公 选 课　可 计 量 的 社 会

概率的定义与公理

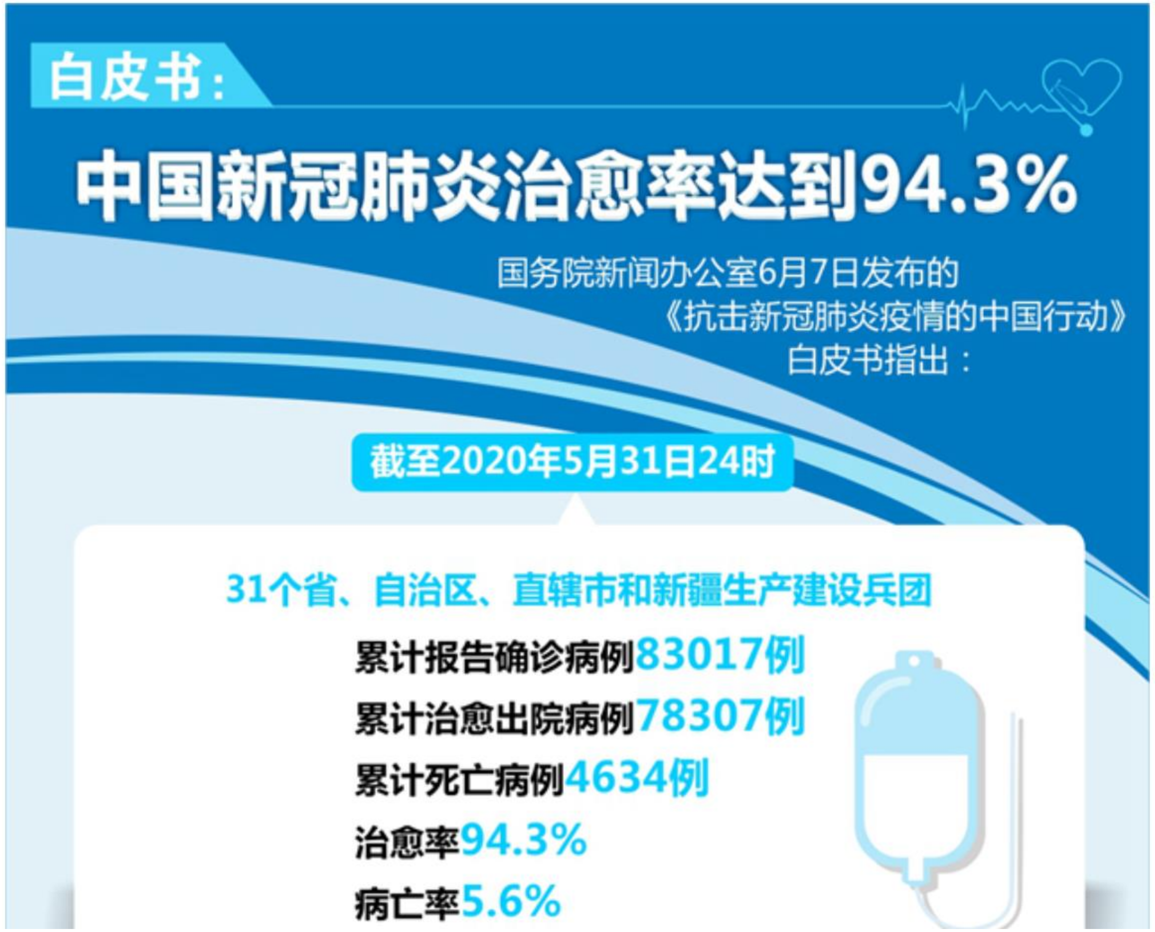
何 濯 羽　二 零 二 四 年 九 月 廿 三 日

生活中的概率

成都
28℃
体感 28℃
阴 23 ~ 30℃
30
28
59%
现在
17:00
19:00
20:07
21:00
07/15 昨天 23 31
07/16 今天 82% 23 30
07/17 明天 77% 23 28
07/18 周四 72% 23 30
07/19 周五 74% 23 30

白皮书：中国新冠肺炎治愈率达到94.3%

来源： 新华社　　字号：默认 大 超大 | 打印



生活中的概率

国内版 国际版

Microsoft Bing

浙江省地震的概率

WEB WORK IMAGES VIDEOS ACADEMIC DICT MAPS MORE TOOLS

About 4,680,000 results

发生大地震的可能性不大

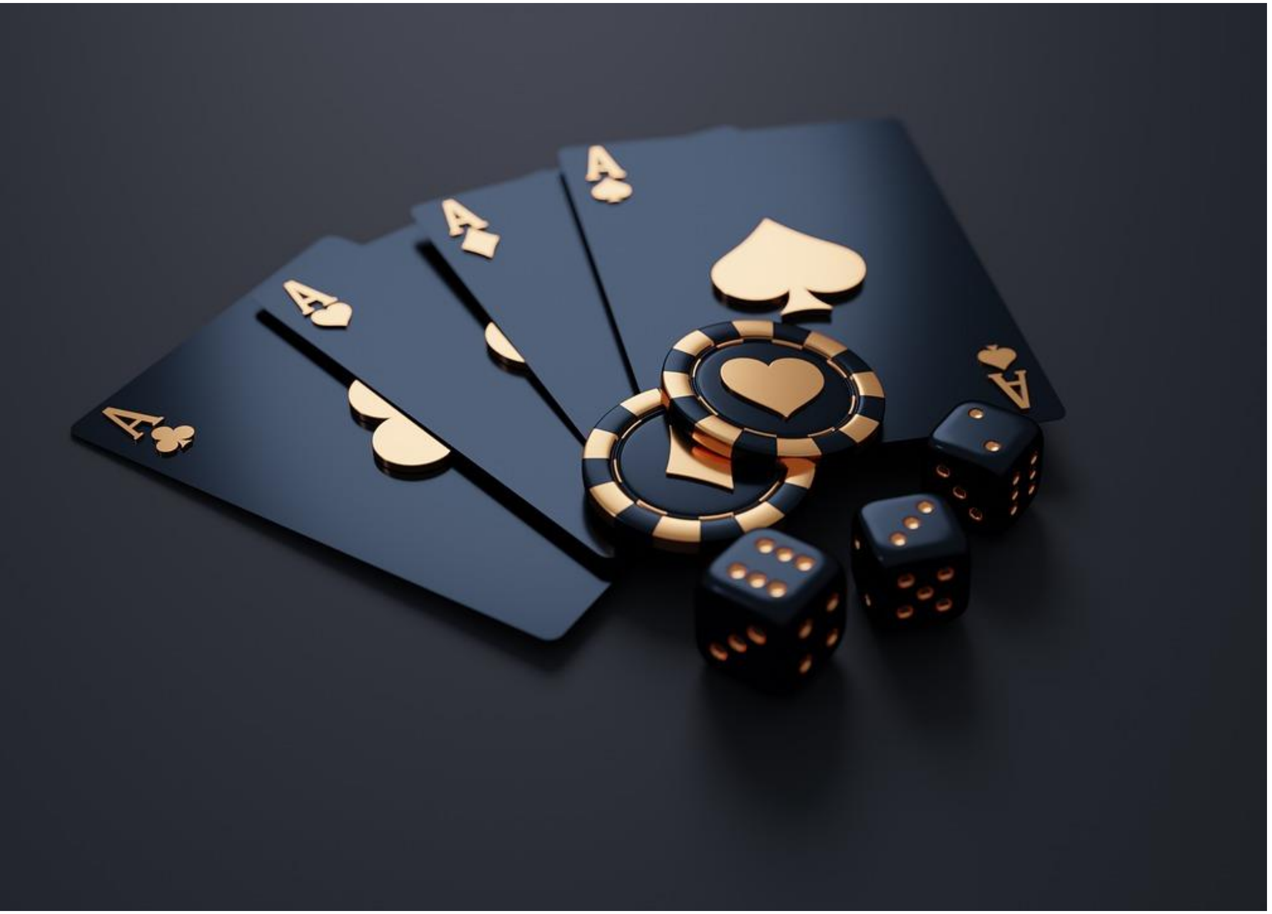
“浙江属于长江中下游南黄海地震带和华南沿海地震带，但这两个地震带在浙江境内的活跃度都比较低。”浙江省地震局高级工程师石树中告诉钱江晚报·小时新闻记者，浙江境内整体地震构造相对稳定，**发生大地震的可能性不大**。近代以来，浙江尚未发生过5级以上地震。不过，在历史上，浙江省11个地市都曾发生过4级以上地震。

浙江近代以来无5级以上地震，两个地方是相对危险区域

new.qq.com/rain/a/20210522A07IIL00

Was this helpful?

目录



Section 1

概率的定义

古典派 VS 频率派 VS 主观派

概率的定义

概率是什么？ What is *probability*?

这不仅是一个数学问题；这也是哲学问题。

根据针对“概率”的定义方式，我们可以将统计学家们分为三个主流学派：

- 古典派（古典概率）
- 频率派（试验概率）
- 主观派（主观概率）

古典学派

The probability of an event is the ratio of the number of cases favorable to it, to the number of all cases possible when nothing leads us to expect that any one of these cases should occur more than any other, which renders them, for us, equally possible.

—— Pierre-Simon Laplace, *Théorie Analytique des Probabilités* (1812)

Jacob Bernoulli
(1655-1705)

Pierre-Simon Laplace
(1749-1827)

古典派认为：

未知的概率均为“等概率”。

例如，因为我们不知道掷出一枚普通硬币后哪一面更可能朝上，所以我们应该赋予正面和反面相同的概率，即

$$\Pr(\text{正面朝上}) = \Pr(\text{反面朝上}) = \frac{1}{2}$$

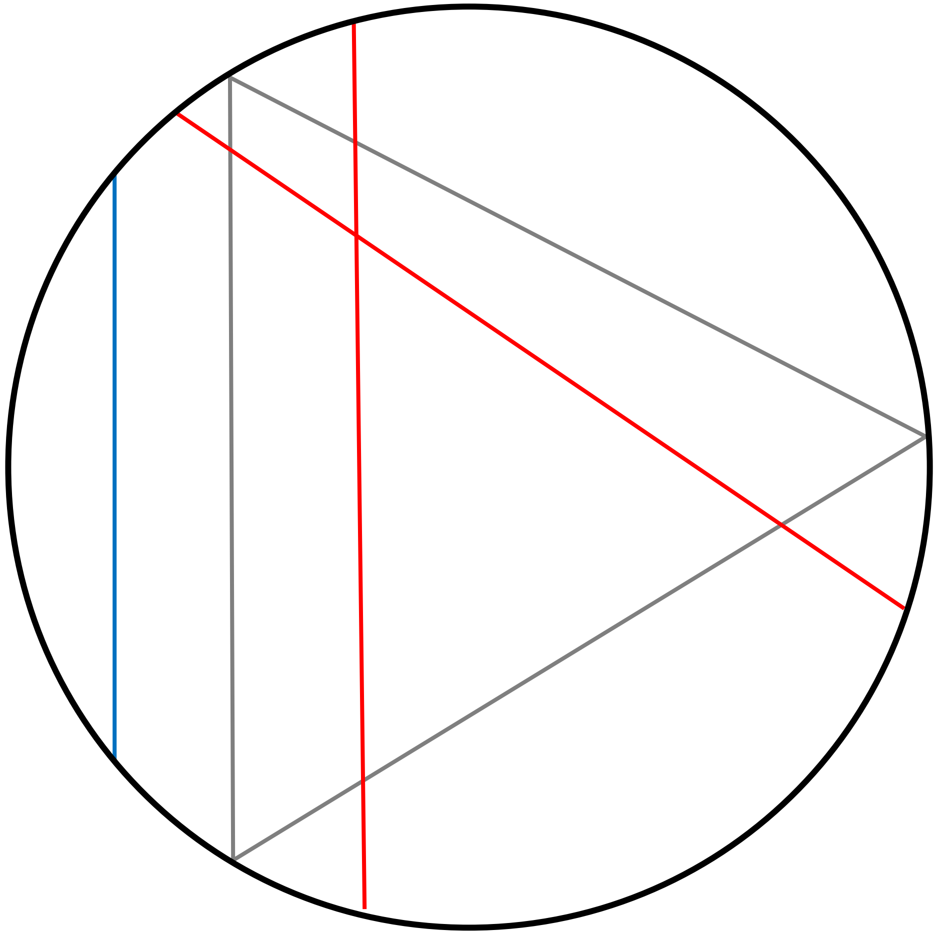
缺 陷

古典概率（classical probability，或称“古典概型”）的**缺陷**：

- “未知的概率均为等概率”是循环定义，并没有回答“概率是什么？”
- 无法分析已知的非等概率的情况。例如：已知这是一枚不均匀的硬币，那么正面朝上的概率是多少?
- 贝特朗悖论（Bertrand’s paradox）

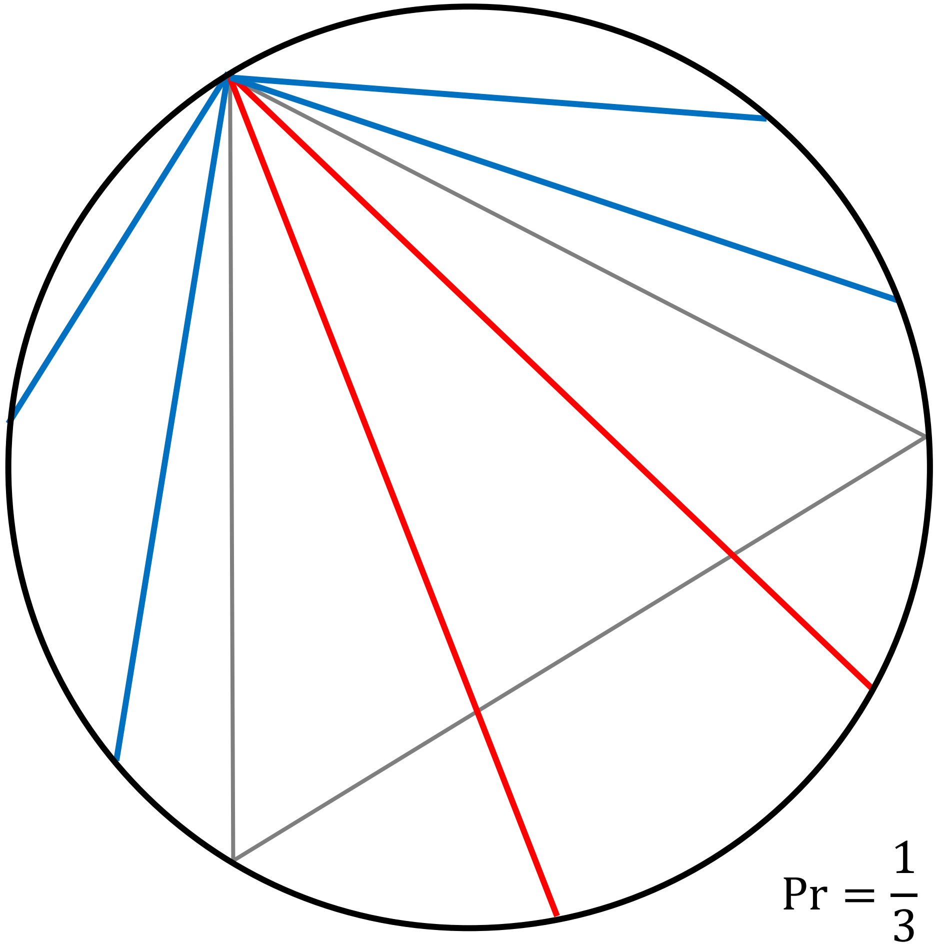
贝特朗悖论

问：在圆上任意画一条弦，这条弦的长度超过圆的内接正三角形的边长的概率是多少？

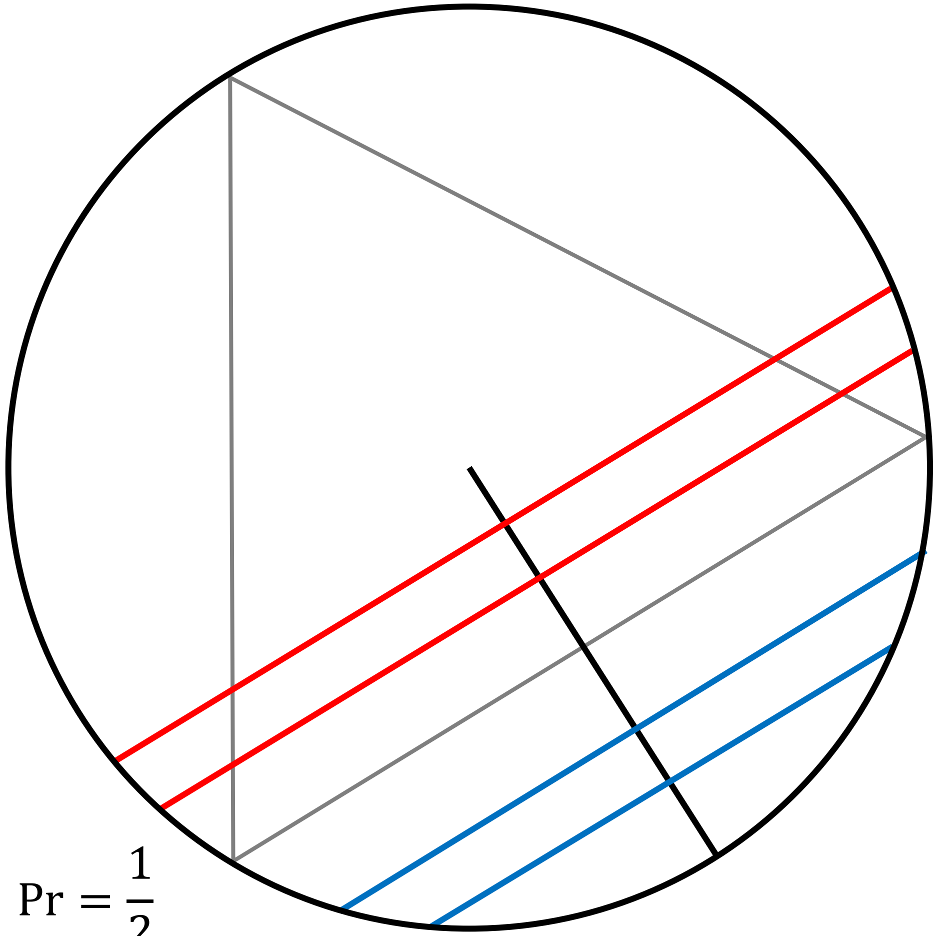


贝特朗悖论

问：在圆上任意画一条弦，这条弦的长度超过圆的内接正三角形的边长的概率是多少?



$$\Pr = \frac{1}{3}$$



$$\Pr = \frac{1}{2}$$

频率学派

The probable is that which for the most part happens.

—— Aristotle, *Rhetoric* (4th century BCE)

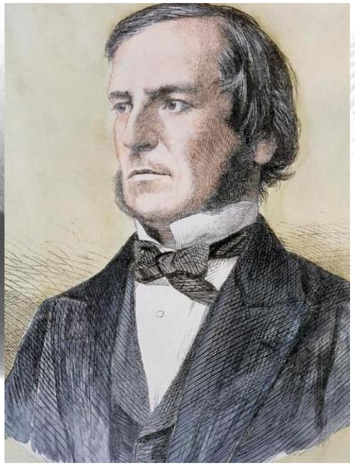
19世纪30年代，频率派崛起，开始对古典概率进行全面轰炸。

频率派认为：我们只能通过记录试验结果，然后计算各种结果的频率，去**估计**概率。当试验次数增大到无穷大时，各种结果的频率无限逼近它们各自的概率，即“**频率存在稳定性**”。

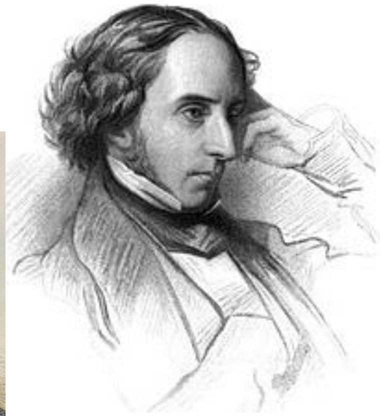
Ronald Aylmer Fisher
(1890-1962)

Jakob Friedrich Fries
(1773-1843)

Antoine Augustin Cournot
(1801-1877)

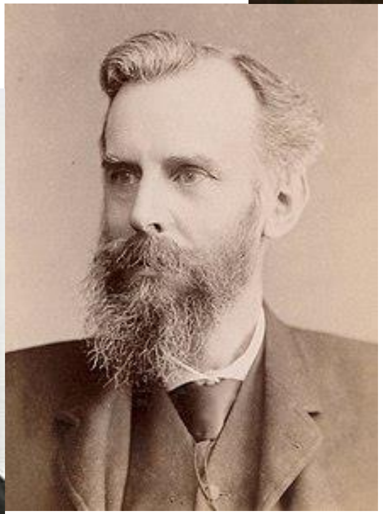


George Boole
(1815-1864)



Robert Leslie Ellis
(1817-1859)

Joseph Bertrand
(1822-1900)



John Venn
(1834-1923)

频率学派

The probable is that which for the most part happens.

—— Aristotle, *Rhetoric* (4th century BCE)

频率派的观点有三个基本特征：

- **试验性**：我们必须通过试验才能计算频率。
- **重复性**：每次重复试验的程序必须完全相同。
- **误差性**：频率只是概率的估计值，存在误差。

例如，一枚硬币落地时正面朝上的概率是

$$\Pr(\text{正面朝上}) = \boxed{\lim_{n\to\infty} \mathrm{Fr}_n(\text{正面朝上})}$$

投掷 ∞ 次相同硬币后计算得到的正面朝上的频率

频率学派

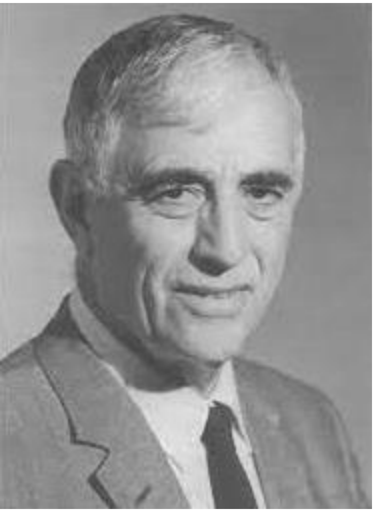
The probable is that which for the most part happens.

—— Aristotle, *Rhetoric* (4th century BCE)

试验概率（frequentist probability）的**缺陷**：

- X 为了准确估计概率，我们需要重复试验 ∞ 次。现实中，∞ 次如何实现?
- X 每次试验的过程必须完全相同，这在现实中难以实现。例如，每次抛掷硬币时，硬币沾上的手汗都不相同。
- X 一些事件无法进行重复试验。例如，如何估计富士火山喷发的概率?

主观学派


Bernard Osgood Koopman
(1900-1981)

Frank Plumpton Ramsey
(1903-1930)


主观派以个人信念强度（degree of belief）为基础，认为“概率”是人们根据有效的证据对某一事件发生的可能性的**主观评价**。简言之，事件 A 发生的概率等于个人对事件 A 发生的信念强度。

例：高考咨询会上，家长问我480分能不能上成都东软学院，我回答：我个人觉得您孩子考上成都东软学院的概率为80%。

主观学派

主观学派提出了**贝叶斯概率**（Bayesian probability），强调了概率并非固定值，而会随着获取信息的增加而变化。

$$\Pr(\text{李主任是男性}) = \frac{1}{2} \xrightarrow{\text{信息：李主任从不穿裙子}} \Pr(\text{李主任是男性}|\text{李主任从不穿裙子}) = ?$$

新信息对概率的影响反映了个人的主观想法

注 主观学派并非支持纯粹的主观臆断，而是强调符合逻辑的个人信念在概率计算中扮演着至关重要的角色！

主观学派

主观概率（subjective probability）的**缺陷**：

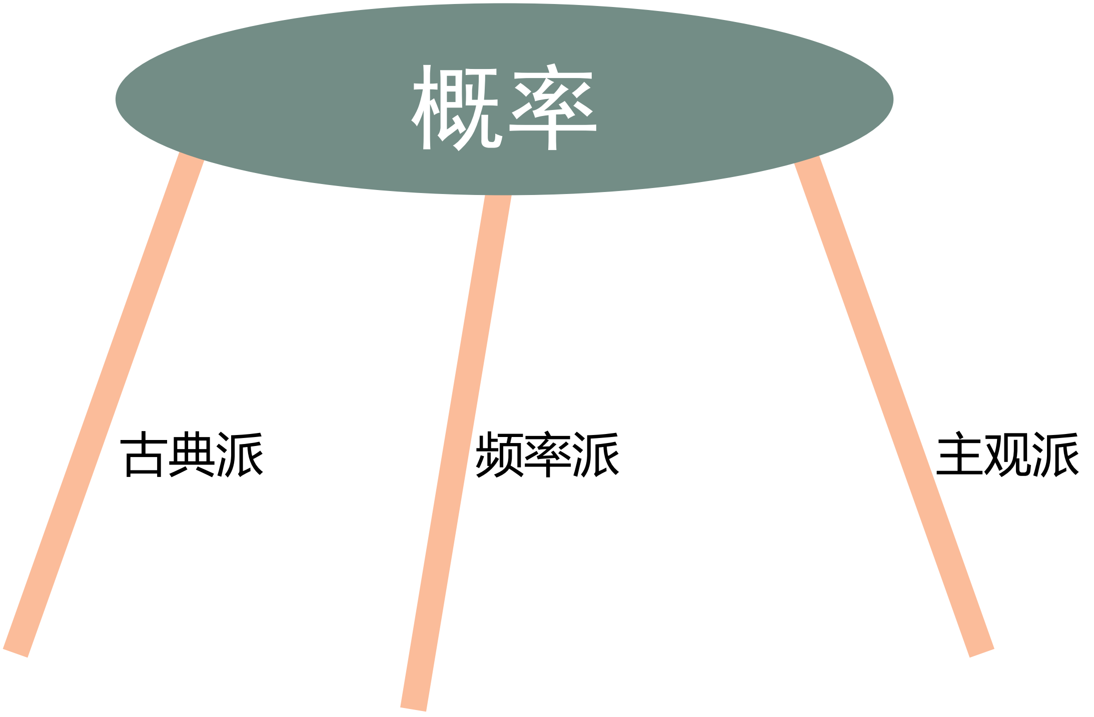
- 主观概率充满了个人偏见，与科学的客观性相冲突。
- 人们很难对某个主观概率达成共识。

概率的定义

三个主流概率学派的分歧是“解释上”的分歧，而非“数学上”的分歧。

实际上，概率论与统计学发展至今，三个学派已经不再泾渭分明。

例如，我们在日后将要学习的线性回归模型与假设检验中，既可以看见频率派的身影（使用实际数据），也可以看见主观派的思想（选择置信度）。

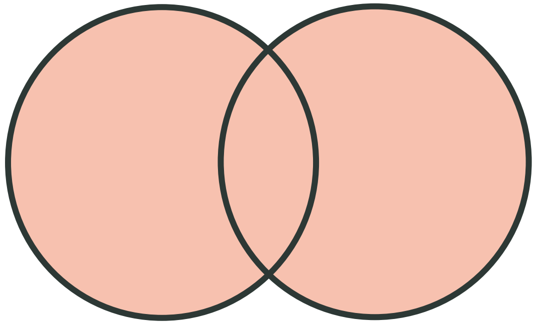


Section 2

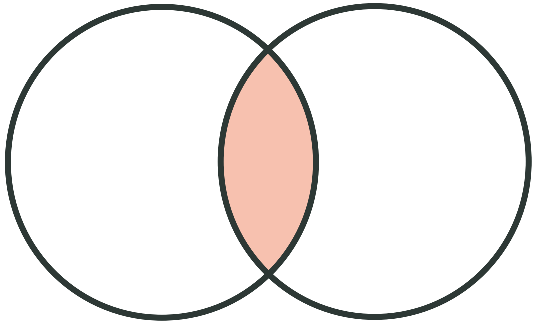
概率论的基石

试验 & 结果 & 样本空间 & 事件 & 概率函数

集合关系的复习

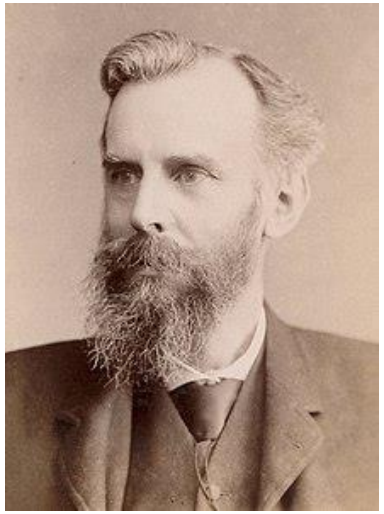


$A \cup B$



$A \cap B$

韦恩图



John Venn
(1834-1923)

如果集合 A 与集合 B 满足 $A \cap B = \emptyset$，我们称集合 A 与集合 B 为**不相交集合**（**disjoint sets**，也称“并查集”）。如果集合 $A_1, A_2, \dots, A_n$ 满足 $A_i \cap A_j = \emptyset, \forall i \neq j$，我们称集合 $A_1, A_2, \dots, A_n$ 为**不相交集合。**

不相交集合

A_1	A_2	A_3	A_4	……	A_{n-1}	A_n

概率论的基石

- **试验**（experiment）：可以在相同条件下重复进行且能够产生随机现象的行为。

- **结果**（outcome）：一次试验产生的特定现象。

- **样本空间**（sample space）：一个以某个试验所有的可能结果为元素的集合，一般用拉丁字母 S 或希腊字母 Ω 表示。

- **事件**（event）：样本空间的任意子集，一般用大写拉丁字母（如 A、B、C 等）表示。

例 将投掷一颗均匀的骰子当作“试验”，落地后朝上的点数为“结果”。那么，样本空间为

$$S = \{1, 2, 3, 4, 5, 6\}$$

针对该样本空间，我们有多少事件?

$$N = \mathrm{C}_6^0 + \mathrm{C}_6^1 + \mathrm{C}_6^2 + \mathrm{C}_6^3 + \mathrm{C}_6^4 + \mathrm{C}_6^5 + \mathrm{C}_6^6 = 2^6 = 64$$

具体如下:

$$\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\},$$

$$\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{1, 6\}, \{2, 3\}, \ldots$$

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}, \{1, 3, 4\}, \ldots$$

$$\{1, 2, 3, 4\}, \{1, 2, 3, 5\}, \{1, 2, 3, 6\}, \{1, 2, 4, 5\}, \ldots$$

$$\{1, 2, 3, 4, 5\}, \{1, 2, 3, 4, 6\}, \{1, 2, 3, 5, 6\}, \{1, 2, 4, 5, 6\}, \ldots$$

$$\{1, 2, 3, 4, 5, 6\}$$

注意：空集是所有集合的子集，空集代表着“**不可能事件**”。

事件的关系

事件（event）：样本空间的子集，一般用大写拉丁字母（如 A、B、C 等）表示。

互斥事件（disjoint events）：当事件 A 和事件 B 不可能同时发生时，我们称 A 和 B 为互斥事件。互斥事件的数学定义为

$$A \cap B = \emptyset$$ **集合 A 和集合 B 不相交**

因此，互斥事件有如下特点：

$$\Pr(A \cup B) = \Pr(A) + \Pr(B)$$

互斥事件的相反情况是“相容事件”——即事件 A 和事件 B 有可能同时发生。

概率论的基石

- **试验**（experiment）：可以在相同条件下重复进行且能够产生随机现象的行为。
- **结果**（outcome）：一次试验产生的特定现象。
- **样本空间**（sample space）：一个以某个试验所有的可能结果为元素的集合，一般用拉丁字母 S 或希腊字母 Ω 表示。
- **事件**（event）：样本空间的任意子集，一般用大写拉丁字母（如 A、B、C 等）表示。
- **概率函数**（probability function）：自变量为事件，因变量为实数，并且满足**概率三大公理**的函数，一般用 P( ) 或 Pr( ) 表示。

Section 3

概率的三大公理

非负性 & 归一性 & 可加性

公理（axiom）是推导出一套特定演绎知识的基本假设。公理**不证自明**。其它所有的命题（例如定理、推论、法则）都必须借助这些基本假设才能被证明。

概率 三大公理

① 非负性

$\Pr(A) \geq 0, \forall A \subseteq S$

② 归一性

$\Pr(S) = 1$

③ 可加性

如果 $A_1 \cap A_2 = \emptyset$，那么

$\Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2)$

概率

三大公理

① **非负性**：$\Pr(A) \geq 0, \forall A \subseteq S$

② **归一性**：$\Pr(S) = 1$

③ **可加性**：如果 $A_1 \cap A_2 = \emptyset$，那么 $\Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2)$

如果 $\cap_{i=1}^{\infty} A_i = \emptyset$，那么 $\Pr(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{n} \Pr(A_i)$

根据概率三大公理，我们可以推导出：

- $\Pr(A) \leq 1, \forall A \subseteq S$；
- $\Pr(\emptyset) = 0$；
- $\Pr(A) = 1 - \Pr(\complement_S A), \forall A \subseteq S$；
- 如果 $A \subseteq B$，那么 $\Pr(A) \leq \Pr(B)$；
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$；
- 布尔不等式（Boole's Inequality）：$\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$；
- 邦佛洛尼不等式（Bonferroni's Inequality）：$\Pr(A \cap B) \geq \Pr(A) + \Pr(B) - 1$

感谢聆听

THANK YOU!