

## 第一章: 概率 (Probability)

李立颖

[lily@sustech.edu.cn](mailto:lily@sustech.edu.cn), 理学院 M622

2023 秋季

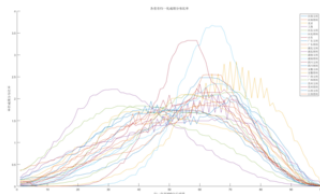
# 第一章: 概率 (Probability)

- ① 引言
- ② 样本空间 (Sample space)
- ③ 概率测度 (Probability measure)
- ④ 概率计算: 计数方法
  - 古典概型
  - 几何概型

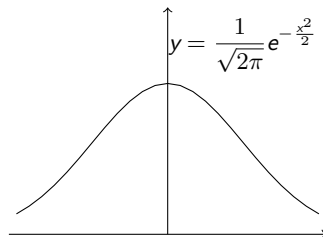
# 为什么学习概率论与统计?

- 现实生活中的随机现象主要来源于复杂系统:
  - 人口普查
  - 分子热运动
  - 天气预报
  - 量子力学?
- 概率论与统计提供了研究随机现象的工具
  - 概率论: 建立于现代分析学基础上, 主要提供理论基础, 得到的很多结果简洁而深刻, 如大数定律、中心极限定理、遍历定理等
  - 统计学: 解决与复杂系统相关的实际问题 (估计、判断、决策), 如参数估计、假设检验、分类、机器学习等

# 例 1: 正态分布



(a) 高考成绩<sup>1</sup>



(b) 标准正态分布

<sup>1</sup>图片来源: [https://blog.csdn.net/HugoChen\\_cs/article/details/107650258](https://blog.csdn.net/HugoChen_cs/article/details/107650258)

## 例 2: 核物理、随机矩阵与黎曼猜想

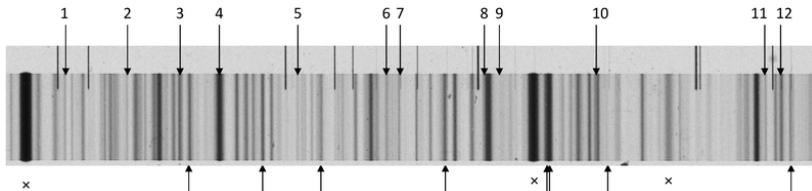


图: 铯原子能量光谱<sup>2</sup>

- 薛定谔方程

$$i\frac{\partial}{\partial t}\phi = \hat{H}\phi.$$

- 能级 = 特征值

$$\hat{H}\phi = \lambda\phi, \quad \lambda \in \mathbb{R}.$$

- 重原子对应的算子  $\hat{H}$  类似于一个高维对称矩阵  $X$

<sup>2</sup><https://www.mdpi.com/2218-2004/5/3/24>

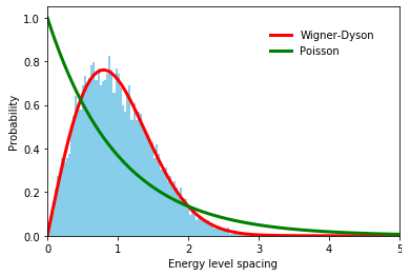
## 例 2: 核物理、随机矩阵与黎曼猜想, II

- 对称随机矩阵

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1N} \\ X_{21} & X_{22} & \cdots & X_{2N} \\ \cdots & & \ddots & \cdots \\ X_{N1} & X_{N2} & \cdots & X_{NN} \end{bmatrix},$$

在满足  $x_{ij} = x_{ji}$  条件下, 矩阵元素随机选取

- 特征值  $\lambda$  满足  $Xv = \lambda v$ . 对称矩阵的特征值都是实数.
- 相邻特征值  $\lambda_{i+1} - \lambda_i$  的分布是什么?



## 例 2: 核物理、随机矩阵与黎曼猜想, III

- 黎曼  $\zeta$ -函数,  $\operatorname{Re} z > 1$

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}.$$

以上函数可延拓至  $z \in \mathbb{C} = \{a + b \cdot i : a, b \in \mathbb{R}\}$ ,  $i = \sqrt{-1}$ .

- 黎曼猜想: 黎曼函数的全部零点在  $\{a = \frac{1}{2}\}$  上.
- Montgomery 猜想:  $\{a = \frac{1}{2}\}$  上相邻零点的距离分布由 Dyson–Wigner 分布给出
- 小结:
  - 概率问题: 怎么定义随机性? 怎么刻画极限过程?
  - 统计问题: 怎么比较两个分布?

# 如何学习概率统计？

- 学思想：概率统计特殊的研究对象包含了许多独特的思维方法和思想方法，特别是如何看待和处理随机规律，这是其他学科中没有的。例如，以比较各种事件出现的可能性的进行决策的思想。
- 学方法：定量描述随机现象及其规律的方法，收集、整理、分析数据，从而建立统计模型的方法。
- 学应用：尽可能多地了解各种概念的背景、各种方法和模型的实际应用。不仅要学课程中提及的，也要自己收集、寻找各种实例。
- 课前预习、课中认真听讲、课后多做练习。



## 定义

试验: 科学实验, 或者对某一事物的某一特征进行观察

## 例

- 抛一枚硬币, 观察正面 H, 反面 T 出现的情况
- 将一枚硬币连抛三次, 观察正面 H 出现的次数
- 掷一颗骰子, 观察出现的点数
- 从一批产品中抽取  $n$  件, 观察次品出现的数量
- 对某厂生产的电子产品进行寿命测试
- 观察某地区的日平均气温和日平均降水量

## 问题:

这些实验都有什么特点?

试验前无法预知结果.

# 随机实验与样本空间

## 试验的特征

- 试验可以在相同的条件下重复进行
- 试验的结果可能不止一个, 但试验前知道所有可能的全部结果
- 在每次试验前无法确定会出现哪个结果

## 定义

具有上述特征的称为随机试验, 或简称试验

## 例

实验  $E$ : 掷一颗骰子, 观察出现的点数.

$E$  的结果

- “1 点”、“2 点”、...、“6 点” (基本结果, 不可分)
- “出现的点数不超过 3”、“至少出现 4 点” (复合结果, 可分解)

## 定义

- 称基本结果为样本点、基本事件
- 称试验的全部样本点构成的集合为样本空间.

## 例

- 离散样本空间:
  - 掷一颗骰子, 观察出现的点数, 其样本空间为  $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - 抛两枚硬币, 观察正、反两面出现的情况, 其样本空间为

$$\Omega = \{(H, T), (H, H), (T, H), (T, T)\}.$$

- 连续样本空间:
  - 记录深圳地区的日平均气温, 其样本空间为  $\Omega = (-60, 60)$
  - 播种飞机对位置为  $(x_0, y_0)$  的目标进行播种, 观察其所覆盖的范围  $(x, y)$ , 其样本空间为

$$\Omega = \{(x, y) \mid (x - x_0)^2 + (y - y_0)^2 < \infty\}.$$

# 随机事件

基本结果	不可分	样本点、基本事件
复合结果	可分解	随机事件、事件

表: 试验的结果

- 从集合看: 事件是样本空间的子集
- 从试验看: 事件是基本事件的复合

## 定义

满足一定条件的样本点的集合称为随机事件, 简称事件. 事件用大写字母  $A, B, C, \dots$  等表示.

## 例

掷一颗骰子, 观察出现的点数, 其样本空间为  $\Omega = \{1, 2, 3, 4, 5, 6\}$

- 事件  $A$  “至少出现 3 点”:  $A = \{3, 4, 5, 6\}$
- 事件  $B$  “出现最小或最大点的点”:  $B = \{1, 6\}$

几个特殊事件:

- 基本事件: 一个样本点构成的单点集  $\{\omega\}$ .
- 必然事件: 每次试验都总发生的事件  $\Omega \subset \Omega$ .
- 不可能事件: 每次试验都不会发生的事件  $\emptyset$  (空集  $\emptyset \subset \Omega$ ).

## 定义

记

$$\mathcal{A} = \{A \mid A \subset \Omega, A \text{ 是事件}\}.$$

称  $\mathcal{A}$  为试验的事件域, 即试验产生的所有事件为元素构成的集合.

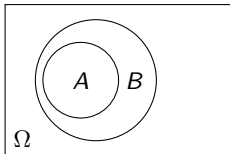
事件语言	基本结果	全体基本试验	事件 $A$	事件 $A$ 发生
集合语言	样本点 $\omega$	样本空间 $\Omega$	子集 $A$	$\omega \in A$

表: 小结 — 随机试验的数学描述

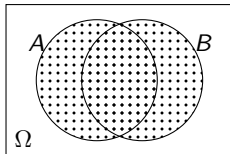
# 事件间的关系与运算 I

设  $A, B$  以及  $A_k, k = 1, 2, \dots$  为事件

- $A$  是  $B$  的子集, 记作  $A \subset B$ :  $A$  发生必然导致  $B$  发生
- 集合  $A$  与  $B$  相等, 记作  $A = B$ :  $A \subset B, B \subset A$
- 集合  $A$  与  $B$  的并, 记作  $A \cup B = \{\omega \mid \omega \in A \text{ 或 } \omega \in B\}$ :  $A$  发生或  $B$  发生, 即  $A, B$  至少有一个发生, 称为事件  $A$  和  $B$  的和.



(a)  $A \subset B$



(b)  $A \cup B$

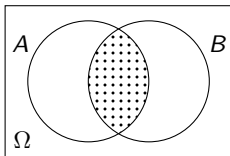
## 事件间的关系与运算 II

- 集合  $A$  与  $B$  的交, 记作  $A \cap B = \{\omega \mid \omega \in A, \omega \in B\}$ :  $A, B$  同时发生, 称为事件  $A, B$  的积, 也记作  $AB$ .
- 类似地, 可以定义  $n$  个事件或者可列个事件的积

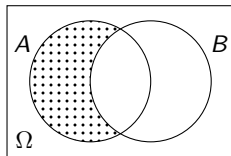
$$\bigcap_{i=1}^n A_i = \{\omega \mid \omega \in A_i, i = 1, 2, \dots, n\}$$

$$\bigcap_{i=1}^{\infty} A_i = \{\omega \mid \omega \in A_i, i = 1, 2, 3, \dots\}$$

- 集合  $A$  和  $B$  的差集, 记作  $A - B = A \setminus B = \{\omega \mid \omega \in A, \omega \notin B\}$ :  $A$  发生而  $B$  不发生



(a)  $A \cap B$

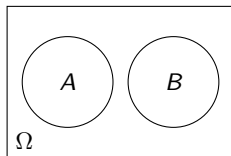


(b)  $A \setminus B$

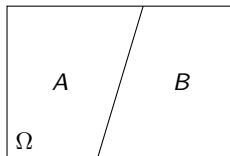
## 事件间的关系与运算 III

- 若  $A \cap B = \emptyset$ , 则称  $A, B$  互不相容 (互斥).
- 若  $A \cup B = \Omega$ , 且  $A \cap B = \emptyset$ , 则称  $A, B$  为逆事件或对立事件, 记为

$$A = \Omega \setminus B = B^c, \quad B = \Omega \setminus A = A^c.$$



(a)  $A \cap B = \emptyset$ ,  $A$  与  $B$  互斥



(b)  $A$  与  $B$  为对立事件



# 事件的运算定律

- 交换律:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A.$$

- 结合律:

$$A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C.$$

- 分配律:

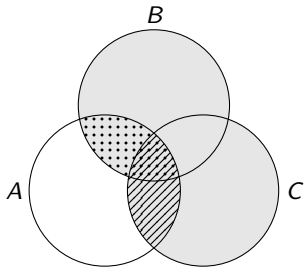
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

- De Morgan 律:

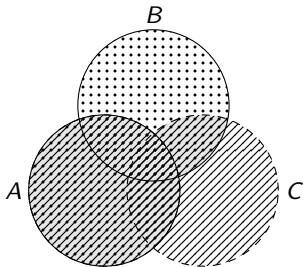
$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c,$$

$$\left( \bigcup_{k=1}^n B_k \right)^c = \bigcap_{k=1}^n B_k^c, \quad \left( \bigcap_{k=1}^n B_k \right)^c = \bigcup_{k=1}^n B_k^c.$$

# 分配律



(a)  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$



(b)  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

# 如何用定义进行证明

## 命题

$$\left(\bigcup_{k=1}^n B_k\right)^c = \bigcap_{k=1}^n B_k^c.$$

## 证明

令  $P = \left(\bigcup_{k=1}^n B_k\right)^c$ ,  $Q = \bigcap_{k=1}^n B_k^c$ . 欲证  $P = Q$ , 我们只需分别证明  $P \subset Q$  与  $Q \subset P$ . 我们这里只示范前者的证明. 由包含关系的定义, 对任意的  $\omega \in P$ , 我们要推出  $\omega \in Q$ . 事实上,

$$\begin{aligned}\omega \in P &= \left(\bigcup_{k=1}^n B_k\right)^c && \xRightarrow{\text{(补集的定义)}} \omega \notin \bigcup_{k=1}^n B_k \\ &\xRightarrow{\text{(并集的定义)}} \omega \notin B_k, \forall k && \xRightarrow{\text{(补集的定义)}} \omega \in B_k^c, \forall k \\ &\xRightarrow{\text{(交集的定义)}} \omega \in \bigcap_{k=1}^n B_k^c = Q.\end{aligned}$$

# 可列

- 可列集：指一个无穷集  $S$ ，其元素可与自然数形成一一对应，因此可表为  $S = \{s_1, s_2, \dots\}$ .
- 至多可列：指可列或有限
- 可以证明：可列是“最小的”无穷，即任何一个无穷集合均含有可列子集

# 作业

- P20: 5, 6

- 补充题:

- ① 设随机事件  $A, B$  满足条件  $AB = A^c B^c$ , 试求  $A \cup B$ .
- ② 试把事件  $A_1 \cup A_2 \cup \dots \cup A_n$  表示成  $n$  个两两互不相容事件之并.

# 概率论公理化的三种学派

## 1921, J. M. Keynes, “主观概率学派”

凯恩斯主张把任何命题都看作事件，例如“明天将下雨”，“土星上有生命”等等都是事件，人们对这些事件的可信程度就是概率，而与随机试验无关，通常称为主观概率。

## 1928, von Mises, “客观概率学派”

米泽斯定义事件的概率为该事件出现的频率的极限，而作为公理就必须把这一极限的存在作为第一条公理，通常称为客观概率。

## 1933, Kolmogorov (柯尔莫哥洛夫), “以测度论为基础的概率公理化体系”

目前，绝大多数教科书都是采用柯尔莫哥洛夫的概率公理化体系。

概率论是研究随机现象的统计规律性的数学学科

### 问题一：什么是统计规律性？

统计规律性是指在大量试验中呈现出的数量规律 (用频率来刻画).

### 问题二：什么是概率？

概率是指刻画随机事件在一次试验中发生的可能性大小的数量指标, 这个数量指标应该满足:

- 它是事件固有的, 不随人们主观意愿而改变; 可以在相同条件下通过大量重复试验予以识别和检验
- 符合常情: 事件发生可能性大, 该值就大, 反之就小; 不可能事件的值最小 (0); 必然事件的值最大 (1)

# 频率

## 频率的定义

设  $A$  为一随机事件, 在相同条件上进行  $n$  次重复实验. 令

$n_A = n$  次实验中  $A$  发生的事件

$$f_n(A) = \frac{n_A}{n}$$

称  $n_A$  为事件  $A$  的频数,  $f_n(A)$  为事件  $A$  的频率 (*frequency*).

## 频率的一般特性

- 一般地,  $n$  越大, 则  $n_A$  越大.
- $n_A$ 、 $f_n(A)$  的值是“随机的”.
- $0 \leq f_n(A) \leq 1$ .

## 问题

频率是否有统计规律性?



## 实例 1: “抛硬币实验”

将一枚硬币连续抛  $n$  次, 记  $H = \{ \text{出现正面} \}$ .

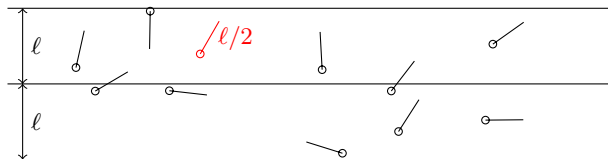
### 问题

$f_n(H)$  有什么规律?

实验者	$n$	$n_K$	$f_n(H)$
德摩根	2048	1061	0.5181
蒲丰	4048	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

表: 历史上有名的抛硬币实验

## 实例二：“蒲丰投针实验”



记投针的总数为  $n$ , 针与平行线相交的次数为  $n_A$ , 则

$$\frac{n_A}{n} \approx \frac{1}{\pi}, \quad \text{或} \quad \frac{n}{n_A} \approx \pi.$$

## 实例 3: 英文字母 (密码破译)

考察英语文章中 26 个字母出现的频率, 当观察次数  $n$  较大时, 每个字母出的频率呈现稳定性, 下面是 Deway 统计了 438023 个字母得到的统计表.

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
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		

# 随机事件的统计规律性

## 频率的稳定性 (大数定律)

当  $n$  很大时, 事件  $A$  的频率  $f_n(A)$  接近一个常数, 即有

$$f_n(A) \rightarrow p, \quad n \rightarrow \infty$$

- 常数  $p$  就是事件  $A$  发生的可能性大小, 即概率.
- 由于频率的取值是“随机的”, 那么上述极限是在什么意义下成立呢? (第五章研究此问题)

## 实例四：“掷骰子”实验

### 问题

记事件

$$A_i = \{ \text{出现 } i \text{ 点} \}, \quad i = 1, 2, \dots, 6.$$

将一颗骰子连续掷  $n$  次,  $f_n(A_i)$  有什么规律?

如果一颗骰子六个面是均匀的, 则当  $n$  很大时应有

$$f_n(A_i) = \frac{n_{A_i}}{n} \approx \frac{1}{6}, \quad n = 1, 2, \dots, 6.$$

## 频率的基本性质

- ①  $0 \leq f_n(A) \leq 1$
- ②  $f_n(\Omega) = 1$
- ③ (有限可加性) 若  $A_1, \dots, A_m$  两两不相容 ( $A_i \cap A_j = \emptyset, \forall i \neq j$ ), 则

$$f_n(\cup_{i=1}^m A_i) = \sum_{i=1}^m f_n(A_i).$$

## 概率的公理化定义

设  $\mathcal{A}$  为样本空间  $\Omega$  上的事件域. 对任意  $A \in \mathcal{A}$ , 若存在实数  $P(A)$  与之对应, 且满足

- ① 非负性:  $P(A) \geq 0, \forall A \in \mathcal{A}$ ;
- ② 规范性:  $P(\Omega) = 1$ ;
- ③ 可列可加性: 对两两不相容的事件列  $\{A_k\}_{k=1}^{\infty}$ , 有  $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$ ;

则称  $P(A)$  为事件  $A$  的概率, 称  $(\Omega, \mathcal{A}, P)$  为概率空间 (*probability space*).

## 概率的公理化定义

设  $\mathcal{A}$  为样本空间  $\Omega$  上的事件域. 对任意  $A \in \mathcal{A}$ , 若存在实数  $P(A)$  与之对应, 且满足

① 非负性:  $P(A) \geq 0, \quad \forall A \in \mathcal{A};$

② 规范性:  $P(\Omega) = 1;$

③ 可列可加性: 对两两不相容的事件列  $\{A_k\}_{k=1}^{\infty}$ , 有  $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k);$

则称  $P(A)$  为事件  $A$  的概率, 称  $(\Omega, \mathcal{A}, P)$  为概率空间 (*probability space*).

注:

- 1933 年苏联的 Kolmogorov 测度论基础上提出的概率论公理化体系.
- 概率是定义在事件域上的特殊函数.
- 物体的长度、区域的面积都具有“非负性”与“可加性”, 故“概率”实际上是对“事件”发生可能性大小的一种“度量”.

### 问题 (贝特朗悖论, Bertrand's Paradox)

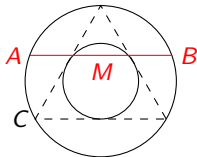
在半径为  $r$  的圆  $C$  内“任意”作一弦. 试求此弦长度  $\ell$  大于圆内接等边三角形边长  $\sqrt{3}r$  的概率  $p$ . (何为正确的概率空间  $(\Omega, \mathcal{A}, \mathbb{P})$ ?)

#### 解一

作半径为  $r/2$  的同心圆  $C_1$ . 设弦  $AB$  的中点  $M$  “任意”落于圆  $C$  内.

若  $M$  落于圆  $C_1$  内, 则  $\ell > \sqrt{3}r$ , 于是

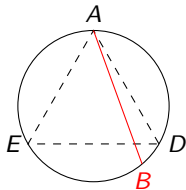
$$p = \frac{\pi(r/2)^2}{\pi r^2} = \frac{1}{4}.$$



#### 解二

设弦  $AB$  的一端  $A$  固定于圆周上, 另一端任意. 考虑等边  $\triangle ADE$ , 如  $B$  落于角  $A$  所对应的弧  $DE$  上, 则  $\ell > \sqrt{3}r$ . 于是

$$p = \frac{\text{DE 的弧长}}{\text{圆周长}} = \frac{1}{3}$$





# 概率的基本性质

## 性质 1

$$P(\emptyset) = 0.$$

证明: 因为  $\emptyset = \emptyset \cup \emptyset$ , 所以  $P(\emptyset) = P(\emptyset) + P(\emptyset)$ , 故  $P(\emptyset) = 0$ .

## 性质 2 (有限可加性)

若  $A_1, A_2, \dots, A_n$  是两两不相容事件, 则  $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$ .

证明: 因为  $\bigcup_{k=1}^n A_k = A_1 \cup \dots \cup A_n \cup \emptyset \cup \emptyset \cup \dots$ . 由可列可加性,

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) + P(\emptyset) + P(\emptyset) + \dots = \sum_{k=1}^n P(A_k).$$

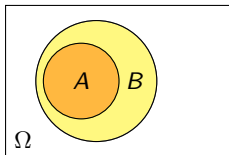
### 性质 3 (单调性)

若  $A \subset B$ , 则  $P(A) \leq P(B)$ .

证明: 因为  $A \subset B$ , 故  $B$  为不相容事件  $A$  与  $B \setminus A$  之和. 由有限可加性,

$$P(B) = P(A) + P(B \setminus A).$$

再由非负性,  $P(B \setminus A) \geq 0$ . 证毕.



## 性质 4

$$0 \leq P(A) \leq 1.$$

由  $\emptyset \subset A \subset \Omega$  及单调性.

## 性质 5

$$P(A^c) = 1 - P(A).$$

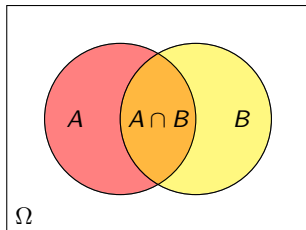
由  $A \cup A^c = \Omega$ ,  $A \cap A^c = \emptyset$  及有限可加性.

## 性质 6: 容斥原理 (inclusion-exclusion principle)

### 两事件容斥原理

对任何事件  $A, B$  有

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

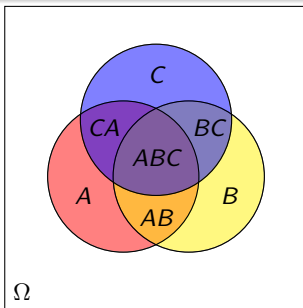


## 性质 6: 容斥原理 (inclusion-exclusion principle)

### 三事件容斥原理

对任何事件  $A, B, C$  有

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(AB) - P(CA) - P(BC) \\ &\quad + P(ABC). \end{aligned}$$



## 性质 6: 容斥原理 (inclusion-exclusion principle)

### 多事件容斥原理

对于  $n$  个事件, 有

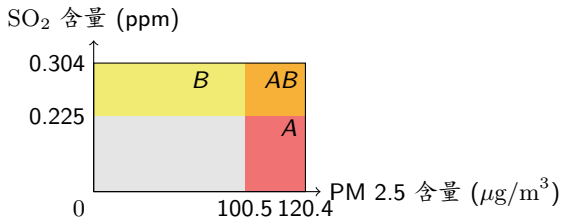
$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) = & \sum_{i=1}^n P(A_i) \\ & - \sum_{1 \leq i < j \leq n} P(A_i A_j) \\ & + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) \\ & + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n). \end{aligned}$$

### 规律

奇加偶减.

### 例题

已知空气中 PM2.5 含量一般在  $0.0-120.4 (\mu\text{g}/\text{m}^3)$  之间,  $\text{SO}_2$  含量一般在  $0.000-0.304 (\text{ppm})$  之间, 假设在上述范围内取值为等可能的. 一般认为, PM2.5 含量在  $100.5 \mu\text{g}/\text{m}^3$  以上或  $\text{SO}_2$  含量在  $0.225 \text{ ppm}$  以上为对人体有害. 问空气质量为有害的概率是多少?



### 解

$$P(A) = 0.165, \quad P(B) = 0.260, \quad P(AB) = 0.043.$$

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0.382.$$

或利用对立事件计算:  $(A \cup B)^c = A^c B^c$ .

# 作业

- P20: 4, 7

- 补充题:

① 设  $A, B, C$  是三个随机事件, 且  $P(A) = P(B) = P(C) = \frac{1}{4}$ ,

$P(AB) = P(BC) = 0, P(AC) = \frac{1}{8}$ . 求  $A, B, C$  至少发生一个的概率.

② 已知  $A, B$  两个事件满足条件  $P(AB) = P(\bar{A}\bar{B})$ , 且  $P(A) = p$ . 求  $P(B)$ .



# 古典概型

“抛硬币”、“掷骰子”等随机试验的特征:

- 只有有限个基本结果 (样本空间为有限集)
- 每个基本结果的出现是等可能的.

## 古典概型

设随机试验的样本空间为  $\Omega$ . 若

- ①  $\Omega$  只含有有限个样本点, 即  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ ;
- ② 每个样本点的出现是等可能的, 即

$$P(\omega_1) = P(\omega_2) = \dots = P(\omega_n) = \frac{1}{n},$$

则称该实验为等可能概型, 也称为古典概型.

## 问题:

怎样计算古典概型中一般事件的概率?

## 古典概型的概率计算

设事件  $A$  含  $k$  个样本点, 记为  $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\}$ . 则由有限可加性:

$$P(A) = P(\omega_{i_1}) + P(\omega_{i_2}) + \dots + P(\omega_{i_k}) = \frac{1}{n} \times k = \frac{k}{n}.$$

$$P(A) = \frac{\text{导致 } A \text{ 发生的方式个数}}{\text{所有试验结果个数}} = \frac{k}{n}.$$

$$P(A) = \frac{A \text{ 的有利场合数}}{\text{样本点总数}} = \frac{k}{n}.$$

## 例

抛两枚硬币，求出现一个正面一个反面的概率。

## 解

该实验的样本空间为

$$\Omega = \{HT, HT, TH, TT\}.$$

这是一个古典概型. 事件  $A = \{\text{一个正面, 一个反面}\}$  的有利场合是  $HT, TH$ . 因此

$$P(A) = \frac{2}{4} = \frac{1}{2}.$$

## 历史趣闻

18 世纪著名法国数学家达朗贝尔取样本空间  $\Omega = \{HH, HT, TT\}$  得到

$$P(A) = \frac{1}{3}.$$

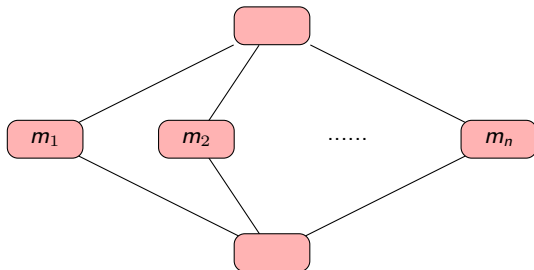
他错在了哪里？这不是古典概型！

# 加法原理 (Addition Principle)

## 加法原理

做一件事一共有  $n$  类方法, 第一类有  $m_1$  种方法, 第二类有  $m_2$  种方法, ..., 第  $n$  类有  $m_n$  种方法, 则完成这件事的方法总数为

$$N = m_1 + m_2 + \cdots + m_n.$$



# 乘法原理 (Multiplication principle, basic principle of counting)

## 乘法原理

做一件事有  $n$  个步骤, 第一步有  $m_1$  种方法, 第二步有  $m_2$  种方法, ..., 第  $n$  步有  $m_n$  种方法, 则完成这件事的方法总数为

$$N = m_1 \cdot m_2 \cdot \cdots \cdot m_n.$$

# 排列与组合

## 选排列 (Arrangement)

从  $n$  个不同的元素中, 任取  $k(\leq n)$  个元素, 按照一定的顺序排成一列, 全部排列个数为

$$A_n^k = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1).$$

## 全排列 (Permutation)

当  $k = n$  时, 称为全排列.

$$A_n^n = n!$$

## 组合 (Combination)

从  $n$  个不同的元素中, 任取  $k(\leq n)$  个元素并成一组, 全部组合数为

$$\binom{n}{k} = C_n^k = \frac{A_n^k}{A_k^k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

## 组合数的推广

一般地, 把  $n$  个球随机地分成  $r$  组 ( $n > r$ ), 要求第  $i$  组恰好有  $n_i$  个球,  $i = 1, \dots, r$ , 共有分法

$$\frac{n!}{n_1! n_2! \cdots n_r!} = \binom{n}{n_1 \ n_2 \ \cdots \ n_r}.$$

## 排列与组合的区别

- 排列: 方法数与次序有关
- 组合: 方法数与次序无关.

## 例题

袋中有  $a$  只红球,  $b$  只白球, 从袋中随机地将球一个一个取出. 求第  $k$  次取出的是红球的概率 ( $1 \leq k \leq a + b$ ).

## 思路一

假设除颜色外, 球是可区分的, 将取出的球依次排成一行



把每一种排法作为一个样本点. 有利场合为第  $k$  个位置为红球, 其它位置的  $(a + b - 1)$  个球可任意放置.

## 解

样本点总数为  $(a + b)!$ , 有利场合数为  $C_a^1 \cdot (a + b - 1)!$ . 故所求概率为

$$p_k = \frac{a(a + b - 1)!}{(a + b)!} = \frac{a}{a + b}, \quad 1 \leq k \leq a + b.$$



## 例题

袋中有  $a$  只红球,  $b$  只白球, 从袋中随机地将球一个一个取出. 求第  $k$  次取出的是红球的概率 ( $1 \leq k \leq a + b$ ).

## 思路二

假设除颜色外, 球是不可的, 将取出的球依次排成一行,



把每一种红球可能出现的位置作为一个样本点. 有利场合为第  $k$  个位置为红球, 其它位置  $(a + b - 1)$  个位置中出现  $(a - 1)$  个红球

## 解

样本点总数为  $C_{a+b}^a$ , 有利场合数为  $C_{a+b-1}^{a-1}$ . 故所求概率为

$$p_k = \frac{(a+b-1)!/(a-1)!b!}{(a+b)!/a!b!} = \frac{a}{a+b}, \quad 1 \leq k \leq a+b.$$

## 总结

### 结果

$$p_k = \frac{a}{a+b}, \quad 1 \leq k \leq a+b.$$

- 思路一用了排列的思路, 假设除颜色外球是可区分的
- 思路二用了组合的思路, 假设除颜色外球是不可区分的
- 确定了样本空间的结构后, 有利场合的构造必须与样本空间结构一致.
- 概率与  $k$  无关: 抽签结果与顺序无关.
- 很多实际问题都可以归结为“摸 (或扔) 球模型”.

## 例题

将  $n$  只 (可区分的) 球随机放入  $N$  ( $N \geq n$ ) 个 (可区分的) 盒子中去, 试求每个盒子至多有一个球的概率.

## 分析

任意一只球放进任一盒子是等可能的. 我们把盒子编号为  $1, 2, \dots, N$ , 并把每个球放入的盒子编号  $(a_1, a_2, \dots, a_n)$  作为样本点, 并用古典概型求解.

## 解

我们样本点总数为  $N^n$ . 每个盒子至多有一个球, 意味着  $(a_1, a_2, \dots, a_n)$  互不相同, 因此有利场合数为选排列数  $A_N^n = N(N-1) \cdots (N-n+1)$ . 故所求概率为

$$p = \frac{A_N^n}{N^n} = 1 \cdot \left(1 - \frac{1}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right).$$

# 生日问题

## 问题

参加某次聚会共  $n$  个人, 求没有两人生日相同的概率.

## 分析

此即为  $n$  个球 (人) 放入 365 个盒子中的问题.

## 解

$$p = 1 \cdot \left(1 - \frac{1}{365}\right) \cdot \left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right).$$

$n$	20	25	30	40	50	55	100
$1 - p$	0.41	0.57	0.71	0.89	0.97	0.99	0.999997

## 例题

从数字 0, 1, ..., 9 中随机地 (可重复) 抽 5 个数字, 则抽出的 5 个数字都不相同的概率是

$$p = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{10^5} = 0.3024$$

## 无理数的随机特征

考虑无理数  $e = 2.71828 \dots$ .

将  $e$  的前 800 位小数分成 160 组, 每组 5 个数字视为从 0, 1, 2, ..., 9 中随机抽出. 数得 5 个数字都不相同的共有 52 组, 其频率为

$$\frac{52}{160} = 0.325 \approx 0.3024.$$

这仍然只是一个猜想! (Normal numbers)

## 例题

某接待站在某周接待了 12 次来访, 已知这 12 次来访都是在周二和周四进行的. 问是否可以推断接待站的接待时间是有规定的?

## 解

假设接待站的时间没有规定, 且认为来访者每周任一天到达是等可能的. 则

$$P(\{12 \text{ 次来访都在周二和周四}\}) = \frac{2^{12}}{7^{12}} = 3 \times 10^{-7}.$$

## 定义

概率非常小的事件, 称为小概率事件.

实际推断原理: 小概率事件在一次试验中是几乎不可能发生的.

由实际推断原理, 可推断接待站的时间是有规定的.

## 古典概型的特点

有限样本点, 等可能性

## 问题

如何将样本空间推广到“无限”个样本点, 同时又有某种“等可能性”?

# 几何概型

## 定义

**随机试验** 向平面有界区域投掷一个点

**样本空间**  $\Omega$

**事件** 点落在可测量面积的平面区域  $A$

**事件概率**  $P(A) = \frac{A \text{ 的面积}}{\Omega \text{ 的面积}}.$

称上述试验为几何概型.

## 注记

- 事件  $A$  发生的概率与位置无关, 只与  $A$  的面积有关, 这体现了某种“等可能性”.
- 如果样本空间为有界区间、空间有界区域 (或更高维的几何体), 则“面积”改为“长度”、“体积” (或高维推广).



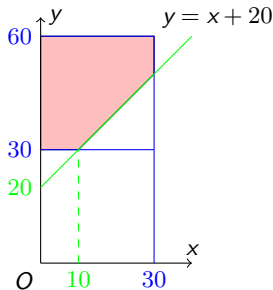
## 例题

在一次演习中, 某部队  $A$  接到命令要赶到某小河  $D$  岸为行进中的  $B$  部队架设浮桥. 假设  $A$  部队将于 7 点到 7 点 30 分之间到达  $D$  岸, 架桥需要 20 分钟时间;  $B$  部队将于 7 点 30 分至 8 点到达  $D$  岸. 试求  $B$  部队到达  $D$  岸时能立即过河的概率.

## 解

用  $x, y \in [0, 60]$  表示  $A$  与  $B$  部队到达  $D$  岸的时间, 则  $(x, y) \in \Omega = [0, 30] \times [30, 60]$ .  $B$  部队能立即过河的充要条件是  $x + 20 \leq y$ . 我们用几何概型求解, 所求概率为

$$p = \frac{30^2 - 20^2/2}{30^2} = \frac{7}{9}.$$



## 例题

三个人去参加某个集合的概率均为 0.4, 其中至少有两个参加的概率为 0.3, 都参加的概率为 0.05. 求 3 人中至少有一个参加的概率.

## 解

以  $A, B, C$  分别记 3 个人参加的事件. 由条件知,

$$P(A) = P(B) = P(C) = 0.4, \quad P(AB \cup CA \cup BC) = 0.3, \quad P(ABC) = 0.05.$$

由容斥原理,

$$P(A \cup B \cup C) = 0.4 \times 3 - 0.3 + 0.05 = 0.85.$$

## 例题

设甲、乙两人同时向同一目标进行射击. 已知甲击中的概率为 0.7, 乙击中目标的概率为 0.6, 两人同时击中目标的概率为 0.4. 求

- (1) 目标不被击中的概率;
- (2) 甲击中目标而乙未击中的概率.

## 解

记事件  $A, B$  为甲、乙击中目标的事件. 于是由条件知

$$P(A) = 0.7, \quad P(B) = 0.6, \quad P(AB) = 0.4.$$

于是,

$$P(\text{目标不被击中}) = P(\bar{A}\bar{B}) = P(\overline{A \cup B}) = 1 - [P(A) + P(B) - P(AB)] = 0.1.$$

$$P(\text{甲击中而乙未击中}) = P(A\bar{B}) = P(A) - P(AB) = 0.7 - 0.4 = 0.3.$$

## 伯努利信封问题

$n$  封信随机送到  $n$  个人手中. 求所有人都没收到自己的信的概率.

### 解

用  $A_i, i = 1, 2, \dots, n$  表示第  $i$  个人拿到自己的信的事件. 于是由容斥原理, 至少一个人拿到信的概率为

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} A_{i_2}) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n).$$

由乘法原理, 我们有  $P(A_{i_1} A_{i_2} \dots A_{i_k}) = \frac{(n-k)!}{n!}$ . 这样的项一共有  $C_n^k$  项. 所以前述概率为

$$\sum_{k=1}^n (-1)^{k-1} C_n^k \frac{(n-k)!}{n!} = \sum_{k=1}^n (-1)^{k-1} \frac{1}{k!}.$$

故没有人拿到自己的信的概率为

$$1 - \sum_{k=1}^n (-1)^{k-1} \frac{1}{k!} = \sum_{k=0}^n (-1)^k \frac{1}{k!} \rightarrow e^{-1}, \quad n \rightarrow \infty.$$

# 作业

- P21: 28, 29
- 补充题:
  - ① 从  $n$  双尺码不同的鞋子中任到  $2r$  ( $2r < n$ ) 只, 求下列事件的概率:
    - a 所取  $2r$  只鞋子中没有两只成对;
    - b 所取  $2r$  只鞋子中只有两只成对;
    - c 所取  $2r$  只鞋子恰好配成  $r$  对.
  - ② (匹配问题) 将 4 把能打开 4 间不同房门的钥匙随机发给 4 个人. 试求至少有一人能打开门的概率.