XI'AN JIAOTONG UNIVERSITY

第一章：随机事件与概率

1.2 概率的定义

赵俊舟

junzhou.zhao@xjtu.edu.cn

2025 年 2 月 20 日

目录

目录

概率的古典定义

- 称具有以下两个特征的随机试验的数学模型为**古典概型**：
  - **有限样本**：只有有限个试验结果；
  - **等可能假设**：每个试验结果发生的可能性相等。
- 古典概型是概率论史上研究最早的情形，不需要重复试验。

定义（概率的古典定义）

设古典概型试验 E 的样本空间为 $\Omega = \{\omega_1, \ldots, \omega_n\}$，若事件 $A \subseteq \Omega$ 包含其中 m 个结果，则事件 A 的概率 $P(A)$ 定义为

$$P(A) \triangleq \frac{|A|}{|\Omega|} = \frac{m}{n}$$

对概率 $P(A)$ 的计算，转化为对样本空间 Ω 和事件 A 两个集合大小的计算。

概率的古典定义

例 (掷硬币)

连续掷两次硬币，求出现一个正面一个反面的概率。

- 样本空间为 $\Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$
- 事件“出现一正一反”为 $A = \{\mathrm{HT}, \mathrm{TH}\}$
- 因此 $P(A) = 1/2$

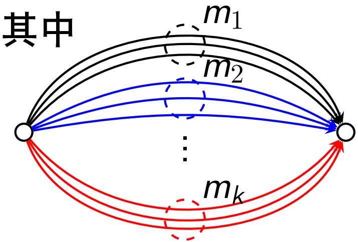
例 (掷骰子)

连续掷两次骰子，求点数之和是 9 的概率。

- 样本空间为 $\Omega = \{(i,j) | i,j \in \{1, 2, \ldots, 6\}\}$
- 事件“点数和是 9”为 $A = \{(3,6), (6,3), (4,5), (5,4)\}$
- 因此 $P(A) = 1/9$

回顾：加法原理

- 如果某件事情可由 k 类不同途径完成，其中
  - 第一类途径有 m_1 种方法
  - 第二类途径有 m_2 种方法
  - $\cdots\cdots$
  - 第 k 类途径有 m_k 种方法
- 那么完成这件事情共有 $m_1 + m_2 + \cdots + m_k$ 种方法。

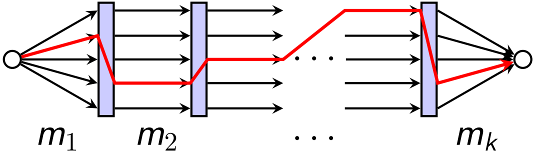


例 (加法原理)

由甲城到乙城去旅游有三类交通工具：汽车、火车和飞机。而汽车有 5 个班次，火车有 3 个班次，飞机有 2 个班次，那么从甲城到乙城共有 $5 + 3 + 2 = 10$ 个班次供旅游者选择。

回顾：乘法原理

- 如果完成某件事情需要经过 k 个步骤，其中
  - 第一步有 m_1 种做法
  - 第二步有 m_2 种做法
  - ⋯⋯
  - 第 k 步有 m_k 种做法



- 那么完成这件事情共有 $m_1 \times m_2 \times \cdots \times m_k$ 种方法。

例 (乘法原理)

由甲城到乙城有 3 条旅游线路，由乙城到丙城有 2 条旅游线路，那么从甲城经乙城去丙城旅游共有 $3 \times 2 = 6$ 条旅游线路可供选择。

回顾：排列组合 I

- 全排列：n 个不同编号的球排成一列，共有 $n!$ 种排列。
- 部分排列：从 n 个不同编号的球中不放回地取出 m 个排成一列，共有 $n(n-1)\cdots(n-m+1) = n!/(n-m)!$ 种排列。
- 组合：从 n 个不同编号的球中取出 m 个作为一组（不考虑顺序），共有 $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ 种取法。
  - $\binom{n}{m}$ 称为二项式系数，也可以写作 C_n^m
- 分组：把 n 个不同编号的球分成 k 组，第 i 组有 n_i 个球，$\sum_{i=1}^k n_i = n$，共有 $\binom{n}{n_1,\ldots,n_k} = \frac{n!}{n_1!\cdots n_k!}$ 种分组方法。
  - $\binom{n}{n_1,\ldots,n_k}$ 称为多项式系数。

回顾：排列组合 II

- 有部分重复的排列：n 个球，其中 m 个球有相同编号，则这 n 个球共有 $n!/m!$ 种排列。
- 分组排列：有 n 个球，分为 2 组，属于同一组的球具有相同编号，其中第 1 组有 m 个球，第 2 组有 $n-m$ 个球，则这 n 个球共有 $\frac{n!}{m!(n-m)!}$ 种排列。
- 分组排列的一般形式：有 n 个球，分为 k 组，同一组的球都具有相同编号，其中第 i 组有 n_i 个球，$\sum_{i=1}^{k} n_i = n$，则这 n 个球共有 $\frac{n!}{n_1!\cdots n_k!}$ 种排列。

举例：扑克同花色

例 (扑克同花色)

将一副扑克（去掉大小王）共 52 张牌均分给四个玩家，问刚好每人拿到一手同花色牌的概率是多少？

- 将 52 张牌分成 4 组，共有 $\binom{52}{13,13,13,13}$ 种分法。
- 每个人拿到同花色共有 $4!$ 种分发。
- 故

$$P(A)=\frac{|A|}{|\Omega|}=\frac{4!}{\binom{52}{13,13,13,13}}\approx 4.4739\times 10^{-28}$$

举例：抽样问题

例 (抽样问题)

从装有 r 个红球和 b 个蓝球的口袋里依次取出两个球，求取出的两个球均为红球的概率？

用 A 表示事件“取出的两个球均为红球”。

- 有放回抽样：

$$P(A) = \frac{|A|}{|\Omega|} = \frac{r^2}{(r+b)^2}$$

- 无放回抽样：

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{r}{2}}{\binom{r+b}{2}}$$

举例：抽样问题

例（抽样问题推广）

将取两个球推广为取 n 个球（$n \le r+b$），其中包含 k 个红球（$k \le r$）和 $n-k$ 个蓝球的情况。

用 A_k 表示事件"取出的 n 个球中有 k 个红球"。

- 有放回抽样：

$$P(A_k) = \frac{|A_k|}{|\Omega|} = \frac{\binom{n}{k} r^k b^{n-k}}{(r+b)^n}$$

- 无放回抽样：

$$P(A_k) = \frac{|A_k|}{|\Omega|} = \frac{\binom{r}{k}\binom{b}{n-k}}{\binom{r+b}{n}}$$

概率的古典定义：几何概型

- 将等可能有限样本假设推广至无限试验结果的情形，即是几何概型。
- 若某试验 E 满足：
  - 样本空间为有限区域 Ω，其测度（例如长度、面积、体积）用 S_Ω 表示；
  - 样本点落入 Ω 内任何区域 A 中的可能性与区域 A 的测度 S_A 成正比。
- 那么与区域 A 相关的事件发生的概率定义为

$$P(A) \triangleq \frac{S_A}{S_\Omega}$$

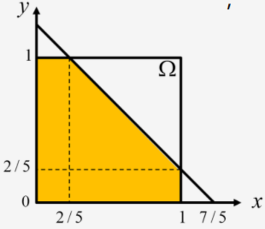
举例：几何概型

例

在区间 $(0,1)$ 中随机取两个数，求两数之和小于 $7/5$ 的概率。

- 在区间 $(0,1)$ 中随机取两个数分别记为 x,y，则 (x,y) 可能取值形成如下的正方形 $\Omega=\{(x,y)|0<x,y<1\}$，其面积 $S_\Omega=1$。
- 而事件“两数之和小于 $7/5$”可表示为 $A=\{(x,y)|0<x,y<1,x+y<7/5\}$，即图中阴影部分，其面积 $S_A=0.82$。
- 所以得到

$$P(A)=\frac{S_A}{S_\Omega}=0.82$$



举例：几何概型

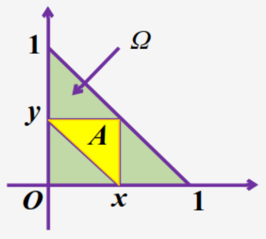
例（三角形概率）

将一单位长木棒随机折成三段，求刚好能构成三角形的概率。

记其中两个边的长度分别为 x 和 y，能形成三角形应满足：

$$A:\begin{cases} x+y>1-(x+y) \\ x+1-(x+y)>y \\ y+1-(x+y)>x \end{cases} \Rightarrow \begin{cases} x+y>\frac{1}{2} \\ y<\frac{1}{2} \\ x<\frac{1}{2} \end{cases}$$

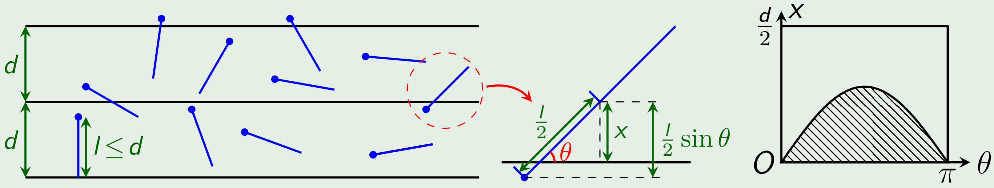
得到 $P(A)=1/4$.



举例：蒲丰投针试验

例 (蒲丰投针试验，1777 年)

平面上画有间隔为 d 的等距平行线，向平面中任意投掷一枚长为 $l < d$ 的针，求针与任一平行线相交的概率。



令针与平行线的夹角为 θ，针的中点与离它比较近的线的距离为 x，则 $0 \le \theta \le \pi, 0 \le x \le d/2$。所以样本空间 Ω 是 θ 与 x 围成的矩形区域。针与线相交的条件为 $x \le \frac{l}{2}\sin\theta$，所以

$$P(A) = \frac{S_A}{S_\Omega} = \frac{\int_0^\pi \frac{l}{2}\sin\theta \, \mathrm{d}\theta}{\frac{d}{2}\pi} = \frac{2l}{d\pi}$$

目录

概率两大学派

- **频率学派（频率说）**：坚持概率的频率解释的统计学家形成的学派，反对贝叶斯学派先验概率的主观性，代表人物为奈曼（1894-1981）、“伯克利们”，上世纪 30 年代学派形成。
- **贝叶斯学派（主观说）**：认为先验分布可以是主观的，它没有也不需要有频率解释，代表人物为贝叶斯（1701-1761），上世纪 60 年代学派形成。
- 两个学派的争论是战后数理统计学发展的一个特色，对今后数理统计学的发展还将产生影响。
- 贝叶斯学派在实际应用上取得的成功慢慢改变了人们的观点，逐渐受到重视，已经成为统计学中的热门研究课题。

频率学派

频率学派（频率说）

在一系列重复随机试验中考察随机事件发生的频率 $\Leftrightarrow$ 样本信息

例 (考察 26 个英文字母在文献中出现的频率)

字母	频率	字母	频率	字母	频率	字母	频率
E	0.1268	R	0.0594	M	0.0244	K	0.0060
T	0.0978	H	0.0573	W	0.0214	X	0.0016
A	0.0788	L	0.0394	Y	0.0202	J	0.0010
O	0.0776	D	0.0389	G	0.0187	Q	0.0009
I	0.0707	U	0.0280	P	0.0186	Z	0.0006
N	0.0706	C	0.0268	B	0.0156		
S	0.0634	F	0.0256	V	0.0102		

贝叶斯学派

贝叶斯学派（主观说）

根据以往的资料或经验，形成的关于随机事件发生可能性的印象。$\Leftrightarrow$ 先验信息

- 例如，考察某人是某案件嫌疑人的概率；推测山洞里隐藏的某种类型动物的概率。

统计概率

例：抛硬币试验

实验者	抛硬币次数	正面朝上次数	正面出现频率
蒲丰	$4,048$	$2,048$	0.5069
德 · 摩根	$2,048$	$1,061$	0.5181
皮尔逊	$12,000$	$6,019$	0.5016
皮尔逊	$24,000$	$12,012$	0.5005
罗曼诺夫斯基	$80,640$	$39,699$	0.4923

频率的“稳定性”就是统计规律性，即频率稳定于概率。

定义 (统计概率或经验概率)

设有随机试验 E，当试验的重复次数 n 充分大时，若事件 A 的发生频率 $f_n(A)$ 稳定在 p 附近，则称 p 为事件 A 发生的概率。这样定义的概率称为统计概率（或经验概率），记为 $P(A) = p$.

目录

概率的公理化定义

- 概率论发展史上有：古典定义、几何定义、频率定义、主观定义，这些定义只适合一类随机现象，如何给出适用于一切随机现象的概率的最一般定义呢？
- 1900 年，希尔伯特提出要建立概率的公理化定义，即从最少的几条本质特性出发去刻画概率。1933 年，前苏联数学家柯尔莫哥洛夫提出了概率的公理化定义。
- 问题：从日常生活经验出发，概率应该满足什么特性？
- 非负：随机事件发生可能性的大小应当非负；
- 规范：概率值属于 $[0, 1]$ 之间，虽然不是本质的，但很自然；
- 可加：两个互斥事件之和发生的可能性大小，应该是各自可能性大小之和。

概率的公理化定义

定义 (概率的公理化定义)

概率空间定义在三元组 $(\Omega, \mathcal{F}, P)$ 上，其中

- Ω 为样本空间；
- $\mathcal{F} \subseteq 2^{\Omega}$ 为事件域且满足公理 F1–F3；
- $P : \mathcal{F} \mapsto \mathbb{R}$ 为将事件映射为概率的实值函数且满足公理 P1–P3。

事件域 F 满足以下公理：

F1 $\Omega \in \mathcal{F}$;

F2 若 $A \in \mathcal{F}$，则 $\bar{A} \in \mathcal{F}$;

F3a 若 $A, B \in \mathcal{F}$，则 $A \cup B \in \mathcal{F}$;

F3b 若 $A_i \in \mathcal{F}$，$i \geq 1$，则 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

概率的公理化定义

概率 P 满足以下公理：

P1 $\forall A \in \mathcal{F}$，都有 $P(A) \geq 0$;

P2 $P(\Omega) = 1$;

P3a 若 $A, B \in \mathcal{F}$ 且 $A \cap B = \emptyset$，则 $P(A \cup B) = P(A) + P(B)$;

P3b 若 $A_i \in \mathcal{F}$，$i \geq 1$，且 $A_i \cap A_j = \emptyset$，$i \neq j$，则

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

- 事实上，F3a 和 P3a 可以分别由 F3b 和 P3b 得到，但反之不成立。
- P3 称为概率的**加法公理**，其中 P3b 也称为概率的**无限可列可加性**。

事件域的性质

性质 1：不可能事件属于事件域

$$\emptyset \in \mathcal{F}$$

由 F1 和 F2 可以直接得到，进而可以证明 F3b $\Rightarrow$ F3a。

性质 2：乘积事件属于事件域

若 $A, B \in \mathcal{F}$，则 $A \cap B \in \mathcal{F}$.

由 F2 得 $\bar{A}, \bar{B} \in \mathcal{F}$，再由 F3a 得 $\bar{A} \cup \bar{B} \in \mathcal{F}$，所以 $A \cap B = \overline{\bar{A} \cup \bar{B}} \in \mathcal{F}$.

概率的性质

性质 1：不可能事件的概率

$$P(\emptyset) = 0$$

由 P2 和 P3a 得到 $P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset)$.

性质 2：有限可列可加性

设 $A_i \in \mathcal{F}, 1 \leq i \leq n$ 且 $A_i A_j = \emptyset, i \neq j$，则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1 \cup \cdots \cup A_n \cup \emptyset \cup \cdots)$

概率的性质

性质 3：单调性

若 $A, B \in \mathcal{F}$ 且 $A \subseteq B$，则 $P(A) \le P(B)$ 且 $P(B-A) = P(B) - P(A)$.

$$P(B) = P(A \cup (B-A)) = P(A) + P(B-A)$$

性质 4：有界性

若 $A \in \mathcal{F}$，则 $0 \le P(A) \le 1$.

$\emptyset \subseteq A \subseteq \Omega$ （或由 P1 和单调性得到）

性质 5：对立事件的概率

若 $A \in \mathcal{F}$，则 $P(\bar{A}) = 1 - P(A)$.

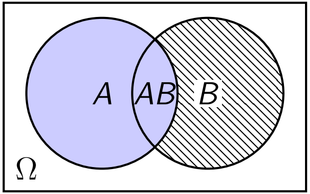
$$P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$$

概率的性质

性质 6：和事件的概率

设 $A, B \in \mathcal{F}$，则 $P(A \cup B) = P(A) + P(B) - P(AB)$.

- 因为 $A \cup B = A \cup (B \cap \overline{AB})$，所以由 P3a 得 $P(A \cup B) = P(A) + P(B \cap \overline{AB})$.
- 又因为 $B \cap \overline{AB} = \overline{\overline{B} \cup AB}$，所以由性质 5 和 P3a 得 $P(B \cap \overline{AB}) = 1 - P(\overline{B} \cup AB) = 1 - P(\overline{B}) - P(AB) = P(B) - P(AB)$.



推论

若 $A_1, \ldots, A_n \in \mathcal{F}$，则

挖补规律：加奇减偶

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j) + \cdots + (-1)^{n-1} P(A_1 \cdots A_n)$$

概率的性质

定义（极限事件）

- 对事件域 $\mathcal{F}$ 中任一单调不减事件列 $A_1 \subseteq A_2 \subseteq \cdots$，称可列并 $\bigcup_{i=1}^{\infty} A_i$ 为事件列 $\{A_n\}$ 的极限事件，记为

$$\lim_{n\to\infty} A_n \triangleq \bigcup_{i=1}^{\infty} A_i$$

- 对事件域 $\mathcal{F}$ 中任一单调不增事件列 $A_1 \supseteq A_2 \supseteq \cdots$，称可列交 $\bigcap_{i=1}^{\infty} A_i$ 为事件列 $\{A_n\}$ 的极限事件，记为

$$\lim_{n\to\infty} A_n \triangleq \bigcap_{i=1}^{\infty} A_i$$

概率的性质

性质 7：概率的连续性

设 P 为事件域 $\mathcal{F}$ 上的概率，则 P 满足下面的性质：

- 对事件域 $\mathcal{F}$ 中任一单调不减事件列 $A_1 \subseteq A_2 \subseteq \cdots$，有
$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n)$$
- 对事件域 $\mathcal{F}$ 中任一单调不增事件列 $A_1 \supseteq A_2 \supseteq \cdots$，有
$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n)$$

- 对于事件域中的单调不减（增）事件列，求极限运算和求概率运算可以交换次序；
- 分别称概率 P 是下连续和上连续的。

对概率下连续性的证明

证明.

令 $B_1 = A_1, B_i = A_i - A_{i-1}, i = 2, 3, \dots$. 显然 $B_iB_j = \emptyset, \forall i \neq j$, 且

$$A_n = \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i.$$

由可列可加性得

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

$$= P(\bigcup_{i=1}^{\infty} B_i) = P(\bigcup_{i=1}^{\infty} A_i) = P(\lim_{n\to\infty} A_n)$$

□

举例：概率的性质

例 (和事件的概率)

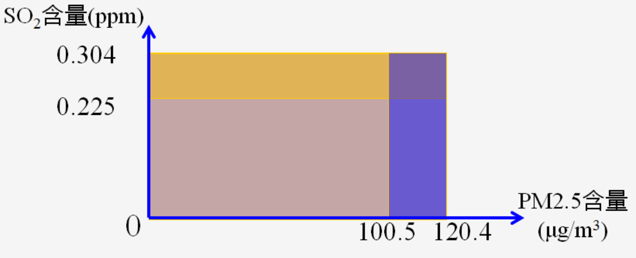
在 $1 \sim 1000$ 的整数中随机取一个数，问取到的整数能被 4 整除或者能被 6 整除的概率是多少?

- 设 A 为事件“取到的数能被 4 整除”，B 为事件“取到的数能被 6 整除”
- 则所求事件概率为 $P(A \cup B) = P(A) + P(B) - P(AB)$
- $P(A) = \frac{\lfloor 1000/4 \rfloor}{1000}$, $P(B) = \frac{\lfloor 1000/6 \rfloor}{1000}$, $P(AB) = \frac{\lfloor 1000/12 \rfloor}{1000}$
- 故 $P(A \cup B) = 0.333$

举例：概率的性质

例

已知空气中 PM2.5 含量一般在 $0 \sim 120.4\mu g/m^3$ 之间，SO_2 含量一般在 $0 \sim 0.304\text{ppm}$ 之间。一般认为，PM2.5 含量在 $100.5\mu g/m^3$ 以上或 SO_2 含量在 0.225ppm 以上为对人体有害。问空气质量为有害的概率是多少？



举例：概率的性质

例 (配对问题)

旅社管理员共管理 n 间客房，钥匙标牌丢失，随机将这 n 个钥匙分给 n 个旅客，求至少有一人能打开房门的概率。

记 A_i 为事件“第 i 个房门被打开”，A 为事件“至少有一人能打开房门”。利用加法公式

$$P(A) = P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) + \cdots$$
$$+(-1)^{n-1}P(A_1\cdots A_n)$$

易知 $P(A_i) = \frac{1}{n}, P(A_iA_j) = \frac{1}{n(n-1)}, \ldots, P(A_1\cdots A_n) = \frac{1}{n!}$

举例：概率的性质

故

$$
\begin{aligned}
P(A) &= \binom{n}{1}\frac{1}{n} - \binom{n}{2}\frac{1}{n(n-1)} + \cdots + (-1)^{n-1}\binom{n}{n}\frac{1}{n!} \\
&= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!}
\end{aligned}
$$

当 $n \to \infty, P(A) \to 1 - e^{-1}$.

举例：概率的性质

例 (掷双骰子)

反复掷两个骰子，观察其点数和，若首次出现 7 点或 8 点，则试验结束。求试验结束的概率。

- 设 A_i 表示事件“前 $i-1$ 次投掷不出现 7 点或 8 点，第 i 次投掷出现 7 点”，A 表示事件“试验结果出现 7 点而结束”
- 设 B_i 表示事件“前 $i-1$ 次投掷不出现 7 点或 8 点，第 i 次投掷出现 8 点”，B 表示事件“试验结果出现 8 点而结束”
- 显然 A_i, B_i 均互斥，如果可以算出 $P(A_i), P(B_i)$，则可算出 $P(A), P(B)$

$$P(A) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),\quad P(B) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i)$$

举例：概率的性质

- A_i：前 $i-1$ 次不能出现 7 点或 8 点，第 i 次出现 7 点，所以

$$P(A_i) = (1 - \frac{11}{36})^{i-1} \cdot \frac{1}{6}$$

所以

$$P(A) = \sum_{i=1}^{\infty} P(A_i) = \frac{6}{11}$$

- B_i：前 $i-1$ 次不能出现 7 点或 8 点，第 i 次出现 8 点，所以

$$P(B_i) = (1 - \frac{11}{36})^{i-1} \cdot \frac{5}{36}$$

所以

$$P(B) = \sum_{i=1}^{\infty} P(B_i) = \frac{5}{11}$$

小结

1. 概率的古典定义
2. 概率的统计定义
3. 概率的公理化定义