§1.2 概率的定义及计算

历史上概率的三次定义

① 古典定义 —— 概率的最初定义

② 统计定义 —— 基于频率的定义

③ 公理化定义 —— 1930年后由前

苏联数学家柯尔莫哥洛夫给出

频率

设在 n 次试验中，事件 A 发生了m次，

则称 $f_n = \dfrac{m}{n}$ 为事件 A 发生的 **频率**

频率的性质

- $0 \le f_n(A) \le 1$ —— 非负性
- $f_n(\Omega) = 1$ —— 归一性
- 事件 A, B 互斥，则

  $f_n(A \cup B) = f_n(A) + f_n(B)$ —— 可加性

  可推广到有限个两两互斥事件的和事件
- $\lim\limits_{n \to \infty} f_n(A) = P(A)$ —— 稳定性

  某一定数

频率稳定性的实例

蒲丰(*Buffon*)投币 ——

投一枚硬币观察正面向上的次数

$$n = 4040, \quad n_H = 2048, \quad f_n(H) = 0.5069$$

皮尔森(*Pearson*) 投币

$$n = 12000, \quad n_H = 6019, \quad f_n(H) = 0.5016$$

$$n = 24000, \quad n_H = 12012, \quad f_n(H) = 0.5005$$

例 Dewey G. 统计了约438023个英语单词中各字母出现的频率，发现各字母出现的频率不同：

A: 0.0788	B: 0.0156	C: 0.0268	D: 0.0389
E: 0.1268	F: 0.0256	G: 0.0187	H: 0.0573
I: 0.0707	J: 0.0010	K: 0.0060	L: 0.0394
M: 0.0244	N: 0.0706	O: 0.0776	P: 0.0186
Q: 0.0009	R: 0.0594	S: 0.0634	T: 0.0987
U: 0.0280	V: 0.0102	W: 0.0214	X: 0.0016
Y: 0.0202	Z: 0.0006		

近百年世界重大地震

"重大"的标准
- ① 震级 7 级左右
- ② 死亡 5000人以上

时间	地点	级别	死亡
1905.04.04	印度克什米尔地区	8.0	88
1906.08.17	智利瓦尔帕莱索港地区	8.4	2 万
1917.01.20	印度尼西亚巴厘岛		1.5 万
1920.12.16	中国甘肃	8.6	10 万
1923.09.01	日本关东地区	7.9	14.2 万
1935.05.30	巴基斯坦基达地区	7.5	5 万

时间	地点	级别	死亡
1948.06.28	日本福井地区	7.3	0.51 万
1970.01.05	中国云南	7.7	1 万
1976.07.28	中国河北省唐山	7.8	24.2
1978.09.16	伊朗塔巴斯镇地区	7.9	1.5 万
1995.01.17	日本阪神工业区	7.2	0.6 万
1999.08.17	土耳其伊兹米特市	7.4	1.7 万
2003.12.26	伊朗克尔曼省	6.8	3 万
2004.12.26	印尼苏门答腊岛附近海域	9.0	15 万

世界每年发生大地震概率约为14%

近百年世界重大流感

1918年 西班牙型流感 H1N1亚型

4亿人感染 5000万人死亡

20天传遍美国 半年席卷全球

1957年 亚洲型流感 H2N2 亚型

1968年 香港型流感 H3N2 亚型

世界性大流感每30-40年发生一次

2005年8月26日“超女”决赛

李宇春
3528308票

周笔畅
3270840票

张靓颖
1353906票

手机投票总数 8153054

李宇春	得票频率	43.27%
周笔畅	得票频率	40.12%
张靓颖	得票频率	16.61%

得票频率可被视为获胜概率

概率的定义

概率的统计定义

在相同条件下重复进行的 n 次试验中, 事件 A 发生的频率稳定地在某一常数 p 附近摆动, 且随 n 越大摆动幅度越小, 则称 p 为事件 A 的概率, 记作 $P(A)$.

对本定义的评价

优点:	直观易懂	缺点:	粗糙模糊	不便使用

概率的公理化定义

概率的公理化理论由前苏联数学家柯尔莫哥洛夫(А.Н. Колмогоров)1933年建立.

设Ω是随机试验E的样本空间，若能找到某一个法则，使得对于E的每一事件A赋于一个实数，记为$P(A)$，称之为事件A的概率，这种赋值满足下面的三条公理:

- 非负性： $\forall A \subset \Omega,\ P(A) \geq 0$
- 归一性： $P(\Omega) = 1$
- 可列可加性： $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

其中 $A_1, A_2, \cdots$ 为两两互斥事件，

概率的性质

- $P(\varnothing)=0$
- 有限可加性：设 $A_1, A_2, \cdots A_n$ 两两互斥

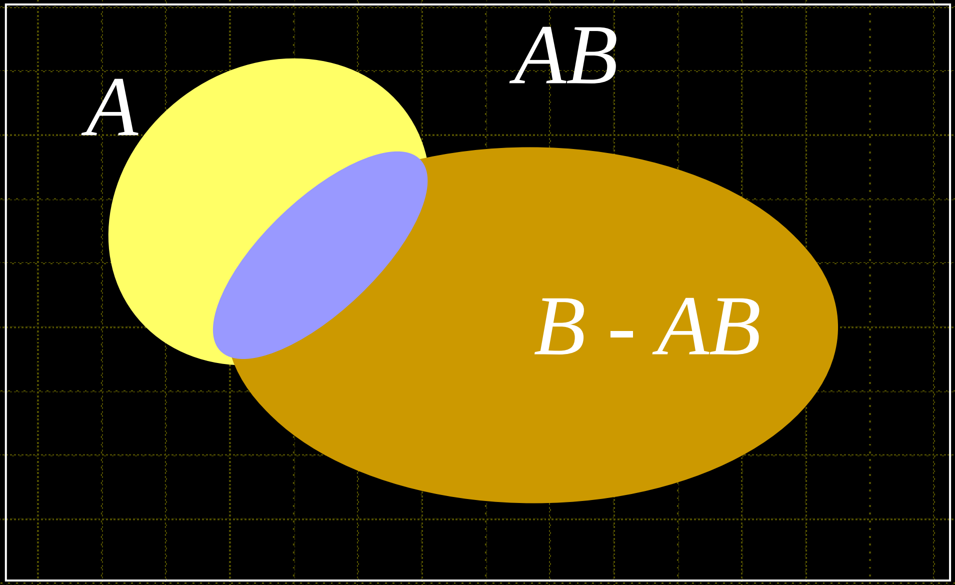
$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)$$

- $P(\bar{A})=1-P(A) \quad \Rightarrow P(A)\le 1$
- 若 $A\subset B \Rightarrow P(B-A)=P(B)-P(A)$

$$\Rightarrow P(A)\le P(B)$$

- 对任意两个事件A, B, 有

$$P(B-A)=P(B)-P(AB)$$



$$B=AB+(B-A)$$

$$P(B)=P(AB)+P(B-AB)$$

- 加法公式：对任意两个事件A, B, 有

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

$$P(A\cup B)\le P(A)+P(B)$$

推广：

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)$$
$$-P(AB)-P(AC)-P(BC)$$
$$+P(ABC)$$

一般:

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)-\sum_{1\le i<j\le n} P(A_iA_j)+$$

$$+\sum_{1\le i<j<k\le n}^{n} P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$$

右端共有 2^n-1 项.

例1 小王参加"智力大冲浪"游戏，他能答出甲、乙二类问题的概率分别为0.7和0.2，两类问题都能答出的概率为0.1．求小王

(1) 答出甲类而答不出乙类问题的概率

(2) 至少有一类问题能答出的概率

(3) 两类问题都答不出的概率

解 事件A, B分别表示"能答出甲,乙类问题"

(1) $P(A\overline{B}) = P(A) - P(AB) = 0.7 - 0.1 = 0.6$

(2) $P(A \cup B) = P(A) + P(B) - P(AB) = 0.8$

(3) $P(\overline{A}\,\overline{B}) = P(\overline{A \cup B}) = 0.2$

课后同学问：

例1 中小王他能答出第一类问题的概率为0.7，答出第二类问题的概率为0.2，两类问题都能答出的概率为0.1. 为什么不是 0.7×0.2 ？

若是的话, 则应有 $P(A_1A_2)=P(A_1)P(A_2)$

而现在题中并未给出这一条件.

在§1.4中将告诉我们上述等式成立的条件是：事件 A_1,A_2 相互独立.

例2 设A , B满足 $P(A)=0.6$, $P(B)=0.7$, 在何条件下， $P(AB)$ 取得最大(小)值？最大(小)值是多少？

解 $$P(A\cup B)=P(A)+P(B)-P(AB)$$

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

$$\ge P(A)+P(B)-1=0.3 \text{ —— 最小值}$$

最小值在 $P(A\cup B)=1$ 时取得

$$P(AB)\le P(A)=0.6 \text{ —— 最大值}$$

最大值在 $P(A\cup B)=P(B)$ 时取得

课上有同学提问

例2 中回答当 $A \cup B = \Omega$ 时, $P(AB)$ 取得最小值是否正确?

这相当于问如下命题是否成立

$$A \cup B = \Omega \Leftrightarrow P(A \cup B) = 1 \quad ⊛$$

答: 不成立!

⊛式是“羊肉包子打狗”——有去路,没回路

为什么呢? 学了几何概型便会明白.

古典（等可能）概型

概率的古典定义

设 随机试验E 具有下列特点:

- 基本事件的个数有限
- 每个基本事件等可能性发生

则称 E 为 古典(等可能)概型

古典概型中概率的计算:

记 $n=\Omega$中所包含的基本事件的个数

$k=$ 组成 A的基本事件的个数

则 $P(A)=\dfrac{k}{n}$

例3 袋中有a只白球，b只红球，从袋中按不放回与放回两种方式取m个球（$m \le a+b$），求其中恰有k个（$k \le a, k \le m$）白球的概率

解（1）不放回情形

E: 球编号，任取一球，记下颜色，放在一边，重复m次

$$\Omega: \ n_{\Omega} = A_{(a+b)}^{m} = (a+b)(a+b-1)\cdots(a+b-m+1)$$

记事件A为m个球中有k个白球，则

$$n_A = C_m^k A_a^k A_b^{m-k}$$

$$= \frac{m!}{k!(m-k)!} \cdot \frac{a!}{(a-k)!} \cdot \frac{b!}{(b-m+k)!}$$

则 $P(A)=\dfrac{C_m^k A_a^k A_b^{m-k}}{A_{a+b}^m}$ $\quad k\le a, k\le m$

又解 E_1: 球编号, 一次取 m 个球,记下颜色

Ω_1: $\quad n_{\Omega_1}=C_{a+b}^m$

记事件 A 为m个球中有k个白球， 则

$$n_A=C_a^k C_b^{m-k}$$

因此 $\boxed{P(A)=\dfrac{C_a^k C_b^{m-k}}{C_{a+b}^m} \quad k\le a, k\le m}$ 称超几何分布

不放回地逐次取 m 个球, 与一次任取 m 个球算得的结果相同.

（2）放回情形

E_2: 球编号, 任取一球, 记下颜色, 放回去, 重复 m 次

Ω_2:

$$n_{\Omega_2}=(a+b)^m$$

记 B 为取出的 m 个球中有 k 个白球, 则

$$P(B)=\frac{C_m^k a^k b^{m-k}}{(a+b)^m}=C_m^k\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{m-k}$$ 记 $p=\dfrac{a}{a+b}$

$$\boxed{P(B)=C_m^k p^k(1-p)^{m-k}\quad k=1,2,\cdots,\min(a,m)}$$

称**二项分布**

例4（分房模型）设有 k 个不同的球，每个球等可能地落入 N 个盒子中（$k \le N$），设每个盒子容球数无限，求下列事件的概率：

（1）某指定的 k 个盒子中各有一球；

（2）某指定的一个盒子恰有 m 个球($m \le k$)

（3）某指定的一个盒子没有球；

（4）恰有 k 个盒子中各有一球；

（5）至少有两个球在同一盒子中；

（6）每个盒子至多有一个球.

解 $n = N^k$

设 (1) ~ (6)的各事件分别为 $A_1 \to A_6$

则 $m_{A_1} = k!$ $\longrightarrow$ $P(A_1) = \dfrac{m_{A_1}}{n} = \dfrac{k!}{N^k}$

$m_{A_2} = C_k^m (N-1)^{k-m}$ $\longrightarrow$ $P(A_2) = \dfrac{C_k^m (N-1)^{k-m}}{N^k}$

$m_{A_3} = (N-1)^k$ $\longrightarrow$ $P(A_3) = \dfrac{(N-1)^k}{N^k}$

$m_{A_4} = C_N^k k!$ $\longrightarrow$ $P(A_4) = \dfrac{C_N^k k!}{N^k}$

$m_{A_5} = N^k - C_N^k k!$ $\longrightarrow$ $P(A_5) = \dfrac{N^k - C_N^k k!}{N^k} = 1 - P(A_4)$

$m_{A_6} = C_N^k k!$ $\longrightarrow$ $P(A_6) = P(A_4)$

例5 “分房模型”的应用

生物系二年级有 n 个人，求至少有两人生日相同（设为事件A）的概率.

解 本问题中的人可被视为“球”，365天为365只“盒子”

$\overline{A}$ 为 n 个人的生日均不相同,这相当于每个盒子至多有一个球. 由例4（6）

$$P(\overline{A})=\frac{C_{365}^{n}\cdot n!}{365^{n}}\Rightarrow P(A)=1-P(\overline{A})=1-\frac{C_{365}^{n}\cdot n!}{365^{n}}.$$

若 $n=64$, $\Rightarrow P(A)\approx 0.997.$

例6 在0,1,2,3, $\cdots$,9中不重复地任取四个数,求它们能排成首位非零的四位偶数的概率.

解 设 A为"能排成首位非零的四位偶数"

$$n_\Omega = A_{10}^4 = 5040.$$

四位偶数的末位为偶数,故有C_5^1种可能

而前三位数有A_9^3种取法,由于首位为零的四位数有$C_4^1 A_8^2$种取法,所以有利于A发生的取法共有 $n_A = C_5^1 A_9^3 - C_4^1 A_8^2 = 2296$ 种.

$$\therefore P(A) = \frac{2296}{5040} = \frac{41}{90}$$

例7 在1,2,3, ⋯,9中重复地任取 $n\,(\geq 2)$个数,求 n 个数字的乘积能被10整除的概率.

解 $n_\Omega = 9^n$

设 A 表示事件 “n 次取到的数字的乘积能被10整除”

设 A_1 表示事件 “n 次取到的数字中有偶数”

A_2表示事件 “n 次取到的数字中有5”

$$A = A_1 A_2$$

$$\overline{A} = \overline{A_1 A_2} = \overline{A_1} \cup \overline{A_2}$$

$$P\left(\overline{A_1}\right)=\frac{5^n}{9^n} \qquad P\left(\overline{A_2}\right)=\frac{8^n}{9^n} \qquad P\left(\overline{A_1}\,\overline{A_2}\right)=\frac{4^n}{9^n}$$

$$P\left(\overline{A}\right)=P\left(\overline{A_1}\cup\overline{A_2}\right)$$

$$=P\left(\overline{A_1}\right)+P\left(\overline{A_2}\right)-P\left(\overline{A_1}\,\overline{A_2}\right)$$

$$=\frac{5^n+8^n-4^n}{9^n}$$

$$P(A)=1-\frac{5^n+8^n-4^n}{9^n}.$$

计算古典概率注意事项

1° 明确所作的试验是等可能概型,有时需设计符合问题要求的随机试验, 使其成为等可能概型.

2° 同一题的样本空间的基本事件总数 n_{Ω} 随试验设计的不同而不同, 如 例3不放回试验的两种不同设计. 一般 n_{Ω} 越小越好.

3° 计算古典概率时须注意应用概率计算的有关公式, 将复杂问题简单化. 如例7.

小概率事件——

若$P(A) \le 0.01$，则称A为小概率事件.

小概率原理——（即实际推断原理）

一次试验中小概率事件一般是不会发生的. 若在一次试验中居然发生了，则可怀疑该事件并非小概率事件.

例8 区长办公室某一周内曾接待过9次来访，这些来访都是周三或周日进行的，是否可以断定接待时间是有规定的？

解 假定办公室每天都接待，则

$$P(9\text{次来访都在周三、日})=\frac{2^9}{7^9}=0.0000127$$

这是小概率事件，一般在一次试验中不会发生。现居然发生了，故可认为假定不成立，从而推断接待时间是有规定的。

作业 P 46习题一

7 8 10 12

15 17 19

补充作业

设事件A, B, C 同时发生必导致事件 D 发生，则

$$P(A)+P(B)+P(C)\le 2+P(D).$$

柯尔莫哥洛夫

(А. Н. К о л м о г о р о в 1903-1987)

俄国数学家

1939年任苏联科学院院士.先后当选美,法,意,荷,英,德 等国的外籍院士 及皇家学会会员.为 20 世纪最有影响的俄国数学家.

柯尔莫哥洛夫为开创现代数学的一系列重要分支作出重大贡献.

他建立了在测度论基础上的概率论公理系统, 奠定了近代概率论的基础.

他又是随机过程论的奠基人之一, 其主要工作包括:

20年代 关于强大数定律、重对数律的基本工作;

1933年在《概率论的基本概念》一文中提出的概率论公理体系(希尔伯特第6问题)

30年代建立的马尔可夫过程的两个基本方程;

用希尔伯特空间的几何理论建立弱平稳序列的线性理论;

40年代完成独立和的弱极限理论，经验分布的柯尔莫哥洛夫统计量等;

在动力系统中开创了关于哈密顿系统的微扰理论与K系统遍历理论;

50年代中期开创了研究函数特征的信息论方法，他的工作及随后阿诺尔德的工作解决并深化了希尔伯特第13问题——用较少变量的函数表示较多变量的函数；

60年代后又创立了信息算法理论;

1980年由于它在调和分析, 概率论, 遍历理论 及 动力系统方面 出色的工作获沃尔夫奖;

他十分重视数学教育,在他的指引下,大批数学家在不同的领域内取得重大成就.其中包括 и .M.盖尔范德,B. и . 阿诺尔德, Я . Г .西奈依等人.

他还非常重视基础教育, 亲自领导了中学 数学教科书的编写工作.

第 2 周

问 题

已知 $P(A) = P(B) = P(C) = 1/4$,

$P(AB) = 0$, $P(AC) = P(BC) = 1/6$

则事件A, B, C全不发生的概率为______.

通过做此题 你能发现什么问题?

一般会解出

$$P(\overline{A}\,\overline{B}\,\overline{C})=1-P(A\cup B\cup C)$$

$$=1-P(A)-P(B)-P(C)$$

$$+P(AB)+P(AC)+P(BC)-P(ABC)$$

$$=1-3/4+2/6=7/12.$$

由题设得

$$P(A\cup B)=P(A)+P(B)-P(AB)=1/2.$$

另一方面又可得

$$P(A\cup B\cup C)=5/12.$$

于是得矛盾

$$P(A\cup B)=1/2>5/12=P(A\cup B\cup C).$$

若将条件修改为 $P(AC) = P(BC) = 1/9$
便无矛盾

$$P(A\cup B)=1/2<19/36=P(A\cup B\cup C).$$

几何概型 (等可能概型的推广)

例9 某人的表停了，他打开收音机听电台报时，已知电台是整点报时的，问他等待报时的时间短于十分钟的概率

10分钟

9点 10点

$$P(A)=\frac{10}{60}=\frac{1}{6}$$

几何概型

设样本空间为有限区域 Ω, 若样本点落入 Ω 内任何区域 G 中的概率与区域 G 的测度成正比, 则样本点落入 G 内的概率为

$$P(A)=\frac{G\text{的测度}}{\Omega\text{的测度}}$$

例10 两船欲停靠同一个码头, 设两船到达码头的时间各不相干， 而且到达码头的时间在一昼夜内是等可能的. 如果两船到达码头后需在码头停留的时间分别是1 小时与2 小 时, 试求在一昼夜内，任一船到达时， 需 要等待空出码头的概率.

解 设船1 到达码头的瞬时为 x ，$0 \le x < 24$

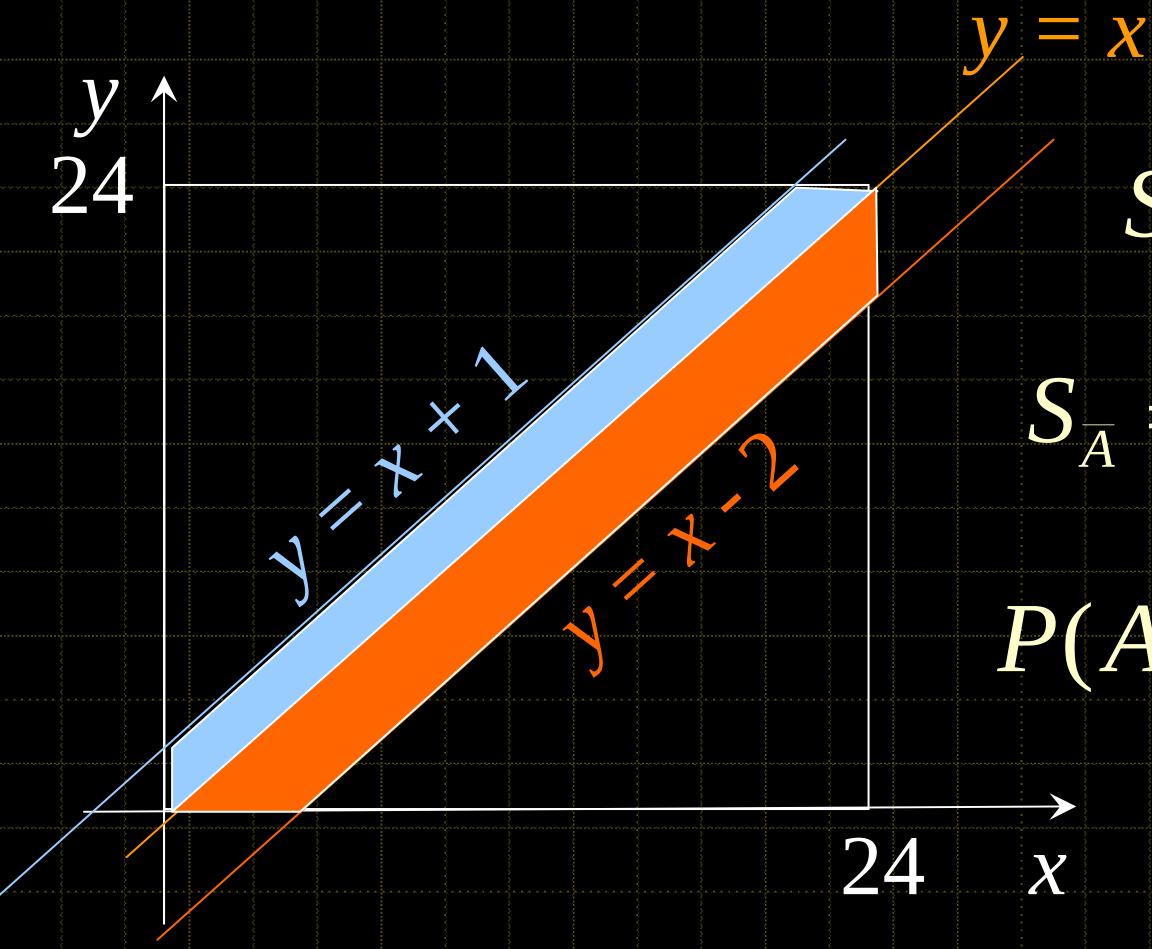
船2 到达码头的瞬时为 y ，$0 \le y < 24$

设事件 A 表示任一船到达码头时需要等待空出码头

$$\Omega=\{(x,y)\mid 0\le x<24, 0\le y<24\}$$

$$A=\{(x,y)\mid (x,y)\in\Omega,$$

$$0\le y-x\le 1, 0\le x-y\le 2\}$$

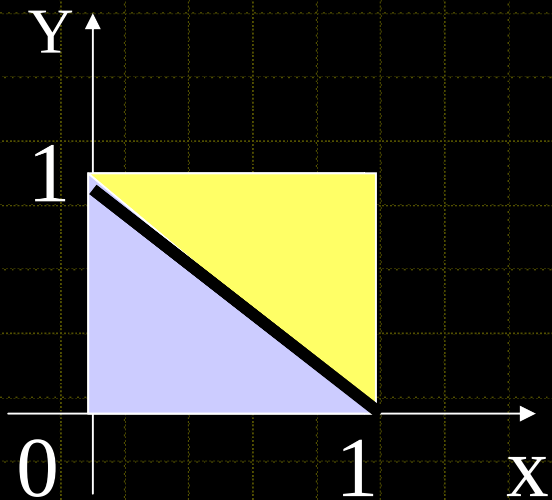


$$S_\Omega=24^2$$

$$S_{\bar{A}}=\frac{1}{2}\left(23^2+22^2\right)$$

$$P(A)=1-\frac{S_{\bar{A}}}{S_\Omega}=0.1207$$

用几何概型可以回答例2中提出的"概率为 1 的事件为什么不一定发生?"这一问题.



如图，设试验E 为" 随机地向边长为1 的正方形内投点" 事件A为"点投在黄、蓝两个三角形内"，求 $P(A)$

$$P(A)=\frac{S_{黄三角形}+S_{蓝三角形}}{S_{正方形}}=\frac{1\times\frac{1}{2}+1\times\frac{1}{2}}{1\times 1}=1$$

由于点可能投在正方形的对角线上, 所以事件A未必一定发生.

作业 P 46习题一

21，22 (1) (2)

附录 完全可加性

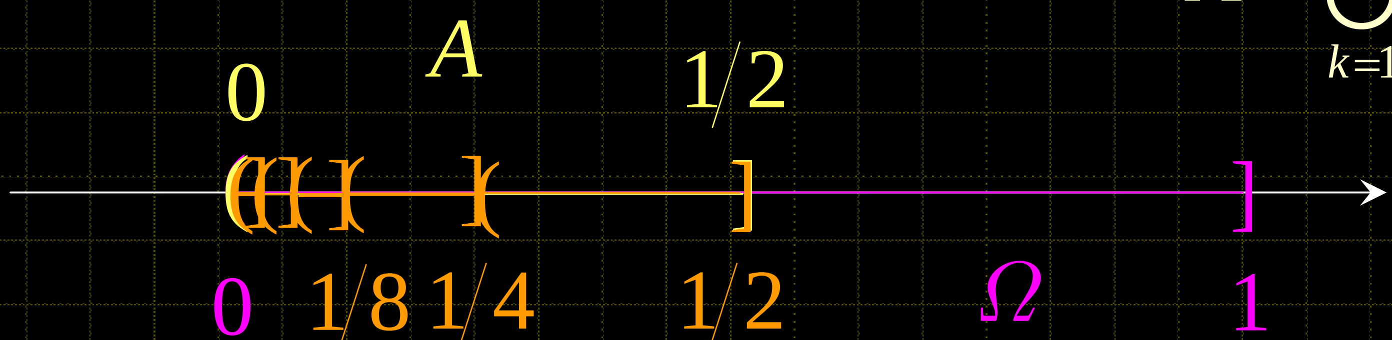
随机地向区间 $(0,1]$ 投掷一个质点,

令事件 A 为该质点落入区间 $\left(0,\dfrac{1}{2}\right]$

事件 A_k 为该质点落入区间

$$\left(\frac{1}{2^{k+1}},\frac{1}{2^k}\right],\ k=1,2,\cdots$$

$$A=\bigcup_{k=1}^{\infty}A_k$$



$$\Rightarrow \quad P(A)=\frac{1}{2},$$

$$P(A_k)=\frac{1}{2^{k+1}}, k=1,2,\cdots$$

$$\Rightarrow \quad \sum_{k=1}^{\infty} P(A_k)=\sum_{k=1}^{\infty}\frac{1}{2^{k+1}}=\frac{1}{2}$$

排列组合有关知识复习

加法原理：完成一件事情有 n 类方法，第 i 类方法中有 m_i 种具体的方法，则完成这件事情共有

$$\sum_{i=1}^{n} m_i$$

种不同的方法

乘法原理：完成一件事情有 n 个步骤，第 i 个步骤中有 m_i 种具体的方法，则完成这件事情共有

$$\prod_{i=1}^{n} m_i$$

种不同的方法

排列 从 n 个不同的元素中取出 m 个 (不放回地）按一定的次序排成一排不同的排法共有

$$A_n^m = n(n-1)(n-2)\cdots(n-m+1)$$

全排列 $A_n^n = n!$

可重复排列 从 n 个不同的元素中可重复地取出 m 个排成一排, 不同的排法有

$$n^m$$

种

不尽相异元素的全排列 n 个元素中有 m 类，第 i 类中有 k_i 个相同的元素，

$$k_1 + k_2 + \cdots + k_m = n,$$

将这 n 个元素按一定的次序排成一排，

不同的排法共有

$$\frac{n!}{k_1!k_2!\cdots k_m!}$$

种

组合 从 n 个不同的元素中取出 m 个(不放回地）组成一组， 不同的分法共有

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

多组组合 把 n 个元素分成 m 个不同的组（组编号）， 各组分别有 $k_1, k_2, \cdots, k_m$ 个元素，$k_1 + k_2 + \cdots + k_m = n$，不同的分法共有

$$C_n^{k_1} C_{n-k_1}^{k_2} \cdots C_{k_n}^{k_n}$$

种

例11 (类似于教材 P.22 例10)

将15 名同学(含3 名女同学), 平均分成三组. 求

(1) 每组有1 名女同学(设为事件A)的概率;

(2) 3 名女同学同组(设为事件B)的概率

解 $n_\Omega = C_{15}^5 C_{10}^5 C_5^5$

（1） $n_A = C_{12}^4 C_8^4 C_4^4 C_3^1 C_2^1 C_1^1$ $\quad P(A) = \dfrac{25}{91}$

（2） $n_B = C_3^1 C_{12}^2 C_{10}^5 C_5^5$ $\quad P(B) = \dfrac{6}{91}$

例12 把标有 1,2,3,4 的 4 个球随机地放入标有1,2,3,4 的 4 个盒子中，每盒放一球，求至少有一个盒子的号码与放入的球的号码一致的概率

解 设 A 为所求的事件

设 A_i 表示 i 号球入 i 号盒， $i=1,2,3,4$

则 $$A=\bigcup_{i=1}^{4} A_i$$

$$P(A_i)=\frac{3!}{4!}=\frac{1}{4}, \qquad i=1,2,3,4$$

$$P(A_iA_j)=\frac{2!}{4!}=\frac{1}{12},\qquad 1\le i<j\le 4$$

$$P(A_iA_jA_k)=\frac{1!}{4!}=\frac{1}{24},\qquad 1\le i<j<k\le 4$$

$$P(A_1A_2A_3A_4)=\frac{1}{24}$$

由广义加法公式

$$P(A)=\sum_{1\le i\le 4}P(A_i)-\sum_{1\le i<j\le 4}P(A_iA_j)$$

$$+\sum_{1\le i<j<k\le 4}P(A_iA_jA_k)-P(A_1A_2A_3A_4)=\frac{5}{8}$$