
古典概型与几何概型

第 2 讲 · 概率论与数理统计

复旦大学经济学院 ECON130001

上一讲我们做了什么

把「事件」变成了集合

样本空间 Ω 是全部可能结果，事件是它的子集；并、交、差、补就是事件之间的运算。一句话里的歧义，写成集合就消失了。

概率的三条公理

非负性、 $P(\Omega) = 1$ 、可列可加性。六条性质全部由它们推出，没有额外假设。

★ 留下的问题

三条公理只规定了概率**要满足什么约束**，没有规定**基本事件的概率是多少**。掷一枚骰子， $P(\{1\}) = 1/6$ ——这个数不是从公理推出来的，是我们另外定的。

为什么需要这一讲

背景 2016 年 3 月，AlphaGo 对李世石。赛前，围棋界和不少 AI 研究者的判断是——**计算机还要十年才能赢职业九段。**

转折 理由听上去无懈可击：国际象棋能被攻克（1997 年深蓝赢卡斯帕罗夫），是因为合法局面只有约 4.8×10^{44} 种，机器搜得动。而围棋的合法局面约 2.08×10^{170} 种——**比可观测宇宙的原子总数（约 10^{80} ）还多得多。**

这个推理没有错。 10^{170} 个局面，再快的机器也数不完——这不是算力不够，是**数量级根本不在一个层面上。**

问题 可 AlphaGo 赢了 4:1。它没有把局面数完，而是**随机试一部分，用赢的比例去估**——**上一讲的布丰投针也是这么干的。**

这一招有个前提：每种情况被抽到的机会一样。可「一样」是对什么说的？换一种抽法，同一道题会不会算出不同的概率？

来源

围棋 19×19 合法局面数 2.08×10^{170} ：John Tromp & Gunnar Farneback, *Combinatorics of Go*, 2016 年 1 月完成计算 (tromp.github.io/go/legal.html)。国际象棋合法局面数 4.8×10^{44} ：同作者 2021 年 7 月计算结果。可观测宇宙原子数 10^{80} 为天体物理学常用量级估计。访问日期 2026-08-01。

本讲学习目标

- ① 在**结果有限且等可能**的场合，用四种计数工具（加法、乘法、排列、组合）把概率算成一个除法
- ② 结果有**无限多个**时改用**几何概型**：分子分母由「个数」换成长度、面积、体积
- ③ 说清这两种模型共用的那条假设——**等可能**——它从哪来，以及为什么它可能不唯一

古典概型：把概率变成除法

定义

当试验的基本事件**有限个**，且每个基本事件出现的**可能性相同**时，称为**古典概型**，其样本空间称为**简单样本空间**。

Ω 含 n 个结果，每个结果概率 $1/n$ 。事件 A 含 m 个结果，则

$$P(A) = \frac{\text{有利于 } A \text{ 的基本事件数}}{\text{基本事件总数}} = \frac{\#A}{\#\Omega} = \frac{m}{n}$$

这就是上一讲缺的那块

「等可能」是一条额外的假设，不是从公理推出来的。 一旦承认它，每个基本事件的概率就只能是 $1/n$ ；有了这 n 个数，任何事件的概率都能由可加性算出来。

例题：从袋中取两个球

例

袋内装有 5 个白球、3 个黑球。任取两个球，求两个都是白球的概率。

解

样本空间：从 8 个球中取 2 个， $\#\Omega = C_8^2 = 28$ 。

事件 A ：从 5 个白球中取 2 个， $\#A = C_5^2 = 10$ 。

$$P(A) = \frac{10}{28} = \frac{5}{14} \approx 0.357$$

难点不在公式，在计数

$P(A) = \#A/\#\Omega$ 只有一行。真正花时间的是**把 $\#A$ 和 $\#\Omega$ 数对**——而且两者必须落在**同一个**样本空间上。所以接下来四页讲怎么数。

计数工具一：分类加法

分类加法计数原理

完成一件事有 n 类方法，第 i 类有 m_i 种，**每一类都能单独把事情办成**，则

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

例

北京到沈阳可乘飞机、火车、长途汽车。一天有 4 个航班、3 列火车、5 趟汽车，共有多少种走法？

$4 + 3 + 5 = 12$ 种。

识别关键词

「**或**」——选了这一类就不用另一类。对应上一讲的**互斥事件之并**。

计数工具二：分步乘法

分步乘法计数原理

完成一件事分 n 步，第 i 步有 m_i 种选择，**每一步都做完才算办成**，则

$$N = m_1 \times m_2 \times \cdots \times m_n$$

例（两种原理混用）

北京到沈阳 4 个航班，沈阳到哈尔滨 5 个航班，北京直飞哈尔滨 3 个航班。北京到哈尔滨有几种走法？

$$\underbrace{4 \times 5}_{\text{转机：分两步}} + \underbrace{3}_{\text{直飞}} = 23 \text{ 种。}$$

识别关键词

「**且**」 / 「**然后**」——每一步都要走。这道题两个原理都用了：先按「转机还是直飞」分类，转机那一类内部再分步。

计数工具三：排列（有重复）

例

罐中装有编号 1 至 8 的小球，摸出一个记下球号后**放回**。摸 4 次并依次记录，最多得到多少种球号排列？

$8^4 = 4096$ 种。

有重复的排列数

从 n 个不同元素中取 m 个排列，**允许重复**，总数为

$$n^m$$

为什么是乘法

每一次摸球是一「步」，每步都有完整的 8 种选择——放回意味着 **前一步不影响后一步**。这正是分步乘法原理。

计数工具四：排列（无重复）

例

邮政编码由 6 位数字组成，每位可取 $0, 1, \dots, 9$ 。最多能编出多少个**各位数字互不相同**的邮编？

$$\frac{10!}{(10-6)!} = 10 \times 9 \times 8 \times 7 \times 6 \times 5 = 151200 \text{ 种。}$$

排列数

从 n 个不同元素中取 m ($m \leq n$) 个，**按顺序**排成一行，总数为

$$A_n^m = \frac{n!}{(n-m)!}$$

和上一页只差一个字

不放回：选走一个，下一步就少一个可选。 $10 \times 9 \times 8 \dots$ 而不是 $10 \times 10 \times 10 \dots$ 。

计数工具五：组合（不计顺序）

组合数

从 n 个不同元素中取 m ($m \leq n$) 个，**不论次序**构成一组，总数为

$$C_n^m = \frac{n!}{(n-m)!m!} = \frac{A_n^m}{m!}$$

为什么要除以 $m!$

同样一组 m 个元素，排成一行有 $m!$ 种顺序，而组合把它们算作同一种。所以组合数就是排列数除以 $m!$ 。

例

30 人的班级选 5 人组成班委，共有多少种选法？

$$C_{30}^5 = \frac{30!}{25!5!} = \frac{30 \times 29 \times 28 \times 27 \times 26}{120} = 142506 \text{ 种。}$$

怎么判断该用排列还是组合

问一句：**换个顺序，还是同一件事吗？** 班委不分先后 → 组合；邮编换个顺序就是另一个号 → 排列。

小结：四种计数工具怎么选

问自己	答	用
分类还是分步？	分类（「或」）	加法 $m_1 + m_2 + \cdots$
	分步（「且」）	乘法 $m_1 \times m_2 \times \cdots$
取了放不放回？	放回	n^m
	不放回， 讲顺序	$A_n^m = \frac{n!}{(n-m)!}$
	不放回， 不讲顺序	$C_n^m = \frac{n!}{(n-m)!m!}$

用哪个工具都要再查一遍： $\#A$ 与 $\#\Omega$ 必须数在同一个样本空间上 —— 一个讲顺序、一个不讲顺序，比出来的就是错的。

应用：AI 怎么判断一段股评是利好还是利空

词袋模型 (Bag of Words)

把一篇文章看成一个装满词的袋子，**暂时不管语法和顺序**——注意，「不管顺序」就是在说：这是个**组合**问题，不是排列问题。

怎么算

先在海量语料上**统计**：「上涨」「盈利」这些词在利好评论里出现得多频繁。判断新评论时算

$$P(\text{利好} \mid \text{评论}) \propto \frac{\text{含这些词的利好评论数}}{\text{利好评论总数}}$$

看清它的骨架

右边这个分式，就是古典概型的 $\#A/\#\Omega$ 。**看起来最现代的文本分析，起点是数数。**（左边那个条件概率的正式算法在第 3 讲。）

一次把计数用到底：扑克牌型

统一设定

从一副 52 张牌中随机抽 5 张。抽牌**不计顺序**，所以

$$\#\Omega = C_{52}^5 = 2\,598\,960$$

这个数是所有牌型的**公共分母**，只要算一次。

接下来只需要数分子

八种牌型，八个 $\#A$ 。八次都是同一套动作：**先定点数，再定花色，最后减掉重复计入的更大牌型。**



牌型从大到小

先算两个极端

最难的：同花顺

五张连号且同花色。

$$\frac{10 \times 4}{C_{52}^5} = \frac{40}{2598960} \approx 1.54 \times 10^{-5}$$

10：起始点数（A-2-3-4-5 到 10-J-Q-K-A）。

4：花色四选一。

最容易的：一对

一个对子 + 三张互不相同的散牌。

$$\frac{C_4^2 \times 13 \times C_{12}^3 \times 4^3}{C_{52}^5} = \frac{1098240}{2598960} \approx 0.423$$

13 定对子点数， C_4^2 定其花色，
 C_{12}^3 定三张散牌的点数， 4^3 定它们的花色。

差了**四个数量级**。中间还有六种牌型——它们的算法完全一样，下一页交给你们自己在工具里拆。

交互：扑克牌型概率计算器

怎么用

点左侧任一牌型，右边会把计数式**逐个因子**拆开，并说明每个因子在数什么。

重点看这两个

- **同花与顺子**：为什么要减掉 40
- **两对**：为什么是 C_{13}^2 而不是 13×12



交互演示：选牌型，看计数式逐因子拆解

https://fin-tech.fun/courseware/probability-statistics-undergrad/toolkit/poker_odds.html

八个牌型的结果，与两处最容易错的地方

牌型	#A	概率
同花顺	40	0.0015%
四条	624	0.024%
葫芦	3 744	0.144%
同花	5 108	0.197%
顺子	10 200	0.392%
三条	54 912	2.11%
两对	123 552	4.75%
一对	1 098 240	42.3%

① 同花与顺子都要减 40

「五张同花色」里混着 40 个同花顺，而同花顺是更大的牌型。不减就**重复计数**了——这是上一讲互斥没处理好。

② 两对不能写成 $13 \times C_4^2 \times 12 \times C_4^2$

那样会把「A 对 + K 对」和「K 对 + A 对」**数两遍**。两个对子没有先后，选点数必须用 C_{13}^2 。

应用：随机选股的组合概率

例

股票池有 10 只科技股、8 只金融股。**完全随机**选 5 只构成组合，求恰好选出 3 只科技股和 2 只金融股的概率。

解

不考虑选取顺序，所以用组合：

$$\# \Omega = C_{18}^5 = 8568, \quad \# A = C_{10}^3 \times C_8^2 = 120 \times 28 = 3360$$

$$P(A) = \frac{3360}{8568} \approx 0.392$$

注意 $\#A$ 用的是**乘法**：从科技股里选 3 只 **且** 从金融股里选 2 只，是分两步走。

这题的现实含义

「随机选」是一个**基准**。基金经理主动选出的组合若在这个分布里毫不出奇，那他的选股能力就无从谈起——这是业绩归因的基本思路。

应用：一个 6 位密码能扛多久

例

密码 6 位，每位可以是 a-z、A-Z、0-9 中任意一个。黑客每秒能试 100 万个，暴力破解平均需要多久？

解

字符集 $26 + 26 + 10 = 62$ ，每位独立、可重复，属**有重复的排列**：

$$\#\Omega = 62^6 = 56\,800\,235\,584 \approx 568 \text{ 亿}$$

平均要试一半才成功：

$$t \approx \frac{56.8 \times 10^9 / 2}{10^6} \approx 28400 \text{ 秒} \approx 7.9 \text{ 小时}$$

568 亿其实撑不了多久

人对大数没有直觉： 62^6 写出来有十一位，可按每秒一百万次试，它连一天都撑不到。

真正管用的是**位数**——每多一位，可能的密码数就乘以 62：7 位约 20 天，8 位约 3.5 年。**密码强度靠的是指数增长，不是这个数看起来大不大。**

接下来

到这里，只要样本空间是**有限**的，我们都能算了——数分子、数分母、相除。

但如果结果有无穷多个呢？「在 $[0, 1]$ 上随机取一个数」，基本事件有不可数无穷个，每个的概率都是 0。分母是无穷，这个除法做不下去了。

先试两道，看该换成什么

例 1

在 $[0, 1]$ 上随机取一个数，求它小于 0.5 的概率。

$P = 0.5$ 。—— 不是数个数，是**比长度**。

例 2

在 $[0, 1]$ 上随机取**两个**数，求两数之差小于 0.5 的概率。

样本空间是单位正方形，事件是 $|x - y| < 0.5$ 的带状区域。

$$P = 1 - 2 \times \frac{1}{2} \times 0.5^2 = 0.75$$

—— 这次是**比面积**。

规律出来了：把「计数」换成「测量」。一维测长度，二维测面积， r 维测体积。

几何概型

先约定「测度」

R_r 是 r 维向量空间。对 R_r 的子集 A ，定义其测度

$$m(A) = \int_A dx_1 dx_2 \cdots dx_r$$

$r = 1$ 是长度， $r = 2$ 是面积， $r = 3$ 是体积。

几何概型的定义

设样本空间 $\Omega \subset R_r$ 的测度 $m(\Omega)$ 为正数，样本点**等可能**地落在 Ω 中。对 $A \subset \Omega$ ，称

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

为事件 A 发生的概率。

和古典概型是**同一个式子**：分子分母都换成了测度而已。「等可能」这条假设也原样保留下来——**它是不是总能说清楚，本讲后半会专门回来问一次。**

几何概型例题（一）：圆内落点

例

质点等可能地落在半径 1 m 的圆 Ω 中， A 是其中半径 0.5 m 的同心圆。求质点落在小圆内、小圆外的概率。

解

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{\pi \times 0.5^2}{\pi \times 1^2} = \frac{1}{4}$$

落在小圆外：由上一讲的对立事件， $P(\overline{A}) = 1 - \frac{1}{4} = \frac{3}{4}$ 。

一个直觉陷阱

半径缩一半，概率不是缩一半，是**缩到四分之一**——面积按半径的平方走。

几何概型例题（二）：会面问题

例

两人在 1:00 至 2:00 间**独立且随机**地到达某地会面，先到者等 20 分钟即离去。求两人能相遇的概率。

解：把「时刻」变成坐标

设两人到达时刻为 x, y （分钟）：

$$\Omega = \{(x, y) \mid 0 \leq x, y \leq 60\}, \quad A = \{(x, y) \mid |x - y| \leq 20\}$$

Ω 是 60×60 的正方形； A 的补集是两个直角边为 40 的三角形：

$$P(A) = \frac{60^2 - 40^2}{60^2} = \frac{2000}{3600} = \frac{5}{9} \approx 0.556$$

这道题的价值

题面里没有任何几何图形，**「面积」是我们自己造出来的**——把两个时刻当成平面上一个点的两个坐标。这一步才是几何概型真正难的地方。

几何概型例题（三）：再练两道

例

在 $[0, 1]$ 中随机取两点 x, y , 求 $x^2 + y^2 < 1$ 的概率。

Ω 是单位正方形（面积 1）, A 是四分之一单位圆：

$$P(A) = \frac{\pi/4}{1} = \frac{\pi}{4}$$

眼熟吗？ 这正是上一讲蒙特卡洛估计 π 的那个模型——反过来用就能测 π 。

例

两人在 5:00 至 6:00 间独立随机到达, 求**一人至少等另一人半小时**的概率。

$A = \{|x - y| \geq 30\}$, 是两个直角边为 30 的三角形：

$$P(A) = \frac{30^2}{60^2} = \frac{1}{4}$$

与例题（二）是同一张图的**另一块区域**——换个问法, 换块面积。

贝特朗悖论：同一道题，三个答案

题

在半径为 1 的圆内**任取一条弦**，求弦长 $\geq \sqrt{3}$ 的概率。

解法一：端点等可能

固定一端 x_0 ，另一端 y 等可能落在圆周上：

$$\Omega = [0, 2\pi), \quad A = \left[\frac{2\pi}{3}, \frac{4\pi}{3}\right]$$

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

解法二：中点等可能

弦的中点等可能落在圆内。弦长 $\geq \sqrt{3}$ 当且仅当中点落在半径 $1/2$ 的小圆盘内：

$$P(A) = \frac{\pi(1/2)^2}{\pi \cdot 1^2} = \frac{1}{4}$$

两个都没算错

还有第三种：**中点到圆心的距离**等可能落在 $[0, 1]$ 上，得 $P = \frac{1}{2}$ 。三种取法都自称「随机」，答案却是 $\frac{1}{3}$ 、 $\frac{1}{4}$ 、 $\frac{1}{2}$ 。

问题出在哪：回到开场

三个答案都对，因为它们回答的是三个不同的问题

「随机取一条弦」这句话**没有指定样本空间**。端点均匀、中点均匀、距离均匀——这是三个不同的 Ω ，自然给出三个不同的 P 。

所以「等可能」不是一句废话，是一个必须交代清楚的模型假设

古典概型说「每个基本事件等可能」，几何概型说「样本点等可能落在 Ω 中」——**这两句话里的「等可能」都必须先说清楚是对什么等可能**，否则概率根本没有定义。

回到 AlphaGo

开场问的是：什么叫「随机撒点」？现在有答案了——**撒法本身就是模型的一部分**。蒙特卡洛方法的全部技术难点，正在于设计一个采样分布，让估计值收敛到你真正想要的那个数。

应用：双寡头价格竞争

例

A、B 两家公司竞争一份合同，各自**独立且随机**地在 $[100, 200]$ （万元）内报价，**价低者得**。求：(1) A 中标的概率；(2) A 中标**且**报价高于 150 万的概率。

解

设 A 报 x 、B 报 y 。 Ω 是边长 100 的正方形， $m(\Omega) = 100^2 = 10000$ 。

(1) A 中标即 $x < y$ ，是对角线上方的三角形：

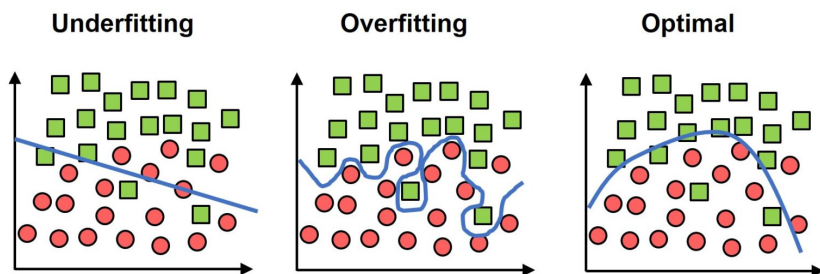
$$P = \frac{\frac{1}{2} \times 100^2}{10000} = \frac{1}{2}$$

——不用算也知道，两家对称。

(2) 加上 $x > 150$ ，区域缩成顶点为 $(150, 150), (150, 200), (200, 200)$ 的小三角形：

$$P = \frac{\frac{1}{2} \times 50^2}{10000} = \frac{1250}{10000} = \frac{1}{8}$$

应用：用面积看懂「过拟合」



扭曲的分类边界

例

一个模型在训练时把所有数据点完美分开了。现在来一个新数据点，随机出现在图中任意位置。**模型分错的概率是多少？**

用几何概型看

- Ω : 整个正方形区域
- A : 被模型划错的所有区域
- $P(A) = m(A)/m(\Omega)$, 就是那些扭曲部分的**总面积占比**

过拟合的一句话定义

边界越扭，训练集上越完美，**划错的面积却越大**。在训练数据上得分高，不等于在新数据上犯错少。

本讲总结

① 回到开场那个问题

围棋的局面数不完，AlphaGo 靠随机采样。而「随机」这个词本身需要被定义——贝特朗悖论证明了：**不说清对什么等可能，概率就没有值。**

上一讲说公理没有给出基本事件的概率；这一讲给了两种给法：等可能地数个数、等可能地量测度。**两种给法都建立在「等可能」这条假设上，而这条假设由建模的人指定，不由公理决定。**指定的对象不同，算出来的概率就不同。

② 核心结论

- 古典概型 $P(A) = \frac{\#A}{\#\Omega}$ ：适用前提是**结果有限且等可能**
- 四种计数：分类用加法、分步用乘法；放回 n^m 、不放回讲顺序 A_n^m 、不讲顺序 C_n^m
- 几何概型 $P(A) = \frac{m(A)}{m(\Omega)}$ ：同一个式子，计数换成测度

本讲总结（续）

③ 易错提醒

- $\#A$ 与 $\#\Omega$ 必须数在同一个样本空间上——一个讲顺序一个不讲，比出来就是错的
- 同一个结果被数了两次——分类时各类必须互不重叠，重叠了就要把重复的部分减掉
- 几何概型里，「等可能」是对哪一个量说的，必须先讲明——换一个量，同一道题会得到不同的答案

④ 下一讲

今天算概率的方式都是一次算完：先定下样本空间，数一次或量一次，答案就出来了。整个过程里，我们手上的信息自始至终没有变过。

可现实中信息是陆续到来的——三门问题里，主持人开门之后，那扇门的概率就变了。已经知道了一部分事实，剩下的概率该怎么更新？

下一讲：条件概率与事件的独立性。