概率论

徐红利

hlxu@nju.edu.cn

工程管理学院

教学安排

2024 概率论（工智+...
群号：913374071

- 上课时间：周一5-7节；
- 上课地点：逸B-403；
- 教学时间：9.2—12.27（共17周）；
- 内容：PPT+教材1-5章+发放的材料；
- 成绩：平时30% + 期末考试 70%；
- 作业：课后习题（在QQ群发布）；

引例：

- 概率论：研究随机现象数量规律的数学分支；
- 公交车门高度设置：根据城市人口高度的分布情况，设置多高能使99.5%的乘客顺利通过？
  - 乘客的身高可能低于车门高度，也可能超过车门高度（随机现象）；超过车门高度的可能性有多大？
- 保险公司设计重大疾病保险产品，如何设置收费标准、收费年限和理赔额度？
  - 投保人可能理赔，也可能不理赔（随机现象）；需要考虑理赔的可能性有多大；

概率论发展简史

- 概率论的起源：（合理分配赌注）17世纪中叶法国数学家帕斯卡与费马在来往信函中讨论“合理分配赌注问题”，概率论的发展被公认为从他们两人开始：
  - 甲、乙两人同掷一枚硬币，规定：正面朝上，甲得一点；若反面朝上，乙得一点；先积满3点者赢取全部赌注；
  - 假定在甲得2点、乙得1点时，赌局由于某种原因中止了，应该怎样分配赌注才算公平合理？

概率论发展简史

➢ 在实践中曲折发展——成为数学分支

- 从17世纪到19世纪，伯努利、棣莫弗、拉普拉斯、高斯、泊松、切贝谢夫、马尔可夫等著名数学家都对概率论的发展做出了杰出的贡献：大数定理，中心极限定理，几何概率、Poisson分布，《概率分析理论》……
- 到20世纪初，概率论的一些基本概念，诸如“概率”等尚没有确切的定义，概率论作为一个数学分支，缺乏严格的理论基础。

概率论发展简史

- 基础理论的建立——公理化阶段
  - 概率论的第一本专著是1713年问世的伯努利的《推测术》。经过二十多年的艰难研究，伯努利表述并证明了著名的“大数定律”，被称为概率论的奠基人。
  - 随着概率论在其他基础学科和工程技术上的应用，由拉普拉斯《概率分析理论》给出的概率定义的局限性很快便暴露了出来；
  - 1933年，苏联数学家柯尔莫哥洛夫发表了著名的《概率论的基本概念》，给出了概率论的公理化结构，为概率论的迅速发展奠定了基础。

马兰·梅森
(1588—1648)

勒内·笛卡儿
(1596—1650)

布莱兹·帕斯卡
(1623—1662)

皮埃尔·费马
(1601—1665)

克里斯蒂安·惠更斯
(1629—1695)

法国贵族梅勒给帕斯卡写信请教赌资分配

↓

帕斯卡与费马讨论赌资分配

↓

惠更斯撰写《论赌博中的计算》

↓

伯努利的大数定理，棣莫弗的中心极限定理，浦丰的几何概率，泊松分布⋯，成为数学分支

↓

波雷尔和勒贝格提出测度论，柯尔莫哥洛夫发表《概率论的基本概念》，形成现代公理化体系；

概率论的应用

- 在最近几十年中，概率论的方法被引入到各个工程技术学科和社会学科。
- 目前，概率论在人口统计、保险理论、误差理论、产品检验和质量控制、近代物理、自动控制、地震预报和气象预报、农业试验和公用事业等方面都得到了重要应用。
- 越来越多的概率论方法被引入到经济、金融和管理科学，概率论成为它们的有力工具。

统计学

- 研究怎样以有效的方式收集、整理和分析带有随机性的数据，以便对所考察的问题作出推断和预测，直至为采取一定的决策和行动提供依据和建议；
- 数理统计应用概率论的结果更深入地分析研究统计资料，通过对某些现象的频率（以及均值、方差）的观察，来发现该现象的内在规律，并作出一定精确程度的估计、判断和预测。
- 后续课程：应用统计、随机过程、大数据分析、多元统计……

第一章 概率论的基本概念

1.1 随机试验
1.2 样本空间、随机事件
1.3 频率与概率
1.4 等可能概型（古典概型）
1.5 条件概率；
1.6 独立性；

随机现象

- 随机现象具有统计规律性：在大量重复试验或者观察下，结果具有统计规律性。
  - ✓ 抛硬币出现正反面的可能性均为二分之一；
  - ✓ 生男孩女孩的可能性均为二分之一；
- 对于随机现象，事情发生的结果具有不确定性；但是，事件发生的可能性的大小是客观存在的，这是不确定性中的确定性；

- 通过随机实验研究随机现象；
- **随机试验：**

E1：抛一枚硬币，观察正面 H、反面 T 出现的情况。

E2：将一枚硬币抛掷三次，观察出现正面的次数。

E3：抛一颗骰子，观察出现的点数。

E4：在一批灯泡中任意抽取一只，测试它的寿命。

随机试验的特点

- 可以在相同的条件下重复进行；
- 每次试验的可能结果不止一个，且能事先明确所有可能的结果；
- 一次试验只出现一个结果，且试验前不能确定哪个结果会出现。

样本空间

- 样本空间：试验的所有可能结果构成的集合
  - 随机试验中，每一个可能结果称为该试验的一个样本点（或基本事件）。全体样本点组成的集合称为该试验的样本空间，记为 S 。

$\mathbf{E_1}$： 抛一枚硬币，观察正面 H、反面 T 出现的情况。

$$S_1=\{H, T\}$$

E_2：将一枚硬币抛掷三次，观察正面H、反面T出现的情况。

$$S_2=\{HHH, HHT，HTH, THH, HTT,THT,TTH,TTT\}$$

E_3：将一枚硬币抛掷三次，观察出现正面的次数。

$$S_3=\{0, 1，2, 3\}$$

E_4：抛一颗骰子，观察出现的点数。

$$S_4=\{1, 2，3, 4，5，6\}$$

E_5：记录某城市120急救电话台一昼夜接到的呼唤次数。

$$S_5=\{0,1,2, \ldots\}$$

E_6: 在一批灯泡中任取一只, 测试它的寿命.

$$S_6=\{t|\ t\geq 0\}$$

E_7: 记录某地一昼夜的最高温度和最低温度。

$$S_7=\{(x,y)\,|\,T_0\leq x\leq y\leq T_1\}$$

这里 x 表示最低温度，y 表示最高温度；并设这一地区的温度不会小于 T_0，也不会大于T_1。

1. 离散样本空间: 样本点为有限多个或可列多个，例 E_1, E_2等。

E_1：抛一枚硬币，观察正面 H、反面 T 出现的情况。

$$S_1=\{H, T\}$$

2. 无穷样本空间: 样本点在区间或区域内取值. 例灯泡的寿命 。

E_6: 在一批灯泡中任取一只, 测试它的寿命.

$$S_6=\{t| \ t \geq 0\}$$

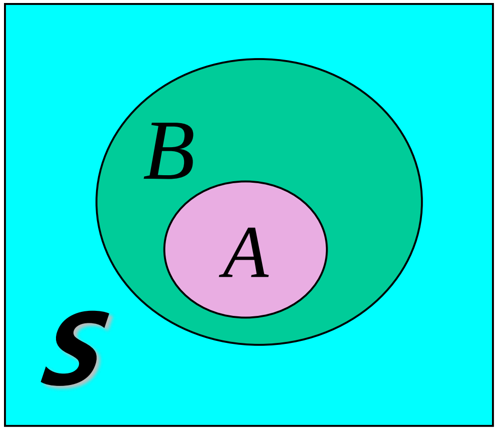
➢ 随机事件

- “在一定条件下可能发生也可能不发生的事情”叫做随机事件，简称事件。
- 如在上面的例子中，“出现正面”、“出现反面”、“点数<4”、“出现偶数点”、$\{t<1000\}$ 等都是随机事件。
- 事件是由样本空间中某些样本点组成的集合（试验 E 的样本空间 S 的子集）；
- 事件发生当且仅当它所包含的某一个样本点出现。

- **基本事件:** 由一个样本点组成的单点集，如：{H},{T}.
- **复合事件:** 由两或两个以上的基本事件复合而成的事件，称为复合事件。如: E_3中 {出现正面次数为偶数}.
- **必然事件:** 样本空间 S 是自身的子集，在每次试验中总是发生的，称为必然事件。
- **不可能事件：** 空集 $\varnothing$，不包含任何样本点，它在每次试验中都不发生，称为不可能事件。

事件间的关系与事件的运算

- **包含关系和相等关系**:
  - 若事件 A 发生必然导致事件 B 发生, 则称事件 B 包含事件A, 记作 $A \subset B$;
  - 若$A \subset B$且 $A \supset B$, 即 $A = B$, 则称 A与 B 相等;



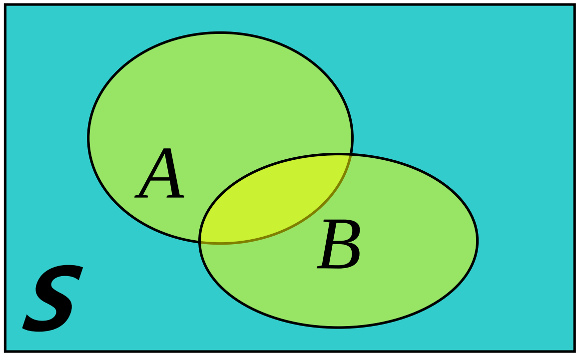
(1) $A \subset B$

➢ 事件的并:

- 事件 C 发生，则 “事件 A 与 B 至少有一个发生” 称为事件 C 是 A 与 B 的并（或和事件），记 $C = A \cup B$ 。

$$C = \{x \mid x \in A \text{ 或 } x \in B\}$$

- 可列个事件$A_1, A_2, \cdots$ 的并事件记为 $\bigcup_{k=1}^{\infty} A_k$.



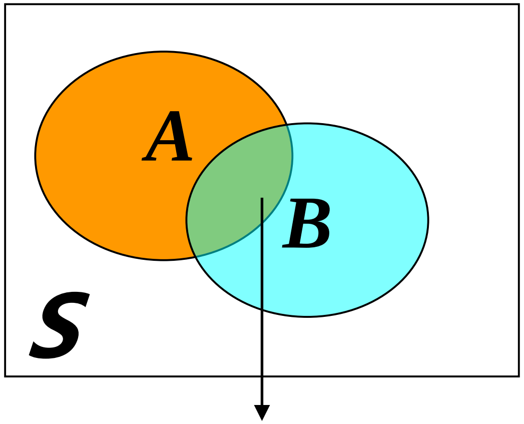
(2) $A \cup B$

➢ 事件的交：

- 事件 C 表示 “事件 A 与 B 同时发生”，事件 C 称为 A 与 B 的交（积事件），记作 $A \cap B$

$$(AB),\ A \cap B=\{x \mid x \in A \text{ 且 } x \in B\}$$

- 类似地，事件 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件A_1，A_2，…的交.



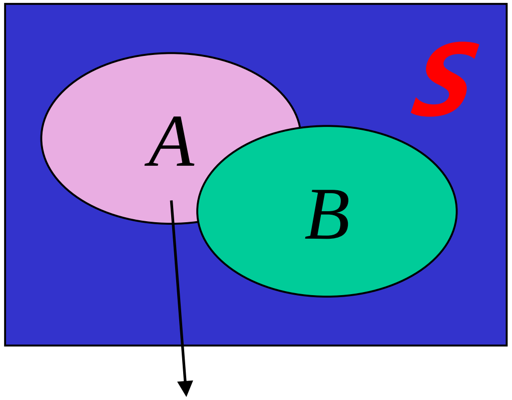
(3)$A \cap B$

➢事件的差:

- 事件 $A\text{-}B=\{x|x\in A$且 $x\notin B\}$ 称为 A 与 B 的差； 当且仅当 A 发生，B 不发生时，事件$A\text{-}B$发生. 即:

$$A\text{-}B=A\overline{B}=A-AB$$

- 显然： $A-A=\varnothing$ ， $A-\varnothing=A$， $A-S=\varnothing$



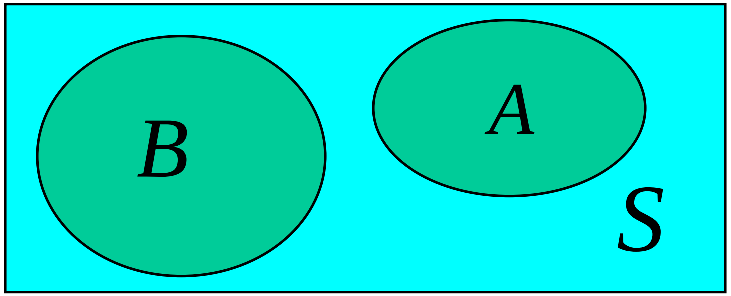
$(4)A-B$

互不相容事件(互斥事件):

- 若 $A\cap B=\varnothing$ ，则称 A 与 B 是互不相容的，或互斥的，即 A 与 B 不能同时发生。

(1) 基本事件是两两互不相容的，即样本点是互不相容的, 事件 A 与 B-A 是互不相容的.

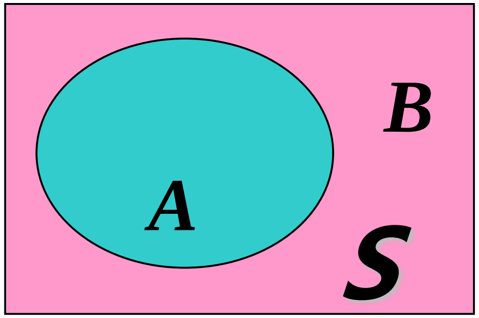
(2) 若用集合表示事件，则 A, B 互不相容即为 A 与 B 是不相交的.



$A\cap B=\varnothing$

➢ 对立事件(逆事件)：

- 若 $A\cup B=S$ 且 $A\cap B=\varnothing$ ，则称 A 与 B 互为逆事件，也称为对立事件。即在一次实验中，事件 A 与 B 中必然有一个发生，且仅有一个发生。
- A 的对立事件记为 $\overline{A}$。若 A 与 B 互为对立事件，则记为 $A=\overline{B}$, 或 $B=\overline{A}$ 。



$$B=\overline{A}$$

➢ 事件的运算律:

交换律: $A \bigcup B = B \bigcup A;\ \ A \bigcap B = B \bigcap A.$

结合律: $(A \bigcup B) \bigcup C = A \bigcup (B \bigcup C),$

$(AB)C = A(BC)$

分配律: $A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C);$

$A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C).$

德·摩根律:

$$\overline{A\cup B}=\overline{A}\cap\overline{B};\quad \overline{A\cap B}=\overline{A}\cup\overline{B}.$$

说明:

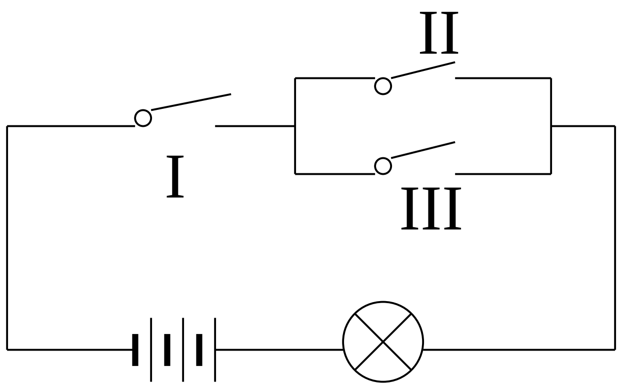
$$\overline{A\cup B}=\overline{\{A、B\text{至少发生一个}\}}$$

$$=\{A、B\text{都不发生}\}=\overline{A}\cap\overline{B}.$$

摩根律推广:

$$\overline{\bigcup_i A_i}=\bigcap_i \overline{A_i},\quad \overline{\bigcap_i A_i}=\bigcup_i \overline{A_i}$$

例1：如右图所示的电路中，以A表示“信号灯亮”这一事件，以B, C, D分别表示事件：继电器接点I, II, III闭合，那么容易知道



$$BC \subset A, BD \subset A, \quad BC \bigcup BD = A, \quad 而 \bar{B}A = \varnothing,$$

即事件$\bar{B}$与事件A互不相容。 又$\overline{B \bigcup C} = \bar{B} \bigcap \bar{C}$

(左边表示事件“I，II至少有一个闭合”的逆事件，也就是I，II都不闭合，即$\bar{B}$，$\bar{C}$同时发生。）

例2. 高射炮对模型飞机射击三次，设 A_i 表示“第 i 次击中飞机”，用 B_i 表示下列事件；

（1）B_1 “只有第一次击中飞机”

（2）B_2 “恰有一次击中飞机”

（3）B_3 “至少有一次击中飞机”

（4）三次击中飞机时，击落了飞机, B_4: “飞机没有被击落”

（1）B_1 “只有第一次击中飞机”；（2）B_2 “恰有一次击中飞机”

（3）B_3 “至少有一次击中飞机”

（4）三次击中飞机时，击落了飞机, B_4: “飞机没有被击落”

解: (1) $B_1 = A_1\overline{A_2}\,\overline{A_3}$

(2) $B_2 = A_1\overline{A_2}\,\overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\,\overline{A_2}A_3$

(3) $\overline{B_3} = \overline{A_1}\,\overline{A_2}\,\overline{A_3}$, 所以

$B_3 = \overline{\overline{A_1}\,\overline{A_2}\,\overline{A_3}} = A_1 \cup A_2 \cup A_3$

(4) 由于 $\overline{B_4} = A_1A_2A_3$

所以 $B_4 = \overline{A_1A_2A_3} = \overline{A_1} \cup \overline{A_2} \cup \overline{A_3}$

频率与概率

➢ 频率

1、将一试验 E 在相同的条件下重复进行 n 次，如果事件A 发生了 n_A 次，则比值 n_A/n 称为 事件 A 发生的频率，记为 $f_n(A)$.

2、频率的基本性质：

(1) $0 \le f_n(A) \le 1$；（非负性）

(2) $f_n(S) = 1$；（规范性）

(3) 可加性：对互斥事件A, B，有

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

推广(有限可加性）：若A_1 ，A_2 ，$\cdots$，A_k 两两互不相容,则

$$f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$$

➢ 频率的特性：波动性和稳定性.

(1) 波动性： 对于同样的试验次数，不同的试验其频率不同； 对于同一试验，不同的试验次数 n， 其频率也不同，当 n 较小时，$f_n(A)$ 随机波动的幅度较大；

(2) 稳定性： 随着 n 逐渐增大，事件 A 的频率总在某一定值$P(A)$ 的附近摆动而逐渐稳定于这个值，这个定值 $P(A)$ 通常称为频率的稳定值。

例 3　考虑“抛硬币”这个实验，我们将一枚硬币抛掷5 次、50次、500次，各做10遍。得到数据如表1所示（其中n_H 表示H 发生的频数，$f_n(H)$ 表示H 发生的频率）。

实验	$n=5$		$n=50$		$n=500$	
序号	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

投币试验

实验者	掷币次数 n	正面次数 $n(A)$	频率 $f_n(A)$
摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

例4 考察英语中特定字母出现的频率，当观察字母的个数 n (试验的次数）较小时，频率有较大幅度的随机波动。但当n 增大时，频率呈现出稳定性。下面就是一份英文字母频率的统计表（从大到小排序）：

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

- 大量验证实，当重复试验的次数n 逐渐增大时，频率$f_n(A)$ 呈现出稳定性，逐渐稳定于某个常数。这种“频率稳定性”即通常所说的统计规律性。

- 我们让试验重复大量次数，计算频率$f_n(A)$，以它来表征事件A 发生可能性的大小是合适的；

➢概率

– 概率有很多种定义，比如，概率的古典定义、几何定义、主观定义，较为常用的是 统计定义 和 公理化定义。

1、统计定义：

频率的稳定值 $P(A)$，反映了事件 A 在一次试验中发生的可能性大小，称 $P(A)$ 为事件 A 的概率。

2、公理化定义：

设 S 为样本空间，A 为事件，对每一事件 A 赋予一实数 $P(A)$，如果 $P(A)$ 满足如下三条公理，则称 $P(A)$ 为 事件 A 的概率。

(1) 非负性：$0 \le P(A) \le 1$;

(2) 规范性（正则性）： $P(S) = 1$;

(3) 对两两互斥事件 $A_i\ (i = 1, 2, \cdots)$，有：

可列可加性：$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

概率的性质

性质1. 对任一事件A, $P(\overline{A})=1-P(A)$.

证：因为 $A\cup\overline{A}=S$, $A\overline{A}=\varnothing$，由概率的规范性及可加性

得$1=P(A\cup\overline{A})=P(A)+P(\overline{A})$，即$P(\overline{A})=1-P(A)$

性质2. $P(\varnothing)=0$.

证：$P(\varnothing)=1-P(S)=1-1=0$

性质3. 有限可加性（A_i 互不相容）:

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

证：令 $A_{n+1} = A_{n+2} = \cdots = \varnothing$, 则有 $A_i A_j = \varnothing$,

由可列可加性可得:

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

可列可加性: $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

性质4. 对任一事件A, $P(A) \le 1.$

证：由于$P(\overline{A}) \ge 0,$

所以$P(A) = 1 - P(\overline{A}) \le 1$

性质5. 若$A \subset B$, 则有：

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

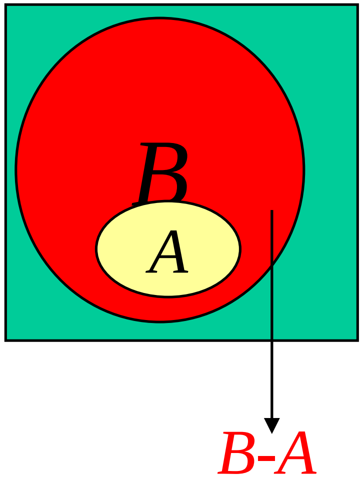
$$P(B) \ge P(A).$$

证： $\because A \subset B$, 有$B = A \cup (B - A)$,

且$A \cap (B - A) = \varnothing$,

$\therefore\ P(B)=P(A)+P(B\text{-}A)$,

即 $P(B\text{-}A)=P(B)\text{-}P(A)$.



由$P(B - A) \geq 0$, 可得：$P(B) \geq P(A)$.

性质6 (加法公式): 对任意两事件 A, B 有

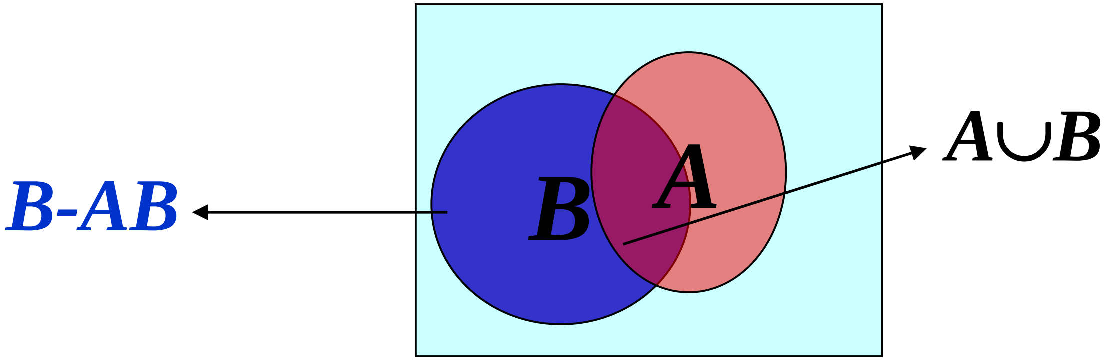
$$P(A\cup B)=P(A)+P(B)-P(AB).$$

证： $\because A\cup B=A\cup(B-AB)$，

且 $A\cap(B-AB)=\varnothing$，

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

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



可以作如下推广：

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

2. $$P(A_1\cup A_2\cup\dots\cup A_n)$$

$$=\sum_{i=1}^{n}P(A_i)-\sum_{1\le i<j\le n}P(A_iA_j)+\sum_{1\le i<j<k\le n}P(A_iA_jA_k)+\dots$$

$$+(-1)^{n-1}P(A_1A_2\dots A_n).$$

这个式子称为“加奇减偶公式”.

例5： 设 A, B 为两个事件, 且 $P(A)=0.5, P(B)=0.3, P(AB)=0.2,$ 求下列各事件的概率.

(1) $\overline{AB}$; (2) $\overline{A}\,\overline{B}$; (3) $\overline{A}B$; (4) $\overline{A}\cup B$.

解：(1) $P(\overline{AB}) = 1 - P(AB) = 1 - 0.2 = 0.8.$

(2) $P(\bar{A}\bar{B}) = P(\overline{A\cup B}) = 1 - P(A\cup B)$

$= 1 - [P(A) + P(B) - P(AB)]$

$= 0.4$

例5： 设 A, B 为两个事件, 且 $P(A)=0.5$, $P(B)=0.3$, $P(AB)=0.2$, 求下列各事件的概率.

(1) $\overline{AB}$; (2) $\overline{A}\,\overline{B}$; (3) $\overline{A}B$; (4) $\overline{A}\cup B$.

解： (3) $P(B)=P(AB\cup\overline{A}B)=P(AB)+P(\overline{A}B)$

$$P(\overline{A}B)=P(B)-P(AB)=0.1.$$

(4) $P(\overline{A}\cup B)=P(\overline{A})+P(B)-P(\overline{A}B)=0.7.$

古典概型

➢ 若随机试验有以下两个特点，这类随机现象的概率模型叫做古典概型：

(1) 样本空间中只有有限个样本点，即 $S=\{e_1, e_2,\ldots, e_n\}$

(2) 试验中每个基本事件 (样本点) 的发生是等可能的，即 $P(e_1)=P(e_2)= \ldots =P(e_n)$.

例如：掷一颗骰子, 观察出现的点数.

计算公式：对古典概型，由概率定义及等可能性，可得：

$$1 = P(S) = P(e_1) + P(e_2) + \cdots + P(e_n) = nP(e_i)$$

故有：$P(e_i) = 1/n \qquad (i = 1,2,\cdots,n)$

若$A = \{e_{i_1}, e_{i_2}, \cdots, e_{i_k}\}$，则有 $P(A) = P(e_{i_1}) + P(e_{i_2}) + \cdots + P(e_{i_k}) = k/n$

称 A 中的样本点为 A 的“有利场合”，于是：

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

例6. 一部5卷的文集随便放在书架上，问：

(1) A： 第三卷刚好放在中间；

(2) B：各卷书自左或自右顺序摆放的概率是多少？

解：5 卷书所有的排列方法数为 $A_5^5 = 5!$

(1) A 所包含的样本点数为 $1 \cdot A_4^4 = 4!$所以

$$P(A) = \frac{4!}{5!} = \frac{1}{5}$$

(2) B所包含的样本点数为2，所以

$$P(B) = \frac{2}{5!} = \frac{1}{60}$$

➢古典概型概率的计算步骤:

(1) 选取适当的样本空间 S, 使它满足有限等可能的要求, 且把事件 A 表示成 S 的某个子集.

(2) 计算样本点总数 n 及事件 A 包含的样本点数 k.

(3) 用下列公式计算:

$$P(A)=\frac{A\text{中的样本点数}k}{S\text{中的样本点总数}n}$$

例7. 袋中装有 4 只白球和 2 只红球. 从中有放回摸球两次, 每次任取一球. 求: (1) A: 两球颜色相同的概率； (2) B: 两球中至少有一只白球的概率.

解： 定义事件： A_1=“两球都是白球”， A_2=“两球都是红球”；

样本空间：取两次球，共有6×6种取法；

A_1包含 4×4 种取法， A_2包含 2×2 种取法, 故：

$P(A_1)=(4\times 4)/(6\times 6)\approx 0.444$，$P(A_2)\approx 0.111$,

由于$A= A_1\cup A_2$， 事件$B=\overline{A_2}$,

所以：$P(A)=P(A_1\cup A_2)=P(A_1)+P(A_2)\approx 0.556$,

$P(B)=1-P(A_2)\approx 0.889$

生日问题： 假定每个人的生日在一年 365 天的任一天都等可能，随机选取 $n(<365)$ 个人，至少有两人生日相同的概率为:

$$p = 1 - \frac{A_{365}^{n}}{365^{n}},$$

取 $n = 32, P(A) \approx 0.75;$

$n = 50, P(A) \approx 0.97;\ \ n = 64, P(A) \approx 0.997.$

- 人们在长期的实践中总结得到“概率很小的事件在一次试验中实际上几乎是不发生的”（称之为实际推断原理)。
- 例. 某接待站在某一周曾接待过12次来访，已知所有这12次接待都是在周二和周四进行的，问是否可以推断接待时间是有规定的？
  - ✓ 假设接待站的接待时间没有规定，而各来访者在一周的任一天中去接待站是等可能的，那么，12次接待来访者都在周二、周四的概率为$2^{12}/7^{12}=0.0000003$.
  - ✓ 现在概率很小的事件在一次试验中竟然发生了，因此，有理由怀疑假设的正确性，从而推断接待站不是每天都接待来访者，即认为其接待时间是有规定的。

抽签与顺序无关？？

- 口袋中有 a 只黑球，b 只白球，现在把球随机地一只只摸出来，求第 k 次摸出的是黑球的概率；
  - 把球依次摸出，共有（a+b）! 种排列法；
  - 第 k 次摸得黑球，有 a 种取法；其他（a+b-1）次摸球共有 $(a+b-1)!$ 种排列法；
  - 因此 $P\{$第 k 次摸到黑球$\}=\dfrac{a\times(a+b-1)!}{(a+b)!}=\dfrac{a}{a+b}$
  - 结果与 k 无关；

例12 一个坛子里共有n个球, 其中一个做了标记. 如果依次从中随机取出k个球, 那么做了标记的球被取出来的概率有多大?

解 – 从 n 个球中选取 k 个球, 一共有$\binom{n}{k}$种选取方法;

– 每一种选取方法都是等可能的;

- 与事件“做标记的球被取出”相关的选法共有 $\binom{1}{1}\binom{n-1}{k-1}$种

- 因此:

$$P\{\text{做标记的球被取出}\}=\frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}}=\frac{k}{n}$$

也可以这样求解（抽签与顺序无关）：

- 设 k 个球是顺序地被取出的,
- 用 A_i 表示做标记的球在第 i 次被取出($i=1,\cdots,k$).
- 既然所有的球在第 i 次被抽取的概率是一样的,

  那么 $P(A_i)=1/n$.
- 而这些事件是彼此互不相容的,

  因此,

$$P\left(\{\text{做标记的球被取出}\}\right)=P\left(\bigcup_{i=1}^{k}A_i\right)=\sum_{i=1}^{k}P(A_i)=\frac{k}{n}$$

另外，$P(A_i)=1/n$ 可以这样推导：

- 考虑到抽球的过程是有顺序的,
- 一共有 $n(n-1)\cdots(n-k+1)=n!/(n-k)!$ 种等可能试验结果,
- 其中有 $(n-1)(n-2)\cdots(n-i+1)(1)(n-i)\cdots(n-k+1)=(n-1)!/(n-k)!$

  种试验结果表示做标记的球被第 i 次取出,

  因此

$$P(A_i)=\frac{(n-1)!}{n!}=\frac{1}{n}.$$

例 13 假设有 $n+m$ 个球，其中 n 个红的，m 个蓝的，将他们随机排成一排，即所有 $(n+m)!$ 种排列都是等可能的．现在记录球的颜色的排列，证明各种可能的颜色排列结果的概率是一样的．

解 - 将$(n+m)$个球的次序排列称为一组球的排列；

- 将$(n+m)$个球的颜色次序排列称为一组球的颜色次序排列；
- 球的排列共有 $(n+m)!$ 种；
- 在红球之间或者蓝球之间作任何一个位置置换，置换的结果并不影响球的颜色次序排列；
- 给定一组球的颜色次序排列，球的排列有多少种情况呢？
- 给定一组球的颜色次序排列，对应于$n!m!$个球的排列；

➢ 这说明球的次序排列也是等可能的,并且每一种颜色次序出现的概率为 $n!m!/(n+m)!$.

- 例如，假设有
  - 2 个红球，记为 r_1, r_2,
  - 2 个蓝球，记为 b_1, b_2,
  - 这样，一共有 4! 种球的排列,
  - 对于每一种颜色次序排列，对应于 2!2! 个球的排列.

- 例如，下面 4 个球的排列对应于相同的颜色次序(红蓝红蓝)排列:

$$r_1, b_1, r_2, b_2 \qquad r_1, b_2, r_2, b_1 \qquad r_2, b_1, r_1, b_2 \qquad r_2, b_2, r_1, b_1$$

因此，每一个颜色次序排列出现的概率都是 **4/24=1/6**.

生活中的等可能概型

- 某人午睡醒来，发现表停转，他打开收音机想听电台报时；求他等待10分钟，听到整点报时的概率是多少？
- 如果在 5 万平方公里的海域里有面积达 40 平方公里的大陆架贮藏着石油，求随意选点钻探，钻到石油的概率是多少？
- 在 400ml 自来水中有一个大肠杆菌，随机取 2ml 水样放到显微镜下观察，发现大肠杆菌的概率是多少？

几何概型

- 古典概型的计算，适用于具有等可能性的有限样本空间；
- 为了克服有限的局限性，利用几何方法，可将古典概型的计算加以推广；

➢ 设试验 E 具有以下特点：

(1) 样本空间 S 是一个几何区域，这个区域的大小是可以度量的（如长度、面积、体积等），并把对 S 的度量记作 $m(S)$;

(2) 向区域 S 内任意投掷，投掷落点在区域内任一个点处都是等可能的，或者设投掷落点在 S 中的区域 A 内的可能性与 A 的度量成正比，而与 A 的位置、状态及形态无关;

- 设事件A：掷点落在区域 A 内，那么事件 A 的概率可用如下公式计算（几何概率公式）：

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

- 可以验证，几何概率公式满足概率的公理化定义（非负性、规范性、可列可加性），因此，它是概率;

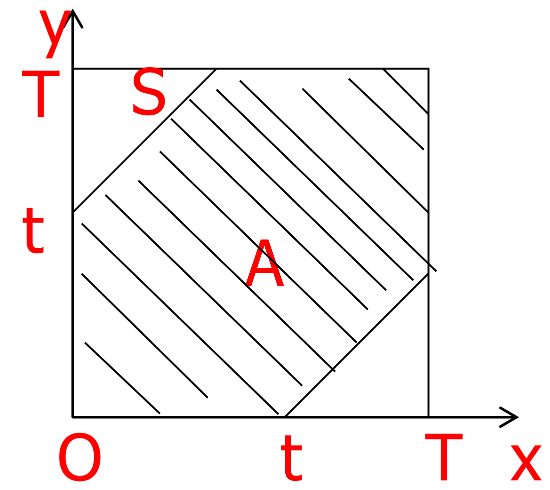
➢ 例16.（约会问题）甲乙两人相约在 0 到 T 这段时间内，在预定的地点会面，他们到达的时间是等可能的，先到的人等候另一个人，经过时间 t $(0<t<T)$ 后离去, 求甲乙两人能会面的概率.

解：设 x，y分别表示甲、乙到达的时刻，则$0 \le x \le T, 0 \le y \le T$，
两人到达的时刻与图中正方形的点是一一对应的，

能会面的充要条件是$|x-y| \le t$

由几何概率公式，所求概率为

$$p = \frac{A\text{的面积}}{S\text{的面积}} = \frac{T^2-(T-t)^2}{T^2} = 1-\left(1-\frac{t}{T}\right)^2$$



- 在约会问题中，一般总希望见到面的概率大一些，这就要求等待时间长一些；
- 而轮船、火车进站等场合却相反，希望不遇见的概率大一些，这就要求等待时间短一些.

条件概率

- 考虑一个有两个孩子的家庭，假定婴儿出生性别男女比例是一样的，求：
  - 这个家庭有一男一女两个孩子的概率？
  - 已知该家庭至少有一个女孩子，有一男一女两个孩子的概率？

条件概率

- 临床数据显示：新冠检测方案对于确诊病例的检测阳性率高达95%，对于健康人士误诊为阳性概率约为8%；
- 抽样数据显示：美国F州检测10000人，确诊率为3%；
  - 1）美国F州居民染上新冠的概率是多少？
  - 2）如果某居民检测结果为阳性，他染上新冠的概率是多少？
  - 3）根据一次检测结果做出判断，误诊概率为多少？
  - 4）如果两次检测结果均显示为阳性，医院将其确诊为新冠肺炎感染者，误诊概率为多少？

条件概率

- 条件概率：设试验 E 的样本空间为 S，A, B 是事件, 要考虑在 A 已经发生的条件下 B 发生的概率, 这就是条件概率, 记为 $P(B|A)$.
  - 例1. 将一枚硬币掷两次, 观察其出现正反面的情况. 设 A—“至少有一次正面”，B—“两次掷出同一面”求：A 发生的条件下 B 发生的概率.

分析： $S=\{HH, HT, TH, TT\}$

$A=\{HH, HT, TH\}$，$B=\{HH，TT\}$

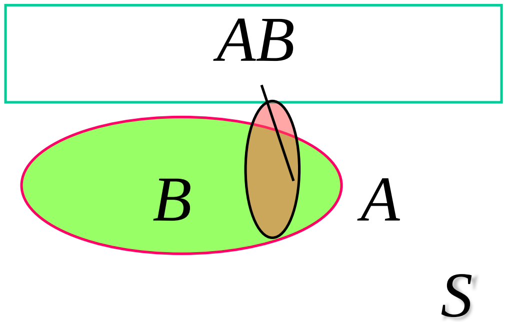
- 已知事件 A 已发生(至少有一次正面)，有了这一信息，知道“TT”不可能发生，即知试验所有可能结果所成的集合就是A。

于是 $P(B|A)=1/3$.

又知, $P(A)=3/4, P(AB)=1/4,$

$$P(B\mid A)=\frac{1/4}{3/4}=\frac{P(AB)}{P(A)}.$$

在古典概型中：样本空间 S 由 n 个样本点组成，若事件 A 包含 n_A 个样本点，AB包含 n_{AB} 个样本点，则



$$P(B \mid A) = \frac{n_{AB}}{n_A} = \frac{n_{AB}/n}{n_A/n} = \frac{P(AB)}{P(A)}$$

直观含义: 求这个条件概率, A发生是一大前提, 在这个空间中求 B 发生的概率，因此 $P(B|A)=P(AB)/P(A)$.

1. 定义： 设A, B是两个事件, 且 $P(A)>0$， 称$P(B|A)$ 为在事件 A 发生的条件下事件 B 发生的条件概率.

$$P(B|A)=\frac{P(AB)}{P(A)}$$

2. 性质: 条件概率符合概率定义中的三条公理，即

1). 非负性：对于每一个事件 B, 有 $P(B|A)\geq 0$.

2). 规范性：$P(S|A)=1$.

3). 可列可加性：设$B_1, B_2, \cdots$两两互不相容, 则

$$P(\bigcup_{i=1}^{\infty} B_i|A)=\sum_{i=1}^{\infty} P(B_i|A).$$

此外，条件概率具有无条件概率类似性质. 例如：

(1) $P(\varnothing \mid A) = 0.$

(2) 设$B_1, B_2, \cdots, B_n$两两互不相容,则

$$P(\bigcup_{i=1}^{n} B_i \mid A) = \sum_{i=1}^{n} P(B_i \mid A).$$

(3) $P(\overline{B} \mid A) = 1 - P(B \mid A).$

(4) $P(B \cup C|A) = P(B|A) + P(C|A) - P(BC|A).$

$$P(B|A)=\frac{P(AB)}{P(A)}$$

特别地，当 $A=S$ 时, $P(B|S)=P(B)$，条件概率化为无条件概率，因此无条件概率可看成条件概率。

例17. 根据长期气象纪录，甲乙两城市一年中雨天的比例分别为20% 和 18%，同时下雨的比例为12%。求条件概率。

例17. 根据长期气象纪录，甲乙两城市一年中雨天的比例分别为20% 和 18%，同时下雨的比例为12%。求条件概率。

解：以 A, B 分别表示甲乙两城市出现雨天。

则 $P(A)=0.2$, $P(B)=0.18$, $P(AB)=0.12$, 于是

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{0.12}{0.18}=\frac{2}{3}$$

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{0.12}{0.2}=0.6$$

(二) 乘法定理:

由条件概率定义，立即可得$P(A)>0$, 则有$P(AB)=P(A)P(B|A)$.

例18. 袋中有某产品 5 件，其中一等品 3 件，二等品 2 件，不放回从中连续抽两件，A 表示第一次抽到一等品，B 表示第二次抽到一等品，求$P(AB)$.

解：$P(A)=\dfrac{3}{5}$, $P(B|A)=\dfrac{2}{4}$

则 $P(AB)=P(A)P(B\,|\,A)=\dfrac{3}{5}\times\dfrac{2}{4}=\dfrac{3}{10}$.

- **推广**: $P(AB)>0$，则有 $P(ABC) = P(A)P(B|A)P(C|AB)$.

  - 可以这样理解：将BC视为一个整体，

$$P(ABC) = P(A)\, P(BC|A)$$

  - $P(BC|A)$是额外条件A发生时BC的联合概率，再用一次乘法公式，有：

$$P(BC|A) = P(B|A)P(C|B|A)=P(B|A)P(C|AB)$$

➢ 推广：$P(AB)>0$，则有 $P(ABC) = P(A)P(B|A)P(C|AB)$.

一般，设 $A_1, A_2, \ldots, A_n$ 是 n 个事件($n\geq 2$), $P(A_1A_2 \ldots A_{n-1})>0$，则有乘法公式：

$$P(A_1A_2\ldots A_n)=P(A_1)P(A_2|A_1)\ldots P(A_{n-1}|A_1A_2\ldots A_{n-2})P(A_n|A_1A_2\ldots A_{n-1}).$$

例19. 设盒中有 a (a>2) 个黑球，b 个白球，连续从盒中取球 3 次，每次取一球，取后不放回，求第 1, 3 次取到黑球第 2 次取到白球的概率。

例19. 设盒中有 a (a>2) 个黑球，b 个白球，连续从盒中取球 3 次，每次取一球，取后不放回，求第1,3次取到黑球第 2 次取到白球的概率。

解: 以A_i 表示事件“第 i 次取到黑球”(i=1, 2, 3),

显然$P(A_1)=\dfrac{a}{a+b}$,若A_1发生，盒中剩下$a-1$个黑球

和b个白球，则得$P(\overline{A}_2\mid A_1)=\dfrac{b}{(a-1)+b}$,类似地

$P(A_3\mid A_1\overline{A}_2)=\dfrac{a-1}{(a-1)+(b-1)}$,于是所求概率为

$$P(A_1\overline{A}_2A_3)=\frac{a}{a+b}\cdot\frac{b}{a+b-1}\cdot\frac{a-1}{a+b-2}$$

复杂事件分解为简单事件

- 送检的两批灯管，每批十支，在运输中各打碎一支，第一批中有1支次品，第二批中有2支次品。现从两批剩下的灯管中各取一支，抽得次品的概率分别是多少？
  - 令A表示从第一批抽得次品，B表示从第二批抽得次品；
  - C表示第一批打碎的是次品；D表示第二批打碎的是次品；
  - $P(A)=\dfrac{9}{10}\times\dfrac{1}{9}=\dfrac{1}{10}$；(分第一批打碎的是次品和正品两种情况)
  - $P(B)=\dfrac{2}{10}\times\dfrac{1}{9}+\dfrac{8}{10}\times\dfrac{2}{9}=\dfrac{1}{10}$；（同上）

全概率公式和贝叶斯公式

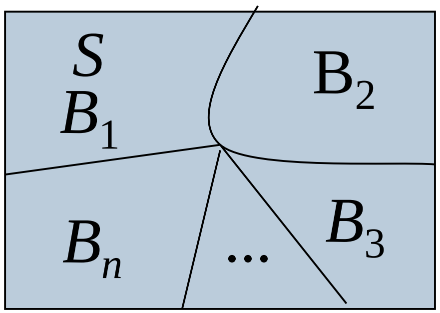
1. 样本空间的划分

定义: 若$B_1, B_2, \cdots, B_n$一组事件满足：

(i) $B_i B_j = \varnothing, i \neq j,\ i, j = 1, 2, \ldots, n,$

(ii) $\bigcup_{i=1}^{n} B_i = S,$

则称$B_1, B_2, \cdots, B_n$为样本空间 S 的一个划分.



(1) 若$B_1, B_2, \ldots, B_n$是样本空间 S 的一个划分, 则每次试验中, 事件$B_1, B_2, \ldots, B_n$中必有一个且仅有一个发生.

(2) 当$n = 2$时, B_1, B_2 为S 的一个划分, 则 B_1, B_2为对立事件, 即$B_1 = \overline{B}_2$.

分析： 由于$A = AS = A\left(\bigcup_{i=1}^{n} B_i\right) = \bigcup_{i=1}^{n} AB_i$,

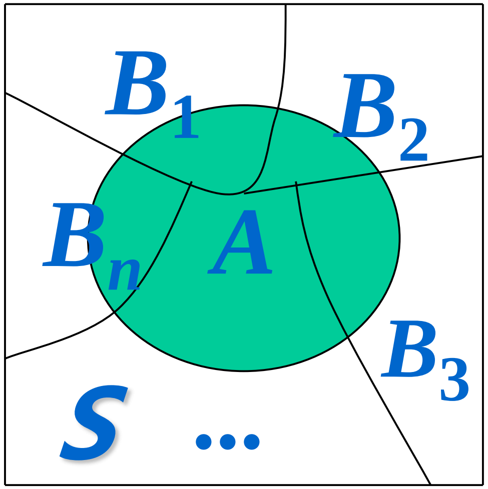
并且$(AB_i)\bigcap(AB_j) = \varnothing \ (i \neq j)$.

由概率的有限可加性, 得 $P(A) = \sum_{i=1}^{n} P(B_i A)$,

再利用乘法定理即得

$$P(A) = \sum_{i=1}^{n} P(B_i A)$$

$$= \sum_{i=1}^{n} P(B_i)P(A \mid B_i).$$



全概率公式:

设 $B_1, B_2, \cdots B_n$ 为 S 的一个划分, $P(B_i) > 0, (i = 1, 2, \cdots, n)$,

对 E 的任一事件A, 有

$$P(A) = \sum_{i=1}^{n} P(B_i) P(A \mid B_i).$$

称上式为全概率公式.

例21. 一批麦种中混有2%的二等种、1%的三等种、1%的四等种。一、二、三、四等种的发芽率为98%、95%、90%、85%，现取一粒种子，问它能发芽的概率是多少？

解: 设表示 B_i “取到一粒种子属 i 等种” $(i=1, 2, \ldots, 4)$，显然 B_i 构成 S 的一个划分；

设 A 表示 “取到一粒种子能发芽”，则由全概率公式

$$P(A)=\sum_{i=1}^{4} P(B_i)P(A|B_i)$$

$$=0.96\times 0.98+0.02\times 0.95+0.01\times 0.90+0.01\times 0.85$$

$$=0.9773$$

例22. 甲箱中装有 3 只红球和 2 只白球, 乙箱中 2 只红球和 2 白球，从甲箱中取两只球放入乙箱中，再从乙箱中取1球，求A: “从乙箱取得白球” 的概率.

例22. 甲箱中装有 3 只红球和 2 只白球, 乙箱中 2 只红球和 2 白球，从甲箱中取两只球放入乙箱中，再从乙箱中取1球, 求 A: “从乙箱取得白球” 的概率.

解: 设B_i={从甲箱中取出 i 只白球}, i=0, 1, 2. 则 B_0, B_1, B_2构成样本空间的一个划分。有:

$$P(B_0)=C_3^2/C_5^2=\frac{3}{10},\quad P(A|B_0)=\frac{2}{6}$$

$$P(B_1)=C_3^1\cdot C_2^1/C_5^2=\frac{6}{10},\quad P(A|B_1)=\frac{3}{6}$$

$$P(B_2)=C_2^2/C_5^2=\frac{1}{10},\quad P(A|B_2)=\frac{4}{6}$$

由全概率公式:

$$P(A)=\sum_{i=0}^{2}P(B_i)P(A|B_i)=\frac{3}{10}\times\frac{2}{6}+\frac{6}{10}\times\frac{3}{6}+\frac{1}{10}\times\frac{4}{6}=\frac{7}{15}$$

贝叶斯公式

设 $B_1, B_2, \ldots, B_n$ 是样本空间S的一个划分，A 是任一随机事件，且P(A)>0，则有：

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n}P(B_j)P(A|B_j)}, i=1, 2, \ldots, n.$$

证明：由条件概率公式 $P(B_i \mid A)=\dfrac{P(AB_i)}{P(A)}$,

由乘法公式: $P(AB_i)=P(A|B_i)P(B_i)$.

由全概率公式: $P(A)=\sum_{j=1}^{n}P(A \mid B_j)P(B_j)$.

于是可得结论.

➢ 贝叶斯公式的直观意义：

✓ 若事件$B_1, B_2, \ldots, B_n$是引起事件 A 发生的 n 个原因，它们的概率 $P(B_i)$ $(i=1, 2, \ldots, n)$ 是在对 A 观察前就已知的，因此通常叫做先验概率。

✓ 如果在一次试验中，事件 A（结果）发生了，那么反过来：A 的发生是由第 i个原因引起的概率 $P(B_i|A)$ 是多少？这就是贝叶斯公式解决的问题。通常称 $P(B_i|A)$ $(i=1, 2, \ldots, n)$为后验概率。

✓ 全概率公式是“由因导果”的一个过程，贝叶斯公式则是“由果溯因”的一个推断公式。

$$P(A)=\sum_{j=1}^{n} P(A|B_j)P(B_j). \qquad P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j)P(A|B_j)}$$

例21. 一批麦种中混有2%的二等种、1%的三等种、1%的四等种。一、二、三、四等种的发芽率为98%、95%、90%、85%，现取一粒种子，问它能发芽的概率是多少？

B_i = "属 i 等种" (i=1, 2, …, 4)，显然 B_i 构成 S 的一个划分，A = "能发芽"

例21.（续）若取一粒种子做发芽实验，结果发芽了，问它是一、二、三、四等种的概率是多大？

解：由贝叶斯公式可得

$$P(B_1 \mid A) = \frac{P(B_1)P(A \mid B_1)}{\sum_{j=1}^{4} P(B_j)P(A \mid B_j)} = \frac{0.96 \times 0.98}{0.9773} = 0.9626$$

同理 $$P(B_2 \mid A) = \frac{0.02 \times 0.95}{0.9773} = 0.0194,$$

$$P(B_3 \mid A) = 0.0092, \qquad P(B_4 \mid A) = 0.0087$$

例24. 对以往数据分析结果表明, 当机器调整得良好时, 产品的合格率为90%, 而当机器发生某一故障时, 其合格率为30%, 每天早晨机器开动时, 机器调整良好的概率为75%, 试求已知某日早上第一件产品是合格品时, 机器调整得良好的概率是多少?

解: 设 A为事件“产品合格”, B 为事件“机器调整良好”,

由已知, $P(A|B)=0.9$, $P(A|\overline{B})=0.3$, $P(B)=0.75$,

$P(\overline{B})=0.25$, 所求的概率为$P(B|A)$.

由Bayes公式:

$$P(B|A)=\frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|\overline{B})P(\overline{B})}$$

$$=(0.9\times0.75)/(0.9\times0.75+0.3\times0.25)$$

$$=0.9.$$

额外条件

- 例：(辩护律师的错误) 一个妻子被谋杀了，其丈夫有嫌疑而被审问，证据显示丈夫有家暴史。
  - 辩护律师认为：数据表明10000个家暴妻子的男子中只有1个会谋杀妻子，因此认为“家暴史”证据与谋杀不相干而应被排除在外。
  - 法官应该接受辩护律师的提议吗？
  - 丈夫有家暴史是否应该被作为其谋杀妻子的证据？
  - “家暴史”作为一个额外条件，是否支持指正丈夫的谋杀？支持的力度有多大？

额外条件

- 由全概率公式：$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$
- 增加一个额外条件F，有：

$$P(A|F) = \sum_{i=1}^{n} P(B_i|F)P\langle A|B_i|F\rangle$$

$$= \sum_{i=1}^{n} P(B_i|F)P(A|B_iF)$$

额外条件

- 定义一个新的概率$\tilde{P}$，$\tilde{P}(A)=P(A|F)$（将条件概率视为特殊的概率模型），进一步用全概率公式，有$\tilde{P}(A)=\sum_{i=1}^{n}\tilde{P}(B_i)\tilde{P}(A/B_i)$，进而有：

$$P(A|F)=\tilde{P}(A)=\sum_{i=1}^{n}\tilde{P}(B_i)\tilde{P}(A|B_i)$$

- 对$\tilde{P}$进行转换，$\tilde{P}(B_i)=P(B_i|F)$，$\tilde{P}(A|B_i)=P(A|B_iF)$

  可以得到：$P(A|F)=\sum_{i=1}^{n}P(B_i|F)P(A|B_iF)$

  结论同前一页；

额外条件

- 由条件概率公式：$P(C|A)=\frac{P(AC)}{P(A)}$；
- 增加一个额外条件F，有：

$$P(C|AF)=\frac{P(AC|F)}{P(A|F)}=\frac{P(A|CF)P(C|F)}{P(A|F)}$$

- 进一步代入有额外条件的全概率公式，有：

$$P(C|AF)=\frac{P(A|CF)P(C|F)}{\sum_{i=1}^{n}P(B_i|F)P(A|B_iF)}$$

➢ 例：(辩护律师的错误) 一个妻子被谋杀了，其丈夫有嫌疑而被审问，证据显示丈夫有家暴史。辩护律师认为，数据表明10000个家暴妻子的男子中只有1个会谋杀妻子，因此认为“家暴史”证据与谋杀不相干而应被排除在外。法官应该接受辩护律师的提议吗？

- 数据显示：10个男子中有一人会家暴，5个被谋杀的已婚妇女中有1个是被丈夫谋杀的；杀害妻子的丈夫中有50%有过家暴行为；
- 用A表示丈夫有家暴行为，用G表示丈夫有罪（谋杀了妻子），用M 表示妻子被谋杀；

- 用A表示丈夫有家暴行为，用G表示丈夫有罪（谋杀了妻子），用M表示妻子被谋杀；
- 由数据：$P(G|A)$=1/10000，$P(G|M)$=0.2，P(A)=0.1，$P(A|GM)$=0.5；$P(A|G^cM)$=0.1；
- 在妻子被谋杀的情况下，有家暴行为的丈夫谋杀了妻子的概率是：
- $P(G|AM)=\frac{P(A|GM)P(G|M)}{P(A|M)}=\frac{P(A|GM)P(G|M)}{P(G|M)P(A|GM)+P(G^c|M)P(A|G^cM)}$

=???

- $P(G|A)$=1/10000；$P(G|M)$=0.2；
- $P(G|AM)$=5/9；
- 辩护律师表示，有家暴行为的丈夫谋杀妻子的概率只有1/10000，这个概率很低；
- 然而，在妻子被谋杀的情况下，“丈夫有家暴行为”这一证据将丈夫谋杀妻子的概率由20%提高到5/9，该证据不应该被排除！
- 本例中，“丈夫有家暴行为”作为“额外条件”，影响了事件发生的概率；
- “额外条件”总是会对事件发生的概率有影响吗？

例25. 设袋中有 a 只红球和 b 只白球 $(b\neq 0)$, 今从袋中取两次球, 每次各取一球, 分为放回和不放回两种情况.

记: A—“第一次取得的是红球”
B—“第二次取得的是红球”

1. 有放回时: $P(A)=\dfrac{a}{a+b}$, $P(B)=\dfrac{a}{a+b}$,

$$P(B|A)=\frac{a}{a+b}=P(B),$$

所以 $P(AB)=P(A)P(B|A)=P(A)P(B)$.

独立性

设 A, B 是试验 E 的两事件，当$P(A)>0$，可以定义 $P(B|A)$

$$P(B \mid A) = \frac{P(AB)}{P(A)}.$$

一般地, $P(B|A) \neq P(B)$， 但当 A 的发生对 B 的发生没有影响时, 有 $P(B|A)=P(B)$。

由乘法公式有 $P(AB)=P(A)P(B|A)=P(A)P(B)$.

例25. 设袋中有 a 只红球和 b 只白球 $(b\neq 0)$, 今从袋中取两次球, 每次各取一球, 分为放回和不放回两种情况.

记: A—"第一次取得的是红球"
B—"第二次取得的是红球"

1. 有放回时: $P(A)=\dfrac{a}{a+b}$, $P(B)=\dfrac{a}{a+b}$,

$$P(B\mid A)=\frac{a}{a+b}=P(B),$$

所以 $P(AB)=P(A)P(B|A)=P(A)P(B)$.

2. 不放回时: B—“第二次取得的是红球”， $P(B)$?

$$P(A)=\frac{a}{a+b},\quad P(AB)=\frac{a(a-1)}{(a+b)(a+b-1)},$$

$$P(\overline{A}B)=\frac{ba}{(a+b)(a+b-1)},$$

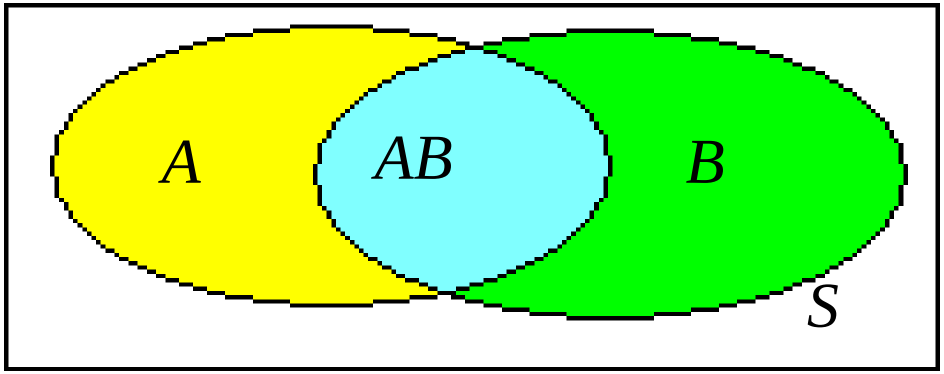
$$P(B)=P(AB)+P(\overline{A}B)=\frac{a}{a+b},$$

得到 $P(B|A)=\dfrac{a-1}{a+b-1}$，显然 $P(B|A)\neq P(B)$,

从而 $P(AB)\neq P(A)P(B)$.

定义1: 设 A, B 是两事件, 如果满足等式 $P(AB)=P(A)P(B)$，则称 事件 A 与事件 B 是相互独立的事件，简称A, B 独立.

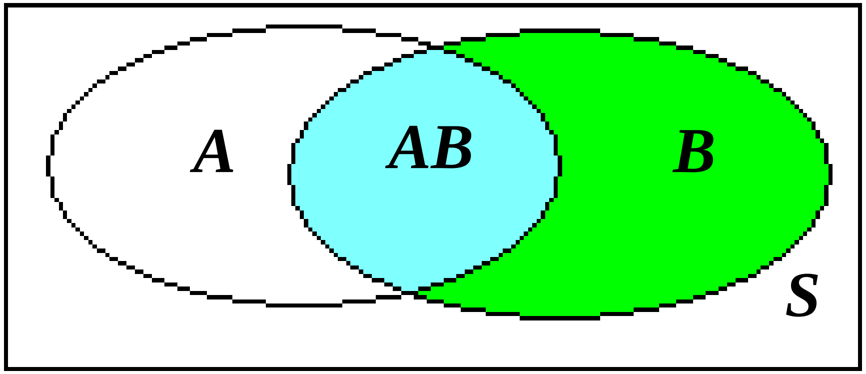
必然事件 S 和不可能事件$\varnothing$与任何事件 A 都独立；



定理： 如果事件 A, B 相互独立，且$P(B)>0$, 则 $P(A|B)=P(A)$，反之亦然.

证：由条件概率及上式定义得：

$$P(A\,|\,B)=\frac{P(AB)}{P(B)}=\frac{P(A)P(B)}{P(B)}=P(A).$$



定理: 若A,B相互独立，则A与$\overline{B}$, $\overline{A}$ 与B, $\overline{A}$ 与$\overline{B}$ 也相互独立.

证明: (1) $\because A\overline{B} = A\text{-}AB$

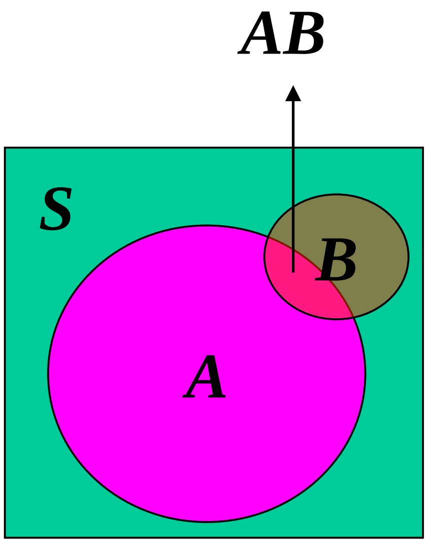
$$\therefore P(A\overline{B}) = P(A) - P(AB)$$

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

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

$$= P(A)P(\overline{B}).$$

故 A 与 $\overline{B}$ 相互独立.



例26. 甲、乙两射手向同一目标各射击一次, 甲击中目标的概率为0.9，乙击中目标的概率为 0.8，求在一次射击中目标被击中的概率。

例26. 甲、乙两射手向同一目标各射击一次, 甲击中目标的概率为0.9，乙击中目标的概率为 0.8，求在一次射击中目标被击中的概率。

解： 记 A: “甲击中目标”， B:“乙击中目标”；
C: “目标被击中”， 这里可认为事件 A, B 独立，则

$$
\begin{aligned}
P(C) &= P(A\cup B) \\
&= P(A)+P(B)-P(AB) \\
&= P(A)+P(B)-P(A)P(B) \\
&= 0.9+0.8-0.9\times 0.8 = 0.98
\end{aligned}
$$

定义2: 设A, B, C 是三个事件, 若满足:

$$P(AB)=P(A)P(B),\ P(AC)=P(A)P(C),$$

$$P(BC)=P(B)P(C),\ P(ABC)=P(A)P(B)P(C)$$

则称 A, B, C 为相互独立的事件.

注意:

1. 定义2中前面三个式子表明 A, B, C三事件两两独立, 并不能说 A, B, C 三事件相互独立.

例：盒中有编号为1,2,3,4的4只球，随机地从中取出一只，定义：

- 事件A：取得的是1号或者2号球；
- 事件B：取得的是1号或者3号球；
- 事件C：取得的是1号或者4号球；

事件A、B、C相互独立吗？

事实上，$P(A)=P(B)=P(C)=1/2$，$P(AB)=P(BC)=P(AC)=1/4$；

$P(ABC)=1/4$；

定义3：对 n 个事件 $A_1, A_2,\ldots,A_n$, 如果对所有可能的组合 $1\le i<j<k<\ldots\le n$ 成立着;

$$P(A_iA_j)=P(A_i)P(A_j)$$

$$P(A_iA_jA_k)=P(A_i)P(A_j)P(A_k)$$

$$\vdots$$

$$P(A_1A_2\ldots A_n)=P(A_1)P(A_2)\ldots P(A_n),$$

则称这 n 个事件 $A_1, A_2, \ldots, A_n$ 相互独立.

推论:

1. 如果 $A_1, A_2, \ldots, A_n$ 相互独立，那么其中任意 m 个事件也相互独立.

2. 如果 $A_1, A_2, \ldots, A_n$ 相互独立，则将其中任意个事件换成其逆事件后也相互独立.

3. $P(A)>0, P(B)>0$, 则A, B相互独立与A,B 互不相容不能同时成立.

证：由A, B 独立知 $P(AB)=P(A)P(B)\neq 0;$

若A, B互不相容, 则有$AB=\varnothing$, 从而 $P(AB)=0$,矛盾!

所以A, B 独立与不相容不能同时成立.

定义4： 设 $A_1, A_2, \ldots, A_n$ 是 n 个事件，如果对任意的$1\le i<j \le n$ 有 $P(A_iA_j)=P(A_i)P(A_j)$，则称 这 n 个事件两两独立.

注意：若 n 个事件相互独立，必蕴含这 n 个事件两两相互独立，反之不真。

例27. 假若每个人血清中有肝炎病毒的概率为0.4%, 混合100个人的血清，求此血清中含有肝炎病毒的概率.

解 以 A_i (i=1,2,…100)记“第i个人的血清含有肝炎病毒”，显然A_i相互独立的. 所求概率为

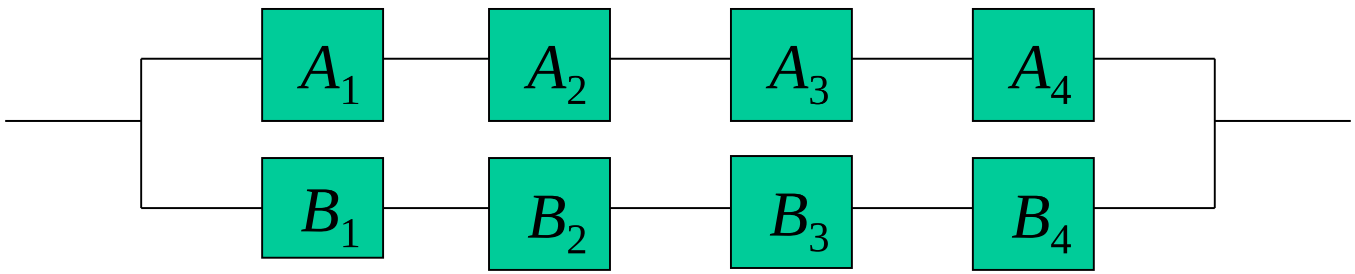
$$P(A_1 \cup \cdots \cup A_{100}) = 1 - P(\overline{A_1}\,\overline{A_2} \cdots \overline{A_{100}})$$

$$= 1 - P(\overline{A_1}) \cdots P(\overline{A_{100}}) = 1 - 0.996^{100} \approx 0.33$$

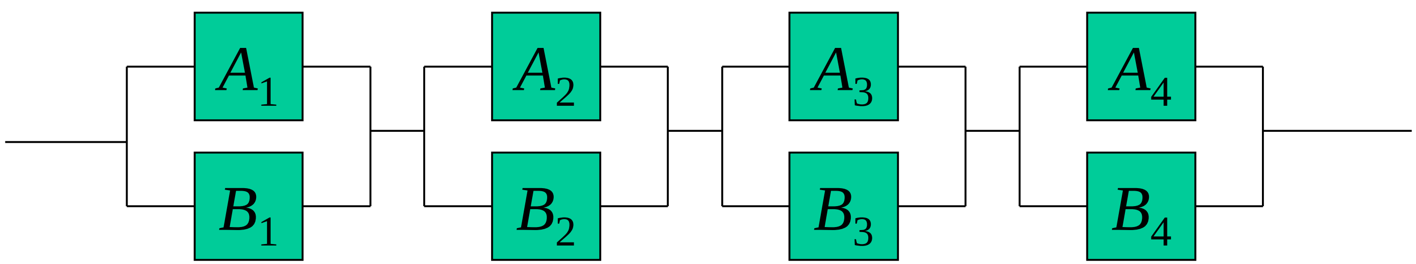
虽然每个人有病毒的概率很小,但是混合后则有很大概率.

例28. 设有8个元件, 每个元件的可靠性均为 p (元件能正常工作的概率), 按如下两种方式组成系统, 试比较两个系统的可靠性.

系统一:先串联后并联



系统二:先并联后串联



解：用A_i, B_i, 表示如图中诸元件能正常工作的事件, i=1, 2, 3, 4.

C_1, C_2表示系统一、二可靠的事件.

则 $C_1=(A_1A_2A_3A_4)\cup(B_1B_2B_3B_4)$

$C_2=(A_1\cup B_1)\ (A_2\cup B_2)\ (A_3\cup B_3)\ (A_4\cup B_4)$

于是$P(C_1)=P(A_1A_2A_3A_4)+P(B_1B_2B_3B_4)-P(A_1A_2A_3A_4B_1B_2B_3B_4)$

$=p^4+p^4-p^8=p^4(2-p^4)$,

于是$P(C_1)=P(A_1A_2A_3A_4)+P(B_1B_2B_3B_4)-P(A_1A_2A_3A_4B_1B_2B_3B_4)$

$=p^4+p^4-p^8=p^4(2-p^4)$,

$P(C_2)=P(A_1\cup B_1)P(A_2\cup B_2)P(A_3\cup B_3)P(A_4\cup B_4)$

$=(p+p-p^2)^4=p^4(2-p)^4$,

当$0<p<1$时, 显然有 $p^4(2-p)^4>p^4(2-p^4)$, 所以 $P(C_2)>P(C_1)$.

一般地 $2n$ 个元件组成以上两个系统, $P(C_1)=p^n(2-p^n)$,

$P(C_2)=p^n(2-p)^n$, 所以有 $P(C_2)>P(C_1)$. $(0<p<1)$

独立试验（伯努利试验）

- 在实际中经常碰到这一类试验，每次试验的结果只有两种，这种概型称为伯努利试验。例如，检验一件产品的质量看其是合格品还是次品；射击一次的结果击中或未击中；考试一次是否通过；等等；
- 有些试验的结果虽然不只有两个，但有时人们在众多的结果中只关心其中一个事件A，而把其余情况都归结为$\bar{A}$，这样又把此试验变成伯努利试验

- 将伯努利试验独立重复进行 n 次，称为 **n 重伯努利试验；**

设在一次伯努利试验中，事件A 发生的概率为$p(0<p<1)$，则在n重伯努利试验中，A 发生k 次的概率为$C_n^k p^k(1-p)^{n-k}$，$(k=0,1,2,\ldots,n)$

例 一张试卷上有10道四选一的单项选择题，某同学投机取巧，随意选答案，试问他至少答对6 道题的概率有多大?

解：这是10重伯努利试验，设事件B表示“至少答对6 道题”，

则$P(B)=\sum_{k=6}^{10} C_{10}^{k}(\frac{1}{4})^{k}(1-\frac{1}{4})^{10-k}=0.01973$

由实际推断原理可知（小概率事件在一次试验中几乎不会发生）能在10道题中答对6 道题以上可能性极小