1.1 随机事件和概率

- 随机现象
- 样本空间
- 概率

1.2 基本概率模型

- 古典概率模型
- 几何概率模型
- 概率空间的公理化定义

1.3 条件概率

- 全概率公式
- 贝叶斯概率公式

1.4 独立性

- 独立事件
- 独立试验
- n-贝努利试验

1.1 随机事件和概率

- 随机现象
- 样本空间
- 概率

(1) 确定性现象

一个标准大气压下，水100度沸腾，0度结冰

向上抛一枚硬币，受地球引力作用，硬币总会落到地上

(2) 随机现象— 不确定性现象

硬币落到地上，观察正面向上或反面向上

从一副扑克中，抽一张扑克，记录其花色

(3) 随机现象的基本属性

(i) 该试验可重复进行或该现象可重复观察

(ii) 试验之前(或现象发生之前)，并不知道会出现何种结果

(iii) 该试验(该现象)所有可能的结果是已知的

(i), (ii) 是随机现象的定性描述

(iii) 是随机现象的定量描述

- 样本空间和样本点

考虑某随机现象。

记所有可能的结果为Ω，并称Ω该随机现象的样本空间

记每一个结果为ω，并称为样本点。

显然$\omega \in \Omega$

不同的样本空间表示不同的随机现象

例如，抛掷一枚硬币，样本空间$\Omega = \{H, T\}$ (通常这样认为)

但如果硬币较厚，样本空间可能为$\Omega = \{H, T, S\}$，S表示硬币垂立

- 事件和事件的发生

具有某种属性的基本结果构成事件。通常用A, B, C等表示，可以写成$A \subseteq \Omega$

例1. 投掷一颗骰子，样本空间$\Omega = \{1, 2, 3, 4, 5, 6\}$.

$A = \{2, 4, 6\}$，$B = \{1, 3, 5\}$，$C = \{4, 5, 6\}$

例2. 考虑某零件的寿命，样本空间$\Omega = [0, \infty)$.

$A = [100, 1000]$，$B = [100, \infty)$.

如果某次试验的结果$\omega \in A$，那么称A发生；

否则，称A不发生。

- 概率

概率：用于描述随机现象的规律
随机现象具有不确定性：每次试验之前，并不知道会出现何种结果。
用某结果出现的频繁程度或一次试验中某结果出现的可能性大小，来作为随机现象规律性的刻画

例. 甲、乙两人下棋。
通过多次反复观察甲、乙两人下棋，可以判定甲、乙两人的棋艺。

事件A发生的概率：
一次试验中事件A发生的可能性大小，用$P(A)$表示。
$P(A)$随机现象固有的属性，与实验者或实验的次数无关。

- 概率的计算

物理方法

(1) 抛掷均匀硬币

(2) 随机抽取一张扑克

统计方法

基本思想：用频率估计概率

通过重复试验N次，计算事件A发生的次数N_A，得频率$f_N(A)=\frac{N_A}{N}$

令$N\to\infty$，$f_N(A)$的极限即为事件A的概率$P(A)$.

the coin to flip it-1.pdf

Click the coin to flip it--or enter a number and click Auto Flip.

1000 Auto Play

Results: Session

Historical

Explain Back

Results:

Flip	Session	Expected
Heads	49.5% (495/1000)	50% (1/2)
Tails	50.5% (505/1000)	50% (1/2)

the coin to flip it-2.pdf

Click the coin to flip it--or enter a number and click Auto Flip.

5000 Auto Play

Results: Session

Historical

Explain Back

Results:

Flip	Session	Expected
Heads	49.06% (2453/5000)	50% (1/2)
Tails	50.94% (2547/5000)	50% (1/2)

the coin to flip it-3.pdf

Click the coin to flip it--or enter a number and click Auto Flip.

10000 Auto Play

Results: Session

Historical

Explain Back

Results:

Flip	Session	Expected
Heads	49.46% (4946/10000)	50% (1/2)
Tails	50.54% (5054/10000)	50% (1/2)

the dice to roll.pdf

Click the dice to roll--or enter a number and click Auto Roll.

Number of dice:

1

1000　Auto Play

Results: Session

Historical

Explain　Back

Results:

Roll	Session	Expected
1	17.1% (513/3000)	16.67% (1/6)
2	16.33% (490/3000)	16.67% (1/6)
3	17.1% (513/3000)	16.67% (1/6)
4	17.2% (516/3000)	16.67% (1/6)
5	16.87% (506/3000)	16.67% (1/6)
6	15.4% (462/3000)	16.67% (1/6)

实验者	抛掷次数(n)	正面向上次数（频数m）	频率(m/n)
棣莫佛（英国1667-1754）	2048	1061	0.5181
蒲丰（法国1707-1788）	4040	2048	0.5069
费勒（美国1906-1970）	10000	4979	0.4979
皮尔逊（英国1857-1936）	12000	6019	0.5016
皮尔逊（英国1857-1936）	24000	12012	0.5005
德-摩根（英国1806-1871）	4092	2048	0.5005

不同时代、不同国家的科学家们为了探求真相，独立重复了成千上万次投掷硬币的试验，唯一目的在于获得第一手资料。

关于统计方法的说明

(1) 统计方法具体、可计算

(2) 统计方法的基本出发点：频率极限存在；并且不依赖于具体的试验环境

(3) 概率论发展的早期，许多人做了大量试验，从而相信这一点。

Bernoulli和Borel 将给出它的数学证明

概率论学科的主要目的：计算随机事件的概率

从简单事件到复杂事件；

涉及事件的运算和概率的运算性质

- 事件的运算

与集合运算类似

注意概率术语的正确使用

$\emptyset$ — 不可能事件

Ω — 必然事件

$A \subseteq B$ — 事件A发生意味着事件B发生

$A \cap B$ — 事件A和B同时发生，有时写作AB

$A \cup B$ — 事件A或者B 发生

$\bar{A}$ — A的对立事件，即A不发生

$A \setminus B$ — 事件A发生，但B不发生

$A \cap B = \emptyset$ — A和B互不相交

当A, B 互不相交时，写$A \cup B = A + B$

De Morgan对偶运算原理

$$\overline{(\cap A_n)} = \cup \bar{A}_n, \quad \overline{(\cup A_n)} = \cap \bar{A}_n$$

- 概率运算的基本性质

由频率和极限的运算性质得到:

(1) $P(\emptyset) = 0, P(\Omega) = 1$

(2)

$$P(\bar{A}) = 1 - P(A)$$

(3) 如果$A \cap B = \emptyset$，那么

$$P(A + B) = P(A) + P(B)$$

如果$A_1, A_2, \cdots, A_m$互不相交，那么

$$P(\sum_{k=1}^{m} A_k) = \sum_{k=1}^{m} P(A_k)$$

问题：如果$A_1, A_2, \cdots, A_m, \cdots$互不相交，那么

$$P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$$

吗？

1.2 概率模型

- 古典概率模型
- 几何概率模型
- 概率空间的公理化定义

概率模型— 随机现象的数学描述，包括样本空间、所关心的事件、每个事件的概率大小

- 古典概率模型
- 模型特征：

(1) 有限个基本结果

(2) 每个结果等可能地发生

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_N\}, \quad N < \infty$$

$$P(\{\omega_i\}) = \frac{1}{N}, \quad i = 1, 2, \cdots N$$

如果A是一个事件，那么

$$P(A) = \frac{|A|}{N}$$

其中，$|A|$表示A中所含基本结果的个数

例. 抛掷一枚均匀硬币；随机抽取一张扑克

对于一个古典概率模型，关键在于计算N和事件A所包含的基本结果个数。通常，需要一些技巧

- 乘法原理、排列、组合

(1) 乘法原理：完成一件任务需要分k步，其中第i步有m_i种方案。那么完成这件任务共有$\prod_{i=1}^{k} m_i$ 种不同方案

(2) 排列: 从N个不同物体随机抽取k个进行排序，共有P_N^k种不同结果

(3) 组合: 从N个不同物体随机抽取k个组成一组，共有$\binom{N}{k}$种不同结果

(4) 一些常用的关系式

$$P_N^0 = 1, \quad \binom{N}{0} = 1, \quad P_N^k = N(N-1)(N-2)\cdots(N-k+1)$$

$$\binom{N}{k} = \frac{P_N^k}{k!}$$

$$\binom{N}{k} + \binom{N}{k-1} = \binom{N+1}{k}$$

$$N! \sim \sqrt{2\pi N} e^{-N} N^N, \quad N \to \infty$$

- 古典概率模型的例子

例1. 一个袋子中装有8个黑球，2个红球。现随机抽出一球，问所得的是红球的概率为多少？

(1) 将球标号：1，2，3，4，5，6，7，8，9，10. 其中号码为9，10的是红球。

(2)随机抽一个球，样本空间为

$$\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

(3) 事件A表示所得的是红球

(4)

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{10} = \frac{1}{5}$$

例2. 一个袋子中装有8个黑球，2个红球。现随机抽出两球，问所得为一个红球，一个黑球的概率为多少？

(1) 将球标号：1，2，3，4，5，6，7，8，9，10. 其中号码为9，10的是红球。

(2)随机抽2个球，样本空间为

$$\Omega = \left\{ \begin{array}{l} \{1,2\},\{1,3\}\{1,4\}\{1,5\}\{1,6\}\{1,7\}\{1,8\}\{1,9\},\{1,10\} \\ \{2,3\},\{2,4\},\{2,5\},\{2,6\},\{2,7\},\{2,8\},\{2,9\},\{2,10\} \\ \{3,4\},\{3,5\},\{3,6\},\{3,7\},\{3,8\},\{3,9\},\{3,10\} \\ \{4,5\},\{4,6\},\{4,7\},\{4,8\},\{4,9\},\{4,10\} \\ \{5,6\},\{5,7\},\{5,8\},\{5,9\},\{5,10\} \\ \{6,7\},\{6,8\},\{6,9\},\{6,10\} \\ \{7,8\},\{7,9\},\{7,10\} \\ \{8,9\},\{8,10\} \\ \{9,10\} \end{array} \right\}$$

(3) 事件A表示所得为一个红球，一个黑球

$$A = \left\{ \begin{array}{c} \{1,9\},\{1,10\} \\ \{2,9\},\{2,10\} \\ \{3,9\},\{3,10\} \\ \{4,9\},\{4,10\} \\ \{5,9\},\{5,10\} \\ \{6,9\},\{6,10\} \\ \{7,9\},\{7,10\} \\ \{8,9\},\{8,10\} \end{array} \right\}$$

(4)

$$P(A) = \frac{|A|}{|\Omega|} = \frac{16}{45}$$

例3 甲乙两人各投掷一颗骰子，问甲的点数比乙大的概率？

一颗骰子有六面，分别刻有1，2，3，4，5，6个点。

(1) 用(x, y)表示甲出现x点，乙出现y点。

(2) 样本空间

$$\Omega = \left\{ \begin{array}{c} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6) \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6) \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6) \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6) \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6) \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{array} \right\}$$

(3) 事件

$$A = \left\{ \begin{array}{l} (2,1) \\ (3,1),(3,2) \\ (4,1),(4,2),(4,3) \\ (5,1),(5,2),(5,3),(5,4) \\ (6,1),(6,2),(6,3),(6,4),(6,5) \end{array} \right\}$$

(4)

$$P(A) = \frac{15}{36}$$

- 几何概率模型

模型特征：样本空间是一个区域，所有基本结果均等可能发生。

(1) 样本空间含有不可数个基本结果，每个基本结果出现的概率为0
严格地说，
Ω 是$\mathbb{R}, \mathbb{R}^2, \cdots, \mathbb{R}^k$上的可求长，可求面积，可求体积的区域。或者说，$\Omega$ 是$\mathbb{R}, \mathbb{R}^2, \cdots, \mathbb{R}^k$ 上的可测区域。

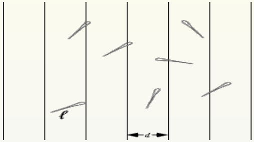
(2) 事件A是Ω的可测子集
(3) 事件A的概率

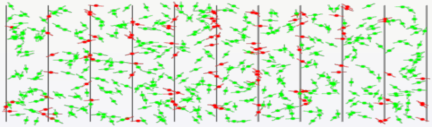
$$P(A) = \frac{|A|}{|\Omega|}$$

对于几何模型，关键在于计算$|\Omega|$和$|A|$.

例1. Buffon (布封) 投针问题(Buffon's Needle Problem)

Buffon's Needle Problem





The above figure shows the result of 500 tosses of a needle of length parameter $x = 1/3$, where needles crossing a line are shown in red and those missing are shown in green. 107 of the tosses cross a line, giving $\hat{\pi} = 3.116 \pm 0.073$.

自然史科学家— Buffon 布封（1707-1788）

Georges-Louis Leclerc, comte de Buffon

- 布封— 中小学生所熟知的

【1】《马》被选入七年级第29课,

【2】《松鼠》被选入小学人教版五年级上册第17课课文,

沪教版六年级上册第17课课文,

苏教版七年级下册第10课课文,

部编版五年级上册第17课课文。

生平简介

1. 1707年出生于法国孟巴尔城一个律师家庭，原名乔治·路易·勒克来克，因继承关系，改姓德·布封；

2. 1728年1728年大学法律本科毕业，随后学习医学两年；

3. 1730年后，游历了法国南方、瑞士和意大利。在德国学者辛克曼的影响下，刻苦研究博物学；

4. 1733年，法国科学院任助理研究员，翻译牛顿的《微积分》；

5. 1739年，被任命为皇家御花园和御书房总管，直到逝世；

6. 1753年，他当选为法兰西学院院士；

7. 1777年，法国政府给他建立了一座铜像，座上用拉丁文写着：

“献给和大自然一样伟大的天才”

- 《自然史》

布封收集了大量的动、植、矿物样品和标本，毕生从事博物学的研究，每天埋头著述，四十年如一日，终于写出三十六册的巨著《自然史》。

这是一部博物志，包括地球史、人类史、动物史、鸟类史和矿物史等几大部分，综合了无数的事实材料，对自然界作了精确、详细、科学的描述和解释，提出许多有价值的创见。破除各种宗教迷信和无知妄说，把上帝从宇宙的解释中驱逐出去，这是布封对现代科学的一大贡献。

处处留心皆学问，行行出状元

- 唯物主义论者

他坚持以唯物主义观点解释地球的形成和人类的起源，指出地球与太阳有许多相似之处，地球是冷却的小太阳；地球上的物质演变产生了植物和动物，最后有了人类；

人类的进化不是如圣经《创世纪》所说的，人类的祖先亚当、夏娃偷吃了禁果才有了智慧，而是在社会实践中获得了知识，增长了才干。

布丰观察、研究大地、山脉、河川和海洋，寻求地面变迁的根源，开了现代地质学的先河。尤其在物种起源方面，他倡导生物转变论，指出物种因环境、气候、营养的影响而变异，对后来的进化论有直接的影响。

达尔文称他“是现代以科学眼光对待这个问题的第一人”（《物种起源》导言）。

注：伽利雷1564 -1642 ；马克思1818 —1883

- 《自然史》的文学价值

《自然史》关于动物活动形态的描绘尤富于艺术性。作者以科学的观察为基础，用形象的语言勾画出各种动物的一幅幅肖像，还通过拟人化的手法，在一定程度上反映了反封建的民主思想倾向。

1749年，《自然史》的头三册一出版，就轰动了欧洲的学术界。

1753年，他当选为法兰西学院院士。入院时发表的著名演说《论风格》，是一篇经典的文论。他针对当时文坛上那种追求绮丽纤巧的风尚，呼吁文章要言之有物、平易近人，提出“风格即人”的名言，强调思想内容对艺术形式的决定作用。

文如其人

先建立模型。

假设$l \le d$

令θ 表示针与平行线的夹角，$\theta \le \frac{\pi}{2}$

令a表示针的中心离平行线的距离，$a < \frac{d}{2}$

$\Omega = [0, \frac{d}{2}] \times [0, \frac{\pi}{2}]$

用A表示针与平行线相交

$$A\text{发生} \Leftrightarrow a \le \frac{l}{2}\sin\theta$$

$$\begin{aligned} P(A) &= \frac{|A|}{|\Omega|} \\ &= \frac{\int_0^{\frac{\pi}{2}} \frac{l}{2}\sin\theta d\theta}{\frac{d}{2}\frac{\pi}{2}} \\ &= \frac{2l}{\pi d} \end{aligned}$$

- π的计算

$$\pi = \frac{2l}{d \cdot P(A)}$$

给定d, l，用频率估计概率，可以计算π。 — 最早的模拟计算

- Mario Lazzarini, an Italian mathematician, performed the Buffon's needle experiment in 1901. Tossing a needle 3408 times, he attained the well-known estimate $\frac{355}{113}$ for π, which is a very accurate value, differing from π by no more than 3×10^{-7}.

This is an impressive result, but is something of a cheat, as follows. Lazzarini chose needles whose length was $\frac{5}{6}$ of the width of the strips of wood. In this case, the probability that the needles will cross the lines is $\frac{5}{3\pi}$. Thus if one were to drop n needles and get x crossings, one would estimate π as

$$\pi \sim \frac{5}{3} \times \frac{n}{x}$$

π is very nearly $\frac{355}{113}$.

in fact, there is no better rational approximation with fewer than 5 digits in the numerator and denominator. So if one had n and x such that:

$$\frac{355}{113} = \frac{5}{3} \times \frac{n}{x}$$

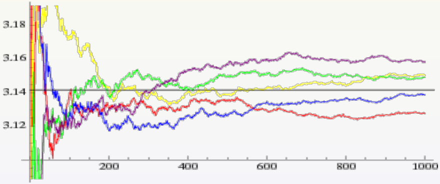
or equivalently,

$$x = \frac{113n}{213}$$

one would derive an unexpectedly accurate approximation to π, simply because the fraction $\frac{355}{113}$ happens to be so close to the correct value. But this is easily arranged.

Buffon's Needle Problem

The above figure shows the result of 500 tosses of a needle of length parameter $x = 1/3$, where needles crossing a line are shown in red and those missing are shown in green. 107 of the tosses cross a line, giving $\hat{\pi} = 3.116 \pm 0.073$.



Several attempts have been made to experimentally determine π by needle-tossing. π calculated from five independent series of tosses of a (short) needle are illustrated above for one million tosses in each trial $x = 1/3$. For a discussion of the relevant statistics and a critical analysis of one of the more accurate (and least believable) needle-tossings, see Badger (1994). Uspensky (1937, pp. 112-113) discusses experiments conducted with 2520, 3204, and 5000 trials.

To do this, one should pick n as a multiple of 213, because then $\frac{113n}{213}$ is an integer; one then drops n needles, and hopes for exactly $x = \frac{113n}{213}$ successes.

If one drops 213 needles and happens to get 113 successes, then one can triumphantly report an estimate of π accurate to six decimal places. If not, one can just do 213 more trials and hope for a total of 226 successes; if not, just repeat as necessary. Lazzarini performed $3408 = 213 \times 16$trials, making it seem likely that this is the strategy he used to obtain his "estimate".

例2. 约会问题

甲、乙两人相约，晚上7:00至8:00在电影院门口见面，每人最多等候20分钟。假定甲、乙二人在7:00至8:00之间随时可能出现，问两人能见面的可能性是多少？

- 先进行如下约化。考虑正方形区域

$$\Omega = [0, 60] \times [0, 60]$$

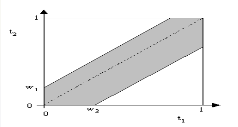
甲、乙二人随机到达，到达时刻为(x, y)，等可能地落在该区域Ω

令A表示两人相见。

$$A\text{发生} \Leftrightarrow |x-y| \le 20$$

因此，

$$P(A) = \frac{|A|}{|\Omega|} = \frac{60^2 - 40^2}{60^2}$$



- 其他概率模型

例1. 抛掷一枚非均匀硬币

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

$$P(\{H\}) \neq P(\{T\})$$

例2. 某口袋里装有中奖彩票50张，其中一等奖5张，二等奖10张，三等奖15张，余下的没有奖。现随机抽取一张彩票，记录中奖情况。

$$\Omega = \{\text{一等奖、二等奖、三等奖、没有奖}\}$$

一等奖概率为$\frac{1}{10}$，二等奖概率为$\frac{1}{5}$，三等奖概率为$\frac{3}{10}$，没有奖概率为$\frac{2}{5}$

- 概率空间公理化体系

19世纪末-20世纪初，在Lebesgue, Hilbert, Frechet, Wiener等人的努力下，抽象的测度和积分理论已经建立起来。
受此启发，俄国数学家Kolmogorov (参考引言) 建立了概率论的公理化体系。

样本空间：Ω
事件类：σ-域$\mathcal{A}$
概率：P
$(\Omega, \mathcal{A}, P)$ 构成概率空间，是随机现象的数学描述— 概率模型

- $\mathcal{A}$ 满足下列条件:

(1) $\emptyset, \Omega \in \mathcal{A}$

(2) 如果$A \in \mathcal{A}$，那么$\bar{A} \in \mathcal{A}$

(3) 如果$A_n \in \mathcal{A}, n \geq 1$，那么$\cup_{n=1}^{\infty} A_n \in \mathcal{A}$

- P 满足下列条件:

(1) $P(\emptyset) = 0, P(\Omega) = 1$

(2) 对任意$A \in \mathcal{A}$，$P(A) \geq 0$

(3) 如果$A_n \in \mathcal{A}, n \geq 1$ 互不相交，那么

$$P(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$$

任何满足上述性质的P都称为是空间$(\Omega, \mathcal{A})$上的概率。

$(\Omega, \mathcal{A}, P)$称为概率空间。

- 概率P的运算性质

(1) 如果$A \subseteq B$，那么$P(A) \leq P(B)$

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

(3) 对任意一列$A_n \in \mathcal{A}, n \geq 1$，

$$P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n), \quad \text{次可加性}$$

(4) 如果$A, B \in \mathcal{A}$，那么

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

一般地，对任意$A_k \in \mathcal{A}, 1 \leq k \leq m$，

$$\begin{aligned} P(\bigcup_{k=1}^{m} A_k) &= \sum_{k=1}^{m} P(A_k) - \sum_{k \neq l}^{m} P(A_k A_l) \\ &\quad + \cdots + (-1)^{m-1} P(\bigcap_{k=1}^{m} A_k) \end{aligned}$$

例1. 匹配问题

某人写n封信，n个信封。现随机地将n封信放入n个信封，每个信封装一封信。求至少有一封信装入正确的信封里的概率？

令A_i表示第i封信装入了正确的信封里。

$$P(\bigcup_{i=1}^{n} A_i) = ?$$

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

$$P(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, i \neq j$$

$$\cdots$$

$$P(\bigcap_{i=1}^{n} A_i) = \frac{1}{n!}$$

$$\begin{aligned}
P(\bigcup_{i=1}^{n} A_i) &= \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) \\
&\quad + \cdots + (-1)^{n-1} P(\bigcap_{i=1}^{n} A_i) \\
&= 1 - \frac{1}{2!} + \cdots + (-1)^{n-1}\frac{1}{n!} \\
&\sim 1 - e^{-1}
\end{aligned}$$

1.3 条件概率

- 全概率公式
- 贝叶斯公式

条件概率是计算概率的一个重要技巧

例1. 从52张扑克种随机抽取一张，让你猜。

(i)如果没有任何提示，那么猜“红桃5”的可能性为$\frac{1}{52}$。
(ii) 如果给出暗示：“抽出的扑克是红色”，那么猜“红桃5”的可能性为$\frac{1}{26}$。

- 暗示“抽出的扑克是红色”意味着这件事已经发生了，它对另一事件的发生产生影响。

- 条件概率

假设$(\Omega, \mathcal{A}, P)$是一个概率空间，A, B是两个事件，$P(B) > 0$。令

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

称$P(A|B)$是在B发生的条件下，A发生的条件概率

- $P(B) > 0$:

(1) 上式中分母不能为0

(2) 零概率事件无法观察到

- 乘法公式

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

可改写成

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

- 链式法则

$$\begin{aligned} P(ABC) &= P(A|BC)P(BC) \\ &= P(A|BC)P(B|C)P(C) \end{aligned}$$

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

上式可推广到多个情形

例1. n张彩票中有一张中奖彩票。求第k个人中奖的概率?
令A_i表示第i个人中奖。

$$P(A_k \cdot \bar{A}_1 \cdots \bar{A}_{k-1}) = ?$$

$$\begin{aligned} P(A_k \cdot \bar{A}_1 \cdots \bar{A}_{k-1}) &= P(\bar{A}_1)P(\bar{A}_2|\bar{A}_1) \\ &\quad \cdot P(\bar{A}_3|\bar{A}_1\bar{A}_2) \cdots P(A_k|\bar{A}_1 \cdots \bar{A}_{k-1}) \\ &= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdots \frac{1}{n-(k-1)} \\ &= \frac{1}{n} \end{aligned}$$

例2. 现有100枚均匀硬币，其中一枚两面都是正面，其余硬币一面是正面，一面是反面。随机从中挑选一枚硬币并连续掷两次，求两次均出现正面的概率？

令A表示两次均出现正面，B表示所选硬币的两面都是正面。

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|\bar{B})P(\bar{B}) \\ &= 1 \cdot \frac{1}{100} + \frac{1}{4} \cdot \frac{99}{100} \\ &= 0.2575 \end{aligned}$$

- 全概率公式

假设$(\Omega, \mathcal{A}, P)$是一个概率空间，$B_k, 1 \le k \le N$是N个互不相交事件，且$\Omega = \sum_{k=1}^{N} B_k$. 那么事件$A$的概率可如下计算

$$P(A) = \sum_{k=1}^{N} P(A|B_k)P(B_k)$$

- N可以是$+\infty$
- 全概率公式可如下理解，事件A可以在N个不同条件下发生，因此其概率大小为各种条件下的加权平均

例3. 播种用的小麦种子中，一等品占95.5%, 二等品占2%, 三等品占1.5%，四等品占1%。经验表明，一、二、三、四等种子长出的穗含50颗以上麦粒的概率分别为0.5、0.15、0.1、0.05。现任选一颗种子，求它所结的穗含50颗以上麦粒的概率？

令A表示所结的穗含50颗以上麦粒

B_i表示所选的种子为第i等品

$$\begin{aligned} P(A) &= \sum_{i=1}^{4} P(A|B_i)P(B_i) \\ &= 0.4825 \end{aligned}$$

- 贝叶斯公式

假设$(\Omega, \mathcal{A}, P)$是一个概率空间，$B_k, 1 \le k \le N$是N个互不相交事件，且$\Omega = \sum_{k=1}^{N} B_k$. 那么

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{k=1}^{N} P(A|B_k)P(B_k)}$$

- 称$P(B_k)$为先验概率；$P(B_k|A)$为后验概率

托马斯·贝叶斯(Thomas Bayes, 1702- 1761.4.7)

1702年生于英国伦敦
英国神学家、数学家、数理统计学家和哲学家
1742年当选英国皇家学会会员
1763 年创立贝叶斯统计理论：归纳推理法用于概率论基础理论
1763年由Richard Price整理发表了贝叶斯的成果《An Essay towards solving a Problem in the Doctrine of Chances》，提出贝叶斯公式。

贝叶斯理论及其应用

数学领域

贝叶斯分类算法(应用：统计分析，测绘学)

贝叶斯风险(应用：统计决策理论)

贝叶斯公式(应用：概率论)

贝叶斯估计(应用：参数估计)

贝叶斯区间估计(应用：人工智能)

贝叶斯序贯决策函数(应用：统计决策理论)

经验贝叶斯方法(应用：统计决策理论)

工程领域

贝叶斯定理(应用：人工智能，心理学，遗传学)

贝叶斯分类器(应用：模式识别，人工智能)

贝叶斯分析(应用：计算机科学)

贝叶斯决策(应用：人工智能)

贝叶斯逻辑(应用：人工智能)

贝叶斯推理(应用：数量地理学，人工智能)

贝叶斯网络(应用：人工智能)

贝叶斯学习(应用：模式识别)

例4. 某公司下设甲、乙、丙、丁四家分厂，他们生产同一种商品，产量分别占18%, 28%, 20%, 34%. 经验表明甲、乙、丙、丁四家分厂次品率分别为0.5%, 1%, 0.8%, 0.5%. 现某顾客购买了该公司产品，发现为次品，向公司索赔10000元。该公司希望追究厂家责任，问各厂家应赔付多少？

次品率为

$$\frac{0.5\times 18+1\times 28+0.8\times 20+0.5\times 34}{10000}=0.70\%$$

甲厂应赔付

$$\frac{0.5\%\times 18\%}{0.70\%}\times 10000=\frac{9000}{7}$$

其它类似

例5. 用血清甲胎蛋白法诊断肝癌。B表示被检验者确实患有肝癌，A表示判断被检验者患有肝癌。已知

$$P(A|B) = 0.95, \quad P(\bar{A}|\bar{B}) = 0.90, \quad P(B) = 0.0004$$

现若有一人被此法诊断为患有肝癌，求此人确实患有肝癌的概率？

- 甲胎蛋白是一种糖蛋白，英文缩写AFP。正常情况下，这种蛋白主要来自胚胎的肝细胞，胎儿出生约两周后甲胎蛋白从血液中消失，因此正常人血清中甲胎蛋白的含量尚不到20微克／升。但当肝细胞发生癌变时，却又恢复了产生这种蛋白质的功能，而且随着病情恶化它在血清中的含量会急剧增加，甲胎蛋白就成了诊断原发性肝癌的一个特异性临床指标。
- http://baike.baidu.com/view/139910.htm

$$
\begin{aligned}
P(B|A) &= \frac{P(A|B)P(B)}{P(A)} \\
&= \frac{0.95 \times 0.0004}{0.95 \times 0.0004 + 0.10 \times 0.9996} \\
&= 0.0038
\end{aligned}
$$

注：该例说明大批健康人群中会有一定数量的人被误诊。原因可能是该方法不够精确。

1.4 独立性

- 两个事件独立
- 多个事件独立
- n-重贝努力试验
- 乘积概率空间

- 独立性是概率论中最重要的概念之一

假设$(\Omega, \mathcal{A}, P)$是一个概率空间，A, B是两个事件。如果$P(B) > 0$，并且

$$P(A|B) = P(A)$$

称A和B 独立

- 上式表明：事件B是否发生对事件A发生的概率大小不产生影响
- 按条件概率定义，上式可写成

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

注：

(1) 即使$P(B) = 0$，乘积公式仍有意义

(2) 容易看出，A, B关系对等，即如果A和B 独立，那么B和A 独立。亦即相互独立

(3) 如果A和B 独立，那么$A, \bar{B}$独立；$\bar{A}, \bar{B}$独立；$\bar{A}, B$独立

(4) 注意与加法的区别：

$$P(A+B) = P(A) + P(B), \qquad A \cap B = \emptyset$$

$$P(AB) = P(A)P(B), \qquad A, B\text{独立}$$

例1. 放回摸球和无放回摸球

一罐子里装有a个白球，b个红球。现依此抽取两个球，用B表示第一次抽取红球，用A表示第二次抽取红球，问A, B事件的独立性如何？

分两种方法：(1) 放回；(2) 无放回

- (1) 放回: 即抽取第一个球，记录下颜色后，将其放回罐子里；并接着抽取第二个球。

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

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

- A, B事件独立。

- (2) 无放回：即抽取第一个球，记录下颜色后，不放回罐子里；并接着抽取第二个球。

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

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|\bar{B})P(\bar{B}) \\ &= \frac{b-1}{a+b-1} \cdot \frac{b}{a+b} + \frac{b}{a+b-1} \cdot \frac{a}{a+b} \\ &= \frac{b}{a+b} \end{aligned}$$

- A, B事件不独立。

- 三个事件独立

假设A, B, C是三个事件，如果

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

$$P(AC) = P(A)P(C)$$

$$P(BC) = P(B)P(C)$$

称A, B, C两两独立

例2. 一个正四面体的三面分别涂成红、黑、白三种颜色，而另一面则涂成三色。现随机一扔，底面涂有红色、黑色、白色的可能性分别是多少？这三个事件两两独立吗？

用A, B, C分别表示底面涂有红色、黑色、白色。那么

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

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

$$P(AC) = P(A)P(C)$$

$$P(BC) = P(B)P(C)$$

所以，A, B, C两两独立

- 但是

$$P(ABC) \neq P(A)P(B)P(C)$$

- 三个事件相互独立

假设A,B,C是三个事件，如果

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

$$P(AC)=P(A)P(C)$$

$$P(BC)=P(B)P(C)$$

$$P(ABC)=P(A)P(B)P(C)$$

称A,B,C相互独立

- 注:

(1) 两两独立不一定相互独立；相互独立一定两两独立

(2) 如果A,B,C相互独立，那么$\bar{A},B,C$相互独立；$A\cup B$ 和C 独立；其它类似成立。

- m个事件相互独立

假设$A_k, 1 \le k \le m$是m个事件，如果$A_k, 1 \le k \le m$中任意$r < m$个都相互独立，并且

$$P\left(\bigcap_{1\le k\le m} A_k\right) = \prod_{1\le k\le m} P(A_k)$$

称$A_k, 1 \le k \le m$相互独立

- 注：$A_k, 1 \le k \le m$中任意$l < m$个事件与另外$m - l$事件相互独立. 其它类似。

例3. 可靠性分析

- n-重Bernoulli试验，也称为二项试验(Binomial Trial)

试验E，包括若干个基本结果。
事件A，具有某种属性的基本结果集合，A发生的概率为

$$P(A) = p_A$$

独立重复进行n次，并观察记录其结果，
判断事件A发生与否。统计事件A发生的次数，记为n_A
n_A是随机数，可为0，1，2，....，n

- Bernoulli试验

固定试验E，事件A。

每次试验，A发生，记为1；A不发生，记为0

这样，每次试验有两个结果，$\{0,1\}$

独立重复n次试验，所得结果为

$$\omega=(\omega_1,\omega_2,\cdots,\omega_n),\qquad \omega_i=0,1$$

用Ω_n表示所有ω的全体

$$\Omega_n=\{\omega=(\omega_1,\omega_2,\cdots,\omega_n),\qquad \omega_i=0,1\}$$

这样，Ω_n中含有2^n个不同的ω

显然，每个ω出现的概率不同，依赖于A发生的次数

$$P_n(\{\omega\})=p_A^{\sum\omega_i}(1-p_A)^{n-\sum\omega_i}$$

这样，我们得到一个新的概率空间

$$(\Omega_n, P_n)$$

这是n-重Bernoulli试验的概率模型。
给定概率空间

$$(\Omega_n, P_n)$$

考虑事件B，如

$$B = \{\omega : n_A(\omega) = k\}$$

那么

$$P_n(B) = \binom{n}{k} p_A^k (1 - p_A)^{n-k}$$

- 乘积概率空间

考虑两个试验E_1, E_2。相应的概率空间分别为

$$(\Omega_1, \mathcal{F}_1, P_1), \qquad (\Omega_2, \mathcal{F}_2, P_2)$$

现独立地做试验E_1和E_2，记录其结果

$$\omega = (\omega_1, \omega_2)$$

考虑E_1, E_2所有基本结果的全体

$$\Omega = \{\omega = (\omega_1, \omega_2), \quad \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

并考虑事件

$$A = A_1 \times A_2 = \{\omega = (\omega_1, \omega_2), \quad \omega_1 \in A_1, \omega_2 \in A_2\}$$

- 乘积概率空间

定义其概率

$$P(A) = P(A_1 \times A_2) = P(A_1)P(A_2)$$

得到新的概率空间：乘积概率空间(Ω, P)

例4：$\Omega = [0,1] \times [0,1]$, $\mathcal{A} = \sigma\{A \times B\}$,

$$P(A \times B) = |A||B|, \quad \text{均匀分布}$$

补充说明

- 概率的连续性

假设$(\Omega, \mathcal{A}, P)$ 是概率空间：

(1) Ω 样本空间

(2) $\mathcal{A}$ σ-域

(3) $P : \mathcal{A} \mapsto [0, 1]$

可加性：如果$A_n, n \geq 1$是一列互不相交的事件，那么

$$P(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$$

简单地说，

$$P(A + B) = P(A) + P(B)$$

除可加性之外，概率还具有连续性。

- 事件的极限

(i) 假设A_n是一列增加事件，

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq$$

那么定义

$$\lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n$$

(ii) 假设A_n是一列递减事件，

$$A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq$$

那么定义

$$\lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n$$

- 概率的极限

假设A_n是一列增加事件，那么

$$P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$$

证明：令$B_1 = A_1$, $B_k = A_k \setminus A_{k-1}$，那么

$$B_1, B_2, \cdots, B_n \qquad ,\text{互不相交}$$

并且

$$\lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} B_n$$

因此，

$$P(\lim_{n\to\infty} A_n) = P(\sum_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} P(B_n)$$

注意到，

$$P(B_n) = P(A_n) - P(A_{n-1})$$

这样，

$$\sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} P(A_n)$$

总结：概率P 是定义在σ域$\mathcal{A}$上的规范化集函数，具有可加性和连续性。

- 条件概率具有概率的运算性质

假设$(\Omega, \mathcal{A}, P)$ 是概率空间，B是一个事件，$P(B) > 0$。
对任意事件$A \in \mathcal{A}$，给定B发生的条件下，事件A发生的概率

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

这样，

$$P(\cdot|B) : \mathcal{A} \mapsto [0, 1] \quad \text{是一个概率}$$

比如，

$$P(A_1 + A_2|B) = P(A_1|B) + P(A_2|B)$$

$$P(A_1 \setminus A_2|B) = P(A_1|B) - P(A_2|B), \quad A_2 \subseteq A_1$$

例. Simpson 悖论

某医生进行临床试验，分析两种药品的治疗效果(有效性)，结果如下

	女		男	
	药 I	药 II	药 I	药 II
成功	200	10	19	1000
失败	1800	190	1	1000

问题：哪一种药物更为有效？

分析1：(药II 有效)

药I 试验人数2020人，治愈219人，成功率为$\frac{219}{2020}$；

药II 试验人数2020人，治愈1010人，成功率为$\frac{1010}{2020}$；

分析2：(药I 有效)

药I 女性试验人数2000人，治愈200 人，成功率为$\frac{200}{2000}$；

药II 女性试验人数20 人，治愈19人，成功率为$\frac{19}{20}$；

药I 男性试验人数20 人，治愈19人，成功率为$\frac{19}{20}$；

药II 男性试验人数2000 人，治愈1000人，成功率为$\frac{1000}{2000}$

Simpson悖论出现在许多问题中。

令A表示药物有效(治疗成功)；B表示药I随机地安排给一个病人；C表示病人为女性。
$\bar{B}$ 表示药II 随机地安排给一个病人；$\bar{C}$表示病人为男性。
Simpson 悖论可以描述为：

$$P(A|BC) > P(A|\bar{B}C), \quad P(A|B\bar{C}) > P(A|\bar{B}\bar{C})$$

$$P(A|B) < P(A|\bar{B})$$

定义

$$a = P(ABC), \quad b = P(\bar{A}BC)$$

$$c = P(A\bar{B}C), \quad b = P(\bar{A}\bar{B}C)$$

$$a = P(AB\bar{C}), \quad b = P(\bar{A}B\bar{C})$$

$$a = P(A\bar{B}\bar{C}), \quad b = P(\bar{A}\bar{B}\bar{C})$$

Simpson悖论可写为

$$ad > bc, \quad eh > fg, \quad (a+e)(d+h) < (b+f)(c+g)$$

约束条件为：

$$a, b, c, d, e, f, g, h > 0, \quad a+b+c+d+e+f+g+h = 1$$

上述不等式存在多个解。比如

$$a = \frac{3}{30}, b = \frac{1}{30}, c = \frac{8}{30}, d = \frac{3}{30}, e = \frac{3}{30}, f = \frac{8}{30}, g = \frac{1}{30}, h = \frac{3}{30}$$