概率论笔记

崔嘉祺
华东师范大学 20 级数学强基拔尖班

2023 年 10 月 18 日

摘要

这是华东师范大学数学专业研究生的基础课“概率论”的课程笔记。
每周三交作业。

目录

1 测度与概率空间

定义 1.0.1. 对集合 Ω, 子集族 $\mathcal{F}\subset 2^{\Omega}$ 称为 Ω 上的一个 σ-代数，若

(1) $\Omega\in\mathcal{F}$.

(2) 若 $A\in\mathcal{F}$, 则 $A^c\in\mathcal{F}$. (补封闭性)

(3) 若 $\{A_i\}_{i=1}^{\infty}\subset\mathcal{F}$, 则 $\bigcup_{i=1}^{\infty}A_i\in\mathcal{F}$. (可列并封闭)

注记 1.0.2. 由定义可推出:

1. $\varnothing\in\mathcal{F}$.
2. 若 $A,B\in\mathcal{F}$, 则 $A\cap B\in\mathcal{F}$. (交封闭性)
3. 若 $A,B\in\mathcal{F}$, 则 $A\cup B\in\mathcal{F}$. (有限并封闭性)

注记 1.0.3. 若把定义中的 *(3)* 改为有限并，则称其为一个 *Borel* 代数。

定义 1.0.4. 对集合 Ω, 子集族 $\mathcal{F}\subset 2^{\Omega}$ 是 Ω 上的一个 σ-代数，$\mu:\mathcal{F}\to[0,+\infty]$, 满足

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

(2) 若 $\{A_i\}_{i=1}^{\infty}\subset\mathcal{F}$, $A_i\cap A_j=\varnothing, i\neq j$, 则有

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

(σ 可加性/可数可加性)

则称三元组 $(\Omega,\mathcal{F},\mu)$ 是一个测度空间。

例 1.0.5. $\mathbb{R}$ 上的 *Lebesgue* 测度：$\Omega = \mathbb{R}$, $\mathcal{F}$ 是 *Lebesgue* 可测集，μ 是 *Lebesgue* 测度。

定义 1.0.6. 称一个测度空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 是一个概率空间，若 $\mathbb{P} : \mathcal{F} \to [0, 1]$, 且 $\mathbb{P}(\Omega) = 1$. 其中称 Ω 是样本空间，$\omega \in \Omega$ 是基本事件，$\mathcal{F}$ 是事件域，$A \in \mathcal{F}$ 是事件，$\mathbb{P}(A)$ 是 A 的概率。

定理 1.0.7. 若 $(\Omega, \mathcal{F}, \mathbb{P})$ 是一个概率空间，则有

(1) 若 $A, B \in \mathcal{F}$, $A \subset B$, 则 $\mathbb{P}(A) \leq \mathbb{P}(B)$. （单调性）

(2) 若 $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, 则有

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

（次可加性）

(3) 设 $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, 若

- $A_1 \supset A_2 \supset A_3 \supset \cdots$, 则记 $\lim_{i\to\infty} A_i = \bigcap_{i=1}^{\infty} A_i = A$, 记为 $A_i \searrow A$.
- $A_1 \subset A_2 \subset A_3 \subset \cdots$, 则记 $\lim_{i\to\infty} A_i = \bigcup_{i=1}^{\infty} A_i = A$, 记为 $A_i \nearrow A$.

则 $A \in \mathcal{F}$, 且 $\mathbb{P}(A) = \lim_{i\to\infty} \mathbb{P}(A_i)$.

(4) $\forall A \in \mathcal{F}$, $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$.

证明. (1) 由 $B = A \cup (B \cap A^c)$, 则 $A \cap (B \cap A^c) = \varnothing$. 由可加性即得。

(2) 记 $B_i = A_i \cap (A_1 \cup \cdots \cup A_{i-1})^c$, 则 $\{B_i\}$ 两两不交，从而由可数可加性与单调性，

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

(3) 只证 $A_i \nearrow A$. 由 $A = \cup_{i=1}^{\infty} A_i$, 由 σ-代数的可数并封闭性，$A \in \mathcal{F}$. 记 $B_1 = A_1$, $B_n = A_n \cap A_{n-1}^c$. 则

$$A_n = B_1 \cup B_2 \cup \cdots \cup B_n,$$

且 $\{B_i\}$ 两两不交，

$$A = B_1 \cup B_2 \cup \cdots .$$

由可数可加性，

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(B_i) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

(4) $A \cap A^c = \varnothing$, 从而由可加性，

$$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1.$$

□

例 1.0.8. $\Omega = [0,1]$, $\mathcal{F}$ 是 *Lebesgue* 可测集，$\mathbb{P}$ 是 *Lebesgue* 测度。

$$\mathbb{P}\left(x \in [0,\frac{1}{3}]\right) = \mathbb{P}\left([0,\frac{1}{3}]\right) = \frac{1}{3}.$$

例 1.0.9 (古典概型)**.** Ω 是有限集，$\mathcal{F} = 2^{\Omega}$, $A \in \mathcal{F}$,

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}.$$

例 1.0.10 (离散概率模型)**.** Ω 是可数集，$\mathcal{F} = 2^{\Omega}$, $p : \Omega \to [0,+\infty]$, 满足 $\sum_{\omega\in\Omega} p(w) = 1$. 对 $A \in \mathcal{F}$, 令 $\mathbb{P}(A) = \sum_{\omega\in A} p(w)$. 则 $(\Omega,\mathcal{F},\mathbb{P})$ 是概率空间。

例 1.0.11 (投均匀硬币)**.** 记 H 是正，T 是反。

- 投一次：$\Omega = \{H,T\}$, $\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = \frac{1}{2}$.

- 投 n 次：$\Omega=\{H,T\}^n=\{(x_1,\cdots,x_n)\mid x_i\in\{H,T\},1\le i\le n\}$, $\mathcal{F}=2^{\Omega}$, $\forall\omega\in\Omega$, $\mathbb{P}(\{\omega\})=\frac{1}{2^n}$.
  - A_1 是第一次投正面，$\mathbb{P}(A_1)=\frac{|A_1|}{|\Omega|}=\frac{2^{n-1}}{2^n}=\frac{1}{2}$. A_1^c 是第一次投正面，$\mathbb{P}(A_1^c)=\frac{1}{2}$.
  - A_i 是第 i 次投正面，$\mathbb{P}(A_i)=\frac{1}{2}$. 若 $1\le i_1\le i_2\le\cdots\le i_k\le n$, $B=B_{i_1}\cap\cdots\cap B_{i_k}$, 其中 $B_{i_j}\in\{A_{i_j},A_{i,j}^c\}$, 则有 $\mathbb{P}(B)=\frac{1}{2^k}$.

例 1.0.12 (Ramsey 数的应用)**.**

定理 1.0.13. 任意六个人中必有三个人互相认识或互相不认识。

证明. 任取一个人 A，考察与其余五个人的关系。有三个人与 A 认识或不认识，不妨设 B,C,D 与 A 认识（不认识的情况类似）。若 B,C,D 中有两人认识（如 B,C），则 A,B,C 互相认识；若 B,C,D 没有两人认识，则三者满足互相不认识。 □

注记 1.0.14. 图论形式：6 阶完全图，边用红蓝染色，则必有一个红色或蓝色三角形。

问题 1.0.15 (一般的 Ramsey 问题)**.** 给定正整数 k,l, 是否存在 n, 满足将 n 阶完全图 K_n 的边红蓝染色后，一定有一个红的 K_k 或蓝的 K_l. 满足这样的条件的最小的 n 记为 $R(k,l)$.

命题 1.0.16. 一些结论：

- $R(3,3)=6$.
- $R(2,l)=l$.
- $R(k,l)=R(l,k)$

定理 1.0.17. 对 $k,l\ge 3$, 有

$$R(k,l)\le R(k-1,l)+R(k,l-1).$$

归纳可证 $R(k,l)$ 是有限数。

证明. 取出一个点 A, 剩下的图记为 W. 记 $n = R(k-1,l) + R(k,l-1)$. 由

$$|W| = n - 1 = R(k-1,l) + R(k,l-1) - 1 \geq R(k-1,l),$$

从而

- 要么有 $k-1$ 个点组成红色完全图，此时再把它们与 A 用红色相连即得。
- 要么有 l 个点组成蓝色完全图，

□

定理 1.0.18. 对 $k \geq 3$, 有

$$R(k,k) \geq \left\lfloor 2^{\frac{k}{2}} \right\rfloor.$$

证明. 对 K_n 的边随机染色，所有可能的染色方案是样本空间 Ω, $|\Omega| = 2^{C_n^2}$. 对指定的 k 个顶点 $v_1, v_2, \cdots, v_k$,

$$\mathbb{P}(v_1, \cdots, v_k \text{ 之间的是红边}) = \frac{1}{2^{C_k^2}}.$$

从而

$$\begin{aligned}
\mathbb{P}(\text{存在 } k \text{ 个点之间的是红边}) &= \mathbb{P}\Big(\bigcup_{v_1,\cdots,v_k} v_1, \cdots, v_k \text{ 之间的是红边}\Big) \\
&\leq \sum_{v_1,\cdots,v_k} \mathbb{P}(v_1, \cdots, v_k \text{ 之间的是红边}) = \frac{C_n^k}{2^{C_k^2}}.
\end{aligned}$$

同理,

$$\mathbb{P}(\text{存在 } k \text{ 个点之间的是蓝边}) \leq \frac{C_n^k}{2^{C_k^2}}.$$

从而

$$
\begin{aligned}
&\mathbb{P}(\text{存在 } k \text{ 个点之间的是红边或 } k \text{ 个点之间的是蓝边})\\
\leq &\mathbb{P}(\text{存在 } k \text{ 个点之间的是红边}) + \mathbb{P}(\text{存在 } k \text{ 个点之间的是蓝边})\\
\leq &C_n^k \cdot 2^{1-C_k^2}.
\end{aligned}
$$

而其中，

$$C_n^k \cdot 2^{1-C_k^2} = \frac{n!}{k!(n-k)!} \cdot 2^{1+\frac{k}{2}} \cdot 2^{-\frac{k^2}{2}} < \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{n^k}{2^{\frac{k^2}{2}}}.$$

易知对 $n \geq 3$, 有 $2^{1+\frac{k}{2}} < k!$, 故

$$C_n^k \cdot 2^{1-C_k^2} < \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{n^k}{2^{\frac{k^2}{2}}} < \frac{n^k}{2^{\frac{k^2}{2}}} = \left(\frac{n}{2^{\frac{k}{2}}}\right)^k.$$

故只要 $n = \left\lfloor 2^{\frac{k}{2}} \right\rfloor$ 时，就有

$$\mathbb{P}(\text{存在 } k \text{ 个点之间的是红边或 } k \text{ 个点之间的是蓝边}) \leq C_n^k \cdot 2^{1-C_k^2} < \left(\frac{n}{2^{\frac{k}{2}}}\right)^k \leq 1.$$

此时，存在一种染色方案，使得不存在 k 个点之间的是红边，也不存在 k 个点之间的是蓝边。 □

例 1.0.19 (投掷硬币无穷多次). $\Omega = \{H, T\}^{\mathbb{N}} = \{(x_1, x_2, \cdots) \mid x_i \in \{H, T\}, i = 1, 2, \cdots\}$. 记 A_i 是第 i 次是 H. 定义

$$\sigma(A_1) = \mathcal{F}_1 = \{\varnothing, A_1, A_1^c, \Omega\},$$

它只能分辨第一次的结果。

定义 1.0.20. 设 Ω 是集合，A 是它的一些子集，记 $\sigma(A)$ 是由 A 生成的 σ-代数，是包含 A 的最小 σ-代数。

定义

$$\sigma(A_1, \cdots, A_n) = \mathcal{F}_n, |\mathcal{F}_n| = 2^{2^n}.$$

它只能分辨前 n 次结果。

问题 1.0.21. 无法处理无穷次的情况。怎么办？

最初的想法：

$$\mathcal{A}=\bigcup_{n=1}^{\infty}\mathcal{F}_n.$$

但 $\mathcal{A}$ 不是 σ-代数，它只对有限并封闭。

目标 1.0.22. 找到一个 σ-代数 $\mathcal{B}$ 包含 $\mathcal{A}$, $\mathbb{P}:\mathcal{B}\to[0,1]$, 使得 $\mathbb{P}|_{\mathcal{F}_n}=\mu_n$.

注意到有

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

对 $i>j$, $\mu_i:\mathcal{F}_i\to[0,1]$, $\mu_i|_{\mathcal{F}_j}=\mu_j$, 拼起来得到 $\mu:\mathcal{A}\to[0,1]$.

定义 1.0.23. 若 $\mathcal{A}$ 是集合 Ω 的子集族，满足

1. $\Omega\in\mathcal{A}$.
2. 补封闭。
3. 有限并封闭。

则称 $\mathcal{A}$ 是一个 *Boole* 代数。

定义 1.0.24. 对 *Boole* 代数 $\mathcal{A}$, 函数 $m:\mathcal{A}\to\mathbb{R}$ 满足

1. $\forall A\in\mathcal{A}$, 有 $m(A)\geq 0$.
2. $\forall A,B\in\mathcal{A}$, 满足 $A\cap B=\varnothing$, 则有 $m(A\cup B)=m(A)+m(B)$.
3. $\forall\{A_n\}\subset\mathcal{A}$, 满足 $A_n\searrow\varnothing$, 则 $m(A_n)\to 0$.（连续性）

则称 m 是 *Boole* 代数 $\mathcal{A}$ 上的测度。

定义 1.0.25. 对 *Boole* 代数 $\mathcal{A}$ 以及它上面的测度 m. $\forall A \subset \Omega$,

$$m^*(A) = \inf\left\{\sum_i m(S_i) \mid S_i \in \mathcal{A}, A \subset \bigcup_i S_i\right\}$$

称为 A 的外测度。

注记 1.0.26. *1.* 在外测度的定义中可以假设 S_i 互不相交：令 $T_1 = S_1, T_i = S_i - (S_1 \cup \cdots \cup S_{i-1})$, 有 $T_i \in \mathcal{A}$, $A \subset \bigcup_i T_i$, $m(T_i) \leq m(S_i)$, 从而 $\sum m(T_i) \leq \sum m(S_i)$.

2. 若 $A \in \mathcal{A}$, 则 $m^*(A) = m(A)$.

定义 1.0.27. 对 $A \subset \Omega$, 称 A 是 m^*-可测的，若 $\forall E \subset \Omega$, 有

$$m^*(A \cap E) + m^*(A^c \cap E) = m^*(E). \tag{1}$$

其中 *(1)* 式称为可测性条件。

定理 1.0.28 (m^* 的次可加性)**.** $\forall A, A_1, A_2, \cdots \subset \Omega$, 满足 $A \subset \bigcup_i A_i$, 则 $m^*(A) \leq \sum_i m^*(A_i)$.

证明. $\forall \varepsilon > 0$, 由 inf 的定义，可取 $\{A_{ij}\}_j \subset \mathcal{A}$, 满足 $A_i \subset \bigcup_j A_{ij}$, 且 $\sum_j m(A_{ij}) \leq m^*(A_i) + \frac{\varepsilon}{2^i}$, 从而有 $A \subset \bigcup_i \bigcup_j A_{ij}$. 此时，

$$m^*(A) \leq \sum_i \sum_j m(A_{ij}) \leq \sum_i \left(m^*(A_i) + \frac{\varepsilon}{2^i}\right) = (m^*(A_i)) + \varepsilon.$$

令 $\varepsilon \to 0$ 即得 $m^*(A) \leq \sum_i m^*(A_i)$. □

注记 1.0.29. 由 m^* 次可加性，$\forall E \subset \Omega$,

$$m^*(E) = m^*((A \cap E) \cup (A^c \cap E)) \leq m^*(A \cap E) + m^*(A^c \cap E).$$

故可测性条件只需要验证

$$m^*(A \cap E) + m^*(A^c \cap E) \leq m^*(E).$$

定理 1.0.30 (Caratheodory 测度扩张定理)**.** 设 $\mathcal{A}$ 是 Ω 上的 *Boole* 代数，m 是 $\mathcal{A}$ 上的测度，m^* 是外测度。记 $\mathcal{M}(m^*)$ 是所有 m^*-可测集构成的子集族。则 $\mathcal{M}(m^*)$ 是包含 $\mathcal{A}$ 的一个 σ-代数，$m^* : \mathcal{M}(m^*) \to \mathbb{R}$ 是一个测度，且 $m^*|_{\mathcal{A}} = m$.

证明. 先证包含关系

引理 1.0.31. $\mathcal{A} \subset \mathcal{M}(m^*)$.

证明. $\forall A \in \mathcal{A}, \forall E \subset \Omega, \forall \varepsilon > 0$, 由 inf 的定义，可以选取一组 $\{A_i\} \subset \mathcal{A}$, 使得 $E \subset \bigcup_i A_i$, 且

$$\sum_i m(A_i) \leq m^*(E) + \varepsilon.$$

注意到有

$$A \cap E \subset \bigcup_i (A \cap A_i),\ A^c \cap E \subset \bigcup_i (A^c \cap A_i).$$

由 Boole 代数的交封闭，有 $\{A \cap A_i\}, \{A^c \cap A_i\} \subset \mathcal{A}$, 从而由 m^* 的定义，

$$m^*(A \cap E) \leq \sum_i m(A \cap A_i),\ m^*(A^c \cap E) \leq \sum_i m(A^c \cap A_i).$$

从而，由于 $(A \cap E) \cap (A^c \cap E) = \varnothing$, 有

$$\begin{aligned} m^*(A \cap E) + m^*(A^c \cap E) &\leq \sum_i (m(A \cap A_i) + m(A^c \cap A_i)) \\ &= \sum_i m((A \cap A_i) \cup (A^c \cap A_i)) \\ &= \sum_i m(A_i) \\ &\leq m^*(E) + \varepsilon. \end{aligned}$$

由 ε 的任意性，得

$$m^*(A \cap E) + m^*(A^c \cap E) \leq m^*(E).$$

从而 A 满足可测性条件，故 $A \in \mathcal{M}(m^*)$. 最后再由 A 的任意性，即得 $\mathcal{A} \subset \mathcal{M}(m^*)$. □

再验证$\mathcal{M}(m^*)$ 是一个 σ-代数

引理 1.0.32. $\mathcal{M}(m^*)$ 是一个 σ-代数。

证明. (1) $\forall E \subset \Omega$,

$$m^*(\Omega \cap E) + m^*(\Omega^c \cap E) = m^*(E) + m^*(\varnothing) = m^*(E \cup \varnothing) = m^*(E).$$

即 Ω 满足可测性条件，有 $\Omega \in \mathcal{M}(m^*)$.

(2) $\forall A \in \mathcal{M}(m^*)$, 由 $(A^c)^c = A$, 在可测性条件式中交换 A 与 A^c 的位置，即知 A^c 也满足可测性条件。事实上，$\forall E \subset \Omega$,

$$m^*(A^c \cap E) + m^*((A^c)^c \cap E) = m^*(A^c \cap E) + m^*(A \cap E) = m^*(E).$$

故 $A^c \in \mathcal{M}(m^*)$. $\mathcal{M}(m^*)$ 满足补封闭性。

(3) 设 $A, B \in \mathcal{M}(m^*)$, $\forall E \subset \Omega$,

$$\begin{aligned}
m^*(E) &= m^*(A \cap E) + m^*(A^c \cap E) \\
&= m^*(B \cap A \cap E) + m^*(B^c \cap A \cap E) + m^*(A^c \cap E) \\
&= m^*(B \cap A \cap E) + m^*(A \cap (A \cap B)^c \cap E) + m^*(A^c \cap (A \cap B)^c \cap E) \\
&= m^*((A \cap B) \cap E) + m^*((A \cap B)^c \cap E).
\end{aligned}$$

这说明 $A \cap B \in \mathcal{M}(m^*)$. 即 $\mathcal{M}(m^*)$ 对交封闭。再由对补封闭，

$$A \cup B = (A^c \cap B^c)^c \in \mathcal{M}(m^*).$$

从而 $\mathcal{M}(m^*)$ 对有限并封闭。此时知 $\mathcal{M}(m^*)$ 是一个 Boole 代数。

设 $\{A_i\}_{i=1}^{\infty} \subset \mathcal{M}(m^*)$. 令 $B_1 = A_1, B_i = A_i - (A_1 \cup \cdots \cup A_{i-1})$, 由补封闭性与有限并封闭性，得到 $\{B_i\}_{i=1}^{\infty} \subset \mathcal{M}(m^*)$, 且 $\{B_i\}_{i=1}^{\infty}$ 两两不交，记 $\bigcup_i A_i = \bigcup_i B_i = A$.

$\forall E \subset \Omega, \forall n,$

$$\begin{aligned}
m^*(E) &= m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)\right) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)^c\right) \\
&= m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right) \cap B_n\right) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right) \cap B_n^c\right) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)^c\right) \\
&= m^*(E \cap B_n) + m^*\left(E \cap \left(\bigcup_{i=1}^{n-1} B_i\right)\right) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)^c\right) \\
&= m^*(E \cap B_n) + m^*(E \cap B_{n-1}) + m^*\left(E \cap \left(\bigcup_{i=1}^{n-2} B_i\right)\right) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)^c\right) \\
&= \cdots \\
&= \sum_{i=1}^{n} m^*(E \cap B_i) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)^c\right).
\end{aligned}$$

因为

$$\bigcup_{i=1}^{n} B_i \subset \bigcup_{i=1}^{\infty} B_i = A,$$

所以

$$A^c \subset \left(\bigcup_{i=1}^{n} B_i\right)^c.$$

于是

$$\begin{aligned}
m^*(E) &= \sum_{i=1}^{n} m^*(E \cap B_i) + m^*\left(E \cap \left(\bigcup_{i=1}^{n} B_i\right)^c\right) \\
&\geq \sum_{i=1}^{n} m^*(E \cap B_i) + m^*(E \cap A^c).
\end{aligned}$$

令 $n\to\infty$, 由次可加性，

$$\begin{aligned}m^*(E) &\geq \sum_{i=1}^{\infty} m^*(E\cap B_i)+m^*(E\cap A^c)\\ &\geq m^*\left(\bigcup_{i=1}^{\infty}(E\cap B_i)\right)+m^*\left(E\cap A^c\right)\\ &= m^*(E\cap A)+m^*(E\cap A^c).\end{aligned}$$

即 A 满足可测性条件，是 m^*-可测的，即 $A\in\mathcal{M}(m^*)$. 从而 $\mathcal{M}(m^*)$ 对可列并封闭。

综上，$\mathcal{M}(m^*)$ 是一个 σ-代数。 □

最后验证<u>m^* 是一个测度</u>

引理 1.0.33. m^* 是 σ-代数 $\mathcal{M}(m^*)$ 上的一个测度。

证明. (1) 由定义显然有 $m^*\geq 0$, $m^*(\varnothing)=0$.

(2) $\forall A,B\in\mathcal{M}(m^*)$, $A\cap B=\varnothing$, 取 $E=A\cup B$, 对 A 使用可测性条件，

$$m^*(A\cup B)=m^*(A\cap(A\cup B))+m^*(A^c\cap(A\cup B))=m^*(A)+m^*(B).$$

即 m^* 有有限可加性。

$\forall\{A_n\}\subset\mathcal{M}(m^*)$, $\{A_n\}$ 两两不交。由次可加性，单调性与有限可加性，

$$\sum_{n=1}^{\infty}m^*(A_n)\geq m^*\left(\bigcup_{n=1}^{\infty}A_n\right)\geq m^*\left(\bigcup_{n=1}^{N}A_n\right)=\sum_{n=1}^{N}m^*(A_n).$$

令 $N\to\infty$, 得上式的等号都成立，从而有

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

这就证明了 m^* 的可数可加性。

综上，m^* 是 σ-代数 $\mathcal{M}(m^*)$ 上的一个测度。 □

□

定义 1.0.34. 设 $\mathcal{C}$ 是集合 Ω 的子集族。若 $\forall\{A_n\} \subset \mathcal{C}$, 满足 $A_n \nearrow A$ 或 $A_n \searrow A$, 则有 $A \in \mathcal{C}$, 则称 $\mathcal{C}$ 是单调类。

定理 1.0.35 (单调类定理)**.** 设 $\mathcal{A}$ 是集合 Ω 上的 *Boole* 代数，$\mathcal{M}^*$ 是包含 $\mathcal{A}$ 的最小单调类，则 $\mathcal{M}^* = \sigma(\mathcal{A})$.

证明. 由定义，σ-代数是单调类，从而 $\sigma(\mathcal{A}) \supseteq \mathcal{M}^*$. 只要证 $\sigma(\mathcal{A}) \subseteq \mathcal{M}^*$.

(1) $\forall A \in \mathcal{A}$, 考虑

$$\mathcal{M}_A = \{B \in \mathcal{M}^* \mid A \cap B, A - B, B - A \in \mathcal{M}^*\} \subseteq \mathcal{M}^*.$$

由于 $\mathcal{A}$ 是 Boole 代数，从而若 $B \in \mathcal{M}_A$, 则 $A \subset \mathcal{M}_A$.

由于 $\mathcal{M}_A$ 是单调类，$\forall\{B_n\} \subset \mathcal{M}_A$, $B_n \nearrow B$（或 $B_n \searrow B$），则有

- $A \cap B_n \nearrow (\searrow) A \cap B$.
- $A - B_n \searrow (\nearrow) A - B$.
- $B_n - A \nearrow (\searrow) B - A$.

由 $\mathcal{M}^*$ 是单调类，从而 $A \cap B, A - B, B - A \in \mathcal{M}^*$. 于是 $B \in \mathcal{A}$. 这说明 $\mathcal{M}_A$ 是单调类。由 $\mathcal{M}^*$ 的最小性，$\mathcal{M}_A = \mathcal{M}^*$.

(2) $\forall A \in \mathcal{M}^*$, 定义

$$\mathcal{M}_A = \{B \in \mathcal{M}^* \mid A \cap B, A - B, B - A \in \mathcal{M}^*\} \subseteq \mathcal{M}^*.$$

由 (1), 有 $\mathcal{A} \subseteq \mathcal{M}_A$. 类似 (1) 的证明，有 $\mathcal{M}_A$ 是单调类，从而 $\mathcal{M}_A = \mathcal{M}^*$. 于是 $\forall A, B \in \mathcal{M}^*$, 有 $A \cap B, A - B, B - A \in \mathcal{M}^*$, 即 $\mathcal{M}^*$ 是 Boole 代数。从而 $\mathcal{M}^*$ 既是 Boole 代数又是单调类，从而它是 σ-代数（习题）。从而 $\sigma(\mathcal{A}) \subseteq \mathcal{M}^*$.

□

定理 1.0.36 (扩张的唯一性定理)**.** 设 $\mathcal{A}$ 是 Ω 上的 *Boole* 代数，m 是 $\mathcal{A}$ 上的测度，m^*, μ 是 $\sigma(\mathcal{A})$ 上的两个测度，且 $m^*|_{\mathcal{A}} = \mu|_{\mathcal{A}} = m$, 则 $m^* = \mu$.

证明. 考虑 $\mathcal{M} = \{A \in \sigma(\mathcal{A}) \mid \mu(A) = m^*(A)\} \subset \sigma(\mathcal{A})$, 则 $\mathcal{A} \subseteq \mathcal{M}$. 由测度的连续性，若 $\{A_n\} \subset \mathcal{M}$, $A_n \nearrow A$ （或 $A_n \searrow A$），由单调类定理（定理 1.0.35），$A \in \sigma(\mathcal{A})$. 则

$$\mu(A) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} m^*(A_n) = m^*(A),$$

从而 $A \in \mathcal{M}$. 这说明 $\sigma(\mathcal{A}) \subseteq \mathcal{M}$. 从而 $\sigma(\mathcal{A}) = \mathcal{M}$. 即 $\forall A \in \mathcal{M} = \sigma(\mathcal{A})$, $m^*(A) = \mu(A)$. □

回到无穷次投硬币。$\Omega = \{H, T\}^{\mathbb{N}} = \{(x_1, x_2, \cdots) \mid x_i \in \{H, T\}, i = 1, 2, \cdots\}$. 记 A_i 是第 i 次是 H. 定义

$$\sigma(A_1) = \mathcal{F}_1 = \{\varnothing, A_1, A_1^c, \Omega\},$$

$$\sigma(A_1, \cdots, A_n) = \mathcal{F}_n,\ |\mathcal{F}_n| = 2^n.$$

有

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

$$\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{F}_n$$

是 Boole 代数。定义 $\mu : \mathcal{A} \to [0, 1]$. 设 $A \in \mathcal{A}$, 则 $\exists n$, 使得 $A \in \mathcal{F}_n$. 令 $\mu(A) = \mu_n(A)$, 其中 μ_n 就是投掷有限次时古典概型中的测度。

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

对 $m < n$,$\mu_n|_{\mathcal{F}_m} = \mu_m$. μ 符合 Boole 代数上测度的定义：

(1) $\mu(A) \geq 0$.

(2) $\forall A \cap B = \varnothing$, $\mu(A \cup B) = \mu(A) + \mu(B)$.

(3) 若 $A_n \searrow \varnothing$, 则 $\mu(A_n) \to 0$.

验证 (3): 设有 $\{A_n\} \subset \mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{F}_n$, $A_n \searrow \varnothing$. 事实上，对充分大的 n, 有 $A_n = \varnothing$.

引理 1.0.37 (Tychonoff 定理)**.** 任意紧拓扑空间的乘积空间是紧空间。

由引理，$\Omega = \{H, T\}^{\mathbb{N}} = \prod_{i=1}^{\infty} \{H, T\}$ 是紧空间，开集与闭集就是 $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{F}_n$. 从而由 $A_n \searrow \varnothing$, 就有 $\exists N, \forall n \geq N$, $A_n = \varnothing$. 从而我们得到了包含 $\mathcal{A}$ 的 σ-代数 $\sigma(\mathcal{A})$ 上的概率测度 $\mathbb{P} = \mu$, 使得 $\mathbb{P}|_{\mathcal{F}_n} = \mu_n$, 就达成了目标 1.0.22, 解决了问题 1.0.21.

2 独立性

定义 2.0.1. 设 $(\Omega,\mathcal{F},\mathbb{P})$ 是概率空间，$C\in\mathcal{F}$, $\mathbb{P}(C)>0$. $\forall A\in\mathcal{F}$,

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

称为在条件 C 下事件 A 的概率。

注记 2.0.2.

- $\mathbb{P}_C(A)$ 也满足概率条件，$(\Omega,\mathcal{F},\mathbb{P}_C)$ 也是一个概率空间。
- 乘法公式：$\mathbb{P}(A\cap C)=\mathbb{P}(A\mid C)\cdot\mathbb{P}(C)$.
- 若 $\{C_i\}$ 是两两不相交的事件，$\mathbb{P}(C_i)>0$, $A\subset\bigcup_i C_i$, 则有如下的全概率公式：

$$\mathbb{P}(A)=\mathbb{P}\left(A\cap\left(\bigcup_i C_i\right)\right)=\mathbb{P}\left(\bigcup_i(A\cap C_i)\right)=\sum_i\mathbb{P}\left(A\cap C_i\right)=\sum_i\mathbb{P}(A\mid C_i)\cdot\mathbb{P}(C_i).$$

定义 2.0.3 (独立性)**.** $A,B\in\mathcal{F}$, 若

$$\mathbb{P}(A\cap B)=\mathbb{P}(A)\cdot\mathbb{P}(B),$$

则称事件 A,B 独立。若有 $A_1,\cdots,A_n\in\mathcal{F}$, $\forall\{i_1,\cdots,i_k\}\subset\{1,\cdots,n\}$, 都有

$$\mathbb{P}(A_{i_1}\cdots A_{i_k})=\mathbb{P}(A_{i_1})\cdots\mathbb{P}(A_{i_k}),$$

则称事件 $A_1,\cdots,A_n$ 互相独立。

注记 2.0.4. n 个事件独立强于两个事件独立，取 $k=2$ 即得。但反之不一定对：$\Omega=\{a,b,c,d\}$, $A_1=\{a,b\}$, $A_2=\{a,c\}$, $A_3=\{a,d\}$. 它们两两独立但三者不独立。

定义 2.0.5. 设 $\mathcal{A},\mathcal{B}\subset\mathcal{F}$, 称事件集 $\mathcal{A},\mathcal{B}$ 独立，若 $\forall A\in\mathcal{A},B\in\mathcal{B}$, 有 A,B 独立。

注记 2.0.6. 若 A,B 独立，则 $\sigma(A)=\{\varnothing,\Omega,A,A^c\}$ 与 $\sigma(B)=\{\varnothing,\Omega,B,B^c\}$ 事件集独立，从而 A,B^c 独立。

问题 2.0.7. 事件集的独立性能否得出它们生成的 σ-代数的独立性？

定义 2.0.8 (π-系)**.** $\mathcal{A}$ 是一个集族，满足

(1) $\varnothing\in\mathcal{A}$.

(2) 若 $A,B\in\mathcal{A}$, 则 $A\cap B\in\mathcal{A}$.

则称 $\mathcal{A}$ 是一个 π-系。

定义 2.0.9 (λ-系)**.** $\mathcal{A}$ 是一个集族，满足

(1) $\Omega\in\mathcal{A}$.

(2) 若 $A,B\in\mathcal{A}$, $A\subset B$, 则 $B-A\in\mathcal{A}$.

(3) 若 $\{A_n\}\subset\mathcal{A}$, $A_n\nearrow A$, 则 $A\in\mathcal{A}$.

则称 $\mathcal{A}$ 是一个 λ-系。

注记 2.0.10. 任意多个 π-系（λ-系）的交仍然是 π-系（λ-系）。

定理 2.0.11 (π-λ 定理)**.** 若 $\mathcal{P}$ 是一个 π-系，$\mathcal{L}$ 是包含 $\mathcal{P}$ 的 λ-系，则 $\sigma(\mathcal{P})\subset\mathcal{L}$.

证明. 可以不妨设 $\mathcal{L}=\lambda(\mathcal{P})$.

1. 首先证明 $\mathcal{L}$ 是 Boole 代数。由定义，$\varnothing,\Omega\in\mathcal{L}$, 且对补封闭。只要证关于交封闭。$\forall A\in\mathcal{P}$, 定义

$$\mathcal{L}_A=\{B\in\mathcal{L}\mid A\cap B\in\mathcal{L}\}\subset\mathcal{L}.$$

   - 由于 $\mathcal{P}$ 是 π-系，$\forall B\in\mathcal{P}$,$A\cap B\in\mathcal{P}$. 从而 $\mathcal{P}\subset\mathcal{L}_A$.

- $\forall B\in\mathcal{L}_A, B\subset C$. 则 $A\cap B, A\cap C\in\mathcal{L}, A\cap C\subset A\cap C$. 由于 $\mathcal{L}$ 是 λ-系，有 $(A\cap C)-(A\cap B)\in\mathcal{L}$. 又 $A\cap(C-B)=(A\cap C)-(A\cap B)\in\mathcal{L}$, 从而 $C-B\in\mathcal{L}_A$.
- 设 $\{B_n\}\subset\mathcal{L}_A$, $B_n\nearrow B$. 有 $A\cap B_n\in\mathcal{L}$, $A\cap B_n\nearrow A\cap B$. 由于 $\mathcal{L}$ 是 λ-系，$A\cap B\in\mathcal{L}$, 从而 $B\in\mathcal{L}_A$.

这说明 $\mathcal{L}_A\subset\mathcal{L}$ 也是 λ-系。由 $\mathcal{L}=\lambda(\mathcal{P})$ 的极小性，必须有 $\mathcal{L}_A=\mathcal{L}$. 即 $\forall A\in\mathcal{P}, B\in\mathcal{L}$, 有 $A\cap B\in\mathcal{L}$. 从而 $\mathcal{L}$ 对交封闭，是一个 Boole 代数。

2. $\forall\{A_i\}\subset\mathcal{L}$, 令 $B_i=\bigcup_{j=1}^{i}$. 有 $B_i\nearrow\bigcup_{i=1}^{\infty}A_i$. 由于 $\mathcal{L}=\lambda(\mathcal{P})$ 是一个 λ-系，$\bigcup_{i=1}^{\infty}A_i\in\mathcal{L}$.

综上，$\mathcal{L}$ 是一个 σ-代数。最后，由于 $\sigma(\mathcal{P})$ 是包含 $\mathcal{P}$ 的最小 σ-代数，有 $\sigma(\mathcal{P})\subset\mathcal{L}$.

□

推论 2.0.12. 若 $\mathcal{P}$ 是一个 π-系，则 $\lambda(\mathcal{P})=\sigma(\mathcal{P})$.

定理 2.0.13. $(\Omega,\mathcal{F},\mathbb{P})$ 是一个概率空间，$\mathcal{B}\subset\mathcal{F}$ 是一个 π-系，$A\in\mathcal{F}$. 若 A 与 $\mathcal{B}$ 独立。则 A 与 $\sigma(\mathcal{B})$ 独立。

证明. 考虑

$$\mathcal{L}=\{B\in\mathcal{F}\mid\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B)\}.$$

是与 A 独立的事件的集合。

下面证明 $\mathcal{L}$ 是一个 λ-系。

1. 全空间与任何事件独立，即 $\Omega\in\mathcal{L}$.
2. 若 $B,C\in\mathcal{L},B\subset C$.

$$\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(B),\ \mathbb{P}(A\cap C)=\mathbb{P}(A)\mathbb{P}(C).$$

从而

$$\mathbb{P}(A\cap(C-B))=\mathbb{P}((A\cap C)-(A\cap B))=\mathbb{P}(A\cap C)-\mathbb{P}(A\cap B)=\mathbb{P}(A)\mathbb{P}(C)-\mathbb{P}(A)\mathbb{P}(B)=\mathbb{P}(A)\left(\mathbb{P}(C)-\mathbb{P}(B)\right)=\mathbb{P}(A)\mathbb{P}(C-B).$$

即 $C-B\in\mathcal{L}$.

3. 若在 $\mathcal{L}$ 中有 $B_n \nearrow B$.

$$\mathbb{P}(A \cap B_n) = \mathbb{P}(A)\mathbb{P}(B_n), \forall n.$$

则有 $A \cap B_n \nearrow A \cap B$. 由概率的连续性，令 $n \to \infty$, 就有

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

从而 $B \in \mathcal{L}$.

综上，$\mathcal{L}$ 是一个 λ-系。由 π-λ 定理（定理 2.0.11），$\sigma(\mathcal{B}) \subset \mathcal{L}$. 即 A 与 $\sigma(\mathcal{B})$ 独立。 □

推论 2.0.14. $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ 是两个 π-系，$\mathcal{A}$ 与 $\mathcal{B}$ 独立。则 $\sigma(\mathcal{A})$ 与 $\sigma(\mathcal{B})$ 独立。

证明. 习题 □

定义 2.0.15. 设 $\{A_n\}_{n=1}^{\infty} \subset \mathcal{F}$. 令

$$B_n = \bigcup_{i=n}^{\infty} A_i, \ C_n = \bigcap_{i=n}^{\infty} A_i.$$

则记 $B_n \searrow B = \limsup_{n\to\infty} A_n$, $C_n \nearrow C = \liminf_{n\to\infty} A_n$. 则有

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\left(\bigcup_{i=n}^{\infty} A_i\right), \ \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\left(\bigcap_{i=n}^{\infty} A_i\right).$$

$\omega \in \limsup_{n\to\infty} A_n$ 当且仅当 ω 属于无穷多个 A_n, $\omega \in \liminf_{n\to\infty} A_n$ 当且仅当不含 ω 的 A_n 只有有限个。

由

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty}\left(\bigcap_{i=n}^{\infty} A_i\right)\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i=n}^{\infty} A_i\right) \le \liminf_{n\to\infty} \mathbb{P}(A_n),$$

由 de Morgan 律，

$$\left(\limsup_{n\to\infty} A_n\right)^c = \left[\bigcap_{n=1}^{\infty}\left(\bigcup_{i=n}^{\infty} A_i\right)\right]^c = \bigcup_{n=1}^{\infty}\left(\bigcap_{i=n}^{\infty} A_i^c\right) = \liminf_{n\to\infty} A_n^c.$$

从而

$$1-\mathbb{P}\left(\limsup_{n\to\infty}A_n\right)=\mathbb{P}\left(\liminf_{n\to\infty}A_n^c\right)\le\liminf_{n\to\infty}(1-\mathbb{P}(A_n))=1-\limsup_{n\to\infty}\mathbb{P}(A_n).$$

即

$$\limsup_{n\to\infty}\mathbb{P}(A_n)\le\mathbb{P}\left(\limsup_{n\to\infty}A_n\right).$$

定理 2.0.16 (Borel-Cantelli)**.** 设 $\{A_n\}_{n=1}^\infty\subset\mathcal{F}$.

1. 若 $\sum_{n=1}^\infty\mathbb{P}(A_n)<\infty$, 则 $\mathbb{P}\left(\limsup_{n\to\infty}A_n\right)=0$.

2. 若 $\{A_n\}_{n=1}^\infty$ 独立，且 $\sum_{n=1}^\infty\mathbb{P}(A_n)=\infty$, 则 $\mathbb{P}\left(\limsup_{n\to\infty}A_n\right)=1$.

证明. 1. $\forall n$,

$$\mathbb{P}\left(\limsup_{n\to\infty}A_n\right)\le\mathbb{P}\left(\bigcup_{i=n}^\infty A_i\right)\le\sum_{i=n}^\infty\mathbb{P}(A_i).$$

由于 $\sum_{n=1}^\infty\mathbb{P}(A_n)<\infty$ 收敛，从而 $\sum_{i=n}^\infty\mathbb{P}(A_i)\to0$. 令 $n\to\infty$, 得 $\mathbb{P}\left(\limsup_{n\to\infty}A_n\right)\le0$, 从而 $\mathbb{P}\left(\limsup_{n\to\infty}A_n\right)=0$.

2. 只要证 $\mathbb{P}\left((\limsup_{n\to\infty}A_n)^c\right)=0$. 由 de Morgan 律, $(\limsup_{n\to\infty}A_n)^c=\left[\bigcap_{n=1}^\infty\left(\bigcup_{i=n}^\infty A_i\right)\right]^c=\bigcup_{n=1}^\infty\left(\bigcap_{i=n}^\infty A_i^c\right)$, 故只要证 $\mathbb{P}\left(\bigcup_{n=1}^\infty\left(\bigcap_{i=n}^\infty A_i^c\right)\right)=0$. 而这只要证 $\forall n$, $\mathbb{P}\left(\bigcap_{i=n}^\infty A_i^c\right)=0$, 从而由次可加性,

$$\mathbb{P}\left(\bigcup_{n=1}^\infty\left(\bigcap_{i=n}^\infty A_i^c\right)\right)\le\sum_{n=1}^\infty\mathbb{P}\left(\bigcap_{i=n}^\infty A_i^c\right)=\sum_{n=1}^\infty0=0.$$

由于 $\{A_n\}_{n=1}^\infty$ 独立，它们的补也独立（注记 2.0.6）。从而 $\forall m>n$,

$$\mathbb{P}\left(\bigcap_{i=n}^m A_i^c\right)=\prod_{i=n}^m\mathbb{P}\left(A_i^c\right)=\prod_{i=n}^m\left(1-\mathbb{P}\left(A_i\right)\right)\le\mathrm{e}^{-\sum_{i=n}^m\mathbb{P}(A_i)}.$$

其中用到了 $\forall x\ge0$, $1-x\le\mathrm{e}^{-x}$. 由于 $\sum_{i=1}^\infty\mathbb{P}\left(A_i\right)=\infty$, 故当 $m\to\infty$ 时，$\sum_{i=n}^m\mathbb{P}\left(A_i\right)\to\infty$. 此时，由概率的连续性,

$$\mathbb{P}\left(\bigcap_{i=n}^\infty A_i^c\right)=\lim_{m\to\infty}\mathbb{P}\left(\bigcap_{i=n}^m A_i^c\right)\le\lim_{m\to\infty}\mathrm{e}^{-\sum_{i=n}^m\mathbb{P}(A_i)}=\mathrm{e}^{-\infty}=0.$$

即有 $\mathbb{P}\left(\bigcap_{i=n}^{\infty} A_i^c\right) = 0$.

$\square$

推论 2.0.17 (Borel 0-1 律)**.** 设 $\{A_n\}_{n=1}^{\infty} \subset \mathcal{F}$. 则 $\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) \in \{0,1\}$.

注记 2.0.18. 在概率论中，有很多某事件的概率是 0 或 1 的定理，这说明该事件要么必然发生，要么不可能发生，称为 0-1 律。

应用 2.0.19. 回到无穷次投硬币的情况。$\Omega = \{H,T\}^{\mathbb{N}}$. A_n 是第 n 次为 T. $\mathbb{P}(A_n) = \frac{1}{2}$. $\{A_n\}_{n=1}^{\infty}$ 独立。由 *Borel-Cantelli* 定理（定理 *2.0.16*），$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1$. 这说明投无穷次硬币出现 T 的概率为 1, 它必然发生。

令 $\varphi : (0,1] \to \Omega = \{H,T\}^{\mathbb{N}}$. $\forall x \in (0,1]$, 将 x 写成二进制小数：$x = 0.a_1a_2a_3\cdots$. 该表示不唯一，因为有限小数也可以写成以 1 循环小数。例如

$$0.10011 = 0.10010\dot{1}$$

规定此时用以 1 循环来表示。

$$x \mapsto \varphi(x) = (a_1, a_2, a_3, \cdots)$$

将 0 换成 H, 1 换成 T. 则 φ 是单射。$\varphi\left((0,1]\right) = \limsup_{n\to\infty} A_n$.

$$\varphi^{-1}(A_1) = \left(\frac{1}{2}, 1\right],\ \varphi^{-1}(A_2) = \left(\frac{1}{4}, \frac{1}{2}\right] \cup \left(\frac{3}{4}, 1\right], \cdots, \varphi^{-1}(A_n) = \bigcup_{j=1}^{2^{n-1}} \left(\frac{2j-1}{2^n}, \frac{2j}{2^n}\right], \cdots$$

设 μ 是 $(0,1]$ 上的 *Lebesgue* 测度，$\mathbb{P}(A_n) = \frac{1}{2} = \mu\left(\varphi^{-1}(A_n)\right)$, 即在 Ω 上取均匀概率就是 $(0,1]$ 上的 *Lebesgue* 测度。

$$\Omega = \{H,T\}^{\mathbb{N}} \supset \Omega' = \limsup_{n\to\infty} \rightleftarrows (0,1].$$

$$\mathbb{P}(\Omega - \Omega') = 0,\ \mathbb{P}(\Omega') = 1.$$

$$\mathbb{P}(A_n \cap \Omega') = \mu(\varphi^{-1}(A_n)).$$

$$\mathcal{F} = \sigma(A_1, A_2, \cdots) \leftrightarrow \textit{Borel} \text{ 代数}.$$

即 $\{A_n\}_{n=1}^{\infty}$ 生成的 σ-代数 $\mathcal{F} = \sigma(A_1, A_2, \cdots)$ 就是 $(0,1]$ 上的 *Borel* 代数。

定理 2.0.20 (Kolmogorov 0-1 律)**.** 设 $\{A_n\}_{n=1}^{\infty}\subset\mathcal{F}$. 令 $\mathcal{T}_n=\sigma(A_{n+1},A_{n+2},\cdots)$. 有 $\mathcal{T}_1\supset\mathcal{T}_2\supset\cdots$. 记 $\mathcal{T}=\bigcap_{n=1}^{\infty}\mathcal{T}_n$. $\mathcal{T}$ 中的事件称为尾事件。若 $\{A_n\}_{n=1}^{\infty}$ 独立。则 $\forall T\in\mathcal{T}$, 有 $\mathbb{P}(T)\in\{0,1\}$.

证明. 记 $\mathcal{F}_n=\sigma(A_1,\cdots,A_n)$, $\mathcal{F}=\sigma(A_1,A_2,\cdots)$, $\mathcal{B}_{n,k}=\sigma(A_{n+1},\cdots.A_{n+k})$.

$\forall n$, $\bigcup_{k=1}^{\infty}\mathcal{B}_{n,k}$ 是 Boole 代数，是 π-系，$\mathcal{T}_n$ 是 $\bigcup_{k=1}^{\infty}\mathcal{B}_{n,k}$ 生成的 σ-代数。由条件，$\mathcal{F}_n$ 与 $\mathcal{B}_{n,k}$ 独立，从而与 $\bigcup_{k=1}^{\infty}\mathcal{B}_{n,k}$ 独立。由 π-λ 定理（定理 2.0.11），$\mathcal{F}_n$ 与 $\sigma(\bigcup_{k=1}^{\infty}\mathcal{B}_{n,k})=\mathcal{T}_n$ 独立。从而 $\mathcal{F}_n$ 与 $\mathcal{T}=\bigcap_{k=1}^{\infty}\mathcal{T}_n$ 独立，从而 $\bigcup_{k=1}^{\infty}\mathcal{F}_n$ 与 $\mathcal{T}=\bigcap_{k=1}^{\infty}\mathcal{T}_n$ 独立。由于 $\bigcup_{k=1}^{\infty}\mathcal{F}_n$ 是 Boole 代数，是 π-系，从而 $\mathcal{F}=\sigma(\bigcup_{k=1}^{\infty}\mathcal{F}_n)$ 与 $\mathcal{T}$ 独立。

特别地，$\forall T\in\mathcal{T}$, T 与 T 独立，

$$\mathbb{P}(T)=\mathbb{P}(T\cap T)=\mathbb{P}(T)\mathbb{P}(T),$$

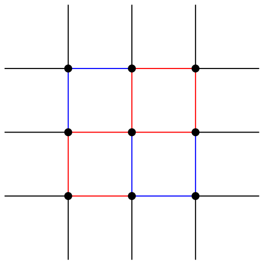
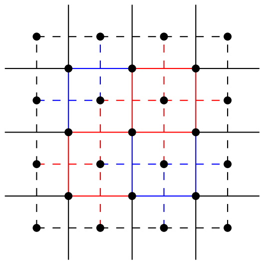
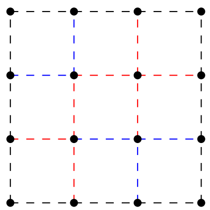
从而 $\mathbb{P}(T)\in\{0,1\}$. □

注记 2.0.21. 有另一种形式的 *Kolmogorov* 0-1 律: $\{\mathcal{A}_n\}_{n=1}^{\infty}$ 是一列独立的 σ-代数, $\mathcal{T}=\bigcap_{n=1}^{\infty}\sigma(\mathcal{A}_{n+1},\mathcal{A}_{n+2},\cdots)$, 则 $\forall T\in\mathcal{T}$, 有 $\mathbb{P}(T)\in\{0,1\}$.

证明是完全类似的：把 A 换成 $\mathcal{A}$ 即可。

应用 2.0.22 (边渗流模型)**.** 考虑平面上的格 $L=\mathbb{Z}^2$. 每一条边随机地，相互独立地，以概率 p 染成红色，以概率 $1-p$ 染成蓝色。L 的边是可数的。记红边构成的图是 G_r, 蓝边构成的图是 G_b. 我们称一个连通分支是巨连通分支，如果它有无穷多个顶点。记事件 $T_{r,p}$ 是 G_r 中有一个巨连通分支，它是一个尾事件：改变有限条边不会影响是否含有巨连通分支的情况，$T_{r,p}\in\bigcap_{n=1}^{\infty}\sigma(A_{n+1},A_{n+2},\cdots)$. 由 *Kolmogorov* 0-1 律（定理 *2.0.20*），$\mathbb{P}(T_{r,p})\in\{0,1\}$.

具体地，考虑 L 的对偶图 L^*. 则 L 上的边渗流模型与 L^* 上的边渗流模型有一一对应：

$L=$  $\longrightarrow$  $\longrightarrow L^*=$ 

断言：若 L 上无红色巨连通分支，则 L^* 上有蓝色巨连通分支。即 $T_{r,p}^c \subset T_{b,1-p}$. 从而有

$$1 - \mathbb{P}(T_{r,p}) = \mathbb{P}(T_{r,p}^c) \leq \mathbb{P}(T_{b,1-p}) = \mathbb{P}(T_{r,1-p}).$$

这说明

1. 若 $\mathbb{P}(T_{r,p}) = 0$, 则 $\mathbb{P}(T_{b,1-p}) \geq 1 - 0 = 1$, 即 $\mathbb{P}(T_{b,1-p}) = 1$.
2. 当 $p = \frac{1}{2}$ 时，有 $1 - \mathbb{P}(T_{r,\frac{1}{2}}) \leq \mathbb{P}(T_{r,\frac{1}{2}})$, 即 $\mathbb{P}(T_{r,\frac{1}{2}}) \geq \frac{1}{2}$. 由 *Kolmogorov* 0-1 律（定理 *2.0.20*），只能有 $\mathbb{P}(T_{r,\frac{1}{2}}) = 1$. 且直观上有 $\mathbb{P}(T_{r,p})$ 关于 p 递增，从而 $\forall p \geq \frac{1}{2}$, $\mathbb{P}(T_{r,p}) = 1$.

3 随机变量

定义 3.0.1. 设两个可测空间 $(\Omega,\mathcal{F})$ 与 $(\Gamma,\mathcal{G})$, 映射 $f:\Omega\to\Gamma$ 称为可测映射，若 $\forall A\in\mathcal{G}$, 有 $f^{-1}(A)\in\mathcal{F}$. 特别地，若 $(\Gamma,\mathcal{G})=(\mathbb{R},Borel$ 可测集$\mathcal{B}(\mathbb{R}))$, 则称 f 是 Ω 上的可测函数。

性质 3.0.2.

- 可测映射的复合还是可测映射。
- $\{f^{-1}(A)\mid A\in\mathcal{G}\}$ 是 $\mathcal{F}$ 的 σ-子代数：
  - $f^{-1}(\varnothing)=\varnothing$, $f^{-1}(\Gamma)=\Omega$.
  - $f^{-1}(\Gamma-A)=\Omega-f^{-1}(A)$.
  - $f^{-1}(\bigcup_{n=1}^{\infty}A_n)=\bigcup_{n=1}^{\infty}f^{-1}(A)$.

  记为 $f^{-1}\mathcal{G}$, 称为 f 在 Γ 上的拉回。它是使 f 可测的最小 σ-代数。
- $f_*\mathcal{F}=\{A\subset\Gamma\mid f^{-1}(A)\in\mathcal{F}\}$ 是 σ-代数，称为 f 在 Ω 上的推出。它是使 f 可测的最大 σ-代数。
- $f:(\Omega,\mathcal{F})\to(\Gamma,\mathcal{G})$ 可测，当且仅当 $f^{-1}\mathcal{G}\subset\mathcal{F}$, 当且仅当 $f_*\mathcal{F}\supset\mathcal{G}$.
- $f:(\Omega,\mathcal{F})\to(\Gamma,\mathcal{G})$. 如果 $\mathcal{G}=\sigma(\{A_\alpha\mid\alpha\in\Lambda\})$, 则 f 可测当且仅当 $f^{-1}(A_\alpha)\in\mathcal{F}$, $\forall\alpha\in\Lambda$.
- 可测函数的四则运算保持可测性。

定义 3.0.3. 可测函数的极限：

- 点点收敛：$f_n\xrightarrow{p.w.}f$: $\forall\omega\in\Omega$, $f_n(\omega)\to f(\omega)$.
- 依测度收敛：$f_n\xrightarrow{m}f$: $\forall\sigma>0$, $m(\{\omega\mid|f_n-f|\geq\sigma\})\to 0$.

定理 3.0.4. 假设 $\mu(\Omega)<\infty$. 若 $f_n\xrightarrow{p.w.}f$, 则 $f_n\xrightarrow{m}f$; 反之，若 $f_n\xrightarrow{m}f$, 则存在子列 $f_{n_k}\xrightarrow{p.w.}f$.

定义 3.0.5. 若一个可测函数只取有限个值，则称它是简单函数：$\Omega = \bigsqcup_{i=1}^n A_i$, $A_i \in \mathcal{F}$, $f = \sum_{i=1}^n a_i \chi_{A_i}$.

定理 3.0.6. $(\Omega, \mathcal{F})$ 上的函数 f 是可测的，当且仅当存在一列 $(\Omega, \mathcal{F})$ 上的简单函数 $\{f_n\}_{n=1}^\infty$, 使得 $f_n \xrightarrow{p.w.} f$.

定义 3.0.7. $(\Omega, \mathcal{F})$ 是一个概率空间，Ω 上的随机变量 X 就是可测函数 $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. 由 X 的拉回确定了一个 σ-子代数 $X^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{F}$, 记为 $\sigma(X)$.

例 3.0.8. $A \subset \Omega$ 是一个事件，特征函数 χ_A 是一个随机变量。$\mathbb{P}(\chi_A = 1) = \mathbb{P}(A)$, $\mathbb{P}(\chi_A = 0) = \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$. $\sigma(\chi_A) = \{\varnothing, A, A^c, \Omega\}$.

定义 3.0.9. $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 是一个随机变量。由 X 可以在 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上定义一个测度 μ_X: $\forall A \in \mathcal{B}(\mathbb{R})$,

$$\mu_X(A) = \mathbb{P}(x \in A) = \mathbb{P}(X^{-1}(A)).$$

容易验证 $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$ 是一个概率空间。μ_X 称为 X 的分布。

$$\Phi_X(x) = \mu_X((\infty, x]) = \mathbb{P}(X \leq x)$$

称为 X 的分布函数。

性质 3.0.10. 分布函数的性质：

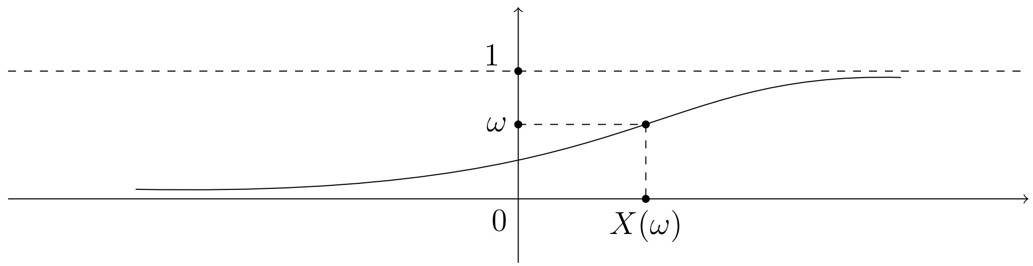
- 单调递增。
- $\lim_{x\to-\infty} \Phi_X(x) = 0$, $\lim_{x\to+\infty} \Phi_X(x) = 1$.
- 右连续性：
$$\lim_{x\to x_0^+} \Phi_X(x) = \Phi_X(x_0).$$

定理 3.0.11. 若 $\mathbb{R}$ 上的函数 $F(x)$ 满足性质 *3.0.10*的全部三条，则存在 $((0,1], \mathscr{L} = \textit{Lebesgue}\ 可测集, \mathrm{d}\omega))$ 上的随机变量 X, 使得 $\Phi_X = F$.

证明. 简单情形：F 连续且严格单调递增。此时可以将 F 看成同胚 $\mathbb{R} \to (0,1)$. 定义

$$\begin{aligned} X : (0,1) &\to \mathbb{R} \\ \omega &\mapsto F^{-1}(\omega). \end{aligned}$$

则 $X(\omega) \le x$ 当且仅当 $\omega \le F(x)$.



从而

$$\Phi_X(x) = \mathbb{P}(X(\omega) \le x) = \mathrm{d}\,\omega(0, F(x)] = F(x).$$

而 $X(1)$ 可以任意定义，不影响分布。

一般情形：$\forall \omega \in (0,1)$, 定义广义逆元

$$F^{-1}(\omega) = \inf\{x \mid F(x) \ge \omega\}.$$

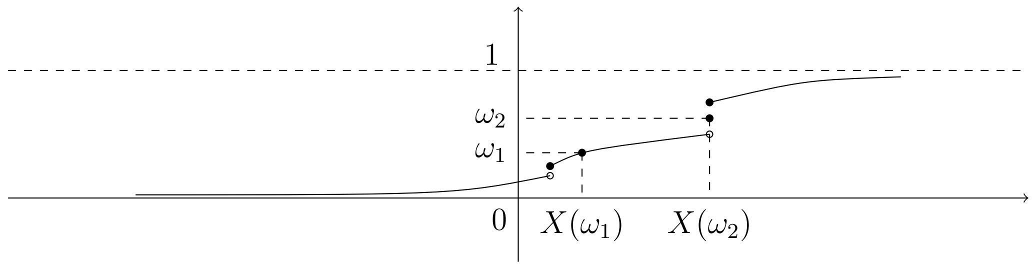
由 F 的右连续性，

$$F^{-1}(\omega) = \inf\{x \mid F(x) \ge \omega\} = \min\{x \mid F(x) \ge \omega\}.$$

定义

$$\begin{aligned} X : (0,1) &\to \mathbb{R} \\ \omega &\mapsto F^{-1}(\omega). \end{aligned}$$

则仍有 $X(\omega) \le x$ 当且仅当 $\omega \le F(x)$.



而 $X(\omega) \le x$ 能得到 $F(X(\omega)) \le F(x)$. 由 $X(\omega) = \min\{x \mid F(x) \ge \omega\}$ 的定义，$X(\omega) \in \{x \mid F(x) \ge \omega\}$. 从而有 $\omega \le F(X(\omega)) \le F(x)$. 反之，若 $F(x) \ge \omega$, 仍由定义，$x \in \{x \mid F(x) \ge \omega\}$. 故 $x \ge \min\{x \mid F(x) \ge \omega\} = X(\omega)$. 从而

$$\Phi_X(x) = \mathbb{P}(X(\omega) \le x) = \mathrm{d}\,\omega(0, F(x)] = F(x).$$

总结：若 F 满足性质 3.0.10的全部三条，在 $((0,1], \mathscr{L}, \mathrm{d}\,\omega))$ 上，X 是均匀分布的随机变量。$A \in \mathscr{L}$, $\mathbb{P}(x \in A) = \mathrm{d}\,\omega(A)$.

$$\mathbb{P}(X \le x) = \begin{cases} 0, & x \le 0; \\ x, & x \in (0,1); \\ 1, & x \ge 1. \end{cases}$$

$F^{-1}(X) \le x$ 当且仅当 $X \le F(x)$, 从而 $\Phi_{F^{-1}(X)}(x) = \mathbb{P}(F^{-1}(X) \le x) = \mathbb{P}(X \le F(x)) = F(x)$. 故 $F^{-1}(X)$ 的分布函数就是 F: $\Phi_{F^{-1}(X)} = F$. □

例 3.0.12. $\Omega = \{H, T\}^{\mathbb{N}}$, $\Omega' \subset \Omega$.

$$\Omega' \leftrightarrow ((0,1], \mathscr{L}, \mathrm{d}\,\omega)$$

$$A_n \leftrightarrow B_n = \bigcup_{i=1}^{2^{n-1}} \left(\frac{2i-1}{2^n}, \frac{2i}{2^n} \right].$$

定义 $R_n=\chi_{B_n}$. 记

$$R(t)=\begin{cases}1, & t\in\left(\frac{1}{2},1\right)\cup\{0\};\\ 0, & t\in\left(0,\frac{1}{2}\right].\end{cases}$$

称为 *Rademacher* 函数。$\omega\in(0,1]$.

$$\chi_{B_1}(\omega)=R(\omega),\ \chi_{B_n}=R(2^{n-1}\omega).$$

定义随机变量

$$X=\sum_{i=1}^{\infty}\frac{\chi_{A_n}}{2^i}.$$

它是 $(0,1]$ 上的均匀分布。

$$\mathbb{P}\left(x\in\left(\frac{i}{2^n},\frac{i+1}{2^n}\right]\right)=\frac{1}{2^n}.$$

$\left(\frac{l_j}{2^{n_j}},\frac{k_j}{2^{m_j}}\right]\nearrow(a,b]$, $\frac{l_j}{2^{n_j}}\searrow a$, $\frac{k_j}{2^{m_j}}\nearrow b$. 例如：$X\in\left(\frac{3}{8},\frac{4}{8}\right]$ 当且仅当 $x\in\overline{A_1}\cap A_2\cap A_3$.

任取分布函数 F,

$$\Omega\xrightarrow{X}(0,1]\xrightarrow{F^{-1}}\mathbb{R}.$$

F^{-1} 服从 $(0,1]$ 上的均匀分布，$F^{-1}(X)$ 服从 F 分布。$F^{-1}(X)$ 看作是 ω 上的随机变量。

定义 3.0.13. 称两个随机变量 X,Y 依分布相等，若从而 $\mu_X=\mu_Y$ 能得到 $\Phi_X=\Phi_Y$. 记为 $X\stackrel{d}{=}Y$.

注记 3.0.14. 依分布相等不要求 X,Y 定义在同一个概率空间上。

定义 3.0.15. $X,\{X_n\}_{n=1}^{\infty}$ 是随机变量。称 $\{X_n\}_{n=1}^{\infty}$ 依分布收敛到 X, 若对任意 Φ_X 的连续点 a, 有 $\Phi_{X_n}(a)\to\Phi_X(a)$. 记为 $X_n\xrightarrow{d}X$.

定理 3.0.16. 若 $X_n\xrightarrow{m}X$, 则 $X_n\xrightarrow{d}X$.

证明. $\forall\varepsilon>0$,

$$\mathbb{P}(X_n\le a)\le\mathbb{P}(X\le a+\varepsilon)+\mathbb{P}(|X_n-X|>\varepsilon).$$

$$\mathbb{P}(X_n \le a-\varepsilon) \le \mathbb{P}(X \le a) + \mathbb{P}(|X_n - X| > \varepsilon).$$

$$\Phi_X(a-\varepsilon) - \mathbb{P}(|X_n - X| > \varepsilon) \le \Phi_{X_n}(a) \le \Phi_X(a+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon).$$

令 $n \to \infty$ 有

$$\Phi_X(a-\varepsilon) \le \liminf_{n\to\infty} \Phi_{X_n}(a) \le \limsup_{n\to\infty} \Phi_{X_n}(a) \le \Phi_X(a+\varepsilon).$$

令 $\varepsilon \to 0^+$, 在 Φ_X 的连续点 a 处，$\Phi_X(a) = \Phi_X(a^-) = \Phi_X(a^+)$, 故上式等号同时成立，即有

$$\lim_{n\to\infty} \Phi_{X_n}(a) = \Phi_X(a).$$

即 $X_n \xrightarrow{d} X$. □

应用 3.0.17. 回到上例。

$$X = \sum_{i=1}^{\infty} \frac{\chi_{A_n}}{2^i}.$$

选取有限子列 $n_1 < n_2 < \cdots < n_k < \cdots$.

$$Y_k = \sum_{i=1}^{k} \frac{\chi_{A_{n_i}}}{2^i}.$$

将 $A_{n_1}, A_{n_2}, \cdots, A_{n_k}$ 的投掷结果重排到第 $a, 2, \cdots, k$ 位上看。

$$\Phi_{Y_k}(x) = \begin{cases} 0, & x \le 0; \\ 2^{-k}\left(1 + \lfloor 2^k x \rfloor\right), & 0 < x \le 1; \\ 1, & x > 1. \end{cases}$$

$Y_k \to Y$ 服从 $(0,1]$ 上的均匀分布。

定义 3.0.18. 称随机变量 X 是（绝对）连续的，若它的分布函数 Φ_X 是（绝对）连续的。

我们从实分析中知道，绝对连续函数几乎处处可以求导。于是有定义

定义 3.0.19. 对一个绝对连续型随机变量 X，存在几乎处处连续函数 f，使得

$$\Phi_X(x)=\int_{-\infty}^{x} f(x)\,\mathrm{d}\,x.$$

$f(x)$ 称为 X 的密度函数。

由于 $\mathbb{P}(X=x)=\Phi_X(x)-\Phi_X(x^-)=f(x)$ 几乎处处成立, 所以 $D=\{x\mid \mathbb{P}(X=x)>0\}$ 至多是可数集。于是有定义:

定义 3.0.20. $\forall x\in\mathbb{R}$,

$$\Phi_X(x)=\sum_{y\le x,y\in D}\mathbb{P}(X=y)$$

称为离散型随机变量。

例 3.0.21. $X\in\{0,1\}$, $\mathbb{P}(X=0)=p$, $\mathbb{P}(X=1)=1-p$. 称为 *Bernoulli* 随机变量。

定义 3.0.22. 设 $X_1,X_2,\cdots,X_n$ 是 n 个随机变量。$X=(X_1,X_2,\cdots,X_n)$ 称为一个随机向量。

$$X:(\Omega,\mathcal{D},\mathbb{P})\to(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$$

是可测映射，成为联合分布。定义

$$\Phi_X(x_1,x_2,\cdots,x_n)=\mathbb{P}(X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n)$$

称为 X 的联合分布函数。

性质 3.0.23. 对联合分布 Φ_X,

1. 对每个变量连续。

2. $\Phi_X(\infty,\infty,\cdots,\infty)=1$; 只要某个 $x_i=-\infty$, 就有 $\Phi_X(x_1,x_2,\cdots,x_n)=0$.

3. 右连续性：$y_1 \to x_1^+, y_1 \to x_1^+, \cdots, y_n \to x_n^+$，则 $\Phi_X(y_1, y_2, \cdots, y_n) \to \Phi_X(x_1, x_2, \cdots, x_n)$.

4. 对任意 $a_1 < b_1, a_2 < b_2, \cdots, a_n < b_n$，要有 $\mathrm{D}^1_{a_1,b_1} \cdots \mathrm{D}^{n-1}_{a_{n-1},b_{n-1}} \mathrm{D}^n_{a_n,b_n} \Phi_X \geq 0$. 其中 $\mathrm{D}^i_{a_i,b_i}$ 是差分：

$$\mathrm{D}^i_{a_i,b_i} \Phi_X = \Phi_X(x_1, \cdots, x_{i-1}, b_i, x_{i+1}, \cdots, x_n) = \Phi_X(x_1, \cdots, x_{i-1}, a_i, x_{i+1}, \cdots, x_n).$$

注记 3.0.24.

$$\mathrm{D}^1_{a_1,b_1} \cdots \mathrm{D}^{n-1}_{a_{n-1},b_{n-1}} \mathrm{D}^n_{a_n,b_n} \Phi_X = \mathbb{P}(X_1 \in (a_1, b_1], X_2 \in (a_2, b_2], \cdots, X_n \in (a_n, b_n]).$$

定理 3.0.25. 若 $\mathbb{R}^n$ 上的函数 Φ 满足性质 *3.0.23*的全部四条，则存在 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathrm{d}\,\omega))$ 上的随机变量 X，使得 $\Phi_X = \Phi$. 事实上，记 X_i 是 X 在第 i 个分量上的投影，则 $X = (X_1, \cdots, X_n)$ 即为所求。

证明. $\mathcal{B}(\mathbb{R}^n)$ 是由 $\prod_{i=1}^n (a_i, b_i]$ 生成的 σ-代数。由长方体生成的 Boole 代数，是所有由有限个这样的长方体的并构成的集合。定义 $m : \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}^+$: $\forall A = \prod_{i=1}^n (a_i, b_i]$,

$$m(A) = m\left(\prod_{i=1}^n (a_i, b_i]\right) = \mathrm{D}^1_{a_1,b_1} \cdots \mathrm{D}^{n-1}_{a_{n-1},b_{n-1}} \mathrm{D}^n_{a_n,b_n} \Phi_X.$$

由性质 4, $m(A) \geq 0$. 对长方体不同的细分（将长方体分解成不交的小长方体有许多种方式）：$A = A_1 \sqcup \cdots A_k$, 要验证定义合理：$m(A) = m(A_1) + \cdots + m(A_k)$.

要验证 m 是 $\mathcal{B}(\mathbb{R}^n)$ 上的测度：若 $A_n \searrow \varnothing$, 则 $m(A_n) \to 0$. 由测度扩张定理，在 $\mathcal{B}(\mathbb{R}^n)$ 上的测度 $m^* = \mathbb{P}$.

$$\mathbb{P}(X_1 \leq a_1, X_2 \leq a_2, \cdots, X_n \leq a_n) = m\left(\prod_{i=1}^n (-\infty, a_i]\right) = \Phi_X(a_1, a_2, \cdots, a_n).$$

从而 $\Phi_X = \Phi$. □

有时候我们要研究定义在两个不同的空间上的随即变量 X, Y。此时我们无法讨论它们的独立性。我们希望找到 $\mathbb{R}$ 上的两个随即变量 $\widehat{X}, \widehat{Y}$, 使得 $\Phi_{\widehat{X}} = \Phi_X$, $\Phi_{\widehat{Y}} = \Phi_Y$, 这样我们就可以通过研究 $\widehat{X}, \widehat{Y}$ 的独立性得到 X, Y 的独立性。

定理 3.0.26. 设 $\{F_n\}_{n=1}^\infty$ 是一列分布函数。则在 $((0,1], \mathcal{B}((0,1]), \mathrm{d}\,\omega)$ 上，存在独立的随机变量序列 $X_1, X_2, \cdots$，满足 $\Phi_{X_n} = F_n$, $n = 1, 2, \cdots$.

证明. 把自然数集 $\mathbb{N}$ 做分化：令 $N_j = \{(2j-1)\cdot 2^k \mid k=1,2,\cdots\}$, 则 $\mathbb{N} = N_1 \sqcup N_2 \sqcup \cdots$. 令

$$Y_j = \sum_{k=1}^{\infty} \frac{\chi_{A_{(2j-1)2^k}}}{2^k},$$

则 $Y_1, Y_2, \cdots$ 相互独立，都是 $(0,1]$ 上的均匀分布。令 $X_n = F_n^{-1}(Y_n)$, 则 $X_1, X_2, \cdots$ 也是相互独立的，且有 $\Phi_{X_n} = F_n$. □

注记 3.0.27. 证明的灵感来自于无穷次投硬币的模型。

例 3.0.28 (二项分布)**.** 设 $X_1, \cdots, X_n$ 是独立同分布的 *Bernoulli* 随机变量，$\mathbb{P}(X_n = 0) = 1-p$, $\mathbb{P}(X_n = 1) = p$, $n = 1, \cdots, n$. 令 $X = X_1 + \cdots + X_n$. $\forall 0 \le k \le n$,

$$\mathbb{P}(X=k) = \mathbb{P}\left(x_1, \cdots, X_n \text{ 中恰好有 } k \text{ 个 } 1, n-k \text{ 个 } 0\right) = \sum_{1\le i_1 \le \cdots \le i_k \le n} \mathbb{P}\left(X_{i_1} = \cdots = X_{i_k} = 1, X_j = 0, j \ne i_1, \cdots, i_k\right) = C_n^k \cdot p^k \cdot (1-p)^{n-k}.$$

称 X 服从二项分布，记为 $X \sim B(n,p)$.

例 3.0.29 (Poisson 分布)**.** 设 $\lambda > 0$ 是参数，定义随机变量 X:

$$\mathbb{P}(X=k) = \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!},\ k = 0,1,2,\cdots.$$

则称 X 服从 *Poisson* 分布，记为 $X \sim P(\lambda)$.

若两个独立的 *Poisson* 分布 $X = P(\lambda_1), Y = P(\lambda_2)$. 则 $Z = X + Y \sim P(\lambda_1 + \lambda_2)$:

$$\begin{aligned}
\mathbb{P}(Z=k) &= \sum_{i=0}^{k} \mathbb{P}(X=i, Y=k=i) = \sum_{i=0}^{k} \mathbb{P}(X=i)\cdot \mathbb{P}(Y=k-i) \\
&= \sum_{i=0}^{k} \mathrm{e}^{-\lambda_1} \cdot \frac{{\lambda_1}^i}{i!} \cdot \mathrm{e}^{-\lambda_2} \cdot \frac{{\lambda_1}^{k-i}}{(k-i)!} = \frac{\mathrm{e}^{-\lambda_1-\lambda_2}}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} {\lambda_1}^i {\lambda_1}^{k-i} = \frac{\mathrm{e}^{-\lambda_1-\lambda_2}}{k!} \sum_{i=0}^{k} C_k^i {\lambda_1}^i {\lambda_1}^{k-i} = \mathrm{e}^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^k}{k!}.
\end{aligned}$$

例 3.0.30 (Bernoulli 分布与 Poisson 分布的比较)**.** 设 $Y \sim P(\lambda)$.

$$F_\lambda(x) = \mathbb{P}(Y \leq x) = \mathrm{e}^{-\lambda} \sum_{i \leq x} \frac{\lambda^i}{i!}$$

是 Y 的分布函数。设 $X \sim (0,1]$ 是均匀分布。

$$Y(\omega) = F_\lambda^{-1}(X)(\omega) = \begin{cases} 0, & \omega \leq \mathrm{e}^{-\lambda} \\ 1, & \omega \in [\mathrm{e}^{-\lambda}, \mathrm{e}^{-\lambda}(1+\lambda)) \\ 2, & \omega \in [\mathrm{e}^{-\lambda}(1+\lambda), \mathrm{e}^{-\lambda}(2+\lambda)) \\ \cdots . \end{cases}$$

令 $G_\lambda = \chi_{[\mathrm{e}^{-\lambda}, \mathrm{e}^{-\lambda}+\lambda)}$, 令 $Z = G_\lambda(X)$. 则 $Z \sim B(1,\lambda)$. 此时,

$$\mathbb{P}(Y \neq Z) = \mathbb{P}\left(\omega \in [\mathrm{e}^{-\lambda}(1+\lambda), 1]\right) = 1 - \mathrm{e}^{-\lambda} - \lambda\,\mathrm{e}^{-\lambda} \leq \lambda^2.$$

一般地, 设 $T_1, \cdots, T_n \sim P\left(\frac{\lambda}{n}\right)$ 互相独立, 则 $T = T_1 + \cdots + T_n \sim P(\lambda)$. 设 $W_1, \cdots, W_n \sim B\left(1, \frac{\lambda}{n}\right)$ 互相独立, 则 $W = W_1 + \cdots + W_n \sim B\left(n, \frac{\lambda}{n}\right)$. 设 $X_1, \cdots, X_n$ 是互相独立的 $(0,1]$ 上的均匀分布。

$$T = T_1 + \cdots + T_n = F_{\frac{\lambda}{n}}^{-1}(X_1) + \cdots + F_{\frac{\lambda}{n}}^{-1}(X_n) \sim P(\lambda).$$

$$W = W_1 + \cdots + W_n = G_{\frac{\lambda}{n}}(X_1) + \cdots + G_{\frac{\lambda}{n}}(X_n) \sim B\left(n, \frac{\lambda}{n}\right).$$

由于 $\{T \neq W\} \subset \bigcup_{i=1}^n \{T_i \neq W_i\}$,

$$\mathbb{P}(T \neq W) = \mathbb{P}(\exists i, T_i \neq W_i) \leq \sum_{i=1}^n \mathbb{P}(T_i \neq W_i) \leq n \cdot \left(\frac{\lambda}{n}\right)^2 = \frac{\lambda^2}{n}.$$

从而

$$|\mathbb{P}(T \leq x) - \mathbb{P}(W \leq x)| = |\Phi_T(x) - \Phi_W(x)| \leq \mathbb{P}(T \neq W) \leq \frac{\lambda^2}{n}.$$

将 W 记为 S_n, 是 n 个独立服从 $B(1, \frac{\lambda}{n})$ 的独立随机变量的和, 就有 $\forall x$, $\Phi_{S_n}(x) \to \Phi_T(x)$, 即 $S_n \xrightarrow{d} T$.

注记 3.0.31. 我们用 $B(n, \frac{\lambda}{n})$ 逼近了 $P(\lambda)$. 这种方法成为耦合。

定义 3.0.32. 设 X, Y 是两个随机变量，若存在某个概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上存在随机变量 $\widehat{X}, \widehat{Y}$, 使得 $\Phi_{\widehat{X}} = \Phi_X, \Phi_{\widehat{Y}} = \Phi_Y$. 则称 $(\widehat{X}, \widehat{Y})$ 是 (X, Y) 的耦合。

定义 3.0.33. 设 X, Y 是两个随机变量，称 X 随机小于 Y, 若存在 (X, Y) 的耦合 $(\widehat{X}, \widehat{Y})$, 使得 $\mathbb{P}((\widehat{X} \leq \widehat{Y})) = 1$, 记为 $X \preceq Y$.

定理 3.0.34. $X \preceq Y$, 当且仅当 $\forall x$, $\Phi_Y(x) \leq \Phi_X(x)$.