \mathrm{d}\omega_2) \in \mathcal{F}_1$. 再根据单调收敛定理可得对任意的非负 $\mathcal{F}_1 \times \mathcal{F}_2$

可测函数 f, $\int_{\Omega_2} f(\omega_1,\omega_2)\lambda(\omega_1,\mathrm{d}\omega_2)$ 是非负且 $\mathcal{F}_1$ 可测的. □

6.1.7 试证定理 6.1.17 及定理 6.1.18.

证明: 定理 6.1.17: 根据定理 6.1.16 知道 $\forall f\in\mathcal{F}_1\times\mathcal{F}_2$ 且 $f\geqslant 0$, $\int_{\Omega_2} f(\cdot,\omega_2)\lambda_2(\omega_1,\mathrm{d}\omega_2)$ 是 $\mathcal{F}_1$ 可测的. 故 $\forall B\in\mathcal{F}_1\times\mathcal{F}_2$, $\lambda(B)=\int_{\Omega_1}\int_{\Omega_2}\mathbb{1}_B(\omega_1,\omega_2)\lambda_2(\omega_1,\mathrm{d}\omega_2)\lambda_1(\mathrm{d}\omega_1)$ 是有意义的. 因此我们只需证明其 σ 有限且 σ 可加.

我们知道 λ_1, λ_2 是 σ 有限的, 因此存在分别两两不交的集列 $\{\Omega_1^{(n)}:n\in\mathbb{N}\}\subset\mathcal{F}_1$, $\{\Omega_2^{(n)}:n\in\mathbb{N}\}\subset\mathcal{F}_2$ 使得 $\lambda_1(\Omega_1^{(n)})<\infty$, $\lambda_2(\Omega_2^{(n)})<\infty$, 且 $\bigcup_{n=1}^{\infty}\Omega_i^{(n)}=\Omega_i$.

注意到集列 $\{\Omega_1^{(m)}\times\Omega_2^{(n)}:m,n\in\mathbb{N}\}$ 仍然是两两不交的且它们的并为 $\Omega_1\times\Omega_2$, 同时

$$\lambda(\Omega_1^{(m)}\times\Omega_2^{(n)})=\lambda_1(\Omega_1^{(m)})\lambda_2(\cdots,\Omega_2^{(n)})<\infty,$$

因此 λ 是 σ 有限的. 下面证明其 σ 可加.

设 $\{A_n:n\in\mathbb{N}\}\subset\mathcal{F}_1\times\mathcal{F}_2$ 且两两不交, 则 $\forall\omega_1\in\Omega_1$, $A_n(\omega_1)\in\mathcal{F}_2$ 同时 $\{A_n(\omega_1):n\in\mathbb{N}\}$ 两两不交. 故

$$\begin{aligned}\lambda\left(\bigcup_{n=1}^{\infty}A_n\right)&=\int_{\Omega_1}\int_{\Omega_2}\mathbb{1}_{\cup_{n=1}^{\infty}A_n}(\omega_1,\omega_2)\lambda_2(\omega_1,\mathrm{d}\omega_2)\lambda_1(\mathrm{d}\omega_1)\\&=\int_{\Omega_1}\lambda_2\left(\omega_1,\left(\bigcup_{n=1}^{\infty}A_n\right)(\omega_1)\right)\lambda_1(\mathrm{d}\omega_1)\\&=\sum_{n=1}^{\infty}\int_{\Omega_1}\lambda_2\left(\omega_1,A_n(\omega_1)\right)\lambda_1(\mathrm{d}\omega_1)\\&=\sum_{n=1}^{\infty}\lambda(A_n).\end{aligned}$$

因此 λ 是 σ 可加的, 故其是 σ 有限测度.

定理 6.1.18: 考虑数学归纳法. 由定理 6.1.17 可知在 $\mathcal{F}^{(n-1)}$ 上存在 σ 有限测度 $\lambda^{(n-1)}$, 以及 $\mathcal{F}^{(n)}$ 上的 σ 有限测度 $\lambda^{(n)}$ 满足 $\forall B^{(n)}\in\mathcal{F}^{(n)}$, 有

$$\lambda^{(n)}(B^{(n)})=\int_{\Omega^{(n-1)}}\int_{\Omega_n}\mathbb{1}_{B^{(n)}}((\omega_1,\omega_2,\cdots,\omega_{n-1}),\omega_n)\lambda_n(\omega_1,\omega_2,\cdots,\omega_{n-1},\mathrm{d}\omega_n)\lambda^{(n-1)}(\mathrm{d}(\omega_1,\omega_2,\cdots,\omega_{n-1})),$$

再根据归纳假设, $\forall B^{(n-1)}\in\mathcal{F}^{(n-1)}$, 有

$$\lambda^{(n-1)}(B^{(n-1)})=\int_{\Omega_1}\cdots\int_{\Omega_{n-1}}\mathbb{1}_{B^{(n-1)}}(\omega_1,\omega_2,\cdots,\omega_{n-1})\lambda_{n-1}(\omega_1,\omega_2,\cdots,\omega_{n-2},\mathrm{d}\omega_{n-1})\cdots\lambda_2(\omega_1,\mathrm{d}\omega_2)\lambda_1(\mathrm{d}\omega)_1,$$

代入便可得证. □

6.1.8 设 $f(x,y)$ 是 $[0,1]\times[0,1]$ 上的非负有界可测函数, $\forall B\in\mathscr{B}[0,1]$, 令

$$\lambda(x,B)=\int_B f(x,y)\mathrm{d}y,$$

其中 $\mathrm{d}y$ 表示对 Lebesgue 测度的积分, 则 λ 是 $[0,1]\times\mathscr{B}[0,1]$ 上的转移测度.

证明: 先证明固定 $B \in \mathscr{B}[0,1]$, $g : x \to \lambda(x,B) = \int_B f(x,y)\,\mathrm{d}y$ 是 $\mathscr{B}[0,1]$ 可测的. 实际上, $g = \int_{[0,1]} \mathbb{1}_B f(x,y)\mathrm{d}y$, 而 $\mathbb{1}_B f(x,y)$ 是 $\mathscr{B}[0,1]$ 可测的, 由引理 6.1.9 我们知道 $g = \lambda(x,B)$ 也是 $\mathscr{B}[0,1]$ 可测的.

其次, 由定理 5.3.15, 我们知道固定 $x \in [0,1]$, $\lambda(x,B)$ 是测度. 因此 λ 是 $[0,1] \times \mathscr{B}[0,1]$ 上的转移测度. □

6.1.9 设 $(\Omega_i, \mathcal{F}_i)$, $i = 1,2$ 是可测空间, λ_1 是 $\mathcal{F}_1$ 上的 σ 有限测度, λ_2 是 $\Omega_1 \times \mathcal{F}_2$ 上的 σ 有限转移测度,

$$\nu(B) := \int_{\Omega_1}\int_{\Omega_2} \mathbb{1}_B(\omega_1,\omega_2)\lambda_2(\omega_1,\mathrm{d}\omega_2)\lambda_1(\mathrm{d}\omega_1), \quad B \in \mathcal{F}_1 \times \mathcal{F}_2,$$

$A \in \mathcal{F}_1 \times \mathcal{F}_2$, 则 $\nu(A) = 0$ 的充分必要条件是存在一个 λ_1 零测集 N, 使 $\forall \omega_1 \in N^c$, $\lambda_2(\omega_1, A(\omega_1)) = 0$.

证明: 我们知道:

$$\begin{aligned}\nu(A) &= \int_{\Omega_1}\int_{\Omega_2} \mathbb{1}_B(\omega_1,\omega_2)\lambda_2(\omega_1,\mathrm{d}\omega_2)\lambda_1(\mathrm{d}\omega_1)\\ &= \int_{\Omega_1} \lambda_2(\omega_1, A(\omega_1))\mathrm{d}\omega_1.\end{aligned}$$

我们知道 $\lambda_2(\omega_1, A(\omega_1))$ 是非负的, 由引理 5.2.3(3), 便得 $\nu(A) = 0$ 当且仅当 $\lambda_2(\omega_1, A(\omega_1)) = 0$, λ_1 -a.e. 因此当且仅当存在一个 λ_1 零测集 N, 使 $\forall \omega_1 \in N^c$, $\lambda_2(\omega_1, A(\omega_1)) = 0$. □

6.1.10 设 $(\Omega_i, \mathcal{F}_i)$, $i = 1,2,3$ 是可测空间, λ 是 $\Omega_2 \times \mathcal{F}_3$ 上的 σ 有限转移测度, f 是 $\mathcal{F}_1 \times \mathcal{F}_3$ 可测函数, 若积分

$$g(\omega_1,\omega_2) := \int_{\Omega_3} f(\omega_1,\omega_3)\,\lambda(\omega_2,\mathrm{d}\omega_3)\,, \omega_i \in \Omega_i, i = 1,2,$$

$\forall \omega_i, i = 1,2$ 存在, 则 g 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测的.

证明: 首先证明 $\forall A = A_1 \times A_2$, $A_1 \in \mathcal{F}_1$, $A_3 \in \mathcal{F}_3$, $g(\omega_1,\omega_2) = \int_{\Omega_3} \mathbb{1}_A(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3)$ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测的. 实际上, $g(\omega_1,\omega_2) = \mathbb{1}_{A_{\omega_1}}(\omega_1)\int_{\Omega_3} \mathbb{1}_{A_{\omega_3}}(\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3) = \mathbb{1}_{A_{\omega_1}}(\omega_1)\lambda(\omega_2, A_3)$ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测的.

下面证明 $\forall A \in \mathcal{F}_1 \times \mathcal{F}_2$, $\int_{\Omega_3} \mathbb{1}_A(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3)$ 也是 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测的. 考虑集类

$$\mathcal{G} = \left\{ A \in \mathcal{F}_1 \times \mathcal{F}_2, \int_{B_{3,n}} \mathbb{1}_A(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3)\text{是}\mathcal{F}_1 \times \mathcal{F}_2\text{可测的} \right\},$$

我们知道 λ 是 σ 有限转移测度, 故对 $i = 2,3$, 存在互不相交的 $\mathcal{F}_1$ 可测集 $\{B_{i,n}, n \in \mathbb{N}\}$, 使 $\Omega_i = \bigcup_{n=1}^{\infty} B_{i,n}$, 且对任意的 $m,n \in \mathbb{N}$ 有 $\sup_{\omega_2 \in B_{3,m}} \lambda(\omega_2, B_{3,n}) < \infty$. 我们知道 $\mathcal{C} = \{A_1 \times A_2, A_i \in \mathcal{F}_i\} \subset \mathcal{G}$, 而 $\sigma(\mathcal{C}) = \mathcal{F}_1 \times \mathcal{F}_2$. 我们只需证明 $\sigma(\mathcal{C}) = \mathcal{G}$ 即可. 显然有 $\mathcal{G} \subset \sigma(\mathcal{C})$, 因此只需证明 $\sigma(\mathcal{C}) \subset \mathcal{G}$. 我们知道 $\mathcal{C}$ 是 π 系, 因此只需证明 $\mathcal{G}$ 是 λ 系.

显然 $\Omega_1 \times \Omega_3 \in \mathcal{G}$, 下面证明 $\mathcal{G}$ 对真差封闭: 考虑 $A, B \in \mathcal{G}$, $B \subset A$, 则

$$\begin{aligned}\int_{\Omega_3} \mathbb{1}_{A\setminus B}(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3) &= \sum_{n=1}^{\infty}\int_{B_{3,n}} \mathbb{1}_{A\setminus B}(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3)\\ &= \sum_{n=1}^{\infty}\left(\int_{B_{3,n}} \mathbb{1}_A(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3) - \int_{B_{3,n}} \mathbb{1}_B(\omega_1,\omega_3)\lambda(\omega_2,\mathrm{d}\omega_3)\right),\end{aligned}$$

因此 $A\backslash B \in \mathcal{G}$.

下面证明 $\mathcal{G}$ 对不降序列的并封闭. 考虑 $\{A_n, n \in \mathbb{N}\} \subset \mathcal{G}$, 且 $A_n \uparrow$. 我们知道 $0 \leqslant \mathbb{1}_{A_n}(\omega_1, \omega_3) \uparrow \mathbb{1}_{\cup_n A_n}(\omega_1, \omega_3)$, 因此由积分的单调收敛定理, 有

$$\int_{\Omega_3} \mathbb{1}_{\cup_n A_n}(\omega_1, \omega_3)\lambda(\omega_2, \mathrm{d}\omega_3) = \lim_{n\to\infty} \int_{\Omega_3} \mathbb{1}_{A_n}(\omega_1, \omega_3)\lambda(\omega_2, \mathrm{d}\omega_3),$$

因此 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{G}$. 故 $\mathcal{G}$ 是 λ 系.

我们知道任意 $\mathcal{F}_1 \times \mathcal{F}_3$ 可测函数 $f(\omega_1, \omega_3)$, 都是某非负简单函数列的极限. 因此由单调收敛定理, 对任意的 $\mathcal{F}_1 \times \mathcal{F}_3$ 可测函数, $\int_{\Omega_3} f(\omega_1, \omega_3)\lambda(\omega_2, \mathrm{d}\omega_3)$ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 可测的. □

6.1.11 设 $(\Omega_i, \mathcal{F}_i), i = 1,2,3,4$ 是可测空间, λ_1, λ_2 分别是 $\Omega_1 \times \mathcal{F}_2, \Omega_2 \times \mathcal{F}_3$ 上的转移概率, 则由 $\lambda_1 \circ \lambda_2(\omega, B) := \int \lambda_2(\omega_2, B)\lambda_1(\omega, \mathrm{d}\omega_2), \omega \in \Omega_1, B \in \mathcal{F}_3$ 定义的 $\lambda_1 \circ \lambda_2$ 是 $\Omega_1 \times \mathcal{F}_3$ 上的转移概率. 若 λ_3 是 $\Omega_3 \times \mathcal{F}_4$ 上的转移概率, 则 $(\lambda_1 \circ \lambda_2) \circ \lambda_3 = \lambda_1 \circ (\lambda_2 \circ \lambda_3)$.

证明: 我们知道, $\forall B \in \mathcal{F}_3$, 根据定理 6.1.16 可知 $\lambda_1 \circ \lambda_2(\omega, B)$ 是 $\mathcal{F}_1$ 可测函数. $\forall \omega \in \Omega_1$, 由定理 6.1.17 可知 $\lambda_1 \circ \lambda_2(\omega, B)$ 是 $\mathcal{F}_3$ 上的测度. 我们知道 $\forall \omega \in \Omega_1$, 有

$$\lambda_1 \circ \lambda_2(\omega, \Omega_3) = \int_{\Omega_2} \lambda_2(\omega_2, \Omega_3)\lambda_1(\omega, \mathrm{d}\omega_2) = \int_{\Omega_2} \lambda_1(\omega, \mathrm{d}\omega_2) = 1,$$

同时 $\lambda_1 \circ \lambda_2(\omega, \varnothing) = 0$, 因此 $\lambda_1 \circ \lambda_2$ 是 $\Omega_1 \times \mathcal{F}_3$ 上的转移概率.

下面证明 $\lambda_1 \circ (\lambda_2 \circ \lambda_3) = (\lambda_1 \circ \lambda_2) \circ \lambda_3$. 考虑 $\forall \omega_1 \in \Omega_1$, $B \in \mathcal{F}_4$, 我们有

$$\begin{aligned}
(\lambda_1 \circ \lambda_2) \circ \lambda_3(\omega, B) &= \int_{\Omega_3} \lambda_3(\omega_3, B)\lambda_1 \circ \lambda_2(\omega, \mathrm{d}\omega_3) \\
&= \int_{\Omega_3} \lambda_3(\omega_3, B) \int_{\Omega_2} \lambda_2(\omega_2, \mathrm{d}\omega_3)\lambda_1(\omega, \mathrm{d}\omega_2) \\
&= \int_{\Omega_2} \int_{\Omega_3} \lambda_3(\omega_3, B)\lambda_2(\omega_2, \mathrm{d}\omega_3)\lambda_1(\omega, \mathrm{d}\omega_2) \\
&= \int_{\Omega_2} \left(\int_{\Omega_3} \lambda_3(\omega_3, B)\lambda_2(\omega_2, \mathrm{d}\omega_3) \right) \lambda_1(\omega, \mathrm{d}\omega_2) \\
&= \int_{\Omega_2} \lambda_2 \circ \lambda_3(\omega_2, B)\lambda_1(\omega, \mathrm{d}\omega_2) \\
&= \lambda_1 \circ (\lambda_2 \circ \lambda_3)(\omega, B).
\end{aligned}$$

□

6.1.12 设 $\lambda_i, i = 1, 2$ 是由转移概率矩阵 $\mathbb{P}_i$ 确定的 $\mathbb{N} \times N, N := \{A : A \subset \mathbb{N}\}$ 上的转移概率, 则 $\lambda_1 \circ \lambda_2$ 是由 $\mathbb{P}_1 \cdot \mathbb{P}_2$(矩阵 $\mathbb{P}_1, \mathbb{P}_2$ 的乘积) 确定的.

证明: 根据习题 6.1.11, 我们知道 $\lambda_1 \circ \lambda_2$ 是 $\mathbb{N} \times N$ 上的转移概率. 用 p_{ij}^k 来表示 $\mathbb{P}_k$ 中第 i 行, j 列的元素. 则 $\lambda_k(i, B) = \sum_{j\in B} p_{ij}^k$. 其中 $i \in \mathbb{N}, B \subset \mathbb{N}$. 我们知道

$$\lambda_1 \circ \lambda_2(i, B) = \int \lambda_2(\omega_2, B)\lambda_1(i, \mathrm{d}\omega_2) = \sum_{j\in B} \sum_{\ell\in\mathbb{N}} p_{il}^1 p_{lj}^2,$$

而 $\mathbb{P}_1 \cdot \mathbb{P}_2$ 的第 i 行, j 列的元素为 $\sum_{\ell \in \mathbb{N}} p_{il}^1 p_{lj}^2$. 因此 $\lambda_1 \circ \lambda_2$ 是 $\mathbb{P}_1 \cdot \mathbb{P}_2$ 确定的. $\square$

6.1.13 设 $p(t;x,A),(t;x,A) \in [0,\infty) \times \mathbb{R}^3 \times \mathscr{B}^3$ 是定义 6.1.15 所定义的, 试证:

$$\forall t,s \in [0,\infty), x \in \mathbb{R}^3, A \in \mathscr{B}^3,$$

$$p(t+s;x,A) = \int_{\mathbb{R}^3} p(t;x,\mathrm{d}y)p(s;y,A).$$

证明: 我们有:

$$\begin{aligned}
\int_{\mathbb{R}^3} p(t;x,\mathrm{d}y)p(s;y,A) &= \int_{\mathbb{R}^3}\int_A (2\pi s)^{-\frac{3}{2}} \exp\left(-\frac{(z-y)^2}{2s}\right)(2\pi t)^{-\frac{3}{2}} \exp\left(-\frac{(y-x)^2}{2t}\right) \mathrm{d}y\mathrm{d}z \\
&= (4\pi^2 ts)^{-\frac{3}{2}} \int_A \int_{\mathbb{R}^3} \exp\left(-\frac{(z-y)^2}{2s} - \frac{(y-x)^2}{2t}\right) \mathrm{d}y\mathrm{d}z \\
&= (4\pi^2 ts)^{-\frac{3}{2}} \int_A \int_{\mathbb{R}^3} \exp\left(-\frac{(z-x)^2}{2(t+s)} - \frac{t+s}{2ts}\left(y - \frac{tz+sx}{t+s}\right)^2\right) \mathrm{d}y\mathrm{d}z \\
&= (4\pi^2 ts)^{-\frac{3}{2}} \int_A \left(\int_{\mathbb{R}^3} \exp\left(-\frac{t+s}{2ts}\left(y - \frac{tz+sx}{t+s}\right)^2\right) \mathrm{d}y\right) \exp\left(-\frac{(z-x)^2}{2(t+s)}\right) \mathrm{d}z \\
&= (4\pi^2 ts)^{-\frac{3}{2}} \left(\frac{2\pi ts}{t+s}\right)^{\frac{3}{2}} \int_A \exp\left(-\frac{(z-x)^2}{2(t+s)}\right) \mathrm{d}z \\
&= (2\pi(t+s))^{-\frac{3}{2}} \int_A \exp\left(-\frac{(z-x)^2}{2(t+s)}\right) \mathrm{d}z \\
&= p(t+s;x,A).
\end{aligned}$$

$\square$

6.1.14 设 $(\Omega,\mathcal{F})$ 为可测空间, $\mathcal{C}$ 为 $\mathcal{F}$ 的子 σ 代数, 称 $\pi(x,A), x \in \Omega, A \in \mathcal{F}$ 为一概率核, 若它对每一 $A \in \mathcal{F}, \pi(\cdot,A)$ 为 $\mathcal{C}$ 可测函数, 对每一 $x \in \Omega, \pi(x,\cdot)$ 是 $\mathcal{F}$ 上的概率. 设 ν 是 $\mathcal{F}$ 上任一概率, 试证

$$\nu\pi(\cdot) := \int \pi(x,\cdot)\nu(\mathrm{d}x)$$

为 $\mathcal{F}$ 上的概率测度.

证明: 我们知道 $\nu\pi(\varnothing) = \int \pi(x,\varnothing)\nu(\mathrm{d}x) = 0$, 而 $\nu\pi(\Omega) = \int \pi(x,\Omega)\nu(\mathrm{d}x) = \int \nu(\mathrm{d}x) = 1$.

同时, 考虑两两不交的 $\{A_n, n \in \mathbb{N}\} \subset \mathcal{F}$, 我们知道 $\sum_{n=1}^{\infty} \pi(x,A_n) = \pi\left(x, \bigcup_{n=1}^{\infty} A_n\right)$, 因此

$$\nu\pi\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu\pi(A_n),$$

因此 $\nu\pi(\cdot)$ 具有可列可加性. 因此是概率测度. $\square$

§ 6.2 Fubini 定理及其应用

6.2.1 应用 Fubini 定理证明：若 $n \in \mathbb{N}$, $\mathbb{E}\xi^n$ 存在, 则

$$\mathbb{E}\xi^n = n\int_0^{\infty} t^{n-1}\left[1 - F(t)\right] \mathrm{d}t - n\int_{-\infty}^{0} t^{n-1} F(t)\, \mathrm{d}t$$

.

证明: 我们有:

$$\begin{aligned}
\mathbb{E}\xi^n &= \int_{\mathbb{R}} x^n F(\mathrm{d}x) \\
&= \int_0^{\infty}\int_0^{x} nt^{n-1}\mathrm{d}tF(\mathrm{d}x) - \int_{-\infty}^{0}\int_x^{0} nt^{n-1}\mathrm{d}tF(\mathrm{d}x) \\
&= \int_0^{\infty}\int_t^{\infty} nt^{n-1}F(\mathrm{d}x)\mathrm{d}t - \int_{-\infty}^{0}\int_{-\infty}^{0} nt^{n-1}F(\mathrm{d}x)\mathrm{d}t \\
&= n\int_0^{\infty} t^{n-1}\left[1 - F(t)\right]\mathrm{d}t - n\int_{-\infty}^{0} t^{n-1}F(t)\,\mathrm{d}t.
\end{aligned}$$

□

6.2.2 设 c 为固定常数, $c > 0$, 则 $\mathbb{E}|X| < \infty$ 的充要条件是 $\sum_{n=1}^{\infty} \mathbb{P}(|X| \geqslant cn) < \infty$.

特别是, 如果对于 c 的某个值上面的级数收敛, 则它对 c 的所有值也都收敛.

证明: 根据习题 6.2.1以及 $\mathbb{E}|X| < \infty$, 我们有

$$\mathbb{E}|X| = c\cdot\mathbb{E}\frac{|X|}{c} = c\int_0^{\infty}\left[1 - \mathbb{P}(|X| \leqslant ct)\right]\mathrm{d}t = c\int_0^{\infty}\mathbb{P}(|X| \geqslant ct)\mathrm{d}t,$$

我们知道

$$\begin{aligned}
1 + \sum_{n=1}^{\infty}\mathbb{P}(|X| \geqslant cn) = \sum_{n=0}^{\infty}\mathbb{P}(|X| \geqslant cn) &\geqslant \sum_{n=1}^{\infty}\int_0^1 \mathbb{P}(|X| \geqslant c(n-1+r))\mathrm{d}r \\
&= \int_0^{\infty}\mathbb{P}(|X| \geqslant ct)\mathrm{d}t \\
&= \sum_{n=0}^{\infty}\int_0^1 \mathbb{P}(|X| \geqslant c(n+r))\mathrm{d}r \geqslant \sum_{n=1}^{\infty}\mathbb{P}(|X| \geqslant cn),
\end{aligned}$$

因此 $\mathbb{E}|X| < \infty$ 的充要条件是 $\sum_{n=1}^{\infty}\mathbb{P}(|X| \geqslant cn) < \infty$.

实际上由前面的推导我们可以计算得 $\frac{1}{c}\mathbb{E}|X| - 1 \leqslant \sum_{n=1}^{\infty}\mathbb{P}(|X| \geqslant cn) \leqslant \frac{1}{c}\mathbb{E}|X|$, 因此对于 c 的某个值该级数收敛, 则 $\forall c$, 级数同样收敛. □

6.2.3 $\forall r > 0$, $\mathbb{E}|X|^r < \infty$ 的充要条件是 $\sum_{n=1}^{\infty} n^{r-1}\mathbb{P}(|X| \geqslant n) < \infty$.

证明: 我们知道 $\mathbb{E}|X|^r = r\int_0^\infty t^{r-1}\mathbb{P}(|X|\geqslant t)\mathrm{d}t$. 若 $r\leqslant 1$, 则

$$\begin{aligned}\sum_{n=1}^\infty n^{r-1}\mathbb{P}(|X|\geqslant n) &\leqslant \sum_{n=1}^\infty\int_{n-1}^n t^{r-1}\mathbb{P}(|X|\geqslant t)\mathrm{d}t = \frac{1}{r}\mathbb{E}|X|^r\\ &= \sum_{n=0}^\infty\int_n^{n+1} t^{r-1}\mathbb{P}(|X|\geqslant t)\mathrm{d}t \leqslant 1+\sum_{n=1}^\infty n^{r-1}\mathbb{P}(|X|\geqslant n)\end{aligned}$$

而 $r>1$ 时, 有

$$\frac{1}{2^{r-1}}\sum_{n=1}^\infty n^{r-1}\mathbb{P}(|X|\geqslant n)\leqslant\sum_{n=1}^\infty(n-1)^{r-1}\mathbb{P}(|X|\geqslant n)\leqslant\sum_{n=1}^\infty\int_{n-1}^n t^{r-1}\mathbb{P}(|X|\geqslant t)\mathrm{d}t,$$

以及

$$\sum_{n=0}^\infty\int_n^{n+1} t^{r-1}\mathbb{P}(|X|\geqslant t)\mathrm{d}t\leqslant\sum_{n=0}^\infty(n+1)^{r-1}\mathbb{P}(|X|\geqslant n)\leqslant 2^{r-1}\sum_{n=1}^\infty n^{r-1}\mathbb{P}(|X|\geqslant n)+2^{r-1},$$

因此 $\frac{r}{2^{r-1}}\sum_{n=0}^\infty n^{r-1}\mathbb{P}(|X|\geqslant n)\leqslant\mathbb{E}|X|^r\leqslant r\cdot 2^{r-1}\sum_{n=0}^\infty n^{r-1}\mathbb{P}(|X|\geqslant n)+r\cdot 2^{r-1}$, 当 $r>1$. 因此是充要条件.

□

6.2.4 (分部积分公式) 设 g_i, $i=1,2$ 为 $\mathscr{B}[a,b]$ 上的可测函数, 而 F_i 为 $[a,b]$ 上的分布函数, $G_i(x)=\int_a^x g_i(u)\mathrm{d}F_i(u)$, $x\in[a,b]$, 且 $\int_a^b|g_i(x)|\,\mathrm{d}F_i(x)<\infty$, $i=1,2$, 则

$$\int_a^b G_1(x)g_2(x)\mathrm{d}F_2(x)=G_1(b)G_2(b)-\int_a^b g_1(x)G_2(x-)\mathrm{d}F_1(x)$$

其中 $\int_a^x$ 理解为 $\int_{(a,x]}$.

证明: 我们有

$$\begin{aligned}\int_a^b G_1(x)g_2(x)\mathrm{d}F_2(x) &= \int_a^b\left(\int_a^x g_1(u)\mathrm{d}F_1(u)\right)g_2(x)\mathrm{d}F_2(x)\\ &= \int_a^b\int_a^b \mathbb{1}_{\{u\leqslant x\}}g_1(u)g_2(x)\mathrm{d}F_1(u)\mathrm{d}F_2(x)\\ &= \int_a^b\int_a^b \mathbb{1}_{\{x\geqslant u\}}g_2(x)g_1(u)\mathrm{d}F_2(x)\mathrm{d}F_1(u)\\ &= \int_a^b\left(\int_u^b g_2(x)\mathrm{d}F_2(x)\right)g_1(u)\mathrm{d}F_1(u)\\ &= \int_a^b\left(G_2(b)-G_2(u-)\right)g_1(u)\mathrm{d}F_1(u)\\ &= G_1(b)G_2(b)-\int_a^b g_1(x)G_2(x-)\mathrm{d}F_1(x).\end{aligned}$$

□

6.2.5 证明推论 6.2.9.

证明: (1) 根据定理 6.2.5, 我们有

$$
\begin{aligned}
&\int_{A_1\times A_2\times\cdots\times A_n} f\,\mathrm{d}(\mu_1\times\mu_2\times\cdots\times\mu_n)\\
=&\int_{\Omega_1\times\Omega_2\times\cdots\times\Omega_n} f\mathbb{1}_{A_1\times A_2\times\cdots\times A_n}\mathrm{d}(\mu_1\times\mu_2\times\cdots\times\mu_n)\\
=&\int_{\Omega_{i_n}}\left(\cdots\left(\int_{\Omega_{i_1}} f\mathbb{1}_{A_1\times A_2\times\cdots\times A_n}(\omega_1,\omega_2,\cdots,\omega_n)\mu_{i_1}(\mathrm{d}\omega_{i_1})\right)\right)\mu_{i_n}(\mathrm{d}\omega_{i_n})\\
=&\int_{A_{i_n}}\left(\cdots\left(\int_{A_{i_1}} f(\omega_1,\omega_2,\cdots,\omega_n)\mu_{i_1}(\mathrm{d}\omega_{i_1})\right)\right)\mu_{i_n}(\mathrm{d}\omega_{i_n}).
\end{aligned}
$$

其中 $(i_1,\cdots,i_n)$ 是 $(1,\cdots,n)$ 的任意一个置换.

(2) 令 $\mathcal{C}:=\{A_1\times A_2\times\cdots\times A_n: A_i\in\mathcal{F}_i\}$, 则 $\sigma(\mathcal{C})=\mathcal{F}_1\times\mathcal{F}_2\times\cdots\times\mathcal{F}_n$. 由于 $\mathcal{C}$ 是 π 系, 我们只需证明

$$
\Lambda:=\{G\in\mathcal{F}_1\times\mathcal{F}_2\times\cdots\times\mathcal{F}_n: G\text{满足推论 6.2.9(2) 的条件}\}=\mathcal{F}_1\times\mathcal{F}_2\times\cdots\times\mathcal{F}_n=\sigma(\mathcal{C}).
$$

根据 (1) 知道 $\mathcal{C}\in\Lambda$, 易证 Λ 是 λ 系, 因此 $\Lambda=\mathcal{F}_1\times\mathcal{F}_2\times\cdots\times\mathcal{F}_n$, 证毕. □

6.2.6 试证推论 6.2.10.

证明: 根据推论 6.2.9, 有

$$
\begin{aligned}
\int_{A_1\times A_2\times\cdots\times A_n} f\,\mathrm{d}(\mu_1\times\mu_2\times\cdots\times\mu_n)&=\int_{A_{i_n}}\left(\cdots\left(\int_{A_{i_1}} f(\omega_1,\omega_2,\cdots,\omega_n)\mu_{i_1}(\mathrm{d}\omega_{i_1})\right)\right)\mu_{i_n}(\mathrm{d}\omega_{i_n})\\
&=\prod_{k=1}^{n}\int_{A_k} f_k\,\mathrm{d}\mu_k,
\end{aligned}
$$

其中 $(i_1,\cdots,i_n)$ 是 $(1,\cdots,n)$ 的任意一个置换. □

§ 6.3 无穷维乘积概率

6.3.1 证明定理 6.3.3 中定义的 $\mathbb{P}_{\mathbb{N}}$ 在 $\varnothing$ 处连续.

证明: 设存在 $A_n\in\mathcal{C}$, $n\in\mathbb{N}$, $A_n\downarrow\varnothing$, 使得 $\lim\limits_{n\to\infty}\mathbb{P}_{\mathbb{N}}(A_n)>0$. 假设 $A_n\in\mathcal{F}^{(n)}$, 因此 $A_n=B^n\times\prod\limits_{j=n+1}^{\infty}\Omega_j$. 我们知道 $A_{n+1}\subset A_n$, 因此 $B_{n+1}\subset B_n$. 此外, $\forall n>1$, $\mathbb{P}_{\mathbb{N}}(A_n)=\displaystyle\int_{\Omega_1} g_n^{(1)}(\omega_1)\mathbb{P}_1(\mathrm{d}\omega_1)$, 其中

$$
g_n^{(1)}(\omega_1)=\int_{\Omega_2}\mathbb{P}_2(\omega_1,\mathrm{d}\omega_2)\cdots\int_{\Omega_n}\mathbb{1}_{B^n}(\omega_1,\cdots,\omega_n)\mathbb{P}_n(\omega_1,\cdots,\omega_{n-1},\mathrm{d}\omega_n),
$$

我们知道 $\mathbb{1}_{B^{n+1}}(\omega_1,\cdots,\omega_n)\leqslant\mathbb{1}_{B^n}(\omega_1,\cdots,\omega_n)$, 因此固定 ω_1, $g_n^{(1)}(\omega_1)$ 单调下降至某极限, 记为 $h_1(\omega)$. 由控制收敛定理, 我们知道

$$
\int_{\Omega_1} h_1(\omega_1)\mathbb{P}_1(\mathrm{d}\omega_1)=\lim_{n\to\infty}\mathbb{P}_{\mathbb{N}}(A_n)>0,
$$

因此存在 $\omega_1' \in \Omega_1$, s.t. $h_1(\omega_1') > 0$. 实际上, $\omega_1' \in B^1$, 否则 $\forall n > 1$, $\mathbb{1}_{B^n}(\omega_1', \omega_2, \cdots, \omega_n) = 0$, $g_n^{(1)}(\omega_1') = 0$. 考虑 $n > 2$, 则

$$g_n^{(1)}(\omega_1') = \int_{\Omega_2} g_n^{(2)}(\omega_2)\mathbb{P}_{\mathbb{N}}(\omega_1', \mathrm{d}\omega_2),$$

其中

$$g_n^{(2)}(\omega_2) = \int_{\Omega_3} \mathbb{P}_3(\omega_1', \omega_2, \mathrm{d}\omega_3) \cdots \int_{\Omega_n}.$$

和前文类似, 固定 ω_2, 存在某极限 $h_2(\omega_2)$, 使得 $g_n^{(2)} \downarrow h_2(\omega_2)$. 如上所证, 知 $(\omega_1', \omega_2') \in B^2$. 由归纳法可得点列 $\{\omega_1', \omega_2', \cdots\}$, 其中 $\omega_j' \in \Omega_j$, 且 $(\omega_1', \omega_2', \cdots, \omega_n') \in B^n$. 因此 $(\omega_1', \omega_2', \cdots) \in \bigcap\limits_{n=1}^{\infty} A_n = \varnothing$, 矛盾! 故 $\mathbb{P}_{\mathbb{N}}$ 在 $\varnothing$ 处连续. $\square$

第七章 不定积分与条件期望

§ 7.1 符号测度的分解

7.1.1 设 φ, μ_1, μ_2 分别是可测空间 $(\Omega, \mathcal{F})$ 上的符号测度和测度, 且 $\varphi = \mu_1 - \mu_2$, 则必有 $\varphi^+ \leqslant \mu_1$, $\varphi^- \leqslant \mu_2$, 其中 φ^+, φ^- 如定理 7.1.5 所定义 (这叫做 Hahn 分解的最小性).

证明: $\forall A \in \mathcal{F}$, 我们有 $\varphi^+(A) = \sup\limits_{B \in A \cap \mathcal{F}} \varphi(B) = \sup\limits_{B \in A \cap \mathcal{F}} (\mu_1(B) - \mu_2(B)) \leqslant \sup\limits_{B \in A \cap \mathcal{F}} \mu_1(B) = \mu_1(A)$. 因此 $\varphi^+ \leqslant \mu_1$, 又 $\varphi^+ - \varphi^- = \mu_1 - \mu_2$, 故 $\varphi^- \leqslant \mu_2$. □

7.1.2 设 φ, μ 分别是可测空间 $(\Omega, \mathcal{F})$ 上的有限可加集函数和有限测度. 若 $\forall A_n:\ n \in \mathbb{N}$, 当 $\mu(A_n) \to 0$ 时有 $\varphi(A_n) \to 0$, 则 φ 是符号测度.

证明: 考虑 $A_n = \varnothing, \forall n \in \mathbb{N}$, 则 $\varphi(\varnothing) = \lim\limits_{n \to \infty} \mu(A_n) = 0$.

考虑两两不交的 $\{B_n : n \in \mathbb{N}\} \subset \mathcal{F}$, 且 $\bigcup\limits_{n=1}^{\infty} B_n \in \mathcal{F}$. 我们知道

$$\mu\left(\bigcup_{n=1}^{\infty} B_n \backslash \bigcup_{k=1}^{n} B_k\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) - \mu\left(\bigcup_{k=1}^{n} B_k\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) - \sum_{k=1}^{n} \mu(B_k) \to 0,$$

因此

$$\varphi\left(B \backslash \bigcup_{k=1}^{n} B_k\right) = \varphi\left(\bigcup_{n=1}^{\infty} B_n\right) - \sum_{k=1}^{n} \varphi(B_k) \to 0,$$

取极限便得 $\varphi\left(\bigcup\limits_{n=1}^{\infty} B_n\right) = \sum\limits_{n=1}^{\infty} \varphi(B_n)$. □

7.1.3 试证定理 7.1.9 中的

$$F^+(x) := \sup_{\substack{t_0 < \cdots < t_n \\ t_n = x,\ n \in \mathbb{N}}} \left\{ \sum_{k=1}^{n} [F(t_k) - F(t_{k-1})]^+ + \frac{F(t_0)}{2} \right\},$$

$$F^-(x) := \sup_{\substack{t_0 < \cdots < t_n \\ t_n = x,\ n \in \mathbb{N}}} \left\{ \sum_{k=1}^{n} [F(t_k) - F(t_{k-1})]^- - \frac{F(t_0)}{2} \right\}.$$

其中

$$[b-a]^+ := \begin{cases} b-a, & b \geqslant a, \\ 0, & b < a; \end{cases} \qquad [b-a]^- := \begin{cases} 0, & b \geqslant a, \\ a-b, & b < a. \end{cases}$$

证明: 我们有

$$\begin{aligned}
F^+(x) &= \frac{V_F + F}{2} \\
&= \frac{1}{2}\left[\sup_{\substack{t_0<\cdots<t_n \\ t_n=x,\ n\in\mathbb{N}}} \sum_{k=1}^{n} |F(t_k) - F(t_{k-1})| + \sum_{k=1}^{n} [F(t_k) - F(t_{k-1})] + F(t_0)\right] \\
&= \sup_{\substack{t_0<\cdots<t_n \\ t_n=x,\ n\in\mathbb{N}}} \left\{\sum_{k=1}^{n} [F(t_k) - F(t_{k-1})]^+ + \frac{F(t_0)}{2}\right\}.
\end{aligned}$$

同理, $F^-(x) := \sup\limits_{\substack{t_0<\cdots<t_n \\ t_n=x,\ n\in\mathbb{N}}} \left\{\sum_{k=1}^{n} [F(t_k) - F(t_{k-1})]^- - \frac{F(t_0)}{2}\right\}$. □

7.1.4 试证下列各函数在其定义域上具有限变差 (以下设 $a<b,\ a,b\in\mathbb{R}$):

(1) $[a,b]$ 上的单调函数 F;

(2) $[a,b]$ 上的 Lipschitz 函数 F

(即, $F:[a,b]\to\mathbb{R}$, 且有一常数 K, 使 $\forall x,y\in[a,b],\ |F(x)-F(y)|\leqslant K|x-y|$);

(3) F 在 $[a,b]$ 上有有界导数.

证明:

(1) 任取 $[a,b]$ 的分割 $\Delta_n := a=t_0<t_1<\cdots<t_n=b$, 定义 $\|\Delta_n\| = \sup\limits_{1\leqslant k\leqslant n}(t_k-t_{k-1})$, 则

$$V_F([a,b]) = \lim_{\|\Delta_n\|\to 0} \sum_{k=1}^{n} |F(t_k) - F(t_{k-1})| = |F(b)-F(a)| < \infty.$$

(2) 我们知道 $\exists K>0$, s.t. $\forall x,y\in[a,b]$, $|F(x)-F(y)|\leqslant K|x-y|$, 故 $V_F([a,b])\leqslant K\lim\limits_{\|\Delta_n\|\to 0}\sum_{k=1}^{n}|t_k-t_{k-1}| = K(b-a)<\infty$.

(3) 不妨设 $|F'(x)|<K$, $\forall x\in[a,b]$, 则根据微分中值定理有

$$\exists\xi, |F(x)-F(y)| = |F'(\xi)||x_y| \leqslant K|x-y|.$$

□

7.1.5 试证有限变差函数有界, 并举一反例说明逆命题不真.

证明: 设 F 为 $[a,b]$ 上的有界变差函数, $\forall x_0,x\in[a,b]$, 有 $|F(x)-F(x_0)|\leqslant V_F([a,b])$, 于是 $|F(x)|\leqslant V_F([a,b])+|F(x_0)| = M<\infty$.

反例: 考虑 $F(x) = \begin{cases} \sin\frac{1}{x}, & x\in\left[-\frac{2}{\pi},0\right); \\ 0, & x=0. \end{cases}$ 取分割 $t_k = -\frac{2}{\pi k}$, $1\leqslant k\leqslant n$, $t_{n+1}=0$. 则

$$\sum_{k=1}^{n} |F(t_{k+1}) - F(t_k)| \geqslant \sum_{k=1}^{n-1} \left|\sin\left(-\frac{\pi(k+1)}{2}\right) - \sin\left(-\frac{\pi k}{2}\right)\right| = n-1 \to \infty.$$

但 $|F(x)| \leqslant 1$, 故其有界但不是有界变差函数. $\square$

7.1.6 设 F, G 是 $[a,b] \subset \mathbb{R}$ 上的有限变差函数, 则它们的和、差、积都是有限变差函数; 若还有 $\inf\limits_{a<x<b} |G(x)| > 0$, 则 $\dfrac{F}{G}$ 也是有限变差函数.

证明: 对任意分割 $a = t_0 < t_1 < \cdots < t_n = b$, 由

$$\sum_{k=1}^{n} |(F(t_k) \pm G(t_k)) - (F(t_{k-1}) - G(t_{k-1}))| \leqslant \sum_{k=1}^{n} |F(t_k) - F(t_{k-1})| + \sum_{k=1}^{n} |G(t_k) - G(t_{k-1})|,$$

因此 $V_{F\pm G}[a,b] \leqslant V_F[a,b] + V_F[a,b]$, 故有界变差函数的和差仍为有界变差函数.

我们知道有界变差函数必然是有界的, 因此 $\exists M > 0$, s.t. $\forall x \in [a.b]$, $|F|, |G| \leqslant M$. 故

$$\begin{aligned}\sum_{k=1}^{n} |F(t_k)G(t_k) - F(t_{k-1})G(t_{k-1})| &\leqslant \sum_{k=1}^{n} |F(t_k)G(t_k) - F(t_{k-1})G(t_k)| + \sum_{k=1}^{n} |F(t_{k-1})G(t_k) - F(t_{k-1})G(t_{k-1})| \\ &\leqslant M\left(\sum_{k=1}^{n} |F(t_k) - F(t_{k-1})| + \sum_{k=1}^{n} |G(t_k) - G(t_{k-1})|\right).\end{aligned}$$

因此 $V_{FG}[a,b] \leqslant M(V_F[a,b] + V_G[a,b]) < \infty$, 故有界变差函数的积也是有界变差函数.

对于 $\inf\limits_{a<x<b} |G(x)| > 0$ 时的 $\dfrac{F}{G}$, 我们只需证明 $\dfrac{1}{G}$ 也是有界变差函数即可. 我们知道 $\exists K > 0$, s.t. $\left|\dfrac{1}{G}\right| \leqslant K$, 故

$$\sum_{k=1}^{n} \left|\frac{1}{G(t_k)} - \frac{1}{G(t_{k-1})}\right| = \sum_{k=1}^{n} \left|\frac{G(t_k) - G(t_{k-1})}{G(t_k)G(t_{k-1})}\right| \leqslant K^2 \sum_{k=1}^{n} |G(t_k) - G(t_{k-1})|,$$

因此 $V_{\frac{1}{G}}[a,b] \leqslant K^2 V_G[a,b]$. 因此 $\dfrac{F}{G}$ 也是有界变差的. $\square$

§ 7.2 Lebesgue 分解定理与 Radon-Nikodym 定理

注意: 本节各题中, 除特殊声明外, φ 表示可测空间 $(\Omega, \mathcal{F})$ 上的符号测度 (习题 7.2.6除外), $\varphi = \varphi^+ - \varphi^-$ 表示 Hahn 分解, $|\varphi| := \varphi^+ + \varphi^-$, μ 表示测度, $A, B, \cdots$ 均为 $\mathcal{F}$ 可测集.

7.2.1 若 $\mu(A_n) \to 0 (n \to \infty)$, $A_n \in \mathcal{F}$ 时 $\varphi(A_n) \to 0 (n \to \infty)$, 则 φ 是 μ 连续的; 若 φ 是有限的, 则反之亦真.

(提示: 首先, 由 φ 的 μ 连续性易知 $|\varphi|$ 也是 μ 连续的, 若其逆不真, 则存在 $\varepsilon > 0$ 与序列 $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, 使得 $\mu(A_n) < \dfrac{1}{2^n}$, 而 $|\varphi(A_n)| \geqslant \varepsilon$, 于是 $B := \limsup\limits_{n\to\infty} A_n$ 使得 $\mu(B) = 0$ 而 $|\varphi|(B) \geqslant \varepsilon$.)

证明: 若 $A \in \mathcal{F}$, 且 $\mu(A) = 0$, 则存在 $\{A_n : n \in \mathbb{N}\}$ 使得 $A_n \uparrow A$ 且 $\mu(A_n) \to 0$. 因此 $\varphi(A_n) \to 0$, 故 $\varphi(A) = 0$, $\varphi \ll \mu$.

假设 φ 有限时命题不真, 则 $\exists \varepsilon > 0$, $\forall n \in \mathbb{N}$, $\exists \{A_n : n \in \mathbb{N}\} \subset \mathcal{F}$, s.t. $\mu(A_n) < 2^{-n}$, 且 $|\varphi(A_n)| \geqslant \varepsilon$. 令 $B = \limsup\limits_{n\to\infty} A_n$, 则

$$\mu(B) = \lim_{n\to\infty} \mu\left(\bigcup_{k=n}^{\infty} A_k\right) \leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(A_n) \leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} 2^{-n} = 0.$$

而根据 Fatou 引理, 有

$$|\varphi|(B) = |\varphi|(\limsup_{n\to\infty} A_n) \geqslant \limsup_{n\to\infty} |\varphi|(A_n) \geqslant \limsup_{n\to\infty} |\varphi(A_n)| \geqslant \varepsilon.$$

不妨设 $\varphi^+(B) \geqslant \varphi(B\cap P) > 0$, 但是 $\mu(B\cap P) \leqslant \mu(B) = 0$, 这与 $\varphi \ll \mu$ 矛盾. □

7.2.2 若 $\{\varphi_n\}$ 是测度序列, 试证: $\forall n \in \mathbb{N}$,

$$\mu_n \ll \mu := \sum_{k=1}^{\infty} \frac{1}{2^k}\mu_k.$$

μ_n 能否换成 φ_n?

证明: 当 $\{\mu_n\}$ 是测度序列时, 考虑 $A \in \mathcal{F}$ 且 $\mu(A) = 0$, 则 $\mu_n(A) \leqslant 2^n\mu(A) = 0$.

当 μ_n 换成 $\varphi(n)$ 时不一定成立, 反例如下:

考虑非零符号测度列 $\{\varphi_n\}$ 满足 $\varphi_1 = -\dfrac{1}{2}\varphi_2$, 以及 $\forall k \geqslant 3$ 有 $\varphi_k = 0$. 则 $\forall A \in \mathcal{F}$, $\varphi(A) = \varphi_1(A) + \dfrac{1}{2}\varphi_2(A) = 0$, 但存在 $A \in \mathcal{F}$ 使得 φ_1, φ_2 不为零, 故 φ_1, φ_2 均不是 φ 连续的. □

7.2.3 R-N 导数的微分公式: 设 $\varphi \ll \nu$, 且 $\nu, \varphi, \varphi' \ll \mu$, 则

$$\frac{\mathrm{d}(\varphi+\varphi')}{\mathrm{d}\mu} = \frac{\mathrm{d}\varphi}{\mathrm{d}\mu} + \frac{\mathrm{d}\varphi'}{\mathrm{d}\mu} \quad \mu \text{ a.e. },$$
$$\frac{\mathrm{d}\varphi}{\mathrm{d}\mu} = \frac{\mathrm{d}\varphi}{\mathrm{d}\nu}\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \quad \mu \text{ a.e. }.$$

证明: 我们知道 $\forall A \in \mathcal{F}$, 存在 $f, f', g \in \mathcal{F}$, 使得

$$\varphi(A) = \int_A f\,\mathrm{d}\mu,\ \varphi'(A) = \int_A f'\mathrm{d}\mu,\ \nu(A) = \int_A g\mathrm{d}\mu,$$

且它们是 μ-a.e. 唯一的. 因此

(1) $\forall A \in \mathcal{F}$, $(\varphi+\varphi')(A) = \displaystyle\int_A f\,\mathrm{d}\mu + \int_A f'\mathrm{d}\mu$. 因此 $\dfrac{\mathrm{d}\varphi}{\mathrm{d}\mu} + \dfrac{\mathrm{d}\varphi'}{\mathrm{d}\mu} = \mathrm{d}(\varphi+\varphi')\mathrm{d}\mu$.

(2) 我们知道 $\forall A \in \mathcal{F}$, $\exists h \in \mathcal{F}$, s.t. $\varphi(A) = \displaystyle\int_A h\mathrm{d}\nu$, 且 $\nu \ll \mu$. 由推论 7.2.7, 我们有

$$\int_A f\,\mathrm{d}\mu = \varphi(A) = \int_A h\mathrm{d}\nu = \int_A h\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu,$$

因此 $f = h\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$, a.e.. 也即 $\dfrac{\mathrm{d}\varphi}{\mathrm{d}\mu} = \dfrac{\mathrm{d}\varphi}{\mathrm{d}\nu}\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$, a.e..

□

7.2.4 设 $\bar{\mu}_n := \displaystyle\sum_{k=1}^{n}\mu_k \to \bar{\mu}(n\to\infty)$, $\bar{\nu}_n := \displaystyle\sum_{k=1}^{n}\nu_k \to \bar{\nu}(n\to\infty)$, 其中带有附标的 μ, ν 都是有限的, $\bar{\mu}, \bar{\nu}$ 是有限的, 并且 $\bar{\nu}_n \ll \bar{\mu}_n$, $\forall n \in \mathbb{N}$, 则

(1) $\dfrac{\mathrm{d}\mu_1}{\mathrm{d}\bar{\mu}_n} \to \dfrac{\mathrm{d}\mu_1}{\mathrm{d}\bar{\mu}}(n\to\infty) \quad \bar{\mu}$ a.e.;

(2) 若 $\bar{\mu}_n \ll \bar{\nu}$, $\forall n \in \mathbb{N}$, 则 $\dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\nu}} \to \dfrac{\mathrm{d}\bar{\mu}}{\mathrm{d}\bar{\nu}}(n\to\infty) \quad \bar{\nu}$ a.e.;

(3) $\bar{\nu} \ll \bar{\mu}$ 且 $\dfrac{\mathrm{d}\bar{\nu}_n}{\mathrm{d}\bar{\mu}_n} \to \dfrac{\mathrm{d}\bar{\nu}}{\mathrm{d}\bar{\mu}}(n \to \infty)\quad \bar{\mu}$ a.e..

(提示: 关于最后一个结论应注意: 若 $\forall n \in \mathbb{N},\ \bar{\mu}_n(A_n) = 0$, 则 $\bar{\mu}(\liminf\limits_{n\to\infty} A_n) = 0$, 由此可知, 只要考察一个特别选择的

$$\frac{\mathrm{d}\bar{\nu}_n}{\mathrm{d}\bar{\mu}_n} = \frac{\sum\limits_{k=1}^{n} f_k}{\sum\limits_{k=1}^{n} g_k},$$

其中 $f_k := \dfrac{\mathrm{d}\nu_k}{\mathrm{d}\bar{\mu}},\ g_k := \dfrac{\mathrm{d}\mu_k}{\mathrm{d}\bar{\mu}}$, 而 $\sum\limits_n f_n = \dfrac{\mathrm{d}\bar{\nu}}{\mathrm{d}\bar{\mu}},\ \sum\limits_n g_n = 1,\ \bar{\mu}$.a.e.)

证明:

(1) 我们知道 $\mu_n \ll \bar{\mu}$, 由 Radon-Nikodym 定理, 存在 f_n 使得 μ_n 是 f_n 关于 $\bar{\mu}$ 的不定积分. 令 $g_n = \sum\limits_{k=1}^{n} f_n$, 则 $\bar{\mu}_n$ 是 g_n 关于 $\bar{\mu}$ 的不定积分, 且 g_n 由 $\bar{\mu}_n$ 几乎唯一决定. 由 $\mu_1 \ll \bar{\mu}_n \ll \bar{\mu}$ 可得 $\dfrac{\mathrm{d}\mu_1}{\mathrm{d}\bar{\mu}} = \dfrac{\mathrm{d}\mu_1}{\mathrm{d}\bar{\mu}_n}\dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\mu}}$. 由于 $\bar{\mu}_n \to \bar{\mu}$, 因此 $\sum\limits_{k=1}^{n} f_k \to 1$. 又 $\dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\mu}} = \sum\limits_{k=1}^{n} f_k \to 1$, $\bar{\mu}$ a.e., 则 $\dfrac{\mathrm{d}\mu_1}{\mathrm{d}\bar{\mu}_n} \to \dfrac{\mathrm{d}\mu_1}{\mathrm{d}\bar{\mu}}$, $m\bar{u}$ a.e..

(2) 类似 (1). 仅需证明 $\dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\nu}} = \dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\mu}}\dfrac{\mathrm{d}\bar{\mu}}{\mathrm{d}\bar{\nu}}$, 由 (1) 知 $\dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\mu}} \to 1$, a.e., 因此 $\dfrac{\mathrm{d}\bar{\mu}_n}{\mathrm{d}\bar{\nu}} \to \dfrac{\mathrm{d}\bar{\mu}}{\mathrm{d}\bar{\nu}}$, $\bar{\nu}$ a.e..

(3) 令 $f_k := \dfrac{\mathrm{d}\nu_k}{\mathrm{d}\bar{\mu}}, g_k = \dfrac{\mathrm{d}\mu_k}{\mathrm{d}\bar{\mu}}$. 注意到 $\sum\limits_{n=1}^{\infty} f_n = \dfrac{\mathrm{d}\bar{\nu}}{\mathrm{d}\bar{\mu}}, \sum\limits_{n=1}^{\infty} g_n = 1$, $\bar{\mu}$ a.e.. 因此

$$\frac{\mathrm{d}\bar{\nu}_n}{\mathrm{d}\bar{\mu}_n} = \frac{\sum\limits_{k=1}^{n} f_k}{\sum\limits_{k=1}^{n} g_k} \to \sum_{n=1}^{\infty} f_n = \frac{\mathrm{d}\bar{\nu}}{\mathrm{d}\bar{\mu}}.$$

□

7.2.5 试证: 要想 $\mathcal{F}$ 上的集函数 φ 是某一 $\mathcal{F}$ 可测函数 f 对 μ 的不定积分, 必须且只需 φ 为 σ 可加, 且对于每一集 $A := \{a \leqslant f \leqslant b\} \cap B,\ B \in \mathcal{F}$, 总有 $a\mu(A) \leqslant \varphi(A) \leqslant b\mu(A)$.

证明: 充分性是显然的, 下面证明必要性.

先证明 $\varphi \ll \mu$. 任取 $B \in \mathcal{F}$ 且 $\mu(B) = 0, \forall n \geqslant 1$, 考虑 $A_n := \{|f| \leqslant n\} \cap B \uparrow B$, 则 $0 = -n\mu(A_n) \leqslant \varphi(A_n) \leqslant n\mu(A_n) = 0$. 因此 $\varphi(B) = \lim\limits_{n\to\infty} \varphi(A_n) = 0$, 也即 $\varphi \ll \mu$.

令 $g =: \dfrac{\mathrm{d}\varphi}{\mathrm{d}\mu}$, 我们只需证明 $f = g$, μ a.e.. 考虑反证法, 若不然, 则 $\mu(\{f \neq g\}) > 0$, 但

$$\mu(\{f \neq g\}) = \mu\left(\bigcup_{r_1, r_2 \in \mathbb{Q}} \{f > r_1 > r_2 > g\} \cup \{f < r_1 < r_2 < g\}\right).$$

因此 $\exists r_1, r_2 \in \mathbb{Q}$, s.t. $\mu\left(\{f > r_1 > r_2 > g\} \cup \{f < r_1 < r_2 < g\}\right) > 0$. 不失一般性, 假设 $\mu(\{f > r_1 > r_2 > g\}) > 0$, 则

$$r_1^{-1}\varphi(\{f > r_1\}) \geqslant \mu(\{f > r_1\}) > \mu(\{f > r_1 > r_2 > g\}) > 0.$$

但是 $\varphi(\{f>r_1\})\leqslant r_2\mu(\{f>r_1\})<r_1\mu(\{f>r_1\})$, 矛盾! □

7.2.6 设 f 对 μ 的积分存在, 令 φ 是 f 关于 μ 的不定积分, 试证:

$$\varphi^+(A)=\int_A f^+\mathrm{d}\mu,\ \varphi^-(A)=-\int_A f^-\mathrm{d}\mu,\ |\varphi|(A)=\int_A|f|\mathrm{d}\mu.$$

证明: 我们知道 φ 是 $\mathcal{F}$ 的符号测度, 由 Hahn 分解定理, 存在 φ^+, φ^- 使得 $\varphi=\varphi^+-\varphi^-$. 我们知道, $\forall A\in\mathcal{F}$, $\varphi^+(A)=\sup\limits_{B\in A\cap\mathcal{F}}\varphi(B)$, $\varphi^-(A)=-\inf\limits_{B\in A\cap\mathcal{F}}\varphi(B)$. 记 $\mu_1=\int_A f^+\mathrm{d}\mu$, $\mu_2=\int_A f^-\mathrm{d}\mu$, 则 $\varphi=\mu_1-\mu_2$. 根据习题 7.1.1, 我们知道 $\varphi^+\leqslant\mu_1$, $\varphi^-\leqslant\mu_2$.

$\forall A\in\mathcal{F}$, 我们有

$$\mu_1=\int_{A\cap\{f\geqslant 0\}}f\,\mathrm{d}\mu\leqslant\sup_{B\in A\cap\mathcal{F}}\int_B f\,\mathrm{d}\mu=\varphi^+(A),$$

以及

$$\mu_2=-\int_{A\cap\{f<0\}}f\,\mathrm{d}\mu\leqslant-\inf_{B\in A\cap\mathcal{F}}\int_B f\,\mathrm{d}\mu=\varphi^-(A).$$

因此 $\varphi^+(A)=\mu_1=\int_A f^+\mathrm{d}\mu$, $\varphi^-(A)=\mu_2=\int_A f^-\mathrm{d}\mu$, $|\varphi|(A)=\int_A|f|\mathrm{d}\mu$. □

7.2.7 设 f 为 $\mathcal{F}$ 可测函数, 且使得下式右边有意义, 则定义

$$\int f\,\mathrm{d}\varphi:=\int f\,\mathrm{d}\varphi^+-\int f\,\mathrm{d}\varphi^-$$

称为 ***f* 对 *φ* 的积分**. 试证: 这种积分具有可测函数对测度的积分的主要性质.

证明: 线性性质是易证的. 我们在此仅证明控制收敛定理. 也即考虑可测函数列 $\{f_n:n\in\mathbb{N}\}$ 以及可积函数 g,h, 当 $g\leqslant f_n\leqslant h$ a.e., 且 $f_n\to f$, a.e. 时, 有 $\int f_n\mathrm{d}\varphi=\int f\,\mathrm{d}\varphi$. 实际上根据定理 5.4.3 可得

$$\lim_{n\to\infty}\int f_n\mathrm{d}\varphi^+\to\int f\,\mathrm{d}\varphi^+,\quad\lim_{n\to\infty}\int f_n\mathrm{d}\varphi^-\to\int f\,\mathrm{d}\varphi^-.$$

因为 g,h 均可积, 所以 $\int g\,\mathrm{d}\varphi^+$, $\int g\varphi^-$, $\int h\,\mathrm{d}\varphi^+$, $\int h\varphi^-$ 均有限, 因此 $\int f\,\mathrm{d}\varphi^+$, $\int f\varphi^-$ 均有限, 故 $\int f_n\mathrm{d}\varphi\to\int f\,\mathrm{d}\varphi$. □

7.2.8 集代数 $\mathscr{A}$ 上的 σ 可加集函数 φ 可以扩张为 $\sigma(\mathscr{A})$ 上的符号测度的充分与必要条件是 φ 有下界或有上界. 若 φ 是 σ 有限测度, 则扩张唯一且扩张所得的测度是 σ 有限的.

证明: 必要性显然, 下面证明充分性. 由定理 7.1.6, 我们知道 φ 存在 Hahn 分解: $\varphi=\varphi^+-\varphi^-$, 且 φ^+, φ^- 均为测度. 再根据测度扩张定理可知, φ^+, φ^- 可以扩张为 $\sigma(\mathscr{A})$ 上的测度 ${\varphi'}^+,{\varphi'}^-$, 而其中任一个在 $\mathscr{A}$ 上有界则 φ 在 $\mathscr{A}$ 上有上界或下界. 又 $\varphi'={\varphi'}^+-{\varphi'}^-$ 在 $\sigma(\mathscr{A})$ 上成立且为符号测度, 故 φ' 是 φ 的扩张. 后一结论由测度扩张定理即得. □

7.2.9 若 μ 不是 σ 有限的, 则即使 φ 有限, Radon-Nikodym 定理也不一定成立.
(提示: 考虑反例: $\Omega=[0,1]$, $\mathcal{F}=\{A:\ A\subset[0,1],\ A\text{或}A^c\text{可数}\}$,$\mu(A)=|A|(A\text{的元数})$,

$$\varphi(A):=\begin{cases}0, & A\text{ 可数},\\ 1, & A^c\text{ 可数}.\end{cases}$$

)

证明: 根据习题 3.1.9, 我们知道不可数集的至多可数子集和它们的余集作成一个 σ 代数. 考虑 $\Omega = [0,1]$, $\mathcal{F} = \{A:\ A \subset [0,1],\ A\text{或}A^c\text{可数}\}$,$\mu(A) = |A|(A\text{的元数})$, $\varphi(A) = \begin{cases} 0, & A \text{ 可数}, \\ 1, & A^c \text{ 可数}. \end{cases}$. 显然 μ 不是 σ 有限的, 而 φ 是有限的. 假设存在有限函数 $f \in \mathcal{F}$ 使得 $\varphi(A) = \int_A f\,\mathrm{d}\mu$, 则 $\forall x \in \Omega$, 有

$$0 = \varphi(\{x\}) = \int_{\{x\}} f\,\mathrm{d}\mu = f(x).$$

因此 $f(x) = 0$. 但考虑 $A = \Omega\backslash\mathbb{Q}$, 则 $1 = \varphi(A) \neq \int_A f\,\mathrm{d}\mu = 0$, 矛盾! 因此不存在这样的函数 f, Radon-Nikodym 定理在这种情况下不成立. □

§ 7.3 条件期望的概念

7.3.1 设随机变量 $X \sim \mathscr{P}(\lambda)$(意指 X 服从参数为 λ 的 Poisson 分布), 随机变量 Y 在给定 $X = n$ 下的条件分布为

$$\mathbb{P}(Y = m|X = n) = C_n^m p^m (1-p)^{n-m},\ m = 0, 1, \cdots, n.$$

试证: $Y \sim \mathscr{P}(p\lambda)$.

证明: 我们有

$$\begin{aligned}
\mathbb{P}(Y = m) &= \sum_{n=m}^{\infty} \mathbb{P}(Y = m|X = n)\mathbb{P}(X = n) \\
&= \sum_{n=m}^{\infty} C_n^m p^m (1-p)^{n-m} \frac{\lambda^n}{n!} \mathrm{e}^{-\lambda} \\
&= \sum_{n=m}^{\infty} \frac{p^m (1-p)^{n-m} \lambda^n \mathrm{e}^{-\lambda}}{m!(n-m)!} \\
&= \sum_{k=0}^{\infty} \frac{\lambda^m p^m}{m!} \frac{[\lambda(1-p)]^k}{k!} \mathrm{e}^{-\lambda} \\
&= \frac{(p\lambda)^m}{m!} \mathrm{e}^{-p\lambda}.
\end{aligned}$$

因此 $Y \sim \mathscr{P}(p\lambda)$. □

7.3.2 用户在单位时间内向电话局要求通话的总时间的平均值称为该电话局的话务量. 设单位时间内用户向电话局呼唤的次数 $N \sim \mathscr{P}(\lambda)$(表示 N 服从参数为 λ 的 Poisson 分布), 而每一用户的通话时间 $X_k \sim \mathscr{E}(\beta)$(表示 X_k 服从参数为 β 的指数分布), $k \in \mathbb{N}$, 则该电话局的话务量是 $\dfrac{\lambda}{\beta}$.

证明: 实际上话务量可以表示为 $\mathbb{E}\left[\sum_{k=1}^{N} X_k\right]$, 由例 7.3.7 可知 $\mathbb{E}\left[\sum_{k=1}^{N} X_k\right] = \mathbb{E}N\mathbb{E}X_1 = \dfrac{\lambda}{\beta}$. □

7.3.3 设二维随机变量 (X,Y) 具有概率分布函数 $p(x,y)$, 且 X 是可积随机变量, 则 $\mathbb{E}[X|X+Y=z]$ 可由下式给出:

$$\frac{\displaystyle\int xp(x,z-x)\mathrm{d}x}{\displaystyle\int p(x,z-x)\mathrm{d}x}.$$

证明: 令 $Y_1=X+Y$, 则

$$\mathbb{P}(X\leqslant t_1,Y_1\leqslant t_2)=\mathbb{P}(X\leqslant t_1,Y\leqslant t_2-X)=\int_{(-\infty,t_1]}\int_{(-\infty,t_2-x]}p(x,y)\mathrm{d}y\mathrm{d}x,$$

则 (X,Y_1) 的概率分布密度为 $p_1(x,y_1):=\dfrac{\partial^2}{\partial t_1\partial t_2}\mathbb{P}(X\leqslant t_1,Y_1\leqslant t_2)=p(x,t_2-x)$. 于是 $\mathbb{E}[X|X+Y=z]=\mathbb{E}[X|Y_1=z]$. 我们知道存在 $g(y_1)\in\mathscr{B}$, 使得 $\mathbb{E}[X|Y_1]=g(Y_1)$, a.e., 则

$$\forall B\in\mathscr{B},\quad \int_{Y_1^{-1}(B)}X\,\mathrm{d}\mathbb{P}=\int_{Y_1^{-1}(B)}g(Y_1)\mathrm{d}\mathbb{P}.$$

我们有

$$\begin{aligned}\int_{Y_1\in B}g(Y_1)\mathrm{d}\mathbb{P}&=\int_\Omega g(Y_1)\mathbb{1}_B(Y_1)\mathrm{d}\mathbb{P}\\&=\int_{\mathbb{R}^2}g(y_1)\mathbb{1}_B(y_1)p_1(x,y_1)\mathrm{d}x\mathrm{d}y_1\\&=\int_B g(y_1)\left(\int_{\mathbb{R}}p_1(x,y_1)\mathrm{d}x\right)\mathrm{d}y_1,\end{aligned}$$

同时

$$\begin{aligned}\int_{Y_1\in B}X\,\mathrm{d}\mathbb{P}&=\int_\Omega X\mathbb{1}_B(Y_1)\mathrm{d}\mathbb{P}\\&=\int_{\mathbb{R}^2}x\mathbb{1}_B(y_1)p_1(x,y_1)\mathrm{d}x\mathrm{d}y_1\\&=\int_B\left(x\int_{\mathbb{R}}p_1(x,y_1)\mathrm{d}x\right)\mathrm{d}y_1.\end{aligned}$$

因此 $g(y_1)=\dfrac{\displaystyle\int_{\mathbb{R}}xp_1(x,y_1)\mathrm{d}x}{\displaystyle\int_{\mathbb{R}}p_1(x,y_1)\mathrm{d}x}$, 故 $\mathbb{E}[X|X+Y=z]=\dfrac{\displaystyle\int xp(x,z-x)\mathrm{d}x}{\displaystyle\int p(x,z-x)\mathrm{d}x}$. □

7.3.4 设二维随机变量 (X,Y) 的概率分布函数

$$p(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-2\rho\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right\},$$

其中 $\sigma_1>0,\ \sigma_2>0,\ \rho\in(0,1)$, 试求 $\mathbb{E}[Y|X=x]$.

证明: 类似例 7.3.8 即有

$$\begin{aligned}\mathbb{E}[Y|X=x] &= \frac{\displaystyle\int_{\mathbb{R}} yp(x,y)\mathrm{d}y}{\displaystyle\int_{\mathbb{R}} p(x,y)\mathrm{d}x}\\ &= \frac{\displaystyle\int_{\mathbb{R}} y\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-2\rho\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right\}\mathrm{d}y}{\displaystyle\int_{\mathbb{R}} \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-2\rho\frac{xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right\}\mathrm{d}y}\\ &= \frac{\displaystyle\int_{\mathbb{R}} \frac{y}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2-\frac{x^2}{2\sigma_1^2(1-\rho^2)}\right\}\mathrm{d}y}{\displaystyle\int_{\mathbb{R}} \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2-\frac{x^2}{2\sigma_1^2(1-\rho^2)}\right\}\mathrm{d}y}\\ &= \frac{\displaystyle\int_{\mathbb{R}} y\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2\right\}\mathrm{d}y}{\displaystyle\int_{\mathbb{R}} \exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2\right\}\mathrm{d}y}.\end{aligned}$$

令 $\displaystyle\int_{\mathbb{R}} \exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2\right\}\mathrm{d}y = I$, 则

$$\int_{\mathbb{R}} y\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2\right\}\mathrm{d}y = \sigma_2\int_{\mathbb{R}}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2\right\}\mathrm{d}y - \frac{\sigma_2\rho x}{\sigma_1}I.$$

注意到

$$\int_{\mathbb{R}}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{y}{\sigma_2}-\rho\frac{x}{\sigma_1}\right)^2\right\}\mathrm{d}y = 0,$$

因此 $\mathbb{E}[Y|X=x] = \dfrac{\sigma_2}{\sigma_1}\rho x$. □

§ 7.4 条件期望的性质

7.4.1 在条件期望的控制收敛性 (IV) 中, 其他条件均成立, 将 $X_n \to X$, $\mathbb{P}$ a.e. 换成 $X_n \xrightarrow{\mathbb{P}} X$, 结果如何?
证明: 见课本推论 8.2.6. □

7.4.2 (1) 若 $\mathcal{C},\mathcal{C}'$ 是 $\mathcal{F}$ 的子 σ 代数, $\mathcal{C}\subset\mathcal{C}'$ 且 X' 是 $\mathcal{C}'$ 可测的, $\mathbb{E}XX',\mathbb{E}X$ 有限, 则

$$\mathbb{E}[XX'|\mathcal{C}] = \mathbb{E}[X'\mathbb{E}[X|\mathcal{C}']|\mathcal{C}],\ \mathbb{P}_{\mathcal{C}}\text{ a.e.}.$$

(2) 若 (1) 成立, 则定理 7.4.3 及定理 7.4.5 成立.
证明: (1) 我们只需证明: $\forall A\in\mathcal{C}\subset\mathcal{C}'$, $\displaystyle\int_A \mathbb{E}[XX'|\mathcal{C}]\mathrm{d}\mathbb{P} = \int_A X'\mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P}$. 也即

$$\forall A\in\mathcal{C}\subset\mathcal{C}',\ \int_A XX'\mathrm{d}\mathbb{P} = \int_A X'\mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P}.$$

(a) 当 X' 为示性函数, $X' = \mathbb{1}_B$, $B \in \mathcal{C}'$, 则

$$\int_A XX' \mathrm{d}\mathbb{P} = \int_{A\cap B} X \,\mathrm{d}\mathbb{P} = \int_{A\cap B} \mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P} = \int_A X'\mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P}.$$

(b) 当 X' 为非负简单函数 $X' = \sum_{n=1}^{\infty} a_n \mathbb{1}_{A_n}$, $A_n \in \mathcal{C}'$. 利用积分的线性性质可得此时成立.

(c) 当 X' 为非负可测函数, 由单调收敛定理, 存在非负简单函数列 $\{X'_n\}$, $0 \leqslant X'_n \uparrow X'$, 且

$$\int_A XX'\mathrm{d}\mathbb{P} = \lim_{n\to\infty}\int_A XX'_n \mathrm{d}\mathbb{P} = \lim_{n\to\infty}\int_A X'_n\mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P} = \int_A X'\mathbb{E}[X|\mathcal{C}'],$$

因此对非负可测的 X' 也成立.

(d) 当 X' 为一般的可测函数, 可将其拆分为正部和负部 X'^{+} 和 X'^{-}, 它们分别是非负可测函数, 则

$$\int_A XX'\mathrm{d}\mathbb{P} = \int_A (XX'^{+} - XX'^{-})\mathrm{d}\mathbb{P} = \int_A (X'^{+} - X'^{-})\mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P} = \int_A X'\mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P},$$

故此时成立.

综上, 对任意的 $X' \in \mathcal{C}'$, 都有 $\mathbb{E}[XX'|\mathcal{C}] = \mathbb{E}[X'\mathbb{E}[X|\mathcal{C}']|\mathcal{C}]$, $\mathbb{P}_{\mathcal{C}}$ a.e..

(2) 令 $\mathcal{C} = \mathcal{C}'$ 便是 $\mathbb{E}[XX'|\mathcal{C}] = \mathbb{E}[X'\mathbb{E}[X|\mathcal{C}]|\mathcal{C}] = X'\mathbb{E}[X|\mathcal{C}]$, 这便是定理 7.4.3.

令 $X' = 1$, 有 $\mathbb{E}[X|\mathcal{C}] = \mathbb{E}[\mathbb{E}[X|\mathcal{C}']|\mathcal{C}]$, 则 $\forall B \in \mathcal{C} \subset \mathcal{C}'$, 有

$$\int_B \mathbb{E}[\mathbb{E}[X|\mathcal{C}']|\mathcal{C}]\mathrm{d}\mathbb{P} = \int_B \mathbb{E}[X|\mathcal{C}']\mathrm{d}\mathbb{P} = \int_B X\,\mathrm{d}\mathbb{P} = \int_B \mathbb{E}[X|\mathcal{C}]\mathrm{d}\mathbb{P} = \int_B \mathbb{E}[\mathbb{E}[X|\mathcal{C}]|\mathcal{C}']\mathrm{d}\mathbb{P}.$$

这便是定理 7.4.5. □

7.4.3 设 $\mathcal{F}_n, n \in \mathbb{N}$ 是 $\mathcal{F}$ 的子 σ 代数列, 且 $\mathcal{F}_n \uparrow$, X_n 是 $\mathcal{F}_n$ 可测的, $n \in \mathbb{N}$. 若 $\forall n, m(m > n)$, 有

$$E[X_m|\mathcal{F}_n] = X_n, \quad \text{a.e.} \tag{*}$$

则称 $\{X_n : n \in \mathbb{N}\}$ 为一 $\mathcal{F}_n$ 鞅 (若将 (*) 式中 $= X_n$ 换为 $\geqslant X_n$ 或者 $\leqslant X_n$, 则对应称为 $\mathcal{F}_n$ 下鞅、$\mathcal{F}_n$ 上鞅). $\sigma(X_1, \cdots, X_n)$ 鞅 (下鞅, 上鞅) 简称为鞅 (相应地: 下鞅, 上鞅). 试证明:

(1) $\{X_n : n \in \mathbb{N}\}$ 为 $\mathcal{F}_n$ 鞅 (相应地: 下鞅, 上鞅) 的充分与必要条件是

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \text{ (相应地: } \geqslant X_n, \leqslant X_n)$$

(2) 设 $Y_n, n \in \mathbb{N}$ 为独立随机变量序列, 若 $\forall n \in \mathbb{N}$, $\mathbb{E}Y_n = 0$(相应地: $\geqslant 0, \leqslant 0$), 则 $\{X_n := \sum_{k=1}^{n} Y_k : n \in \mathbb{N}\}$ 为鞅 (相应地: 下鞅, 上鞅)

证明: (1) 必要性显然. 下面证明充分性, 我们采用归纳法证明:

$$\forall m > n, \mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \text{ a.e.} \Longrightarrow \mathbb{E}[X_m|\mathcal{F}_n] = X_n \text{ a.e..}$$

当 $m = n + 1$ 时显然成立. 考虑对某个 $m > n$, $\mathbb{E}[X_m|\mathcal{F}_n] = X_n$, 则

$$\mathbb{E}[X_{m+1}|\mathcal{F}_n] = \mathbb{E}[[X_{m+1}|\mathcal{F}_m]|\mathcal{F}_n] = \mathbb{E}[X_m|\mathcal{F}_n] = X_n \text{ a.e..}$$

因此 $\forall m > n$, $\mathbb{E}[X_m|\mathcal{F}_n] = X_n$, a.e. 当 $\{X_n, n \in \mathbb{N}\}$ 为 $\mathcal{F}_n$ 上鞅, 下鞅时可以仿照上述过程, 并根据条件期望的单调性立得.

(2) 我们知道,

$$\begin{aligned}\mathbb{E}[X_{n+1}|\sigma(X_1,\cdots,X_n)] &= \mathbb{E}\left[\sum_{k=1}^n Y_k + Y_{n+1}\middle|\sigma(Y_1,\cdots,Y_n)\right] \\ &= \sum_{k=1}^n Y_k + \mathbb{E}[Y_{n+1}|\sigma(Y_1,\cdots,Y_n)] \\ &= \sum_{k=1}^n Y_k = X_n.\end{aligned}$$

因此 $\{X_n\}$ 为鞅. 上鞅, 下鞅的情况类似. □

7.4.4 随机变量序列 $\{X_n : n \in \mathbb{Z}_+\}$ 称为**马尔可夫过程**, 如果 $\forall n \in \mathbb{Z}_+, B \in \mathscr{B}$, 有

(1) $$\mathbb{P}(X_{n+1} \in B|X_1, X_2, \cdots, X_n) = \mathbb{P}(X_{n+1} \in B|X_n) \quad \text{a.e.}$$

试证: (1) 与下列命题等价 (都是指对任何 $n \in \mathbb{Z}_+$ 而言)

(2) $\mathbb{E}[Y|X_1, X_2, \cdots, X_n] = \mathbb{E}[Y|X_n]$, $\forall Y \in \sigma(X_{n+1})$

(3) $\mathbb{E}[Y_1Y_2\cdots Y_m|X_1, X_2, \cdots, X_n] = \mathbb{E}[Y_1Y_2\cdots Y_m|X_n]$, $\forall Y_k \in \sigma(X_{n+k}), k = 1, 2, \cdots m$

(4) $\mathbb{E}[Y|X_1, X_2, \cdots, X_n] = \mathbb{E}[Y|X_n]$, $\forall Y \in \sigma(\{X_m : m > n\})$

(5) $\mathbb{P}(FB|X_n) = \mathbb{P}(F|X_n)\mathbb{P}(B|X_n)$, $\forall B \in \sigma(X_1, X_2, \cdots, X_n), F \in \sigma(\{X_m : m > n\})$

证明: 我们将证明 (1)$\Longrightarrow$(2)$\Longrightarrow$(3)$\Longrightarrow$(4)$\Longrightarrow$(5)$\Longrightarrow$(1).

(1)$\Longrightarrow$(2): 由于 $Y \in \sigma(X_{n+1})$, 因此存在 $f \in \mathscr{B}$ 使得 $Y = f(X_{n+1})$. 因此只需由 (1) 推出

$$\forall f \in \mathscr{B}, \mathbb{E}[f(X_{n+1})|X_1, \cdots, X_n] = \mathbb{E}[f(X_{n+1})|X_n].$$

(i) 当 f 是示性函数 $f = \mathbb{1}_B$, $B \in \mathscr{B}$, 由 (1) 知此时成立;

(ii) 当 f 是任意非负简单函数, 根据条件期望的线性性质知此时成立;

(iii) 当 f 是任意非负可测函数, 根据条件期望的单调收敛定理, 存在非负简单函数列 $\{f_n : n \in \mathbb{N}\}$ 使得 $f_n \uparrow f$. 于是有

$$\begin{aligned}\mathbb{E}[f(X_{n+1})|X_1, \cdots, X_n] &= \lim_{k\to\infty} \mathbb{E}[f_k(X_{n+1})|X_1, \cdots, X_n] \\ &= \lim_{k\to\infty} \mathbb{E}[f_k(X_{n+1})|X_n] = \mathbb{E}[f(X_{n+1})|X_n].\end{aligned}$$

因此此时也成立;

(iv) 对于一般可测函数 f, 可将其分解为正部和负部. 利用 (iii) 和条件期望的线性性质可知此时也成立.

(2)$\Longrightarrow$(3): 采用归纳法证明. 显然对 $m=1$, (3) 成立. 假设对某个 $m\in\mathbb{N}$, (3) 成立, 则:

$$\begin{aligned}\mathbb{E}[Y_1\cdots Y_mY_{m+1}|X_1,\cdots,X_n]&=\mathbb{E}[\mathbb{E}[Y_1\cdots Y_{m+1}|X_1,\cdots,X_{n+m}]|X_1,\cdots,X_n]\\&=\mathbb{E}[Y_1\cdots Y_m\mathbb{E}[Y_{m+1}|X_1\cdots,X_{n+m}]|X_1,\cdots,X_n]\\&=\mathbb{E}[Y_1\cdots Y_m\mathbb{E}[Y_{m+1}|X_{n+m}]|X_1,\cdots,X_n]\\&=\mathbb{E}[Y_1\cdots Y_m\mathbb{E}[Y_{m+1}|X_{n+m}]|X_n]\\&=\mathbb{E}[\mathbb{E}[Y_1\cdots Y_mY_{m+1}|X_{n+m}]|X_n]\\&=\mathbb{E}[Y_1\cdots Y_mY_{m+1}|X_n].\end{aligned}$$

(3)$\Longrightarrow$(4): 先证明

$$\forall A\in\sigma(\{X_m:m>n\}),\quad \mathbb{E}[\mathbb{1}_A|X_1,\cdots,X_n]=\mathbb{E}[\mathbb{1}_A|X_n].$$

考虑集类

$$\Lambda:=\{A\in\sigma(\{X_m,m>n\}),\mathbb{E}[\mathbb{1}_A|X_1,\cdots,X_n]=\mathbb{E}[\mathbb{1}_A|X_n]\}$$

以及

$$\mathscr{A}:=\left\{\bigcap_{i=1}^k\{X_{n+i}\in C_i\},C_i\in\mathscr{B}\right\},$$

则 $\mathscr{A}$ 对交封闭, 是 π 系. 我们知道 $\forall C_i\in\mathscr{B}$, $C:=\bigcap\limits_{i=1}^k\{X_{n+i}\in C_i\}\in\sigma(\{X_m,m>n\})$. 我们知道

$$\mathbb{1}_C=\mathbb{1}_{C_1}(X_{n+1})\mathbb{1}_{C_2}(X_{n+2})\cdots\mathbb{1}_{C_k}(X_{n+k}),$$

且 $\mathbb{1}_{C_i}(X_{n+i})\in\sigma(X_{n+i})$, 根据 (3) 有 $\mathbb{E}[\mathbb{1}_C|X_1,\cdots,X_n]=\mathbb{E}[\mathbb{1}_C|X_n]$, 因此 $\mathscr{A}\subset\Lambda$.

下面证明 Λ 是 λ 系. 显然 $\Omega\in\Lambda$, 考虑 $A_1,A_2\in\Lambda$, $A_2\subset A_1$, 则

$$\begin{aligned}\mathbb{E}[\mathbb{1}_{A_1\backslash A_2}|X_1,\cdots,X_n]&=\mathbb{E}[\mathbb{1}_{A_1}|X_1,\cdots,X_n]-\mathbb{E}[\mathbb{1}_{A_2}|X_1,\cdots,X_n]\\&=\mathbb{E}[\mathbb{1}_{A_1}|X_n]-\mathbb{E}[\mathbb{1}_{A_2}|X_n]=\mathbb{E}[\mathbb{1}_{A_1\backslash A_2}|X_n].\end{aligned}$$

因此 $A_1\backslash A_2\in\Lambda$. 考虑不降序列 $\{A_k:k\in\mathbb{N}\}$, $A_k\uparrow$, 则 $\mathbb{1}_{A_k}\uparrow-\lim\limits_{n\to\infty}\mathbb{1}_{A_k}$, 由单调收敛定理有

$$\mathbb{E}[\lim_{k\to\infty}\mathbb{1}_{A_k}|X_1,\cdots,X_n]=\lim_{k\to\infty}\mathbb{E}[\mathbb{1}_{A_k}|X_1,\cdots,X_n]=\lim_{k\to\infty}\mathbb{E}[\mathbb{1}_{A_k}|X_n]=\mathbb{E}[\lim_{k\to\infty}\mathbb{1}_{A_k}|X_n].$$

因此 Λ 对不降序列的并封闭, Λ 是 λ 系, 又 $\mathscr{A}\subset\Lambda$, 故 Λ 是 σ 代数. 因此

$$\sigma(\{X_m,m>n\})=\sigma(\mathscr{A})\subset\Lambda\subset\sigma(\{X_m,m>n\}).$$

故 $\Lambda=\sigma(\{X_m,m>n\})$, $\forall A\in\sigma(\{X_m,m>n\})$, $\mathbb{E}[\mathbb{1}_A|X_1,\cdots,X_n]=\mathbb{E}[\mathbb{1}_A|X_n]\}$. 与 (1)$\Longrightarrow$(2) 的证明同理可得

$$\forall Y\in\sigma(\{X_m,m>n\}),\mathbb{E}[Y|X_1,\cdots,X_n]=\mathbb{E}[Y|X_n]\}.$$

(4)$\Longrightarrow$(5): 我们有

$$\begin{aligned}\mathbb{P}(F\cap B|X_n)&=\mathbb{E}[\mathbb{P}(F\cap B|X_1,\cdots,X_n)|X_n]\\&=\mathbb{E}[\mathbb{1}_B\mathbb{P}(F|X_1,\cdots,X_n)|X_n]\\&=\mathbb{E}[\mathbb{1}_B\mathbb{P}(F|X_n)|X_n]\\&=\mathbb{P}(F|X_n)\mathbb{E}[\mathbb{1}_B|X_n]\\&=\mathbb{P}(F|X_n)\mathbb{P}(B|X_n).\end{aligned}$$

(5)$\Longrightarrow$(1): 令 $F=\{X_{n+1}\in A_{n+1}\}$, $B=\{X_\ell\in A_\ell, 1\leqslant \ell\leqslant n\}$, 则

$$\mathbb{P}(\{X_\ell\in A_\ell,1\leqslant\ell\leqslant n\}\cap\{X_{n+1}\in A_{n+1}\}|X_n)=\mathbb{P}(X_\ell\in A_\ell,1\leqslant\ell\leqslant n|X_n)\mathbb{P}(X_{n+1}\in A_{n+1}|X_n),$$

得到

$$\mathbb{P}(X_{n+1}\in A_{n+1}|X_1,\cdots,X_n)=\mathbb{P}(X_{n+1}\in A_{n+1}|X_n).$$

证毕. □

7.4.5 设 $\mathcal{C}\subset\mathcal{F}$ 且随机变量 X,Y 满足 $\mathbb{E}[Y^2|\mathcal{C}]=X^2$ a.e., $\mathbb{E}[Y|\mathcal{C}]=X$ a.e., 则 $X=Y$ a.e..

证明: 我们知道 $X,X^2\in\mathcal{C}$, a.e., 于是

$$\begin{aligned}\mathbb{E}[(X-Y)^2|\mathcal{C}]&=\mathbb{E}[X^2-2XY+Y^2|\mathcal{C}]\\&=\mathbb{E}[Y^2|\mathcal{C}]+\mathbb{E}[X^2|\mathcal{C}]-2\mathbb{E}[XY|\mathcal{C}]\\&=\mathbb{E}[Y^2|\mathcal{C}]+\mathbb{E}[X^2|\mathcal{C}]-2X\mathbb{E}[X|\mathcal{C}]\\&=X^2+X^2-2X^2=0.\ \text{a.e..}\end{aligned}$$

因此 $X=Y$ a.e.. □

§ 7.5 条件概率分布

7.5.1 设 X_T 是 $(\Omega,\mathcal{F},\mathbb{P})$ 到 $(\mathbb{R}^n,\mathscr{B}^T)$ 上的可测映射, $\mathcal{C}$ 是 $\mathcal{F}$ 的子 σ 代数, 则 $\mathbb{P}^{\mathcal{C}}(\omega,B)$ ($(\omega,B)\in\Omega\times\mathscr{B}^T$) 是 X_T 在 $\mathcal{C}$ 下的条件概率分布的充分必要条件是下述二条件同时成立:

(1) $\mathbb{P}^{\mathcal{C}}(\omega,B)$ 是 $(\Omega,\mathcal{C})$ 到 $(\mathbb{R}^T,\mathscr{B}^T)$ 的转移概率;

(2) $\forall B^T\in\mathscr{B}^T, C\in\mathcal{C}$,

$$\int_C\mathbb{P}^{\mathcal{C}}(\,\cdot\,,B^T)\mathrm{d}\mathbb{P}=\mathbb{P}(C\cap\{X_T\in B^T\}).$$

证明: 必要性: 如果 $\mathbb{P}^{\mathcal{C}}(\omega,B)$ 是 X_T 在 $\mathcal{C}$ 之下的条件概率分布, 则:

(1): 固定 $B\in\mathscr{B}^T$, $\mathbb{P}^{\mathcal{C}}(\omega,B)\in\mathcal{C}$; 固定 ω, $\mathbb{P}^{\mathcal{C}}(\omega,\bullet)$ 是 $(\Omega,\mathcal{C})$ 上的概率分布.

(2): 由条件期望的定义可以知道, $\forall C\in\mathcal{C}$, 我们有

$$\mathbb{P}(C\cap\{X_T\in B\})=\int_C\mathbb{1}_{X_T\in B}\mathrm{d}\mathbb{P}=\int_C\mathbb{P}^{\mathcal{C}}(\bullet,B^T)\mathbb{P}(\mathrm{d}\omega).$$

因此 $\displaystyle\int_C\mathbb{P}^{\mathcal{C}}(\,\cdot\,,B^T)\mathrm{d}\mathbb{P}=\mathbb{P}(C\cap\{X_T\in B^T\})$.

充分性: 由 (1) 知, 对任意固定的 $B^T\in\mathscr{B}^T$, $\mathbb{P}^{\mathcal{C}}(\omega,B)\in\mathcal{C}$, 且

$$\int_C\mathbb{1}_{\{X_T\in B\}}\mathrm{d}\mathbb{P}=\mathbb{P}(C\cap\{X_T\in B\})=\int_C\mathbb{P}^{\mathcal{C}}(\bullet,B^T)\mathbb{P}(\mathrm{d}\omega),$$

由条件期望的定义可得 $\mathbb{P}(X^T\in B^T|\mathcal{C})=\mathbb{P}^{\mathcal{C}}(\omega,B)$, 再由定义 7.5.4 可知 $\mathbb{P}^{\mathcal{C}}(\omega,B)$ 为 X^T 在 $\mathcal{C}$ 上的条件概率分布. □

7.5.2 设 $X = (X_1, X_2, \cdots, X_n)$ 是 $(\mathbb{R}^n, \mathscr{B}^n, \mathbb{P})$ 上的样本函数, 具有 n 维正态分布 $N(0, D)$; 令 $Y(x) = xD^{-1}x', x \in \mathbb{R}^n$, 试求 X 在 $\sigma(Y)$ 之下的条件概率分布.

证明: (by 229) 设 $D := O^\top \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\} O$, 以及 $Q_n := \operatorname{diag}\{\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}\}$. 则 $D = O^\top Q_n^\top Q_n O$. 我们知道 $X \sim N(0, D)$, 故 $W := XO^\top (Q_n^{-1})^\top \sim N(0, I_n)$. 又 $D^\top = O^\top (Q_n^{-1})^\top Q_n^{-1} O$, 故 $Y := XD^{-1}X = WW^\top \sim \chi^2(n)$.

实际上, $\forall A \in \mathscr{B}^n$, 有

$$\begin{aligned}\mathbb{P}(X \in A | Y = y) &= \mathbb{P}\left(WQ_nO \in A | \sum_{i=1}^n w_i^2 = y\right) = \mathbb{P}\left(WQ_n \in A | \sum_{i=1}^n w_i^2 = y\right) \\ &= \mathbb{P}\left(W \in AQ_n^{-1} | \sum_{i=1}^n w_i^2 = y\right).\end{aligned}$$

考虑变换 $w_n = \sqrt{T} \cos\theta_n \sin\theta_{n-1} \cdots \sin\theta_1$. 则 $(T, \theta_1, \cdots, \theta_{n-1})$ 的联合密度函数为

$$f(T, \theta_1, \cdots, \theta_{n-1}) = \frac{1}{(2\pi)^{n/2}} \mathrm{e}^{-\frac{1}{2}T} T^{\frac{n}{2}-1} \sin^{n-2}\theta_1 \cdots \sin\theta_{n-2}.$$

同时 $(\theta_1, \cdots, \theta_{n-1} | T)$ 的条件密度函数为

$$f(\theta_1, \cdots, \theta_{n-1} | T) = \frac{\Gamma\left(\frac{n}{2}\right)}{\pi^{\frac{n}{2}}} \sin^{n-2}\theta_1 \cdots \sin\theta_{n-2}.$$

因此

$$\mathbb{P}(\omega, A) := \mathbb{P}(X \in A | Y = Y(\omega)) = \mathbb{P}\left(W^\top \in AQ_n^{-1} | \sum_{i=1}^n w_i^2 = Y(\omega)\right) = \frac{\Gamma\left(\frac{n}{2}\right)}{\pi^{\frac{n}{2}}} \Theta(F(A, \omega)).$$

其中 Θ 是球面测度, 也即 $\mathrm{d}\Theta = \sin^{n-1}\theta_1 \cdots \sin\theta_{n-2} \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_{n-1}$. 同时

$$F(A, \omega) = \left\{x \in \mathbb{R}^n \left| \left(\frac{x_1}{\sqrt{\lambda_1}}, \cdots, \frac{x_n}{\sqrt{\lambda_n}}\right) \in A, |x|^2 = Y(\omega)\right.\right\}.$$

对 $\mathbb{R}^n$ 作仿射变换 $\mathscr{T} : y_i = \dfrac{x_i}{\sqrt{\lambda_i}}, 1 \leqslant i \leqslant n$, 则

$$\mathscr{T}(F) = \left\{y \in \mathbb{R}^n \left| y \in A, \sum_{i=1}^n y_i^2 \lambda_i = Y(\omega)\right.\right\}.$$

同时此映射将 n 维球面映射到 $\Gamma(\omega) := \left\{y \in \mathbb{R}^n \left| \sum_{i=1}^n \lambda_i y_i^2 = Y(\omega)\right.\right\}$, 则

$$\mathbb{P}(\omega, A) = \frac{\mathbb{H}(A \cap \Gamma(\omega))}{\mathbb{H}(\Gamma(\omega))}.$$

其中 $\mathbb{H}$ 是球面测度仿射后得到的椭球面测度. 因此 X 在 $\sigma(Y)$ 下的条件概率分布是椭球面 $\Gamma(\omega)$ 上的均匀分布. □

7.5.3 称随机变量序列 $X_n, n \in \mathbb{N}$ 为一马尔可夫序列, 如果对任意 $B \in \mathscr{B}$, $(x_1, x_2, \cdots, x_{n-1}) \in \mathbb{R}^{n-1}$, $n = 2, 3, \cdots$, 有

$$\mathbb{P}(X_n \in B | X_1 = x_1, X_2 = x_2, \cdots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n \in B | X_{n-1} = x_{n-1}), \quad \text{a.e..}$$

若给定一组转移概率序列

$$P_1(B_1),\quad B_1\in\mathscr{B},$$
$$P_n(x_{n-1},B_n),\quad B_n\in\mathscr{B},\ x_{n-1}\in\mathbb{R},\ n=2,3,\cdots$$

则有一定义在概率空间 $(\mathbb{R}^{\mathbb{N}},\mathscr{B}^{\mathbb{N}},\mathbb{P}^{\mathbb{N}})$ 上的马尔科夫序列 $\{X_n,n\in\mathbb{N}\}$ 使得 X_n 在 X_{n-1} 之下的条件概率

$$\mathbb{P}^{\mathbb{N}}(X_n\in B|X_{n-1}=x_{n-1})=P_n(x_{n-1},B)\quad \text{a.e.},$$
$$B\in\mathscr{B},\ x_{n-1}\in\mathbb{R},\ n=2,3,\cdots;$$

而
$$\mathbb{P}^{\mathbb{N}}(X_1\in B)=P_1(B),\ B\in\mathscr{B}.$$

(提示: 将此处的 $P_n(x_{n-1},B_n)$ 看成是定理 7.5.15 证明中的 $P_n(x_1,x_2,\cdots,x_{n-1},B_n)$.)

证明: 定义

$$P(x_1,x_2,\cdots,x_{n-1},B)=P_n(x_{n-1},B),$$

我们知道 $P_n(x_{n-1},B)$ 是 $(\mathbb{R},\mathscr{B})$ 到 $(\mathbb{R},\mathscr{B})$ 的转移概率, 因此 $P(x_1,x_2,\cdots,x_{n-1},B)$ 是从 $(\mathbb{R}^{n-1},\mathscr{B}^{n-1})$ 到 $(\mathbb{R},\mathscr{B})$ 的转移概率. 由 Tulcea 定理, 存在 $(\mathbb{R}^{\mathbb{N}},\mathscr{B}^{\mathbb{N}})$ 上的概率测度 $\mathbb{P}^{\mathbb{N}}$, 使得对任意的 $B^n\in\mathscr{B}^n$, 有

$$\mathbb{P}^{\mathbb{N}}\left(B^n\times\mathbb{R}^{\mathbb{N}\setminus\{1,2,\cdots,n\}}\right)=\int\cdots\int_{B^n}P_n(x_{n-1},\mathrm{d}x_n)\cdots P_2(x_1,\mathrm{d}x_2)P_1(\mathrm{d}x_1).$$

因此 $\forall f\in\mathscr{B}^n$, 有

$$\int f(x_1,\cdots,x_n)\mathrm{d}\mathbb{P}^{\mathbb{N}}=\int\cdots\int_{\mathbb{R}^n}f(x_1,\cdots,x_n)P_n(x_{n-1},\mathrm{d}x_n)\cdots P_2(x_1,\mathrm{d}x_2)P_1(\mathrm{d}x_1).$$

考虑 $(\mathbb{R}^{\mathbb{N}},\mathscr{B}^{\mathbb{N}})$ 上的随机变量 $X=(X_1,X_2,\cdots,X_n,\cdots)$. 对任意的 $A\in\mathscr{B}$, 有

$$\begin{aligned}\int_{X_{n-1}^{-1}(A)}\mathbb{1}_{X_n^{-1}(B)}\mathrm{d}\mathbb{P}^{\mathbb{N}}&=\int\cdots\int_{X_{n-1}^{-1}(A)\cap X_n^{-1}(B)}P_n(x_{n-1},\mathrm{d}x_n)\cdots P_2(x_1,\mathrm{d}x_2)P_1(\mathrm{d}x_1)\\&=\int\cdots\int_{X_{n-1}^{-1}(A)}P_n(x_{n-1},B)\cdots P_2(x_1,\mathrm{d}x_2)P_1(\mathrm{d}x_1)\\&=\cdots\\&=\int_{X_{n-1}^{-1}(A)}P_n(x_{n-1},B)\mathrm{d}\mathbb{P}^{\mathbb{N}},\end{aligned}$$

又 $P_n(x_{n-1},B)$ 关于 $(\mathbb{R},\mathscr{B})$ 可测, 于是 $\mathbb{P}^{\mathbb{N}}(X_n\in B|X_{n-1}=x_{n-1})=P_n(x_{n-1},B)$, a.e..

又对任意的 $A_1,\cdots,A_{n-1}\in\mathscr{B}$, 有

$$\begin{aligned}\int_{\cap_{i=1}^{n-1}X_i^{-1}(A_i)}\mathbb{1}_{X_n^{-1}(B)}\mathrm{d}\mathbb{P}^{\mathbb{N}}&=\int\cdots\int_{(\cap_{i=1}^{n-1}X_i^{-1}(A_i))\cap X_n^{-1}(B)}P_n(x_{n-1},\mathrm{d}x_n)\cdots P_2(x_1,\mathrm{d}x_2)P_1(\mathrm{d}x_1)\\&=\int_{\cap_{i=1}^{n-1}X_i^{-1}(A_i)}P_n(x_{n-1},B)P_{n-1}(x_{n-2},\mathrm{d}x_{n-1})\cdots P_2(x_1,\mathrm{d}x_2)P_1(\mathrm{d}x_1)\\&=\int_{\cap_{i=1}^{n-1}X_i^{-1}(A_i)}P_n(X_{n-1},B)\mathrm{d}\mathbb{P}^{\mathbb{N}},\end{aligned}$$

又 $P_n(x_{n-1},B)$ 关于 $(\mathbb{R},\mathscr{B})$ 可测, 于是

$$\mathbb{P}(X_n\in B|X_1=x_1,X_2=x_2,\cdots,X_{n-1}=x_{n-1})=\mathbb{P}(X_n\in B|X_{n-1}=x_{n-1}),a.e..$$

其初始分布为 $\mathbb{P}^{\mathbb{N}}(X_1\in B)=\int_{X_1^{-1}(B)}P_1(\mathrm{d}x_1)=P_1(B)$. □

第八章 收敛概念

§ 8.1 几乎处处收敛

8.1.1 证明引理 8.1.6.

(提示: 证明 $\forall\varepsilon>0,\exists n_0,s.t.\forall n\geqslant n_0,\sum\limits_{k=n}^{\infty}\delta_k<\varepsilon$且$\forall\nu\geqslant 1,\{d\left(f_{n+\nu},f_n\right)\geqslant\varepsilon\}\subset\bigcup\limits_{k=n}^{\infty}\{d\left(f_{k+1},f_k\right)\geqslant\delta_k\}$.)

证明: 由 $\sum\limits_{k=1}^{\infty}\delta_k<\infty$ 知 $\lim\limits_{n\to\infty}\sum\limits_{k=n}^{\infty}\delta_k=0$, 从而 $\forall\varepsilon>0,\ \exists n_0\in\mathbb{N},\ \text{s.t.}\forall n\geqslant n_0,\ \sum\limits_{k=n}^{\infty}\delta_k<\varepsilon$, 从而

$$\forall n\geqslant n_0,\ \forall\nu\in\mathbb{N},\ \{d(f_{n+\nu},f_n)\geqslant\varepsilon\}\subset\bigcup_{k=n}^{\infty}\{d(f_{k+1},f_k)\geqslant\delta_k\}$$

(这是因为 $\forall\nu\in\mathbb{N},\ \bigcap\limits_{k=n}^{\infty}\{d(f_{k+1},f_k)<\delta_k\}\subset\{d(f_{n+\nu},f_n)<\varepsilon\}$, 即, 如果 $\forall k\geqslant n,d(f_{k+1},f_k)<\delta_k$, 则

$$d(f_{n+\nu},f_n)\overset{\text{距离的三角不等式}}{\leqslant}\sum_{k=n}^{n+\nu-1}d(f_{k+1},f_k)<\sum_{k=n}^{n+\nu-1}\delta_k<\sum_{k=n}^{\infty}\delta_k<\varepsilon,\ \forall\nu\in\mathbb{N}$$

), 从而 $\forall n\geqslant n_0$,

$$\begin{aligned}&\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f_n)\geqslant\varepsilon\}\subset\bigcup_{k=n}^{\infty}\{d(f_{k+1},f_k)\geqslant\delta_k\}\\ \Longrightarrow&\mu\left(\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f_n)\geqslant\varepsilon\}\right)\leqslant\mu\left(\bigcup_{k=n}^{\infty}\{d(f_{k+1},f_k)\geqslant\delta_k\}\right)\leqslant\sum_{k=n}^{\infty}\mu\{d(f_{k+1},f_k)\geqslant\delta_k\}\end{aligned}$$

又, 有前提 $\sum\limits_{n=1}^{\infty}\mu\{d(f_{n+1},f_n)\geqslant\delta_n\}<\infty$, 则 $\lim\limits_{n\to\infty}\sum\limits_{k=n}^{\infty}\mu\{d(f_{k+1},f_k)\geqslant\delta_k\}=0$, 从而

$$\forall\varepsilon>0,\ \mu\left(\bigcup_{\nu=1}^{\infty}\{d(f_{n+\nu},f_n)\geqslant\varepsilon\}\right)\to 0\ (n\to\infty)$$

即定理 8.1.3 中 (4) 式成立, 由定理 8.1.3 知 $f_n,n\in\mathbb{N}$ a.e. 相互收敛. □

8.1.2 若 $(\Omega,\mathcal{F},\mu)$ 为测度空间, μ 有限, $\{B_n,n\in\mathbb{N}\}\subset\mathcal{F}$, 且 $\exists\delta>0,s.t.\forall n\in\mathbb{N},\mu\left(B_n\right)\geqslant\delta$, 试证在 Ω 中至少存在一个点 ω 属于无穷多个 B_n.

证明: 由 μ 具有上方连续性质 (定理 3.3.2) 及 μ 有限知

$$\mu\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}B_k\right)=\lim_{n\to\infty}\mu\left(\bigcup_{k=n}^{\infty}B_k\right)\geqslant\lim_{n\to\infty}\mu\left(B_n\right)\overset{\forall n,\ \mu(B_n)\geqslant\delta}{\geqslant}\delta>0$$

从而 $\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k \neq \varnothing$, 即 $\exists \omega \in \Omega$, s.t.$\forall n \geqslant 1$, $\exists k \geqslant n$, s.t.$\omega \in B_k$, 即 Ω 中至少存在一个点 ω 属于无穷多个 B_n. (或, 即 $\limsup_{n\to\infty} B_n \neq \varnothing$, 由定义 1.1.6 中集合序列上极限的定义即得证.) □

8.1.3 概率空间 $(\Omega, \mathcal{F}, \mu)$ 上的随机变量列 $X_n \xrightarrow{\text{a.e.}} X$ 有限, 试证: $\forall \varepsilon > 0, \exists M(\varepsilon) > 0, s.t.$

$$P\left\{\sup_n |X_n| \leqslant M(\varepsilon)\right\} \geqslant 1 - \varepsilon.$$

注: 教材在定义 4.1.1 中约定: 当我们只说 X 是一个 "随机变量" 时, 默认 X 只能取有限实值.

证明: 由 $X_n \xrightarrow{\text{a.e.}} X$ 及 Egorov 定理 (8.1.7) 知

$$\forall \varepsilon > 0,\ \exists A_\varepsilon \in \mathcal{F},\ \text{s.t.}\mathbb{P}\left(A_\varepsilon^c\right) < \frac{\varepsilon}{2}\ \text{且}\ \sup_{\omega\in A_\varepsilon} d(X_n(\omega), X(\omega)) \to 0\ (n \to \infty)$$

即在 A_ε 上一致收敛, 从而由 X 有限知对于每一点 $\omega \in A_\varepsilon$, $\sup_n |X_n(\omega)| < \infty$, 这是因为由 $\sup_{\omega\in A_\varepsilon} d(X_n(\omega), X(\omega)) \to 0\ (n \to \infty)$ 可知

$$\begin{aligned}
&\exists N \in \mathbb{N},\ \text{s.t.}\forall n \geqslant N,\ \sup_{\omega\in A_\varepsilon} d(X_n(\omega), X(\omega)) < \varepsilon \\
&\Longrightarrow n \geqslant N\text{时}\ \forall \omega \in A_\varepsilon, |X_n(\omega)| - |X(\omega)| \leqslant |X_n(\omega) - X(\omega)| < \varepsilon, \\
&\qquad\qquad\qquad \text{即}\ |X_n(\omega)| \leqslant |X(\omega)| + \varepsilon \\
&\Longrightarrow \forall n \in \mathbb{N}, \forall \omega \in A_\varepsilon, |X_n(\omega)| \leqslant \max\left\{|X_1(\omega)|, |X_2(\omega)|, \cdots, |X_{N-1}(\omega)|, |X(\omega)| + \varepsilon\right\} < \infty
\end{aligned}$$

记 $M(\varepsilon, \omega) := \max\left\{|X_1(\omega)|, |X_2(\omega)|, \cdots, |X_{N-1}(\omega)|, |X(\omega)| + \varepsilon\right\}$, 则此时我们已经证明

$$\forall \omega \in A_\varepsilon,\ \sup_n |X_n(\omega)| \leqslant M(\varepsilon, \omega)$$

但题目要找的 $M(\varepsilon)$ 与 ω 无关, 因此我们还需要证明 $X_1, X_2, \cdots, X_{N-1}, X$ 均关于 ω 具有一致有界性, 或者说是几乎一致有界的.

事实上, 有如下断言成立: 若测度有限, 则对于取有限实值的可测函数而言, 它必然几乎一致有界. 该断言的意义可以参考证明后的附注帮助理解. 对于任意取有限实值的随机变量 X 而言, 此断言的意思是

$$\forall \delta > 0,\ \exists B_\delta \in \mathcal{F},\ \exists M(\delta) > 0,\ \text{s.t.}\mathbb{P}\left(B_\delta^c\right) < \delta\ \text{且}\ \sup_{\omega\in B_\delta} |X(\omega)| \leqslant M(\delta)$$

该断言的证明如下: 由 X 取有限实值知 $\Omega = \bigcup_{k=1}^{\infty} \{|X| \leqslant k\}$, 从而由测度的下连续性 (定理 3.3.2) 知

$$\lim_{n\to\infty} \mathbb{P}\{|X| \leqslant n\} = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=1}^{n} \{|X| \leqslant k\}\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} \{|X| \leqslant k\}\right) = \mathbb{P}(\Omega) = 1$$

从而 $\forall \delta > 0,\ \exists N_0 \in \mathbb{N},\ \text{s.t.}\forall n \geqslant N_0,\ \mathbb{P}\{|X| \leqslant n\} \geqslant 1 - \delta$, 取 $B_\delta = \{|X| \leqslant N_0\}$, $M(\delta) = N_0$ 即证.

由以上断言可知:

$$\begin{aligned}
&\forall \varepsilon > 0,\ \exists B_{i,\varepsilon} \in \mathcal{F},\ \exists M_i(\varepsilon) > 0,\ i = 1, 2, \cdots, N,\ \text{s.t.}\mathbb{P}\left(B_{i,\varepsilon}^c\right) < \frac{\varepsilon}{2N}, i = 1, 2, \cdots, N, \\
&\qquad \text{且}\ \sup_{\omega\in B_{i,\varepsilon}} |X_i(\omega)| \leqslant M_i(\varepsilon), i = 1, 2, \cdots, N-1,\ \sup_{\omega\in B_{N,\varepsilon}} |X(\omega)| \leqslant M_{N,\varepsilon}(\varepsilon)
\end{aligned}$$

从而, 取 $C_\varepsilon = A_\varepsilon \cap \left(\bigcap_{i=1}^{N} B_{i,\varepsilon}\right) \in \mathcal{F}$, $M(\varepsilon) = \max\{M_1(\varepsilon), M_2(\varepsilon), \cdots, M_{N-1}(\varepsilon), M_N(\varepsilon)+\varepsilon\}$, 则有

$$\forall \omega \in C_\varepsilon,\ \sup_n |X_n(\omega)| \leqslant M(\varepsilon,\omega) \leqslant M(\varepsilon)$$

且 $\mathbb{P}(C_\varepsilon^c) = \mathbb{P}\left(A_\varepsilon^c \cup \left(\bigcup_{i=1}^{N} B_{i,\varepsilon}^c\right)\right) \leqslant \mathbb{P}(A_\varepsilon^c) + \sum_{i=1}^{N} \mathbb{P}\left(B_{i,\varepsilon}^c\right) < \frac{\varepsilon}{2} + N\frac{\varepsilon}{2N} = \varepsilon$, 从而

$$\mathbb{P}\left\{\sup_n |X_n| \leqslant M(\varepsilon)\right\} \geqslant \mathbb{P}(C_\varepsilon) \geqslant 1-\varepsilon$$

证毕. □

注: (1) 上述证明中提出的断言可以这样理解: 例如设 $\Omega = (0,1)$, $\mathcal{F} = \mathscr{B}(0,1)$, $\mu = \lambda|_{\mathscr{B}(0,1)}$(勒贝格测度限制在 $\mathscr{B}(0,1)$ 上), 考虑函数 $f(x) = \frac{1}{x}$, 它在 Ω 的每个点上虽然取有限实值, 在 Ω 上整体而言是无界的: $\sup_{x\in\Omega} |f(x)| = \infty$; 但如果挖去 0 的一个小邻域 $B_\delta = (0,\delta)$, 就有 $\sup_{x\in\Omega\setminus B_\delta} |f(x)| = \frac{1}{\delta} < \infty$, 即在 $\Omega \setminus B_\delta$ 上是一致有界的.

(2) 上述断言要求测度是有限测度, 否则未必成立, 例如取 $\Omega = \mathbb{R}$, $\mathcal{F} = \mathscr{B}(\mathbb{R})$, $\mu = \lambda$(Lebesgue 测度), $f(x) = x$, 则不可能仅仅挖去一个测度比较小的集合后 $f(x)$ 一致有界.

(3) 上述证明过程还可以稍作简化, 例如可以先直接限制在 X 有界的集合上使用 Egorov 定理 (8.1.7)(构造一个以 Ω 的子集为全空间的较小的测度空间, 参考习题 3.2.21), 得到一致收敛性, 再去考虑前 $N-1$ 个 X_i 的有界性.

8.1.4 对概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 上的任何随机变量列 $\{X_n, n\in\mathbb{N}\}$, 存在常数序列 $\{A_n, n\in\mathbb{N}\}$, 使得

$$\frac{X_n}{A_n} \xrightarrow{\text{a.e.}} 0.$$

证法一: 令 $a_{n,m} = (n(\sup_{\omega\in\Omega} |X_n(\omega)| \wedge m)) \vee n^2$. 考虑集合 $E_{n,m} = \left\{\left|\frac{X_n}{a_{n,m}}\right| > \frac{1}{n}\right\} = \{|X_n| > m\}$. 我们有

$$\lim_{m\to\infty} \mathbb{P}(|X_n| > m) = 0 \Longrightarrow \exists m_n, \text{s.t. } \mathbb{P}(E_{n,m_n}) < 2^{-n}.$$

注意到, $\forall\varepsilon > 0$, 若取 $k > \frac{2}{\varepsilon}$, 我们有

$$\mathbb{P}\left(\bigcup_{n=k}^{\infty}\left\{\frac{X_n}{a_{n,m}} > \varepsilon\right\}\right) \leqslant \mathbb{P}\left(\bigcup_{n=k}^{\infty} E_{n,m_n}\right) \leqslant \sum_{n=k}^{\infty} 2^{-n} = 2^{-k} \to 0.$$

即满足定理 8.1.3 中的 (4) 式, 由定理 8.1.3 知 $\frac{X_n}{A_n} \xrightarrow{\text{a.e.}} 0$. □

证法二: 首先, 教材在定义 4.1.1 中约定: 当我们只说 X 是一个 “随机变量” 时, 默认 X 只能取有限实值; 因此, 由习题 8.1.3证明过程中提出的断言知

$$\forall\delta > 0,\ \exists B_\delta \in \mathcal{F},\ \exists M(\delta) > 0,\ \text{s.t.} \mathbb{P}\left(B_\delta^c\right) < \delta \text{ 且 } \sup_{\omega\in B_\delta} |X(\omega)| \leqslant M(\delta)$$

取 $\delta=\dfrac{1}{n^2}$, 则 $\exists M_n>0$, $\exists B_n\in\mathcal{F}$, s.t.$\mathbb{P}(B_n^c)<\dfrac{1}{n^2}$ 且 $\sup\limits_{\omega\in B_n}|X_n(\omega)|\leqslant M_n$, 从而 $\{|X_n|>M_n\}\subset B_n^c$; 再取 $A_n=nM_n$, 则 $\left\{\left|\dfrac{X_n}{A_n}\right|>\dfrac{1}{n}\right\}=\{|X_n|>M_n\}\subset B_n^c\Longrightarrow\mathbb{P}\left\{\left|\dfrac{X_n}{A_n}\right|>\dfrac{1}{n}\right\}\leqslant\mathbb{P}(B_n^c)<\dfrac{1}{n^2}$, 从而

$$\sum_{n=1}^{\infty}\mathbb{P}\left\{\left|\frac{X_n}{A_n}\right|>\frac{1}{n}\right\}\leqslant\sum_{n=1}^{\infty}\frac{1}{n^2}<\infty$$

由 Borel-Cantelli 引理 (引理 8.1.4) 知 $\mathbb{P}\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}\left\{\left|\dfrac{X_k}{A_k}\right|>\dfrac{1}{k}\right\}\right)=0$. 记 $N=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}\left\{\left|\dfrac{X_k}{A_k}\right|>\dfrac{1}{k}\right\}$, 则 $\mathbb{P}(N)=0$ 且 $\forall\omega\in N^c$, $\exists n\in\mathbb{N}$, s.t.$\forall k\geqslant n$, $\left|\dfrac{X_k(\omega)}{A_k}\right|\leqslant\dfrac{1}{k}$, 即 $\forall\omega\in N^c$, $\left|\dfrac{X_k(\omega)}{A_k}\right|\to 0$ $(k\to\infty)$, 即 $\left|\dfrac{X_k}{A_k}\right|\xrightarrow{\text{a.e.}}0$. □

8.1.5 举例说明: 当 μ 不是有限测度时, Egorov 定理不成立.

证明: 考虑测度空间 $(\mathbb{R}_+,\mathscr{B}(\mathbb{R}_+),\lambda|_{\mathscr{B}(\mathbb{R}_+)})$(其中 $\mathbb{R}_+$ 表示全体正实数, $\lambda|_{\mathscr{B}(\mathbb{R}_+)}$ 表示 Lebesgue 测度限制在 $\mathscr{B}(\mathbb{R}_+)$ 上), 取 $f_n(x)=\left(x+\dfrac{1}{n}\right)^2$, $f(x)=x^2$, 则 $f_n\to f$, a.e.. 取 $\delta=1$, 则对任意 $\lambda(A_\delta^c)<1$ 的 $A_\delta\in\mathscr{B}(\mathbb{R}_+)$, 有 $\lambda(A_\delta)=\infty$, 即 A_δ 是无界的. 因此

$$\sup_{x\in A_\delta}d(f_n(x),f(x))=\sup_{x\in A_\delta}\left(\frac{1}{n}+\frac{2x}{n}\right)=\infty.$$

□

§ 8.2 依测度收敛

8.2.1 直接证明: 若 $X_n\xrightarrow{\mathbb{P}}X, Y_n\xrightarrow{\mathbb{P}}Y$, 则$X_nY_n\xrightarrow{\mathbb{P}}XY$.

证明: 对任意的 M_1,M_2 以及 $\varepsilon>0$, 我们有

$$\mathbb{P}(|X_nY_n-XY|>\varepsilon)\leqslant\mathbb{P}\left(|X||Y_n-Y|>\frac{1}{2}\varepsilon\right)+\mathbb{P}\left(|X_n-X||Y_n|>\frac{1}{2}\varepsilon\right),$$

因此只需证明

$$\lim_{n\to\infty}\mathbb{P}\left(|X||Y_n-Y|>\frac{1}{2}\varepsilon\right)=0,$$
$$\lim_{n\to\infty}\mathbb{P}\left(|X_n-X||Y_n|>\frac{1}{2}\varepsilon\right)=0.$$

我们有

$$\begin{aligned}\mathbb{P}\left(|X||Y_n-Y|\geqslant\frac{1}{2}\varepsilon\right)&=\mathbb{P}\left(|X||Y_n-Y|\geqslant\frac{1}{2}\varepsilon,|X|\leqslant M_1\right)+\mathbb{P}\left(|X||Y_n-Y|\leqslant\frac{1}{2}\varepsilon,|X|>M_1\right)\\&\leqslant\mathbb{P}\left(|Y_n-Y|\geqslant\frac{\varepsilon}{2M_1}\right)+\mathbb{P}(|X|>M_1),\end{aligned}$$

以及

$$\begin{aligned}\mathbb{P}\left(|X_n-X||Y_n|>\frac{1}{2}\varepsilon\right)&=\mathbb{P}\left(|X_n-X||Y_n|>\frac{1}{2}\varepsilon,|Y_n-Y|\geqslant 1\right)+\mathbb{P}\left(|X_n-X||Y_n|>\frac{1}{2}\varepsilon,|Y_n-Y|<1\right)\\&\leqslant\mathbb{P}(|Y_n-Y|\geqslant 1)+\mathbb{P}\left((|Y|+1)|X_n-X|\geqslant\frac{1}{2}\varepsilon,|Y_n-Y|<1\right)\\&\leqslant\mathbb{P}(|Y_n-Y|\geqslant 1)+\mathbb{P}\left((|Y|+1)|X_n-X|\geqslant\frac{1}{2}\varepsilon\right)\\&=\mathbb{P}(|Y_n-Y|\geqslant 1)+\mathbb{P}\left((|Y|+1)|X_n-X|\geqslant\frac{1}{2}\varepsilon,|Y|\leqslant M_2\right)\\&\quad+\mathbb{P}\left((|Y|+1)|X_n-X|\geqslant\frac{1}{2}\varepsilon,|Y|>M_2\right)\\&\leqslant\mathbb{P}(|Y_n-Y|\geqslant 1)+\mathbb{P}\left(|X_n-X|\geqslant\frac{\varepsilon}{2(M_2+1)}\right)+\mathbb{P}(|Y|>M_2).\end{aligned}$$

又 X,Y 几乎处处有限, 因此我们在两式中令 $n\to\infty$, $M_1\to\infty$, $M_2\to\infty$ 便得

$$\lim_{n\to\infty}\mathbb{P}\left(|X||Y_n-Y|>\frac{1}{2}\varepsilon\right)=\lim_{n\to\infty}\mathbb{P}\left(|X_n-X||Y_n|>\frac{1}{2}\varepsilon\right)=0.$$

□

8.2.2 若 $X_n\xrightarrow{\mathbb{P}}X$, f 为有界且一致连续的函数, 试证: $\mathbb{E}f(X_n)\to\mathbb{E}f(X)$.

证明: 我们知道 $f(x)$ 一致连续, 因此 $\forall\varepsilon>0$, $\exists\delta>0$, s.t. $\forall x_1,x_2$, 当 $|x_1-x_2|<\delta$ 时, $|f(x_1)-f(x_2)|<\varepsilon$. 所以

$$\{|f(X_n)-f(X)|\geqslant\varepsilon\}\subset\{|X_n-X|\geqslant\delta\}.$$

因为 $X_N\xrightarrow{\mathbb{P}}X$, 所以 $n\to\infty$ 时, $\mathbb{P}\{|X_n-X|\geqslant\delta\}\to 0$. 因此 $\lim\limits_{n\to\infty}\mathbb{P}\{|f(X_n)-f(X)|\geqslant\varepsilon\}=0$. 也即 $f(X_n)\xrightarrow{\mathbb{P}}f(X)$. 又因为 f 有界, 根据推论 8.2.6 可得 $\mathbb{E}f(X_n)\to\mathbb{E}f(X)$. □

8.2.3 $X_n\downarrow X$ a.e., 其中每一个 X_n 均可积且 $\inf\limits_n\mathbb{E}X_n>-\infty$, 试证: $\mathbb{E}X_n\to\mathbb{E}X$.

证明: 注意到 $X_1-X_n\uparrow X_1-X$, 且 X_1-X_n 可积, 因此

$$\lim_{n\to\infty}\mathbb{E}[X_1-X_n]=\mathbb{E}[X_1-X].$$

因为 X_1 可积, 故 $\mathbb{E}X_1$ 有限. 于是 $\lim\limits_{n\to\infty}\mathbb{E}X_n=\mathbb{E}X$. □

§ 8.3 L^r 收敛

8.3.1 证明 Hölder 不等式及 Minkowski 不等式.

证明: (1) 我们知道实数域上的 Hölder 不等式: 若 $p>1$, $q>1$, $\dfrac{1}{p}+\dfrac{1}{q}=1$, 则

$$\forall a,b\in\mathbb{R}_+,\ a^{\frac{1}{p}}b^{\frac{1}{q}}\leqslant\frac{a}{p}+\frac{b}{q}.$$

因此

$$\left(\frac{|X|^p}{\mathbb{E}|X|^p}\right)^{\frac{1}{p}}\left(\frac{|Y|^q}{\mathbb{E}|Y|^q}\right)^{\frac{1}{q}} \leqslant \frac{1}{p}\frac{|X|^p}{\mathbb{E}|X|^p}+\frac{1}{q}\frac{|Y|^q}{\mathbb{E}|Y|^q},$$

两边取期望可得

$$\frac{\mathbb{E}|XY|}{\left(\mathbb{E}|X|^p\right)^{\frac{1}{p}}\left(\mathbb{E}|Y|^q\right)^{\frac{1}{q}}} \leqslant 1,$$

也即

$$\mathbb{E}|XY| \leqslant \left(\mathbb{E}|X|^p\right)^{\frac{1}{p}}\left(\mathbb{E}|Y|^q\right)^{\frac{1}{q}}.$$

(2) 考虑 $r \geqslant 1$, 我们有

$$\begin{aligned}\mathbb{E}|X+Y|^r &\leqslant \mathbb{E}|X+Y|^{r-1}|X|+\mathbb{E}|X+Y|^{r-1}|Y| \\ &\leqslant \left(\mathbb{E}|X+Y|^r\right)^{\frac{r-1}{r}}\left(\mathbb{E}|X|^r\right)^{\frac{1}{r}}+\left(\mathbb{E}|X+Y|^r\right)^{\frac{r-1}{r}}\left(\mathbb{E}|Y|^r\right)^{\frac{1}{r}} \\ &\leqslant \left(\mathbb{E}|X+Y|^r\right)^{\frac{r-1}{r}}\left(\left(\mathbb{E}|X|^r\right)^{\frac{1}{r}}+\left(\mathbb{E}|Y|^r\right)^{\frac{1}{r}}\right).\end{aligned}$$

因此

$$\left(\mathbb{E}|X+Y|^r\right)^{\frac{1}{r}} \leqslant \left(\mathbb{E}|X|^r\right)^{\frac{1}{r}}+\left(\mathbb{E}|Y|^r\right)^{\frac{1}{r}}.$$

□

8.3.2 试完成定理 8.3.18 中的 (2) ⇒ (1).

证明: 首先，从书上定理 8.3.18 已有的证明过程中可提取出如下两个引理及其证明:

引理 1: $\forall x, y \in \mathbb{R}$或$\mathbb{C}$, $\forall r>0$, $|x-y|^r \leqslant 2^r\left(|x|^r+|y|^r\right)$.

证明: $|x-y|^r \leqslant (|x|+|y|)^r \leqslant (2(|x| \vee |y|))^r=2^r\left(|x|^r \vee |y|^r\right) \leqslant 2^r\left(|x|^r+|y|^r\right)$.

引理 2: 若 $X_n \xrightarrow{r} X$ (其中 $\forall n \in \mathbb{N}$, $X_n \in L^r(\mathbb{P})$), 则 $X \in L^r(\mathbb{P})$.(此引理表明 L^r 空间是闭的)

证明: 由引理 1 知 $|X|^r \leqslant 2^r\left(|X_n|^r+|X_n-X|^r\right)$, 从而 $\mathbb{E}|X|^r \leqslant 2^r\left(\mathbb{E}|X_n|^r+\mathbb{E}|X_n-X|^r\right)$;

又由 $X_n \xrightarrow{r} X$ 知 $\lim\limits_{n \rightarrow \infty} \mathbb{E}|X_n-X|^r=0$, 即 $\forall \varepsilon>0$, $\exists N \in \mathbb{N}$, s.t.$\forall n \geqslant N$, $\mathbb{E}|X_n-X|^r<\varepsilon$, 从而

$$\mathbb{E}|X|^r \leqslant 2^r\left(\mathbb{E}|X_N|^r+\mathbb{E}|X_N-X|^r\right)<2^r\left(\mathbb{E}|X_N|^r+\varepsilon\right) \overset{X_N \in L^r(\mathbb{P})}{<} \infty.$$

以下分两步完成证明: 1°: 证明 $|X_n-X|^r$ 一致可积; 2° 由 1° 证明 $|X_n|^r$ 一致可积.

1°: 由引理 1 知 $|X_n-X|^r \leqslant 2^r\left(|X_n|^r+|X|^r\right)$, 从而

$$\forall n \in \mathbb{N},\ \mathbb{E}|X_n-X|^r \leqslant 2^r\left(\mathbb{E}|X_n|^r+\mathbb{E}|X|^r\right) \overset{\text{题设 + 引理 2: } \mathbb{E}|X_n|^r, \mathbb{E}|X|^r<\infty}{<} \infty \tag{$*$}$$

又由 $X_n \xrightarrow{r} X$ 知 $\lim\limits_{n \rightarrow \infty} \mathbb{E}|X_n-X|^r=0$, 即

$$\forall \varepsilon>0,\ \exists N \in \mathbb{N},\ \text{s.t.} \forall n \geqslant N,\ \mathbb{E}|X_n-X|^r<\varepsilon \tag{$**$}$$

从而

$$\sup_n \mathbb{E}|X_n-X|^r \leqslant \max\left\{\mathbb{E}|X_1-X|^r, \cdots, \mathbb{E}|X_N-X|^r, \varepsilon\right\}<\infty \tag{①}$$

再由 $(*)$: $\forall n \in \mathbb{N}$, $|X_n-X|^r$ 可积, 以及推论 5.4.6(积分的绝对连续性) 知

$$\forall n \in \mathbb{N},\ \forall \varepsilon>0,\ \exists \delta_n>0,\ \text{s.t.} \forall A \in \mathcal{F}: \mathbb{P}(A)<\delta_n,\ \int_A |X_n-X|^r \, \mathrm{d}\mathbb{P}<\varepsilon$$

取 $\delta = \min\{\delta_1, \cdots, \delta_N\}$, 并注意到 $(**)$ 式, 则当 $A \in \mathcal{F} : \mathbb{P}(A) < \delta$ 时, 有

$$\int_A |X_n - X|^r \, d\mathbb{P} \leqslant \max\left\{\int_A |X_1 - X|^r \, d\mathbb{P}, \cdots, \int_A |X_N - X|^r \, d\mathbb{P}, \varepsilon\right\} \leqslant \varepsilon$$

也就是说

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0,\ \text{s.t.} \forall A \in \mathcal{F} : \mathbb{P}(A) < \delta,\ \int_A |X_n - X|^r \, d\mathbb{P} < \varepsilon \quad (\text{即一致绝对连续条件}) \qquad ②$$

由①、②及引理 8.3.16 知 $\{|X_n - X|^r :\ n \in \mathbb{N}\}$ 一致可积;

2°: 由引理 1 知 $|X_n|^r \leqslant 2^r (|X_n - X|^r + |X|^r) \Rightarrow \mathbb{E}|X_n|^r \leqslant 2^r (\mathbb{E}|X_n - X|^r + \mathbb{E}|X|^r) (***)$;
而由引理 8.3.16 知 $|X_n - X|^r$ 一致可积等价于 1° 中①、②同时成立.
由①知 $\sup_n \mathbb{E}|X_n - X|^r < \varepsilon$, 又由引理 2 知 $\mathbb{E}|X|^r < \infty$, 从而由 $(***)$ 知 $\sup_n \mathbb{E}|X_n|^r < \infty$;
由②知 $\lim\limits_{\mathbb{P}(A)\to 0} \sup\limits_n \int_A |X_n - X|^r \, d\mathbb{P} = 0$; 又由 $\mathbb{E}|X|^r < \infty$ 及推论 5.4.6(积分的绝对连续性) 知 $\lim\limits_{\mathbb{P}(A)\to 0} \int_A |X|^r \, d\mathbb{P} = 0$, 从而由 $(***)$ 知 $\lim\limits_{\mathbb{P}(A)\to 0} \sup\limits_n \int_A |X_n|^r \, d\mathbb{P} = 0$, 即 $\{|X_n|^r :\ n \in \mathbb{N}\}$ 一致绝对连续;
由引理 8.3.16 知 $\{|X_n|^r :\ n \in \mathbb{N}\}$ 一致可积. □

注: (1) 本题目标是证明一致可积, 思路是利用引理 8.3.16: 一致可积等价于一致有界且一致绝对连续, 转而证明这两个条件成立;

(2) 事实上, 当 $X_n \xrightarrow{r} X$ 且 $\forall n \in \mathbb{N},\ X_n \in L^r(\mathbb{P})$ 时, $|X_n - X|^r$ 一致可积 $\Longleftrightarrow |X_n|^r$ 一致可积: 只需再注意到不等式 $\mathbb{E}|X_n - X|^r \leqslant 2^r (\mathbb{E}|X_n|^r + \mathbb{E}|X|^r)$ 成立, 同理即可证明;

(3) 上述证明中引理 1 所给不等式的意义: 提供了 r 阶矩的一个"三角不等式". 也可以这样去构造用于放缩 r 阶矩的"三角不等式": 记 $f(x) = x^r, x \in (0, \infty)$, 则 $r \geqslant 1$ 时 f 是凸函数, 由 $f\left(\dfrac{|x| + |y|}{2}\right) \leqslant \dfrac{1}{2} f(|x|) + \dfrac{1}{2} f(|y|)$ 整理得 $|x - y|^r \leqslant (|x| + |y|)^r \leqslant 2^{r-1}(|x|^r + |y|^r)$; $0 < r < 1$ 时利用 f 增长速度越来越慢的性质, 容易证明 $f(|x| + |y|) - f(|y|) \leqslant f(|x|)$, 即 $|x - y|^r \leqslant (|x| + |y|)^r \leqslant |x|^r + |y|^r$. 也就是说, 我们有比引理 1 放缩得更紧的不等式:

$$\forall x, y \in \mathbb{R}\text{或}\mathbb{C},\ \forall r > 0,\ \ |x - y|^r \leqslant \left(2^{r-1} \vee 1\right)(|x|^r + |y|^r).$$

8.3.3 设 $r \in (0, \infty)$, 则 $X_n \xrightarrow{r} X$ 当且仅当 $X_n \xrightarrow{\mathbb{P}} X$ 及下列两条件之一成立:

(1) $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r < \infty$;

(2) $|X_n|^r$ 一致可积.

证明: "$\Longrightarrow$": 若 $X_n \xrightarrow{r} X$, 则由定理 8.3.2 知 $X_n \xrightarrow{\mathbb{P}} X$, 进而由定理 8.3.18 知条件 (1)、(2) 均成立. (细节: $X_n \xrightarrow{r} X$ 蕴含着 $\forall n \in \mathbb{N},\ X_n \in L^r(\mathbb{P})$, 见定义 8.3.1)

"$\Longleftarrow$": 利用定理 8.3.18 容易证明. 为满足定理 8.3.18 成立的前提, 只需验证 $\forall n \in \mathbb{N},\ X_n \in L^r(\mathbb{P})$.

若条件 (1) 成立, 则可知 $\{X_n\}$ 除去前面有限项之后都属于 $L^r(\mathbb{P})$ (将条件 (1) 用 $\varepsilon - N$ 语言翻译即可得知: $\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \text{s.t.}\ \forall n \geqslant N, |\mathbb{E}|X_n|^r - \mathbb{E}|X|^r| < \varepsilon$, 取 $\varepsilon = 1$, 即得 $\exists N \in \mathbb{N},\ \text{s.t.}\ \forall n \geqslant N,\ \mathbb{E}|X_n|^r < \mathbb{E}|X|^r + \varepsilon < \infty$, 即

从某个有限的 N 开始都有 $\mathbb{E}|X_n|^r<\infty$), 从而由定理 8.3.18 知 $X_n\xrightarrow{r}X$ (也可以照抄定理 8.3.18 中证明 (3)$\Longrightarrow$(2) 的步骤直接证明);

若条件 (2) 成立, 则由引理 8.3.16 的条件 (1) 知 $\forall n\in\mathbb{N}$, $\mathbb{E}|X_n|^r<\infty$, 即 $\forall n\in\mathbb{N}, X_n\in L^r(\mathbb{P})$, 从而由定理 8.3.18 知 $X_n\xrightarrow{r}X$. □

注: 本题与定理 8.3.18 几乎相同, 可视作定理 8.3.18 的一个推论.

8.3.4 若 $\sup\limits_n|X_n|\in L^r(\mathbb{P})$ 且 $X_n\xrightarrow{\text{a.e.}}X$, 则 $X\in L^r(\mathbb{P})$ 且 $X_n\xrightarrow{r}X$.

证明: 由题设及 Fatou 引理 (定理 5.4.2)(或者引理 8.3.14), 我们有

$$\mathbb{E}|X|^r\xlongequal{X_n\xrightarrow{\text{a.e.}}X}\mathbb{E}\left(\lim_{n\to\infty}|X_n|^r\right)\overset{\text{Fatou 引理}}{\leqslant}\liminf_{n\to\infty}\mathbb{E}|X_n|^r\leqslant\liminf_{n\to\infty}\mathbb{E}\left(\sup_n|X_n|\right)^r\overset{\sup_n|X_n|\in L^r(\mathbb{P})}{<}\infty$$

从而 $X\in L^r(\mathbb{P})$;

由 $\sup\limits_n|X_n|\in L^r(\mathbb{P})$ 知 $\left(\sup\limits_n|X_n|\right)^r$ 可积, 又 $|X_n|^r\leqslant\left(\sup\limits_n|X_n|\right)^r$ 且 $|X_n|^r\xrightarrow{\text{a.e.}}|X|^r$, 由控制收敛定理 (定理 5.4.3(2)) 知 $\mathbb{E}|X_n|^r\longrightarrow\mathbb{E}|X|^r<\infty$; 又由 $\sup\limits_n|X_n|\in L^r(\mathbb{P})$ 显然有 $\forall n\in\mathbb{N}$, $X_n\in L^r(\mathbb{P})$, 从而由定理 8.3.18 知 $X_n\xrightarrow{r}X$ (或者也可由定理 8.3.2 知 $X_n\xrightarrow{\mathbb{P}}X$, 从而由习题 8.3.3知 $X_n\xrightarrow{r}X$). □

8.3.5 证明引理 8.3.17 的 (II).

证明: 我们知道存在子列 $\{X_{n_k}\}_{k\geqslant1}$, 使得

$$\limsup_{n\to\infty}\int|X_n-X|\mathrm{d}\mu=\lim_{k\to\infty}\int|X_{n_k}-X|\mathrm{d}\mu.$$

又 $U_{n_k}\xrightarrow{\mu}U$, $X_{n_k}\xrightarrow{\mu}X$, 所以存在子列 $\{X_{n_{k_\ell}}\}_{\ell\geqslant1}$, 使得 $X_{n_{k_\ell}}\xrightarrow{\text{a.e.}}X$, 且

$$\sum_{\ell=1}^{\infty}\mu\left(|X_{n_{k_\ell}}-X|\geqslant\frac{1}{2^\ell}\right)<\infty.$$

同理 $U_{n_{k_\ell}}\xrightarrow{\text{a.e.}}U$. 根据习题 5.4.2知道, $\int|X_{n_{k_\ell}}-X|\to0$, 当 $\ell\to\infty$, 因此

$$0\leqslant\limsup_{n\to\infty}\int|X_n-X|\mathrm{d}\mu=\lim_{k\to\infty}\int|X_{n_k}-X|\mathrm{d}\mu=\lim_{\ell\to\infty}\int|X_{n_{k_\ell}}-X|=0.$$

故 $\lim\limits_{n\to\infty}|X_n-X|=0$. □

§ 8.5 概率测度的收敛

8.5.1 证明淡收敛的极限是唯一的.

证明: 设 $F_n\xrightarrow{v}F$, 且 $F_n\xrightarrow{v}G$. 则由淡收敛的定义知道

$$F_n(x)\to F(x),\quad\forall x\in C(F);$$
$$F_n(x)\to G(x),\quad\forall x\in C(G).$$

于是由数列极限的唯一性有 $F(x)=G(x), \forall x\in C(F)\cap C(G)$. 我们知道单调函数的不连续点至多可数, 因此 $(C(F)\cap C(G))^c$ 至多可数. 于是 $\forall x\in (C(F)\cap C(G))^c$, 存在 $\{x_n\}_{n\geqslant 1}\subset C(F)\cap C(G)$, 使得 $x_n\downarrow x$, 我们知道分布函数是右连续的, 因此

$$F(x)=\lim_{n\to\infty}F(x_n)=\lim_{n\to\infty}G(x_n)=G(x).$$

因此淡收敛的极限是唯一的. □

8.5.2 证明定义 8.5.1 所定义的 $(\mathbb{R},\mathscr{B})$ 上的概率测度的弱收敛与定义 8.5.5 一致.

证明: 定义 8.5.1 $\Longrightarrow$ 定义 8.5.5: 由定义对任意的 $f\in C_b$, 有 $\int f\,\mathrm{d}\mu_n\to\int f\,\mathrm{d}\mu$. 于是 $\forall f\in C_0$, 有 $\int f\,\mathrm{d}\mu_n\to\int f\,\mathrm{d}\mu$. 因此 μ_n 淡收敛于 μ. 特别地, 取 $f=1$, 则由定义 8.5.5 知道 $\mu_n(\mathbb{R})\to\mu(\mathbb{R})$.

定义 8.5.5 $\Longrightarrow$ 定义 8.5.1: 设 F_n 是 μ_n 的分布函数, F 是 μ 的分布函数, 因为 F 的不连续点至多可数, 所以 $\forall\varepsilon>0$, $\exists L>0$, 使得 $F(-L)<\varepsilon$, $1-F(L)<\varepsilon$. 对于上述的 $\varepsilon>0$, 因为 $F_n(L)\to F(L)$, $F_n(-L)\to F(-L)$, 所以存在 $N>0$, 使得

$$F_n(-L)<2\varepsilon,\quad 1-F(L)<2\varepsilon.$$

任取 $f\in C_b$, 设 $|f|\leqslant M$, 有

$$\begin{aligned}\left|\int f\,\mathrm{d}\mu_n-\int f\,\mathrm{d}\mu\right| &\leqslant \left|\int_{(-\infty,-L]\cap\{L_1\}} f\,\mathrm{d}F_n\right|+\left|\int_{(-\infty,-L]\cap\{L_1\}} f\,\mathrm{d}F\right|+\left|\int_{(-L,L]} f\,\mathrm{d}F_n-\int_{(-L,L]} f\,\mathrm{d}F\right|\\ &\leqslant 4M+\left|\int_{(-L,L]} f\,\mathrm{d}F_n-\int_{(-L,L]} f\,\mathrm{d}F\right|.\end{aligned}$$

根据定理 8.5.4(1) 可得当 $n\to\infty$, $\left|\int_{(-L,L]} f\,\mathrm{d}F_n-\int_{(-L,L]} f\,\mathrm{d}F\right|\to 0$, 由 ε 的任意性可得 $\int f\mathrm{d}\mu_n\to\int f\mathrm{d}\mu$. □

8.5.3 证明定理 8.5.14.

证明: $\Longrightarrow$: $\mu_n\xrightarrow{w}\mu$, 则 $\{\mu_n\}$ 相对紧. 在完备可分空间中, 由定理 8.5.12, $\{\mu_n\}$ 胎紧, 因此 $\mu_n\xrightarrow{v}\mu$.

$\Longleftarrow$: 由定理 8.5.12, $\{\mu_n\}$ 胎紧则相对紧, 故 $\{\mu_n\}$ 任何子列都存在弱收敛的子列, 我们知道 $\{\mu_n\}$ 是淡收敛的, 同时 $\{\mu_n(\mathbb{R}^k)\}$ 为常数序列, 故 $\mu_n\xrightarrow{v}\mu$. 再根据推论 8.5.13, $\{\mu_n\}$ 的任意子列都有收敛到同一概率测度的子列, 故 $\{\mu_n\}$ 弱收敛. □

8.5.4 证明: $\mathbb{R}$ 上概率分布函数族 $\{F_\alpha\}$ 所对应的概率测度族相对紧的充分必要条件是当 $x\to-\infty$ 和 $x\to\infty$ 时, $\{F_\alpha\}$ 对 α 一致收敛.

证明: 设 $\{F_\alpha\}$ 对应的概率测度序列是 $\{\mu_\alpha\}$, 因为 $\mathbb{R}$ 完备可分, 所以根据定理 8.5.12, 我们只需证明 $\{\mu_\alpha\}$ 胎紧等价于当 $x\to-\infty$ 和 $x\to\infty$ 时, $\{F_\alpha\}$ 对 α 一致收敛.

必要性: 因为 $\{\mu_\alpha\}$ 胎紧, 所以 $\forall\varepsilon>0$, 存在紧集 K, s.t. $\forall\alpha$, $\mu_\alpha(K^c)<\varepsilon$. 因为紧集 $K\subset\mathbb{R}$, 故有界, 因此 $\exists M>0$, s.t. $K\subset[-M,M]$. 故 $\mu_\alpha([-M,M]^c)<\varepsilon$. 因此 $\forall x>M$, $\forall\alpha$, $F_\alpha(-x)<\varepsilon$, $1-F_\alpha(x)<\varepsilon$. 因此当 $x\to-\infty$ 和 $x\to\infty$ 时, $\{F_\alpha\}$ 对 α 一致收敛.

充分性: 考虑 $\{F_\alpha\}$ 对 α 一致收敛, 且 $\lim\limits_{x\to\infty}F_\alpha(x)=1$, $\lim\limits_{x\to-\infty}F_\alpha(x)=0$, 所以 $\forall\varepsilon>0$, $\exists M>0$, $\forall\alpha$, 当 $x>M$ 有 $F_\alpha(-x)<\varepsilon$, $1-F_\alpha(x)<\varepsilon$. 取 $K=[-M,M]$ 是紧集, 则 $\forall\alpha$, $\mu_\alpha(K^c)<\varepsilon$. 因此 $\{F_\alpha\}$ 胎紧.

□

8.5.5 设与随机变量族 $\{X_\alpha\}$ 相对应的概率分布族是 $\{\mu_\alpha\}$. 如果对某个实数 $r > 0$, $\{\mathbb{E}\,|X_\alpha|^r\}$ 对 α 有界, 则 $\{\mu_\alpha\}$ 相对紧.

证明: 根据定理 8.5.12, 我们只需证明 $\{\mu_\alpha\}$ 是胎紧的. 我们知道 $\exists M > 0$, s.t., $\sup\limits_\alpha \mathbb{E}|X_\alpha|^r < M$, 因此 $\mu_\alpha(\{|X| > A\}) < \dfrac{\mathbb{E}|X|^r}{A^r}$, 又 $\{|X| > A\}$ 是紧集的余集, 故 $\{\mu_\alpha\}$ 胎紧. □

第九章 大数定律、随机级数

§ 9.1 简单的极限定理及其应用

9.1.1 设 $(\Omega,\mathcal{F},\mathbb{P}_\theta),\theta\in\Theta$ 为一族概率空间, Θ 为一有限或无限区间, $X_n,n\in\mathbb{N}$ 为一列取有限实数值的随机变量, 且

$$\mathbb{E}_\theta X_n=\theta,\quad \sigma_n^2(\theta):=\mathbb{E}_\theta(X_n-\theta)^2=\sigma_\theta^2(X_n).$$

设 $u\in C_b(\Theta)$(Θ 上的一切有界连续函数组成的集合), 且 $\forall\theta\in\Theta,\sigma_n^2(\theta)\to0\,(n\to\infty)$, 试证:

$$\mathbb{E}_\theta u(X_n)\to u(\theta)\,(n\to\infty),\quad \theta\in\Theta$$

而且在 $\sigma_n^2(\theta)\to0\,(n\to\infty)$ 一致成立的每个闭区域 Θ_0 上, 上述收敛是一致的.
进一步, 给出上述结论在下列各种特定情形下的具体结论:

(1) $\mathbb{P}_\theta\left(X_n=\dfrac{k}{n}\right)=C_n^k\theta^k(1-\theta)^{n-k},k=0,1,2,\cdots,n;$

(2) $\mathbb{P}_\theta\left(X_n=\dfrac{k}{n}\right)=\mathrm{e}^{-n\theta}\dfrac{(n\theta)^k}{k!},k=0,1,2,\cdots,n,\cdots;$

(3) X_n 在 $\mathbb{P}_\theta$ 下服从参数为 n 和 $\dfrac{n}{\theta}$ 的 Γ 分布, 即

$$\mathbb{P}_\theta(X_n\leqslant x)=\begin{cases}0, & x\leqslant0,\\ \dfrac{1}{(n-1)!}\displaystyle\int_0^x\left(\frac{n}{\theta}\right)\left(\frac{nt}{\theta}\right)^{n-1}\mathrm{e}^{-\frac{nt}{\theta}}\,\mathrm{d}t, & x>0.\end{cases}$$

注: 对于 (2)(3) 两种情形, n 为正实数时亦有相应结论, 其中在 (3) 把 $(n-1)!$ 换成 $\Gamma(n)$ 即可.

证明: 我们知道 $X_n\xrightarrow{L^2(\mathbb{P}_\theta)}\theta$, 因此由定理 8.3.2 知 $X_n\xrightarrow{\mathbb{P}_\theta}\theta$, 并且由 u 是连续的知 $u(X_n)\xrightarrow{\mathbb{P}_\theta}u(\theta)$. 又由于 u 有界, 故由推论 8.2.6(依概率收敛下的控制收敛定理)$\mathbb{E}_\theta u(X_n)\to u(\theta)$ $(n\to\infty)$, $\theta\in\Theta$.

再证一致收敛性, 由 $|\mathbb{E}_\theta u(X_n)-u(\theta)|\leqslant\mathbb{E}_\theta|u(X_n)-u(\theta)|$ 知只需证明 $\sup\limits_{\theta\in\Theta_0}\mathbb{E}_\theta|u(X_n)-u(\theta)|\to0$ $(n\to\infty)$. 首先由 Chebyshev 不等式可知 $\sup\limits_{\theta\in\Theta_0}\mathbb{P}_\theta\{|X_n-\theta|\geqslant\varepsilon\}\leqslant\varepsilon^{-2}\sup\limits_{\theta\in\Theta_0}\sigma_n^2(\theta)\to0$ $(n\to\infty)$; 由此,

下面验证 (1)(2)(3) 均满足题设:

(1) $\mathbb{E}_\theta X_n=\theta$, $\sigma_\theta^2(X_n)=\dfrac{\theta(1-\theta)}{n}\to0.$

(2) $\mathbb{E}_\theta X_n=\theta$, $\sigma_\theta^2(X_n)=\dfrac{\theta}{n}\to0.$

(3) $\mathbb{E}_\theta X_n = \theta$, $\sigma_\theta^2(X_n) = \dfrac{\theta^2}{n} \to 0$.
故对于 (1)(2)(3) 而言，均有 $\forall u \in C_b(\Theta)$，$\mathbb{E}_\theta u(X_n) \to u(\theta)$ $(n \to \infty)$，$\theta \in \Theta$, 而且在 $\sigma_n^2(\theta) \to 0\,(n \to \infty)$ 一致成立的每个闭区域 Θ_0 上, 这种收敛是一致收敛. $\square$

9.1.2 设 $f \in C[0,1]$, $0 \leqslant f \leqslant 1$, $\xi_n \overset{\text{i.i.d.}}{\sim} U[0,1]$ $(n \in \mathbb{N})$, $\forall \varepsilon > 0$, 令 $\eta_\varepsilon(n) := \mathbb{1}_{\{\xi_n \leqslant f(\varepsilon n)\}}, n \in \left[0, \dfrac{1}{\varepsilon}\right]$, 试证:

$$\varepsilon \sum_{n:0 \leqslant \varepsilon n \leqslant 1} \eta_\varepsilon(n) \xrightarrow{\mathbb{P}} \int_0^1 f(r)\mathrm{d}r.$$

证明: 实际上 $\eta_\varepsilon(n)$ 仍是独立的 r.v. 列, 注意到 $\sup\limits_n \sigma^2(\eta_\varepsilon(n)) < \infty$, 故由定理 9.1.3 知

$$\frac{\sum_{0 \leqslant n \leqslant \varepsilon^{-1}} \eta_\varepsilon(n) - \mathbb{E}[\sum_{0 \leqslant n \leqslant \varepsilon^{-1}} \eta_\varepsilon(n)]}{n} \xrightarrow{\mathbb{P}} 0.$$

又注意到 (由黎曼积分的定义)

$$\mathbb{E}\left[\sum_{n:0 \leqslant n \leqslant \varepsilon^{-1}} \eta_\varepsilon(n)\right] = \sum_{n:0 \leqslant n \leqslant \varepsilon^{-1}} \mathbb{P}(\xi_n \leqslant f(\varepsilon n)) = \varepsilon^{-1} \sum_{n:0 \leqslant \varepsilon n \leqslant 1} \varepsilon \cdot f(\varepsilon n) \to \varepsilon^{-1} \int_0^1 f(r)\mathrm{d}r \ (n \to \infty),$$

因此

$$\varepsilon \sum_{0 \leqslant \varepsilon n \leqslant 1} \eta_\varepsilon(n) \xrightarrow{\mathbb{P}} \int_0^1 f(r)\mathrm{d}r.$$

$\square$

9.1.3 f 的假设与习题 9.1.2相同, $\{X_n : n \in \mathbb{N}\}$ 为一列相互独立的随机变量且 $X_n \sim U[0,1]^2$ $(\forall n \in \mathbb{N})$, 设 $X_n = (X_{n1}, X_{n2})$, 试证:

$$\frac{1}{n}\sum_{k=1}^n \mathbb{1}_{\{X_{n2} \leqslant f(X_{n2})\}} \xrightarrow[\mathbb{P}]{\text{a.s.}} \int_0^1 f(r)\,\mathrm{d}r.$$

证明: 我们知道 X_{n1}, X_{n2} 独立且都服从 $(0,1)$ 上的均匀分布, 因此 $\mathbb{1}_{\{X_{n2} \leqslant f(X_{n1})\}}$ 是独立序列, 且

$$\sup_n \sigma^2(\mathbb{1}_{\{X_{n2} \leqslant f(X_{n1})\}}) < \infty, \ \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^n \mathbb{1}_{\{X_{n2} \leqslant f(X_{n1})\}}\right) = \frac{1}{n}\sum_{k=1}^n f(X_{n2}).$$

又 $\sup \sigma^2(f(X_{n2})) \leqslant 1 < \infty$, 因此

$$\frac{1}{n}\sum_{k=1}^n f(X_{n2}) \xrightarrow[\mathbb{P}]{\text{a.s.}} \int_0^1 f(r)\mathrm{d}r.$$

故

$$\frac{1}{n}\sum_{k=1}^n \mathbb{1}_{\{X_{n2} \leqslant f(X_{n2})\}} \xrightarrow[\mathbb{P}]{\text{a.s.}} \int_0^1 f(r)\,\mathrm{d}r.$$

$\square$

9.1.4 $f, \xi_n, \eta_\varepsilon(n)$ 的假设与习题 9.1.2相同, 设 $\varphi \in C[0,1]$, 试证:

$$\varepsilon \sum_{0 \leqslant \varepsilon n \leqslant 1} \varphi(\varepsilon n)\eta_\varepsilon(n) \xrightarrow{\mathbb{P}} \int_0^1 \varphi(r) f(r)\mathrm{d}r.$$

证明: 与习题 9.1.2类似, 实际上 $\varphi(\varepsilon n)\eta_\varepsilon(n)$ 仍然是独立的随机变量列. 且

$$\sup\sigma(\varphi(\varepsilon n)\eta_\varepsilon(n)) = [\varphi(\varepsilon n)]^2\sigma(\eta_\varepsilon(n)) = 0,$$

以及

$$\mathbb{E}\left[\sum_{0\leqslant\varepsilon n\leqslant 1}\varphi(\varepsilon n)\eta_\varepsilon(n)\right] = \sum_{0\leqslant\varepsilon n\leqslant 1}\varphi(\varepsilon n)\mathbb{E}[\eta_\varepsilon(n)] = \sum_{0\leqslant\varepsilon n\leqslant 1}\varphi(\varepsilon n)\mathbb{P}(\xi_n\leqslant f(\varepsilon n)) = \sum_{0\leqslant\varepsilon n\leqslant 1}\varphi(\varepsilon n)f(\varepsilon n).$$

又 $\varepsilon\sum\limits_{0\leqslant\varepsilon n\leqslant 1}\varphi(\varepsilon n)f(\varepsilon n)\to\int_0^1\varphi(r)f(r)\mathrm{d}r$, 故根据定理 9.1.5 即证. □

9.1.5 将定理 9.1.8 和定理 9.1.9 推广到 $f\in C\,[0,1]^d$ 的情形.

证明: 定理 9.1.8 的推广: 设 $f\in C([0,1]^d)$, $x\in[0,1]^d$, $n\in\mathbb{N}$. 令

$$B_n^d(x) = \sum_{\substack{1\leqslant n_k\leqslant n\\ k=1,2,\cdots,d}}\prod_{k=1}^d C_n^{n_k}x_k^{n_k}(1-x_k)^{n-n_k}f\left(\frac{n_1}{n},\frac{n_2}{n},\cdots,\frac{n_d}{n}\right),$$

则 $\sup\limits_{x\in[0,1]^d}|B_n^d(x)-f(x)|\to 0$.

我们知道 $\forall x=(x_1,\cdots,x_d)\in[0,1]^d$, 都有 d 维 i.i.d. 序列 $\{X_n\}=\{(X_{n1},X_{n2},\cdots,X_{nd})\}$, s.t. $\mathbb{P}(X_{nk}=1)=1-\mathbb{P}(X_{nk}=0)=x_k$, $k=1,2,\cdots,d$. 因此

$$\mathbb{E}X_n=(x_1,x_2,\cdots,x_d),\sigma^2(X_n)=\mathrm{diag}\{x_1(1-x_1),x_2(1-x_2),\cdots,x_d(1-x_d)\}.$$

令 $S_n=\left(\sum\limits_{i=1}^n X_{i1},\sum\limits_{i=1}^n X_{i2},\cdots,\sum\limits_{i=1}^n X_{id}\right)$, 则

$$\mathbb{P}(S_n=(n_1,n_2,\cdots,n_d))=\prod_{k=1}^d C_n^{n_k}x_k^{n_k}(1-x_k)^{n-n_k}.$$

因此 $\mathbb{E}\left[f\left(\dfrac{S_n}{n}\right)\right]=\sum\limits_{\substack{1\leqslant n_k\leqslant n\\ k=1,2,\cdots,d}}\prod\limits_{k=1}^d C_n^{n_k}x_k^{n_k}(1-x_k)^{n-n_k}f\left(\dfrac{n_1}{n},\dfrac{n_2}{n},\cdots,\dfrac{n_d}{n}\right)=B_n^d(x)$. 此后的证明与教材定理 9.1.8 类似, 不再赘述.

定理 9.1.9: 教材中此定理并没有出现函数 f, 因此不再讨论. □

§ 9.2 弱大数定律

9.2.1 设 $\{X_n:n\in\mathbb{N}\}$ 为一列 i.i.d. 的随机变量, 且 X_1 服从 Cauchy 分布, 即

$$\mathbb{P}\left(X_1\leqslant x\right)=\int_{-\infty}^x\frac{\mathrm{d}t}{\pi\left(1+t^2\right)},\quad x\in\mathbb{R}$$

试证: $\lim\limits_{x\to\infty}x\mathbb{P}\left(|X_1|>x\right)=\dfrac{2}{\pi}\neq 0$, 因而由定理 9.2.4 知不存在实数列 $\{a_n:n\in\mathbb{N}\}$ 使得

$$\frac{S_n}{n}-a_n\xrightarrow{\mathbb{P}}0.$$

证明: 我们有

$$x\mathbb{P}(|X_1|>x)=2x\int_x^\infty \frac{\mathrm{d}t}{\pi(1+t^2)}=2x\left(\frac{\pi-2\arctan x}{2\pi}\right)\xrightarrow{x\to\infty}\frac{2}{\pi}\neq 0,$$

故不存在 a_n, s.t. $\dfrac{S_n}{n}-a_n\xrightarrow{\mathbb{P}}0$. □

9.2.2 设 $\{X_n:n\in\mathbb{N}\}$ 为一列 i.i.d. 的随机变量, 且

$$\mathbb{P}(X_n=(-1)^{k-1}k)=\frac{c}{k^2\ln k},\quad k\geqslant 3,$$

其中 c 满足 $\displaystyle\sum_{k=3}^\infty\frac{c}{k^2\ln k}=1$, 试证: $\mathbb{E}|X_1|=\infty$, 但有一常数 a 使得 $\dfrac{S_n}{n}\xrightarrow{\mathbb{P}}a$.

证明: 我们有

$$\mathbb{E}|X_1|=\sum_{k=3}^\infty\frac{c}{k\ln k}>\sum_{k=3}^\infty\frac{3}{k}=\infty.$$

同时

$$k\mathbb{P}(|X_1|>k)=k\sum_{n=k+1}^\infty\frac{c}{n^2\ln n}\leqslant k\sum_{n=k}^\infty\int_k^{k+1}\frac{c}{x^2\ln x}\mathrm{d}x\leqslant\frac{c\ln\ln k}{k}\to 0,$$

故根据定理 9.2.4, 存在 a_n, $\dfrac{S_n}{n}\xrightarrow{\mathbb{P}}a_n$. □

9.2.3 令 $p_k=\dfrac{1}{2^k k(k+1)}, l\in\mathbb{N}, p_0=1-\displaystyle\sum_{k=1}^\infty p_k$, 设 $\{X_n:n\in\mathbb{N}\}$ 为一列 i.i.d. 的随机变量, 满足

$$\mathbb{P}(X_n=-1)=p_0,\quad \mathbb{P}(X_n=2^k-1)=p_k,\quad k\in\mathbb{N}$$

则 $\mathbb{E}X_n=0$. 进一步, 设 $S_n=\displaystyle\sum_{k=1}^n X_k$, 试应用定理 9.2.5 证明

$$\frac{S_n}{\frac{n}{\log_2 n}}\xrightarrow{\mathbb{P}}-1,$$

其中 $\log_2 n$ 表示 n 的以 2 为底的对数.

证明: 教材中题目有误, p_k 不应为 $\dfrac{1}{2^k}k(k+1)$, 否则 $\mathbb{E}X_n\neq 0$. 错误已经在本文档中更正.

我们有

$$\mathbb{E}X_n=-p_0+\sum_{k=1}^\infty\frac{2^k-1}{2^k k(k+1)}=0,$$

同时令

$$m(n):=\inf\{m:2^{-m}m^{-3/2}\leqslant n^{-1}\},\quad b_n=2^{m(n)},$$

则

$$\mathbb{P}(X_i>2^m)=\sum_{k=m+1}^\infty\frac{1}{2^k(k+1)}\leqslant\sum_{k=m+1}^\infty\frac{1}{2^k m(m+1)}=\frac{2^{-m}}{m(m+1)},$$

同时

$$n\mathbb{P}(X_i>b_n)\leqslant\frac{n2^{-m(n)}}{m(n)(m(n)+1)}\leqslant(m(n)+1)^{-1/2}\to 0.$$

又令 $X'=X\mathbb{1}_{\{|X|\leqslant b_n\}}$, 则

$$\begin{aligned}
\mathbb{E}{X'}^2 &= p_0+\sum_{k=1}^{m(n)}\frac{(2^k-1)^2}{2^k k(k+1)}\\
&\leqslant 1+\sum_{k=1}^{m(n)}\frac{2^k}{k(k+1)}\\
&\leqslant 1+\sum_{k=1}^{m(n)/2}\frac{2^k}{k(k+1)}+\sum_{k=m(n)/2}^{m(n)}\frac{2^k}{k(k+1)}\\
&\leqslant 1+\sum_{k=1}^{m(n)/2}2^k+\frac{4}{m(n^2)}\sum_{k=m(n)/2}^{m(n)}2^k\\
&\leqslant 1+2\cdot 2^{m(n)/2}+\frac{8}{m(n)^2}\cdot 2^{m(n)}\\
&= O\left(\frac{2^{m(n)}}{m(n)^2}\right).
\end{aligned}$$

故

$$\frac{n\mathbb{E}{X'}^2}{b_n^2}\leqslant\frac{C2^{m(n)}}{m(n)^2}\cdot\frac{n}{2^{2m(n)}}\leqslant Cm(n)^{-1/2}\to 0.$$

因此其满足定理 9.2.5. 我们有

$$\begin{aligned}
a_n &= \mathbb{E}X'=-\mathbb{E}X\mathbb{1}_{\{|X|\geqslant b_n\}}\\
&= -\sum_{m(n)+1}^{\infty}\frac{2^k-1}{2^k k(k+1)}\\
&= -\sum_{k=m(n)+1}^{\infty}\frac{1}{k(k+1)}+\sum_{k=m(n)+1}^{\infty}\frac{1}{2^k k(k+1)}\\
&= -\frac{1}{1+m(n)}+\sum_{k=m(n)+1}^{\infty}\frac{1}{2^k k(k+1)}\sim-\frac{1}{m(n)}\sim-\frac{1}{\log_2 n}.
\end{aligned}$$

又

$$2^{m(n)-1}\leqslant\frac{n}{m(n)^{3/2}}\sim\frac{n}{(\log_2 n)^{3/2}},$$

故

$$\frac{S_n+\frac{n}{\log_2 n}}{\frac{n}{(\log_2 n)^{3/2}}}\xrightarrow{\mathbb{P}}0,$$

也即

$$\frac{S_n}{\frac{n}{\log_2 n}}\xrightarrow{\mathbb{P}}-1,$$

□

9.2.4 证明: 用于证明定理 9.2.4 的引理 (3) 中, $(L,M),(L,-M),(-L,M),(-L,-M)$ 同分布.
(提示: 由于 $X_1,X_2,\cdots,X_n$ 都是对称的取有限实值的随机变量且相互独立, 于是对任何 Borel 可测函数 $f(x_1,x_2,\cdots,x_n)$ 来说, $f\left(\widetilde{X_1},\widetilde{X_2},\cdots,\widetilde{X_n}\right),\widetilde{X_n}=X_n$或$-X_n,k=1,2,\cdots,n$ 都是同分布的.)

证明: 我们将证明, 对于独立同分布的对称随机变量列 $\{X_n : n \in \mathbb{N}\}$, 任意的 Borel 可测函数 f 都有

$$f\left(\widetilde{X_1}, \widetilde{X_2}, \cdots, \widetilde{X_n}\right), \quad \widetilde{X_n} = \pm X_n$$

和 $f(X_1, X_2, \cdots, X_n)$ 同分布.

我们采用归纳法证明. 当 $n = 1$,

$$\mathbb{P}(f(X_1) \in B) = \mathbb{P}(X_1 \in f^{-1}(B)) = \mathbb{P}(-X_1 \in f^{-1}(B)) = \mathbb{P}(f(-X_1) \in B), \forall B \in \mathscr{B}.$$

不妨设 $n \leqslant k$ 时结论成立, 当 $n = k + 1$ 时, 定义

$$g(X_1, X_2, \cdots, X_k) := f(X_1, X_2, \cdots, \widetilde{X_{k+1}}), h(X_{k+1}) = f(X_1, X_2, \cdots, X_{k+1}).$$

则

$$g\left(\widetilde{X_1}, \widetilde{X_2}, \cdots, \widetilde{X_k}\right) \stackrel{d}{=} g(X_1, X_2, \cdots, X_k), \quad h(X_{k+1}) \stackrel{d}{=} h\left(\widetilde{X_{k+1}}\right).$$

也即

$$f\left(\widetilde{X_1}, \widetilde{X_2}, \cdots, \widetilde{X_{k+1}}\right) \stackrel{d}{=} f\left(X_1, X_2, \cdots, X_k, \widetilde{X_{k+1}}\right) = h\left(\widetilde{X_{k+1}}\right) \stackrel{d}{=} h(X_{k+1}) = f(X_1, X_2, \cdots, X_{k+1}).$$

因此 L, M 为独立对称随机变量. 故

$$(L, M),\ (L, -M),\ (-L, M),\ (-L, -M)$$

同分布. □

§ 9.3 随机级数的收敛

9.3.1 设 $\{X_n : n \in \mathbb{N}\}$ 为一列 i.i.d. 的随机变量, 满足

$$\mathbb{P}(X_n = -1) = \mathbb{P}(X_n = 1) = \frac{1}{2} \quad (\forall n \in N),$$

则级数 $\sum\limits_{n=1}^{\infty} \dfrac{X_n}{n^\theta}$ 当 $\theta \in \left(\dfrac{1}{2}, 1\right]$ 时 a.s. 收敛, 当 $\theta \in \left[0, \dfrac{1}{2}\right]$ 时 a.s. 发散.

证明: 由我们知道 $\mathbb{E}X_n = 0$, $\mathbb{D}X_n = 1$, $|X_n| \leqslant 1$, 从而 $\theta \geqslant 0$ 时 $\left|\dfrac{X_n}{n^\theta} - \mathbb{E}\left(\dfrac{X_n}{n^\theta}\right)\right| \leqslant n^{-\theta} \leqslant 1$, 可以使用引理 9.3.6. 又 $\sum\limits_{n=1}^{\infty} \sigma\left(\dfrac{X_n}{n^\theta}\right) = \sum\limits_{n=1}^{\infty} \dfrac{1}{n^{2\theta}}$, 故由引理 9.3.6 知当 $\theta \in \left(\dfrac{1}{2}, 1\right]$ 时 a.s. 收敛, 当 $\theta \in \left[0, \dfrac{1}{2}\right]$ 时 a.s. 发散. □

注: 利用三级数定理 (定理 9.3.7), 事实上还可以证明 $\theta > \dfrac{1}{2}$ 是 $\sum\limits_{n=1}^{\infty} \dfrac{X_n}{n^\theta}$a.s. 收敛的充要条件, $\theta \leqslant \dfrac{1}{2}$ 是 $\sum\limits_{n=1}^{\infty} \dfrac{X_n}{n^\theta}$a.s. 发散的充要条件 (注意此时 $\left|\dfrac{X_n}{n^\theta} - \mathbb{E}\left(\dfrac{X_n}{n^\theta}\right)\right| \leqslant n^{-\theta} \leqslant 1$ 未必有界, 仅使用引理 9.3.6 无法证明).

9.3.2 证明推广的 Kolmogorov 不等式 (即定理 9.3.5 的推广): 若 $\{X_n : n \in \mathbb{N}\}$ 为一列相互独立的随机变量, $\mathbb{E}X_n = 0$; 记 $S_n := \sum_{k=1}^{n} X_k$; 对 $c > 0$, 设 $C := \left\{\max_{1 \leqslant k \leqslant n} |S_k| \geqslant c\right\}$, 则

$$c^r \mathbb{P}(C) \leqslant \mathbb{E}\left(|S_n|^r \mathbb{1}_C\right) \leqslant \mathbb{E}|S_n|^r, \quad \forall r \geqslant 1.$$

进一步, 应用此不等式证明: 若 $S_n \xrightarrow{r} S$ (S 有限), 则 $S_n \xrightarrow{\text{a.s.}} S$.

证明: 记 $\Lambda_k = \{\omega : \max_{1 \leqslant j \leqslant k-1} |S_j(\omega)| < c, S_k(\omega) \geqslant c\}$. 则 $C = \bigcup_{k=1}^{n} \Lambda_k$. 又 Λ_k 两两不交, 故

$$\mathbb{E}|S_n|^r = \int |S_n|^r \mathrm{d}\mathbb{P} \geqslant \int_C |S_n|^r \mathrm{d}\mathbb{P} = \mathbb{E}[\mathbb{1}_C |S_n|^r] \geqslant \sum_{k=1}^{n} \int_{\Lambda_k} |S_k|^r \mathrm{d}\mathbb{P} \geqslant c^r \sum_{k=1}^{n} \mathbb{P}(\Lambda_k) = c^r \mathbb{P}(C).$$

□

§ 9.4　强大数律

9.4.1 设 $\{X_n : n \in \mathbb{N}\}$ 为一列 i.i.d. 的随机变量, $\mathbb{E}X_1^+ = \infty, \mathbb{E}X_1^- < \infty$, $S_n = \sum_{k=1}^{n} X_k$, 证明: $\frac{S_n}{n} \xrightarrow{\text{a.s.}} \infty$. 进一步可证: 只要 $\mathbb{E}X_1$ 存在, 就有

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} \mathbb{E}X_1$$

证明: 令 $S_n^+ = X_1^+ + X_2^+ + \cdots + X_n^+$, $S_n^- = X_1^- + X_2^- + \cdots + X_n^-$. 则 $\frac{S_n}{n} = \frac{S_n^+}{n} - \frac{S_n^-}{n}$. 我们知道 $\{X_n^+ : n \in \mathbb{N}\}$ 和 $\{X_n^- : n \in \mathbb{N}\}$ 也是独立同分布序列, 且 $\mathbb{E}X_1^- < \infty$, $\mathbb{E}X_1^+ = \infty$. 故

$$\frac{S_n^-}{n} \xrightarrow{\text{a.s.}} \mathbb{E}X_1^-, \quad \limsup_{n \to \infty} \frac{S_n^-}{n} = \infty, \text{a.s.}.$$

因此我们只需证明 $\liminf_{n \to \infty} \frac{S_n^+}{n} = \infty$, a.s..

令 $Y_n = X_n \mathbb{1}_{\{|X_n| \leqslant M\}}$, $M \in \mathbb{N}$, 则 $\mathbb{E}Y_n < \infty$. 由强大数定律知道 $\frac{S_n'}{n} := \frac{1}{n}\sum_{k=1}^{n} Y_k \to \mathbb{E}Y_1$. 又 $X \geqslant Y$, 故

$$\liminf_{n \to \infty} \frac{S_n}{n} \geqslant \lim_{n \to \infty} \frac{S_n'}{n} = \mathbb{E}Y_1.$$

注意到 $\mathbb{E}Y_1 \uparrow \mathbb{E}X_1 = \infty$, 再根据单调收敛定理可证. □

9.4.2 设 $\{X_n : n \in \mathbb{N}\}$ 为一列 i.i.d. 的随机变量, 应用注 9.4.3 证明:

(1) 若存在某个 $p \in [1, 2]$ 使得 $\sum_{n=1}^{\infty} \frac{1}{n^p} \mathbb{E}|X_n|^p < \infty$, 则 $\sum_{k=1}^{n} X_k \xrightarrow{\text{a.s.}} 0$;

(2) 若存在某个 $\delta \in (0, 1]$ 和 $M < \infty$ 使得 $\forall n \in \mathbb{N}, \mathbb{E}|X_n|^{1+\delta} \leqslant M$, 则 $\frac{1}{n}\sum_{k=1}^{n} c_k X_k \xrightarrow{\text{a.s.}} 0$.

证明: (1) 取 $\varphi(x) = |x|^p$, $p \in [1,2]$. 又 $\sum_{n=1}^{\infty} \frac{\mathbb{E}\varphi(X_n)}{\varphi(n)} = \sum_{n=1}^{\infty} \frac{\mathbb{E}|X_n|^p}{n^p} < \infty$, 则其满足注 9.4.3 之条件, 故 $\frac{1}{n}\sum_{j=1}^{n} X_j \xrightarrow{\text{a.s.}} 0$.

(2) 取 $\varphi(x) = |x|^{1+\delta}$, 类似地有 $\frac{1}{n}\sum_{j=1}^{n} X_j \xrightarrow{\text{a.s.}} 0$. □

9.4.3 若 $\{X_n : n \in \mathbb{N}\}$ 为一列 i.i.d. 的随机变量且 $\mathbb{E}X_1 = 0$, $\{c_n : n \in \mathbb{N}\}$ 是有界实数序列, 则 $\frac{1}{n}\sum_{k=1}^{n} c_k X_k \xrightarrow{\text{a.s.}} 0$.

证明: 不妨设 $m \leqslant c_n \leqslant M$, 则由强大数律, $\frac{1}{n}\sum_{j=1}^{n} X_j \xrightarrow{\text{a.s.}} 0$. 故 $\frac{m}{n}\sum_{j=1}^{n} X_j \leqslant \frac{1}{n}\sum_{j=1}^{n} c_j X_j \leqslant \frac{M}{n}\sum_{j=1}^{n} X_j$. 由夹逼定理即证. □

第十章 特征函数和中心极限定理

§ 10.1 特征函数的定义及简单性质

10.1.1 试求均值为 $\dfrac{1}{\lambda}$ 的指数分布的特征函数.

证明: 均值为 $\dfrac{1}{\lambda}$ 的指数分布的参数为 λ. 考虑服从参数为 λ 的指数分布的随机变量 ξ, 我们有:

$$f(u)=\mathbb{E}\mathrm{e}^{\mathrm{i}u\xi}=\int_0^{\infty}\lambda\mathrm{e}^{\mathrm{i}ut}\mathrm{e}^{-\lambda t}\mathrm{d}t=\frac{\lambda}{\lambda-\mathrm{i}u}.$$

□

10.1.2 试求在 $[-a,a]$ 上分布的三角分布, 即分布函数 $p\left(x\right)=\dfrac{a-|x|}{a^2}$ 的特征函数.

证明:

$$\begin{aligned}f(u)=\mathbb{E}\mathrm{e}^{\mathrm{i}ux}&=\int_{-a}^{a}\mathrm{e}^{\mathrm{i}ux}p(x)\mathrm{d}x\\&=\frac{1}{a^2}\int_{-a}^{a}(a-|x|)(\cos ux+\mathrm{i}\sin ux)\mathrm{d}x\\&=\frac{2}{a^2}\int_0^a(a-x)\cos ux\mathrm{d}x\\&=\frac{2(1-\cos au)}{a^2u^2}.\end{aligned}$$

□

10.1.3 如果 $f_k,k=1.2,\cdots,n$ 是特征函数, $\lambda_k>0,k=1,2,\cdots,n,\sum\limits_{k=1}^{n}\lambda_k=1$, 证明 $\sum\limits_{k=1}^{n}\lambda_kf_k$ 也是特征函数.

证明: 令 $f_k=\displaystyle\int\mathrm{e}^{\mathrm{i}ux}\mu_k(\mathrm{d}x)$, 则

$$\sum_{k=1}^{n}\lambda_kf_k=\sum_{k=1}^{n}\lambda_k\int\mathrm{e}^{\mathrm{i}ux}\mu_k(\mathrm{d}x)=\int\sum_{k=1}^{n}\lambda_k\mathrm{e}^{\mathrm{i}ux}\mu_k(\mathrm{d}x)=\int\mathrm{e}^{\mathrm{i}ux}\left(\sum_{k=1}^{n}\lambda_k\mu_k\right)(\mathrm{d}x).$$

注意到 $\sum\limits_{k=1}^{n}\lambda_k\mu_k$ 是概率测度, 因此 $\sum\limits_{k=1}^{n}\lambda_kf_k$ 也是特征函数. □

10.1.4 试由特征函数的定义, 找出下列各个特征函数对应的随机变量的分布:

(i) $\mathrm{e}^{\mathrm{i}au}$;

(ii) $\cos u$;

(iii) $\cos^2 u$;

(iv) $\sum_{k=0}^{\infty}\lambda_k \mathrm{e}^{\mathrm{i}ku}, \lambda_k \geqslant 0, \sum_{k=0}^{\infty}\lambda_k = 1$;

(v) $\sum_{k=0}^{\infty}\lambda_k \cos ku, \lambda_k \geqslant 0, \sum_{k=0}^{\infty}\lambda_k = 1$;

(vi) $\dfrac{1}{1+\mathrm{i}u}$.

(提示: 利用习题 10.1.1及命题 10.1.2.)

证明: 实际上, 回忆初等概率论的内容可以知道, 特征函数可以唯一决定分布函数, 因此只需找到一个特征函数为题中函数的随机变量即可. 我们有

(i) 若 $\mathbb{P}(X=a)=1$, 则 $f(u)=\mathbb{E}[\mathrm{e}^{\mathrm{i}uX}]=\mathrm{e}^{\mathrm{i}au}$.

(ii) 若 $\mathbb{P}(X=1)=\mathbb{P}(X=-1)=\dfrac{1}{2}$, 则 $f(u)=\mathbb{E}[\mathrm{e}^{\mathrm{i}uX}]=\cos u$.

(iii) 注意到 $\cos^2 u=\dfrac{\cos 2u+1}{2}=\dfrac{1}{2}\mathrm{e}^{\mathrm{i}0u}+\dfrac{1}{2}\cos 2u$, 因此取 $\mathbb{P}(X=0)=\dfrac{1}{2}$, $\mathbb{P}(X=2)=\mathbb{P}(X=-2)=\dfrac{1}{4}$ 即可.

(iv) 只需令 $\mathbb{P}(X=a_k)=\lambda_k$ 即可.

(v) 只需令 $\mathbb{P}(X=k)=\mathbb{P}(X=-k)=\dfrac{\lambda_k}{2}$ 即可.

(vi) 根据习题 10.1.1可知, 参数为 1 的指数分布的特征函数为 $\dfrac{1}{1-\mathrm{i}u}$. 因此根据命题 10.1.2, $\dfrac{1}{1-\mathrm{i}u}$ 是参数为 -1 的指数分布的特征函数.

□

10.1.5 试证: 若 f 是特征函数, 则 $|f|^2$ 也是特征函数.

(提示: 构造相互独立的随机变量 X_1, X_2, 使得 X_1 与 X 同分布, X_2 与 $-X$ 同分布.)

证明: 考虑相互独立的随机变量 X_1, X_2, 其中 $X_1 \stackrel{d}{=\!=} X$, $X_2 \stackrel{d}{=\!=} -X$, 则 X_1, X_2 的特征函数分别为 $f, \bar{f}$. 再根据命题 10.1.2, $|f|^2 = f\bar{f}$ 是 X_1+X_2 的特征函数. □

10.1.6 设 X 的 n 阶绝对矩有限, 试证

$$\mathbb{E}\left(X-\mathbb{E}X\right)^n=\mathrm{i}^{-n}\frac{\mathrm{d}^n}{\mathrm{d}u^n}\left[\mathrm{e}^{-\mathrm{i}u(\mathbb{E}X)}f_X(u)\right]_{u=0}.$$

证明: 令 $Y=X-\mathbb{E}X$, 则根据命题 10.1.2 有 $f_Y(u)=f_X(u)\mathrm{e}^{-\mathrm{i}u\mathbb{E}X}$. 因此再根据命题 10.1.4 可以得到

$$\mathbb{E}(X-\mathbb{E}X)^n=\mathrm{i}^{-n}\frac{\mathrm{d}^n}{\mathrm{d}u^n}\left[\mathrm{e}^{-\mathrm{i}u\mathbb{E}X}f_X(u)\right]_{u=0}.$$

□

10.1.7 试证: 如果 $\{f_n : n \in \mathbb{N}\}$ 是 $\mathbb{R}$ 上的特征函数序列, 且 $\forall u \in \mathbb{R}, f_n(u) \to g(u)$, 且 g 在零点处连续, 则 g 在 $\mathbb{R}$ 上一致连续.

证明: 实际上根据命题 10.1.2, 我们只需证明 g 是特征函数. 由于 g 在零点连续, 故 $\forall \varepsilon > 0$, $\exists \delta > 0$, s.t. $\forall |u| < \delta$, $|1 - g(u)| < \varepsilon$. 设 $\{f_n\}$ 对应的概率测度序列是 $\{\mu_n\}$, 由 Fubini 定理, 当 $u > 0$ 时,

$$\begin{aligned}\frac{1}{u}\int_{-u}^{u}(1-f_n(t))\mathrm{d}t &= \int_{-\infty}^{\infty}\frac{1}{u}\int_{-u}^{u}(1-\mathrm{e}^{\mathrm{i}tx})\mathrm{d}t\mu_n(\mathrm{d}x)\\ &= 2\int_{-\infty}^{\infty}\left(1-\frac{\sin(ux)}{ux}\right)\mu_n(\mathrm{d}x)\\ &\geqslant 2\int_{\{|x|\geqslant \frac{2}{n}\}}\left(1-\frac{1}{|ux|}\right)\mu_n(\mathrm{d}x)\\ &\geqslant \mu_n\left(\left\{x : |x| > \frac{2}{u}\right\}\right).\end{aligned}$$

故

$$\limsup_{n\to\infty}\mu_n\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \leqslant \limsup_{n\to\infty}\frac{1}{u}\int_{-u}^{u}(1-f_n(t))\mathrm{d}t = \frac{1}{u}\int_{-u}^{u}(1-g(t))\mathrm{d}t < 2\varepsilon.$$

根据引理 8.5.11, 任何 $(\mathbb{R}, \mathscr{B})$ 上的概率测度都是胎紧的, 因此 $\forall n < n_0$, 存在紧集 K_n, s.t. $\mu_n(K_n^c) < \varepsilon$. 令 $K = \left(\bigcup_{n=1}^{n_0-1} K_n\right) \cup \left[-\frac{2}{u}, \frac{2}{u}\right]$, 则 K 仍是紧集. 又 $\forall n \in \mathbb{N}$, $\mu_n(K^c) < \varepsilon$, 故概率测度列 $\{\mu_n, n \in \mathbb{N}\}$ 是胎紧的, 再根据定理 8.5.12, $\{\mu_n, n \in \mathbb{N}\}$ 在 $\mathbb{R}$ 上相对紧, 故其任一子列均存在弱收敛子列. 不妨设 $\{\mu_{n_k}\}$ 是其本身的一个弱收敛子列, 则 $\mu_{n_k} \xrightarrow{w} \mu$, 而 $f_{n_k} \to g$. 故特征函数列的极限等于测度列极限的特征函数, 因此 g 是特征函数. □

§ 10.2 逆转公式及连续性定理

10.2.1 试证: $\lim\limits_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\mathrm{e}^{-\mathrm{i}ux_0}f(u)\mathrm{d}u = \mu(\{x_0\})$.

证明: 我们知道 $f(u) = \mathbb{E}\mathrm{e}^{\mathrm{i}ux} = \int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}ux}\mu(\mathrm{d}x)$, 故

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\mathrm{e}^{-\mathrm{i}ux_0}f(u)\mathrm{d}u &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\mathrm{e}^{-\mathrm{i}ux_0}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}ux}\mu(\mathrm{d}x)\,\mathrm{d}u\\ &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}u(x-x_0)}\mu(\mathrm{d}x)\,\mathrm{d}u\\ &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{\mathbb{R}}\cos(u(x-x_0))\mu(\mathrm{d}x)\,\mathrm{d}u\\ &= \int_{\mathbb{R}}\left(\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\cos(u(x-x_0))\mathrm{d}u\right)\mu(\mathrm{d}x)\end{aligned}$$

我们知道 $\cos(u(x-x_0)) \equiv 1$, 当 $x = x_0$. 而 $x \neq x_0$ 时, $|\cos(u(x-x_0))| < 1$, a.s. 故

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\cos(u(x-x_0))\mathrm{d}u = \mathbb{1}_{\{x=x_0\}},$$

因此 $\lim\limits_{T\to\infty}\dfrac{1}{2T}\displaystyle\int_{-T}^{T}\mathrm{e}^{-\mathrm{i}ux_0}f(u)\mathrm{d}u=\int_{\mathbb{R}}\mathbb{1}_{\{x=x_0\}}\mu(\mathrm{d}x)=\mu(\{x_0\})$. □

10.2.2 试证: $\lim\limits_{T\to\infty}\dfrac{1}{2T}\displaystyle\int_{-T}^{T}|f(u)|^2\,\mathrm{d}u=\sum_{x\in\mathbb{R}}\mu(\{x\})^2$.

证明: 根据习题 10.1.5, 我们知道 $|f(u)|^2$ 是 $X-Y$ 的特征函数, 其中 X,Y 独立同分布. 根据习题 10.2.1有

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|f(u)|^2\,\mathrm{d}u=\mu(\{x-y=0\})=\sum_{x\in\mathbb{R}}\mu(\{x\})^2.$$

□

10.2.3 若 X 的特征函数为 f, 概率分布为 μ, 则 f 是实函数 $\Leftrightarrow$ X 与 $-X$ 同分布 $\Leftrightarrow$ (其中 $-B:=\{-x:x\in B\}$).

证明: 第二个等价关系是显然的, 因此我们只需证明 f 是实函数等价于 X 与 $-X$ 同分布. 实际上

$$f_{-X}(u)=\mathbb{E}[\mathrm{e}^{\mathrm{i}u(-X)}]=-\mathrm{i}\mathbb{E}[\sin uX]+\mathbb{E}[\cos uX],f_X(u)=\mathbb{E}[\mathrm{e}^{\mathrm{i}uX}]=\mathrm{i}\mathbb{E}[\sin uX]+\mathbb{E}[\cos uX].$$

因此 f 是实函数 $\Leftrightarrow$ $\mathbb{E}[\sin uX]=0\Leftrightarrow f_X(u)=f_{-X}(u)$. □

10.2.4 证明唯一性定理对于 $(\mathbb{R},\mathscr{B})$ 上的有限符号测度也成立, 即若 μ,ν 是有限符号测度且

$$\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\mu(\mathrm{d}x)=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\nu(\mathrm{d}x),\quad \forall u\in\mathbb{R},$$

则 $\mu=\nu$.

证明: 令 μ^+,μ^-,ν^+,ν^- 是 μ,ν 的 Hahn 分解, 则

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

根据定理 7.1.6, $\forall u\in\mathbb{R}$, 有

$$\begin{aligned}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\mu(\mathrm{d}x)&=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\mu^+(\mathrm{d}x)+\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\mu^-(\mathrm{d}x),\\ \int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\nu(\mathrm{d}x)&=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\nu^+(\mathrm{d}x)+\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\nu^-(\mathrm{d}x).\end{aligned}$$

因此由逆转公式, $\forall x_1,x_2$, $x_1<x_2$, 有

$$\begin{aligned}&\mu((x_1,x_2))+\frac12\mu(\{x_1\})+\frac12\mu(\{x_2\})\\ =&\mu^+((x_1,x_2))+\frac12\mu^+(\{x_1\})+\frac12\mu^+(\{x_2\})-\mu^-((x_1,x_2))-\frac12\mu^-(\{x_1\})-\frac12\mu^-(\{x_2\})\\ =&\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-\mathrm{i}ux_1}-\mathrm{e}^{-\mathrm{i}ux_2}}{\mathrm{i}u}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\mu^+(\mathrm{d}x)+\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\mu^-(\mathrm{d}x)\right)\\ =&\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-\mathrm{i}ux_1}-\mathrm{e}^{-\mathrm{i}ux_2}}{\mathrm{i}u}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\nu^+(\mathrm{d}x)+\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\nu^-(\mathrm{d}x)\right)\\ =&\nu^+((x_1,x_2))+\frac12\nu^+(\{x_1\})+\frac12\nu^+(\{x_2\})-\nu^-((x_1,x_2))-\frac12\nu^-(\{x_1\})-\frac12\nu^-(\{x_2\})\\ =&\nu((x_1,x_2))+\frac12\nu(\{x_1\})+\frac12\nu(\{x_2\}).\end{aligned}$$

所以当

$$x_1, x_2 \in \Lambda := \{y \in \mathbb{R} : \mu(\{y\}) = \nu(\{y\}) = 0\}$$

时, $\mu((x_1, x_2)) = \nu((x_1, x_2))$, 同时也有

$$\mu^+((x_1, x_2)) = \nu^+((x_1, x_2)), \quad \mu^-((x_1, x_2)) = \nu^-((x_1, x_2)).$$

再令 $x_1 \to -\infty$ 且 $x_1 \in \Lambda$, 则有 μ^+, ν^+, μ^-, ν^- 的分布函数分别在它们共同的连续点上相同, 而不连续点至多可数, 所以 $\mu^+ = \nu^+$, $\mu^- = \nu^-$, 也即 $\mu = \nu$. □

10.2.5 设一族概率测度 $\{\mu_n : n \in \mathbb{N}\}$ 满足 $\mu_n(\{0\}) = \dfrac{1}{2}\mu_n(\{n\})\,(\forall n \in \mathbb{N})$, 试讨论 $\{\mu_n\}$ 及其特征函数的收敛性.

证明: 题目表述不清, 此时只考虑 $\{y \in \mathbb{R} : \mu_n(\{y\}) \neq 0\} = \{0, n\}$ 的情况. 我们有

$$f_n(u) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}xu}\mu_n(\mathrm{d}x) = \frac{2}{3}\mathrm{e}^{\mathrm{i}un} + \frac{1}{3} = \frac{1}{3} + \frac{2(\cos \mathrm{i}un + \sin \mathrm{i}un)}{3},$$

因此 f_n 不收敛, μ_n 也并非淡收敛. □

10.2.6 设 X_λ 是均值为 λ 的服从 Poisson 分布的随机变量, 试证: 当 $\lambda \to \infty$ 时, $\dfrac{X_\lambda - \lambda}{\sqrt{\lambda}}$ 按分布率收敛向标准正态分布.

证明: 设 f_λ 为 X_λ 的特征函数, 则 $f_\lambda(u) = \exp(\lambda(\mathrm{e}^{\mathrm{i}u} - 1))$. 根据命题 10.1.2, $\dfrac{x_\lambda - \lambda}{\sqrt{\lambda}}$ 的特征函数为

$$\mathrm{e}^{\mathrm{i}(-\sqrt{\lambda})u} f_\lambda\left(\frac{u}{\sqrt{\lambda}}\right) = \exp\left(\lambda\left(\mathrm{e}^{\frac{\mathrm{i}u}{\sqrt{\lambda}}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right)\right).$$

再根据 Taylor 展开可得

$$\lim_{\lambda\to\infty} \lambda\left(\mathrm{e}^{\frac{\mathrm{i}u}{\sqrt{\lambda}}} - 1 - \frac{\mathrm{i}u}{\sqrt{\lambda}}\right) = \lim_{\lambda\to\infty} \lambda\left(-\frac{u^2}{2\lambda} + o\left(\left(\frac{\mathrm{i}u}{\sqrt{\lambda}}\right)^2\right)\right) = -\frac{u^2}{2}.$$

而 $\mathrm{e}^{-\frac{u^2}{2}}$ 是标准正态分布的特征函数且在零点连续, 故由推论 10.2.5, 当 $\lambda \to \infty$ 时, $\dfrac{X_\lambda - \lambda}{\sqrt{\lambda}}$ 按分布律收敛向标准正态分布. □

§ 10.3 中心极限定理

10.3.1 证明推论 10.3.13 和推论 10.3.14.

证明: 设 X_{nk} 的特征函数为 $f_{nk}(u)$, $X_{nk} - a_{nk}$ 对应的特征函数为 $g_{nk(u)}$, 则

$$\prod_{k=1}^{k_n} f_{nk}(u) = \exp\left\{\sum_{k=1}^{k_n} \mathrm{i}a_{nk}u + \sum_{k=1}^{k_n} \ln g_{nk}(u)\right\} = \exp\left\{\sum_{k=1}^{k_n} \mathrm{i}a_{nk}u + \sum_{k=1}^{k_n} \int_{\mathbb{R}} \frac{\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux}{x^2}\mu_k(\mathrm{d}x)\right\}.$$

其中 μ_k 是有限 L-S 测度, 且 $\mu_k(\{0\}) = \mathbb{D}X_{nk}$, $\mu_k(\mathbb{R}\backslash\{0\}) = 0$.

推论 10.3.13 的证明: 由连续性定理, $\sum_k X_{nk} \xrightarrow{D} \xi$ 的充要条件是 $\prod_{k=1}^{k_n} f_{nk}(u) \to f_\xi(u)$, 令 $\xi \sim N(a,\delta)$, 则

$$f_\xi(u) = \exp\left\{\mathrm{i}au + \int_{\mathbb{R}} \frac{\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux}{x^2}\mu_\delta(\mathrm{d}x)\right\},$$

其中 $\mu_\delta(\{0\}) = \delta$, $\mu_\delta(\mathbb{R}\backslash\{0\}) = 0$. 故 $\sum_k X_{nk} \xrightarrow{D} \xi$ 的充要条件为 $\sum_k a_{nk} \to a$ 且 $\sum_k \mu_k \xrightarrow{v} \mu_\delta$. 后者即对一切不以 0 为端点的连续区间 $I \subset \mathbb{R}$, $\sum_k \mu_k(I) \to \mu_\delta(I)$, 也即

$$\sum_k \int_{|x-a_{nk}|\geqslant\varepsilon} |x-a_{nk}|^2 \mathbb{P}_{X_{nk}}(\mathrm{d}x) \to \delta \mathbb{1}_{\{|a_{nk}|\geqslant\varepsilon\}}.$$

再令 $\delta \to 0$, 则 $\xi \xrightarrow{D} a$, 此时上式变为

$$\sum_k \int_{|x-a_{nk}|\geqslant\varepsilon} |x-a_{nk}|^2 \mathbb{P}_{X_{nk}}(\mathrm{d}x) \to 0.$$

推论 10.3.14 的证明: 令 $\eta \sim \mathscr{P}(\lambda)$, 由连续性定理, $\sum_k X_{nk} \xrightarrow{D} \eta$ 的充要条件是 $\prod_{k=1}^{k_n} f_{nk}(u) \to f_\eta(u)$, 其中

$$f_\eta(u) = \exp\left\{\mathrm{i}\lambda u + \int_{\mathbb{R}} \frac{\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux}{x^2}\mu_\eta(\mathrm{d}x)\right\},$$

其中 $\mu_\eta(\{1\}) = \lambda$, $\mu_\eta(\mathbb{R}\backslash\{1\}) = 0$. 故 $\sum_k X_{nk} \xrightarrow{D} \eta$ 的充要条件为 $\sum_k a_{nk} = \sum_k \mathbb{E}X_{nk} \to \lambda$ 且 $\sum_k \mu_k \xrightarrow{v} \mu_\eta$. 后者即对一切不以 1 为端点的连续区间 $I \subset \mathbb{R}$, $\sum_k \mu_k(I) \to \mu_\eta(I)$, 也即

$$\sum_k \int_{|x-1|\geqslant\varepsilon} x^2 \mathrm{d}\mathbb{P}_{X_{nk}-\mathbb{E}X_{nk}} \to 0.$$

(实际上书中推论 10.3.14 的积分区域有误, λ 应为 1). □

10.3.2 若 X_n 在 $[-n,n]$ 上均匀分布, $n \in \mathbb{N}$, 试证 $\{X_n\}$ 满足 Lindeberg 条件.
证明: 我们有

$$\mathbb{E}X_n = 0, \mathbb{D}X_n = \frac{n^2}{3}, s_n^2 = \sum_{k=1}^n \mathbb{D}X_n = \sum_{k=1}^n \frac{k^2}{3}.$$

下面验证其满足 Lindeberg 条件. 我们有

$$\begin{aligned}
g_n(\varepsilon) &= \frac{1}{s_n^2}\sum_{k=1}^n \int_{|x-\mathbb{E}X_k|\geqslant\varepsilon s_n} |x-\mathbb{E}X_k|^2 \mathrm{d}\mathbb{P}_{X_k} \\
&= \frac{1}{s_n^2}\sum_{k=1}^n \int_{|x|\geqslant\varepsilon s_n} |x|^2 \mathrm{d}\mathbb{P}_{X_k} \\
&= \frac{2}{s_n^2}\sum_{k=1}^n \int_{\varepsilon s_n}^n x^2 \mathrm{d}x \\
&= \frac{2}{3s_n^2}\sum_{k=1}^n \max\{k^3 - (\varepsilon s_n)^3, 0\}.
\end{aligned}$$

由于 $s_n^2 \to \infty$ 当 $n \to \infty$, 故上式仅有有限项非零项, 因此 $g_n(\varepsilon) \to 0$, 故 Lindeberg 条件成立. □

10.3.3 设 $\{X_n : n \in \mathbb{N}\}$ 是一列相互独立的随机变量, $S_n = \sum_{k=1}^n X_k$, $s_n^2 = \sum_{k=1}^n \mathbb{D}X_k$, 判断

$$\frac{S_n - \mathbb{E}S_n}{s_n} \xrightarrow{D} N$$

在下列哪种情况下成立:

(i) $\mathbb{P}\left(X_k = -2^k\right) = \mathbb{P}\left(X_k = 2^k\right) = \frac{1}{2}, k = 1, 2, \cdots;$

(ii) $\mathbb{P}\left(X_k = -2^k\right) = \mathbb{P}\left(X_k = 2^k\right) = 2^{-(k+1)}, \mathbb{P}\left(X_k = 0\right) = 1 - 2^{-k}, k = 1, 2, \cdots;$

(iii) $\mathbb{P}\left(X_k = -2^k\right) = \mathbb{P}\left(X_k = 2^k\right) = \frac{1}{2}k^{-\frac{1}{2}}, \mathbb{P}\left(X_k = 0\right) = 1 - k^{-\frac{1}{2}}, k = 1, 2, \cdots;$

证明: 我们只需逐一验证 Lindeberg 条件.

(i) 我们有 $\mathbb{E}X_k = 0$, $\mathbb{D}X_k = 4^k$, 则 $s_n^2 = \frac{4^{n+1}-4}{3}$. 只需取 $\varepsilon < \frac{3}{4}$, 就有

$$\frac{1}{s_n^2}\sum_{k=1}^n \mathbb{E}[|X_k - \mathbb{E}X_k|^2; |X_k - \mathbb{E}X_k| \geqslant \varepsilon s_n] = \frac{1}{s_n^2}\sum_{k=1}^n 4^k \mathbb{1}_{\{4^k \geqslant \varepsilon s_n\}} \nrightarrow 0.$$

因此不满足 Linderberg 条件.

(ii) 我们有 $\mathbb{E}X_k = 0$, $\mathbb{D}X_k = 2^k$, 则 $s_n^2 = 2^{n+1} - 2$. 只需取 $\varepsilon < \frac{1}{2}$, 就有

$$\frac{1}{s_n^2}\sum_{k=1}^n \mathbb{E}[|X_k - \mathbb{E}X_k|^2; |X_k - \mathbb{E}X_k| \geqslant \varepsilon s_n] = \frac{1}{2^{n+1}-1}\sum_{k=1}^n 2^k \mathbb{1}_{\{2^k \geqslant \varepsilon s_n\}} \nrightarrow 0.$$

因此不满足 Linderberg 条件.

(iii) 我们有 $\mathbb{E}X_k = 0$, $\mathbb{D}X_k = k^{\frac{3}{2}}$, $s_n^2 = \sum_{k=1}^n k^{\frac{3}{2}} = O(n^{\frac{5}{2}})$. 我们有

$$g_n(\varepsilon) = \frac{1}{s_n^2}\sum_{k=1}^n \mathbb{E}[|X_k - \mathbb{E}X_k|^2 : |X_k - \mathbb{E}X_k| \geqslant \varepsilon s_n] = \frac{1}{s_n^2}\sum_{k=1}^n k^{\frac{3}{2}} \mathbb{1}_{\{X_k \geqslant \varepsilon s_n\}}.$$

又 $\frac{n}{s_n} \sim O(n^{-\frac{1}{4}})$, 故求和项至多有限. 因此 $g_n(\varepsilon) \to 0$.

0 因此第 (iii) 种情况满足 $\frac{S_n - \mathbb{E}S_n}{s_n} \xrightarrow{D} N$. □

10.3.4 设 $\{X_n : n \in \mathbb{N}\}$ 是一列 i.i.d. 的随机变量, X_1 服从参数为 1 的 Poisson 分布, 对这一随机变量序列应用定理 10.3.12 (Lyapunov 定理), 证明:

$$\lim_{n\to\infty} \mathrm{e}^{-n}\sum_{k=0}^n \frac{n^k}{k!} = \frac{1}{2}.$$

证明: 我们知道 $\mathbb{E}X_k=1$, $\mathbb{D}X_k=1$, 故 $s_n=n$. 因此

$$\frac{1}{s_n^3}\sum_{k=1}^{n}\mathbb{E}|X_k-\mathbb{E}X_k|^3=n^{-\frac{3}{2}}\cdot n\mathbb{E}|X_1-\mathbb{E}X_1|^3=\frac{\mathbb{E}[X_1-\mathbb{E}X_1]^3+2\mathbb{P}(X_1=0)}{\sqrt{n}}\to 0,$$

故 $\dfrac{\sum_{k=1}^{n}(X_k-\mathbb{E}X_k)}{s_n}=\dfrac{\sum_{k=1}^{n}X_k-n}{\sqrt{n}}\xrightarrow{D}N$. 因此

$$\mathrm{e}^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\mathbb{P}\left(\sum_{k=1}^{n}X_k\leqslant n\right)=\mathbb{P}\left(\frac{\sum_{k=1}^{n}X_k-n}{\sqrt{n}}\leqslant 0\right)\to\frac{1}{2}.$$

□

10.3.5 证明形如 $\mathrm{e}^{\psi(u)}$ (其中 $\psi(u)=\displaystyle\int\frac{(\mathrm{e}^{\mathrm{i}ux}-1-\mathrm{i}ux)}{x^2}\mu(\mathrm{d}x)$, 即教材本节的 (*) 式) 的特征函数是无穷可分的.

证明: 考虑 $\forall 0<\varepsilon<1$, 我们有

$$I_1:=\int_{\varepsilon}^{1/\varepsilon}\frac{\mathrm{e}^{\mathrm{i}ux}-1-\mathrm{i}ux}{x^2}\mu(\mathrm{d}x)=\lim_{n\to\infty}\sum_{k=0}^{n-1}\frac{\mathrm{e}^{\mathrm{i}u\xi_k}-1-\mathrm{i}u\xi_k}{\xi_k^2}\mu((x_k,x_{k+1}]),$$

其中 $\varepsilon=x_0<x_1<\cdots<x_n=\dfrac{1}{\varepsilon}$, $x_k\leqslant\xi_k<x_{k+1}$, $k=0,1,\cdots,n-1$, 且 $\sup\limits_k\{x_{k+1}-x_k\}\to 0$. 令

$$b_{n,k}=\xi_k,\quad \lambda_{n,k}=\frac{\mu((x_k,x_{k+1}])}{\xi_k^2},\quad a_{n,k}=-\frac{\mu((x_k,x_{k+1}])}{\xi_k},$$

再注意到, 若考虑独立的 Poisson 分布随机变量列 $\{Y_{n,k}:k\in\mathbb{N}\cap\{0,n-1\}\}$, 其中 $Y_{n,k}$ 服从参数为 $\lambda_{n,k}$ 的 Poisson 分布. 令 $X_{n,k}=b_{n,k}Y_{n,k}$, 则其特征函数的对数为

$$\ln f_{X_{n,k}}(u)=\lambda_{n,k}(\mathrm{e}^{\mathrm{i}b_{n,k}u}-1)=\frac{\mathrm{e}^{\mathrm{i}u\xi_k}-1}{\xi_k^2}\mu((x_k,x_{k+1}]).$$

再令 $I_{1,n}:=\displaystyle\sum_{k=0}^{n-1}\frac{\mathrm{e}^{\mathrm{i}u\xi_k}-1-\mathrm{i}u\xi_k}{\xi_k^2}\mu((x_k,x_{k+1}])$, 则 $\exp\{I_{1,n}\}$ 是 $\displaystyle\sum_{k=0}^{n-1}(X_{n,k}+a_{n,k})$ 的特征函数. 且 $\exp\left\{\lim\limits_{n\to\infty}I_{1,n}\right\}=\exp\{I_1\}$.

类似地令 $I_2:=\displaystyle\int_{-\varepsilon}^{-1/\varepsilon}\frac{\mathrm{e}^{\mathrm{i}ux}-1-\mathrm{i}ux}{x^2}\mu(\mathrm{d}x)$, 则 $\exp\{I_2\}$ 也是某个由有限个独立的 Poisson 型特征函数的乘积的极限.

再令 $\varepsilon\to 0$, 现在我们有 $\psi(u)=\lim\limits_{\varepsilon\to 0}(I_1+I_2)-\dfrac{1}{2}u^2\mu(\{0\})$. 显然特征函数 $\exp\left\{-\dfrac{1}{2}u^2\mu(\{0\})\right\}$ 是退化分布的特征函数, 其无穷可分. 下面考虑 I_1,I_2.

我们知道 Poisson 分布也是无穷可分的, 那么 $X_{n,k}+a_{n,k}$ 无穷可分, 进而 $\displaystyle\sum_{k=0}^{n-1}(X_{n,k}+a_{n,k})$ 无穷可分.

因此只要能够证明无穷可分律的极限仍是无穷可分律即可 (这实际上是定理 10.3.16, 但书中未给出证明). 考虑收敛到某个特征函数 $f(u)$ 的无穷可分的特征函数序列 $\{f_n(u)\}$, 记 $(f_n(u))^{\mathbb{1}/m}=\exp\left\{\dfrac{1}{m}\,\mathrm{Ln}\,f_n(u)\right\}$, 我们有 $\mathrm{Ln}\,f_n(0)=0$, 因此 $\forall m$, 当 $n\to\infty$ 时有 $(f_n(u))^{1/m}\to(f(u))^{1/m}$. 由于 f_n 是无穷可分的, 因此 $\forall n$, $(f_n(u))^{1/m}$ 是特征函数. 由于 $(f(u))^{1/m}$ 在 0 处连续, 根据习题 10.1.7可知 $(f(u))^{1/m}$ 也是特征函数, 因此 $f(u)$ 无穷可分.

因此无穷可分律的极限仍是无穷可分律. 我们知道 e^{I_1} 是无穷可分特征函数, 故 $\lim\limits_{\varepsilon\to 0} e^{I_1}$ 也是无穷可分特征函数, 类似地, $\lim\limits_{\varepsilon\to 0}\mathrm{e}^{I_2}$ 也是无穷可分特征函数. 故

$$\mathrm{e}^{\psi(u)} = \left(\lim_{\varepsilon\to 0}\mathrm{e}^{I_1}\right)\times\left(\lim_{\varepsilon\to 0}\mathrm{e}^{I_2}\right)\times\exp\left\{-\frac{1}{2}u^2\mu(\{0\})\right\}$$

是无穷可分的. □

10.3.6 设特征函数 f 是无穷可分的, 试证: $\forall u, f(u)\neq 0$.
(提示: 考虑 $g(u)=|f(u)|^2$, 证明它仍是无穷可分的, 且 $g(u)$ 恒不为零.)
证明: 由于 f 是无穷可分的, 故 $\forall n\in\mathbb{N}$, $\exists f_n$, s.t. $f=f_n^n$. 令 $g=|f|^2$, $g_n=|f_n|^2$, 以及 $h:=\lim\limits_{n\to\infty} g_n = \mathbb{1}_{\{g>0\}}$. 我们知道 g, g_n 均是实特征函数, 则它们连续且 $g^{\frac{1}{n}}=g_n$ 唯一. 又 $g(0)=1$ 且 g 在 0 处连续, 故 h 在 0 处连续. 根据推论 10.2.5 可得 h 是特征函数, 因此其连续. 又 $h(0)=1$, 因此 $h\equiv 1$. 故对任意 u 有 $g_n\neq 0$, 因此 $\forall u$, $f(u)\neq 0$. □

参考文献

[Du] Durrett R. Probability: Theory and Examples[M]. Cambridge University Press, 2019.

[GI] Wacker P. Please, not *another* note about Generalized Inverses[J]. arXiv preprint arXiv:2306.06989, 2023.