

概率论与数理统计



第 4 讲

独立性、全概率公式、贝叶斯公式

一般说来 $P(A|B) \neq P(A)$ 这反映了 A 和 B 之间存在着一定的概率关系。

若 $P(A|B) > P(A)$ ，说明事件 B 的发生使事件 A 发生的可能性增大了，也就是说事件 B 促进了 A 的发生。

那么有没有 $P(A) = P(A|B)$ 的情形。

例1. 盒子中有3个黑球，2个白球，每次取一个，有放回地取两次，记 $B=\{\text{第一次取到白球}\}$ ， $A=\{\text{第二次取到白球}\}$ ，求 $P(A)$ 和 $P(A|B)$ 。

解： $P(A) = \frac{A_5^1 \times A_2^1}{5^2} = \frac{5 \times 2}{5^2} = \frac{2}{5}$ $B: 5 \begin{cases} 3 \text{ 黑球} \\ 2 \text{ 白球} \end{cases}$ $A|B: 5 \begin{cases} 3 \text{ 黑球} \\ 2 \text{ 白球} \end{cases}$

所以 $P(A|B) = 2/5$ 即： $P(A) = P(A|B)$

这就是说，事件 B 的发生，并不影响事件 A 发生的概率，此时在概率上就称事件 A 、 B 独立 (Independence)。

事件 A 、 B 独立，应该包含两层含义：

(1) 事件 B 的发生，不影响 A 发生的概率。→

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

(2) 事件 A 的发生，不影响 B 发生的概率。→

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

独立定义

用两个公式作为独立性的定义显得比较繁琐，同时由于两个公式中条件概率的存在，必然要求 $P(A) > 0$ 且 $P(B) > 0$ ，对于概率为0的事件无法定义独立性。

不难证明，当 $P(B) > 0$ 时，有 $P(A|B) = P(A) \Leftrightarrow P(AB) = P(A)P(B)$

当 $P(A) > 0$ 时，有 $P(B|A) = P(B) \Leftrightarrow P(AB) = P(A)P(B)$

所以有
$$\left. \begin{aligned} P(A|B) &= P(A) \\ P(B|A) &= P(B) \end{aligned} \right\} \Leftrightarrow P(AB) = P(A)P(B)$$

同时, $P(AB)=P(A)P(B)$ 不受 $P(A)>0$, $P(B)>0$ 的限制, 因此可以作为两个事件独立性的定义。

定 义

对任意的事件 A 和 B , 若 $P(AB)=P(A)P(B)$, 则称事件 A 和 B 是相互独立的。

特别地: S 与任意事件 A 相互独立; ϕ 与任意事件 A 相互独立。

性质1. 若 A 与 B 独立, 则 \bar{A} 与 B , A 与 \bar{B} , \bar{A} 与 \bar{B} 相互独立

证明: 仅证 A 与 \bar{B} 独立

$$\begin{aligned} P(A\bar{B}) &= P(A - AB) \\ &= P(A) - P(AB) \\ &= P(A) - P(A)P(B) \\ &= P(A)[1 - P(B)] = P(A)P(\bar{B}) \end{aligned}$$

A, B 独立

故 A 与 \bar{B} 独立。

实际应用中，经常根据实际意义加以判断两个事件的独立性，如果从事件发生概率的角度去分析判断他们**没有关联**或者**关联很弱**，就可以认为事件是独立的。然后就可以利用**独立性的**性质进行相关的计算。

例如

甲、乙两人向同一目标射击，而且彼此互不影响，则甲、乙两人是否击中目标相互独立。

在一个城市中相距很远的两个地段是否发生交通事故，也是互相独立的。

例2. 据某日天气预报，甲城市明天下雨的概率为0.8，乙城市明天下雨的概率为0.7，且甲乙两个城市相距非常遥远。求 (1) 甲乙两个城市明天都下雨的概率； (2) 甲乙两个城市至少有一个下雨的概率。

解：设 A 表示甲城市明天下雨， B 表示乙城市明天下雨。

由于两个城市相距遥远，可以认为两个城市明天是否下雨关系不大，即认为 A 和 B 是相互独立的。

$$(1) P(AB) = P(A)P(B) = 0.8 \times 0.7 = 0.56$$

$$(2) P(A \cup B) = P(A) + P(B) - P(AB) = 0.8 + 0.7 - 0.56 = 0.94$$

例3. 设事件A和B独立，且A和B都不发生的概率为1/9，A发生B不发生的概率与B发生A不发生的概率相等，求P(A)。

解: 易知 $P(\bar{A}\bar{B}) = 1/9$ $P(A\bar{B}) = P(\bar{A}B)$

因为事件A、B独立 所以有 \bar{A} 和 \bar{B} \bar{A} 和B A 和 \bar{B} 独立

所以有 $P(\bar{A}\bar{B}) = P(\bar{A})P(\bar{B})$ $P(\bar{A}B) = P(\bar{A})P(B)$ $P(A\bar{B}) = P(A)P(\bar{B})$

$$\therefore [1 - P(A)][1 - P(B)] = \frac{1}{9} \quad P(A)[1 - P(B)] = [1 - P(A)]P(B)$$

舍掉

$$\therefore P(A) = P(B) \quad \therefore 1 - 2P(A) + [P(A)]^2 = \frac{1}{9} \quad \text{解得 } P(A) = \frac{2}{3} \text{ 或 } P(A) = \frac{4}{3}$$

多个事件的独立性

对任意三个事件 A, B, C , 如何定义独立性呢?

首先需要满足任意一个事件的发生不影响其他任何一个事件发生的概率,
即任意两个事件都是独立的, 即有以下等式成立

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

其次, 任意两个事件同时发生, 不影响第三个事件发生的概率, 即有

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

$$P(ABC) = P(A)P(B)P(C) \longleftarrow \uparrow$$

多个事件的独立性

定义：对任意三个事件 A, B, C ，若

$$\begin{cases} P(AB) = P(A)P(B) \\ P(AC) = P(A)P(C) \\ P(BC) = P(B)P(C) \\ P(ABC) = P(A)P(B)P(C) \end{cases}$$

则称事件 A, B, C 相互独立，简称 A, B, C 独立。

多个事件的独立性

对任意 n 个事件 A_1, A_2, \dots, A_n , 若

$$(1) P(A_i A_j) = P(A_i)P(A_j) \quad 1 \leq i < j \leq n$$

$$(2) P(A_i A_j A_k) = P(A_i)P(A_j)P(A_k) \quad 1 \leq i < j < k \leq n$$

.....

$$(n-1) P(A_1 A_2 \cdots A_n) = \prod_{i=1}^n P(A_i)$$

则称事件 A_1, A_2, \dots, A_n 相互独立, 简称 A_1, A_2, \dots, A_n 独立。

也可以表示为: 若对于任意的 k ($2 \leq k \leq n$) 和任意一组数 $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$, 有

$$P(A_{i_1} A_{i_2} \cdots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

多个事件的独立性

请注意多个事件两两独立与相互独立的区别与联系

对 $n (n > 2)$ 个事件 相互独立 \longleftrightarrow ? 两两独立

例2. 一个均匀的正四面体，其第一面染成红色，第二面染成白色，第三面染成黑色，而第四面同时染上红、白、黑三种颜色. 现以 A 、 B 、 C 分别记投一次四面体出现红、白、黑颜色朝下的事件，问 A 、 B 、 C 是否相互独立？

解：由于在正四面体中红、白、黑分别出现在两个面，因此

$$P(A)=P(B)=P(C)=1/2$$

又由题意知 $P(AB)=P(BC)=P(AC)=1/4$; $P(ABC)=1/4$

$$P(A)=P(B)=P(C)=1/2; P(AB)=P(BC)=P(AC)=1/4; P(ABC)=1/4。$$

$$P(AB)=1/4=1/2 \times 1/2 = P(A)P(B)$$

从而有 $P(BC)=1/4=1/2 \times 1/2 = P(B)P(C)$ 即 A, B, C 两两独立。

$$P(AC)=1/4=1/2 \times 1/2 = P(A)P(C)$$

但是 $\frac{1}{4} = P(ABC) \neq P(A)P(B)P(C) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ 即 A, B, C 不互相独立。

性质1 若事件 $A_1, A_2, \dots, A_n (n \geq 2)$ 相互独立, 则其中任意 $k (2 \leq k \leq n)$ 个事件也相互独立, 即 $\forall \{j_1, \dots, j_k\} \subset \{1, \dots, n\}, A_{j_1} \cdots A_{j_k}$ 相互独立。

性质2 若事件 $A_1, A_2, \dots, A_n (n \geq 2)$ 相互独立, 则将 A_1, A_2, \dots, A_n 中任意多个事件换成它们的对立事件, 所得 n 个事件仍相互独立, 也即 B_1, B_2, \dots, B_n 相互独立, 其中 $B_i = A_i$ 或 \bar{A}_i 。如: $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$

性质3 若事件 A_1, A_2, \dots, A_n ($n \geq 2$)相互独立, 则将这 n 个事件分成 k 组, 同一个事件不能同时属于两个不同的组, 则对每组事件进行和、差、积、逆等运算所得到的新的 k 个事件也是互相独立的

如: $A_1 \cup A_2, A_3, \dots, A_n$ 相互独立; $A_1 \cap A_2, A_3 \cup A_4, A_5, \dots, A_n$ 相互独立。

性质4 若 A_1, A_2, \dots, A_n 相互独立, 则 $P(\bigcup_{k=1}^n A_k) = 1 - \prod_{k=1}^n [1 - P(A_k)]$

证明: $P(\bigcup_{k=1}^n A_k) = 1 - P(\overline{\bigcup_{k=1}^n A_k}) = 1 - P(\bigcap_{k=1}^n \bar{A}_k) = 1 - \prod_{k=1}^n P(\bar{A}_k) = 1 - \prod_{k=1}^n [1 - P(A_k)]$

例3.甲、乙、丙三人同时独立向同一目标射击，他们射中目标的概率分别为0.4、0.5、0.7 求至少有一人射中目标的概率；

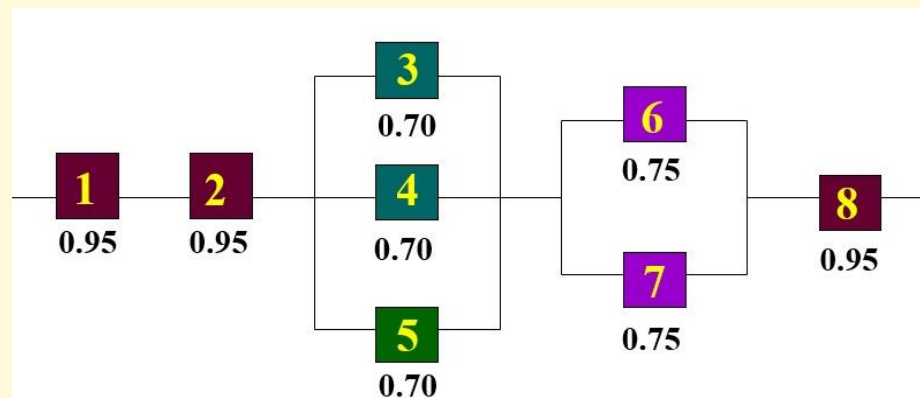
解：设 A 、 B 、 C 分别表示甲、乙、丙射中目标，则 A 、 B 、 C 互相独立。

所求为 $P(A \cup B \cup C)$

$$\therefore P(A \cup B \cup C) = 1 - P(\bar{A})P(\bar{B})P(\bar{C}) = 1 - 0.6 \times 0.5 \times 0.3 = 0.91$$

例4.下面是一个串并联电路示意图。

1、2、3、4、5、6、7、8是8个独立工作的元件。它们正常工作的概率分别为0.95, 0.95, 0.70, 0.70, 0.70, 0.75, 0.75, 0.95。求电路正常工作的概率



解：设 A_i 分别表示上述8个元件中第 i 个元件正常工作， B 表示电路正常工作，则

$$B = A_1 A_2 (A_3 \cup A_4 \cup A_5) (A_6 \cup A_7) A_8$$

由于各元件独立工作，所以

$$P(B)=P\{A_1 A_2(A_3\cup A_4\cup A_5)(A_6\cup A_7) A_8\} =P(A_1)P(A_2)P(A_3\cup A_4\cup A_5)P(A_6\cup A_7)P(A_8)$$

其中

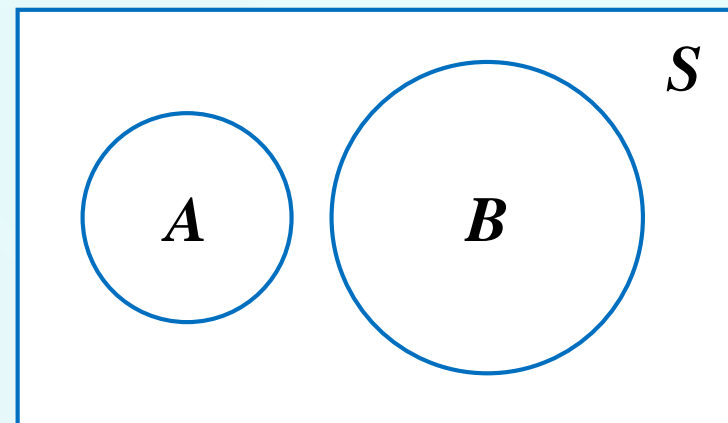
$$P(A_3\cup A_4\cup A_5) = 1 - P(\bar{A}_3)P(\bar{A}_4)P(\bar{A}_5) = 0.973$$

$$P(A_6\cup A_7) = 1 - P(\bar{A}_6)P(\bar{A}_7) = 0.9375$$

代入得 $P(B) \approx 0.782$



思考：如图的两个事件是独立的吗？



若 A 、 B 互斥，且 $P(A)>0$ ， $P(B)>0$ ，则 A 与 B 不独立。

由互斥性得： $P(AB)=0$ ；

由于 $P(A)>0$ ， $P(B)>0$ ，所以 $P(A)P(B)\neq 0$ ，即： A 、 B 不独立。

反之，若 A 与 B 独立，且 $P(A)>0$ ， $P(B)>0$ ，则： A 、 B 不互斥。

易知： $P(AB)=P(A)P(B)>0$ ，故 $AB\neq\phi$ ，即 A 、 B 不互斥。

若 A 与 B 互斥, 且 $P(A)$ 和 $P(B)$ 至少有一个为0时

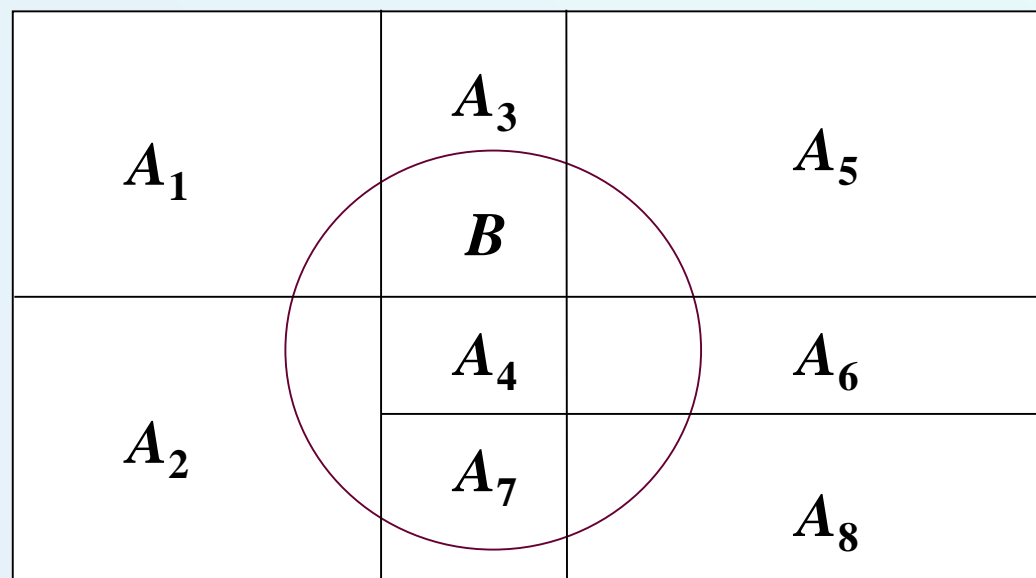
由于 $P(AB)=P(\phi)=0$, 且 $P(A)P(B)=0$, 所以 A 与 B 独立。

但若 A 、 B 独立, 且 $P(A)$ 和 $P(B)$ 至少一个为0时

虽然 $P(AB)=P(A)P(B)=0$, 但 AB 不一定为 ϕ , 所以 A 、 B 不一定互斥。

我们在数学上可以举出两两独立但不互相独立的例子, 但是现实生活中, 是很难想象这样的例子的。

在概率论中常常会遇到一些较复杂的事件。提出如下问题：复杂事件B的概率如何求？



定义： 设 S 为试验 E 的样本空间， A_1, \dots, A_n 为 E 的一组事件。若

$$(1) A_1, \dots, A_n \text{ 互不相容}; \quad (2) \bigcup_{i=1}^n A_i = S$$

则称 A_1, \dots, A_n 为样本空间 S 的一个划分，也称 A_1, \dots, A_n 是**完备事件组**。

可见若 A_1, A_2, \dots, A_n 为样本空间 S 的一个划分，则在每次试验中必有一个且只有一个 A_i 发生。

易知对任意随机事件 A ， A 和 \bar{A} 就组成了一个样本空间 S 的划分。

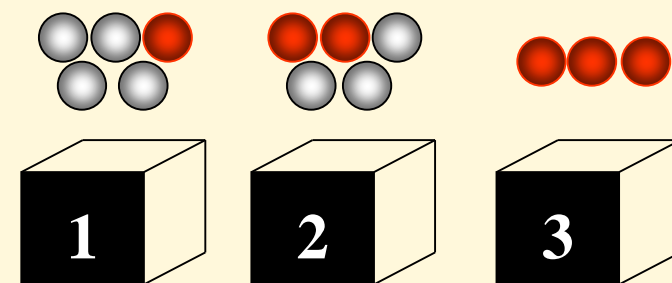
定理： 设 S 为试验 E 的样本空间， B 为 E 的事件， A_1, \dots, A_n 为 S 的一个划分，且 $P(A_i) > 0, i=1, \dots, n$ ，则 $P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$
上式称为全概率公式。

证明： $\because B = B \cap S = B \cap (A_1 \cup A_2 \cup \dots \cup A_n) = BA_1 \cup BA_2 \cup \dots \cup BA_n$

又由于 BA_1, BA_2, \dots, BA_n 互不相容

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

例5. 有三个编号为1, 2, 3的箱子, 1号箱装有1个红球4个白球, 2号箱装有2个红3个白球, 3号箱装有3个红球。某人从三个箱子中任取一箱, 从中任意摸出一球, 求取得红球的概率。



解: 记 $B = \{\text{任意摸出一球为红球}\}$ $A_i = \{\text{任意摸出的球取自} i \text{ 号箱}\}$, $i = 1, 2, 3$

若将1号箱子中的球编号为1-5, 二号箱子中的球编号为6-10, 3号箱子中的球编号为11,12,13。

摸到第*i*号球的可能性不全相同

则样本空间为: $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$

则 $A_1=\{1,2,3,4,5\}, A_2=\{6,7,8,9,10\}, A_3=\{11,12,13\}$ 即 A_1, A_2, A_3 构成了 S 的一个划分

$$\therefore \text{由全概率公式可得 } P(B) = \sum_{i=1}^3 P(A_i)P(B|A_i) = \frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{2}{5} + \frac{1}{3} \times 1 = \frac{8}{15}$$

在例5中，当摸出的球取自不同的箱子时，取得红球的概率的是不同的，因此，可以认为诸 $A_i, i=1, 2, 3$ 是引起事件 B 的概率发生变化的原因，也就是说，所有引起事件 B 的概率发生变化的原因就是**样本空间的划分**，这为寻找样本空间的划分提供了一个简单的途径。

例6.发报机发出 “.” 的概率为0.60, 发出 “—” 的概率为0.40; 收报机将 “.” 收为 “.” 的概率为0.99, 将 “—” 收为 “.” 的概率为0.02。求收报机将任一信号收为 “.” 的概率。

解：记 $B = \{\text{收报机将任一信号收为 “.”}\}$

$A_1 = \{\text{发报机发出 “.”}\}, A_2 = \{\text{发报机发出 “—”}\}$

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) = 0.6 \times 0.99 + 0.4 \times 0.02 = 0.602$$

例7.据美国一份资料报导,在美国总的来说患肺癌的概率是0.1%, 在人群中
有20%是吸烟者, 他们患肺癌的概率约为0.4%。
求不吸烟者患肺癌的概率。

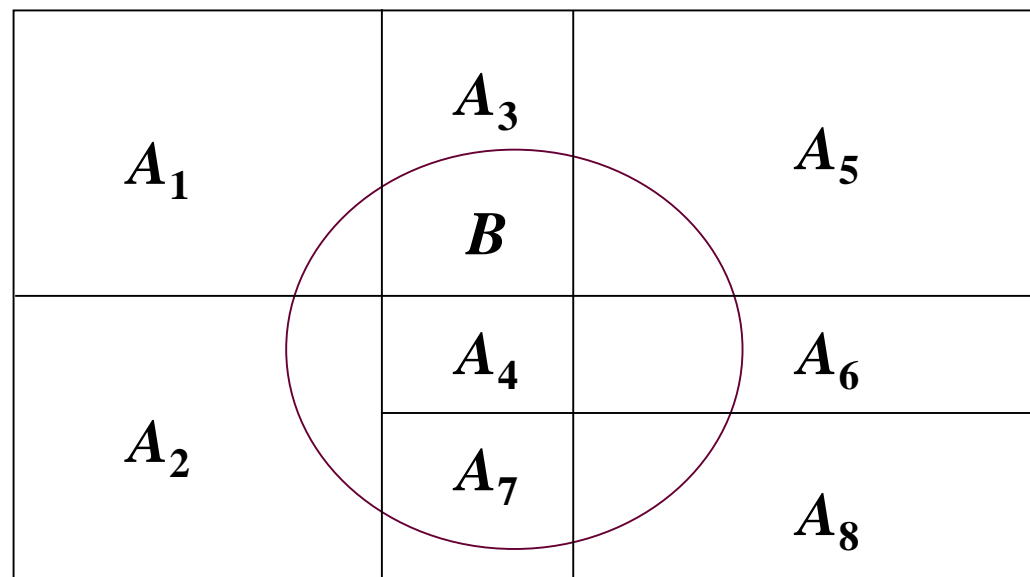
解: 设 B 表示任意找一个人患肺癌, A 表示任意找一个人且此人吸烟

则 $P(B)=0.001, P(A)=0.20, P(B|A)=0.004$

由全概率公式有 $P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$

代入数据有 $0.001 = 0.20 \times 0.004 + 0.80 \times P(B|\bar{A})$ 于是 $P(B|\bar{A}) = 0.00025$

可以形象地把全概率公式看成为“由原因推结果”，每个原因对结果的发生有一定的“作用”，即结果发生的可能性与各种原因的“作用”大小有关。全概率公式表达了它们之间的关系。



诸 A_i 是原因 B 是结果

应用：关于敏感问题的调查

敏感问题：“你吸大麻吗？” “你是同性恋吗？” “你是单亲家庭吗？” 等等

做法，提出两个问题， S ：敏感问题，如你是同性恋吗？ T ：你的电话号码的最后一位数是偶数吗？（无关紧要）

然后要求被提问者投掷一枚硬币，出现正面要求正确回答 S ，出现反面要求正确回答 T 。这时提问者并不知道被提问者回答的是哪一个问题，这个信息是保密的，每个被提问者只是回答“是”或“否”。从这些结果，可以估计出是同性恋人群的比例。

设 B 表示回答是, A_1 表示掷出正面朝上, A_2 表示掷出反面朝上;

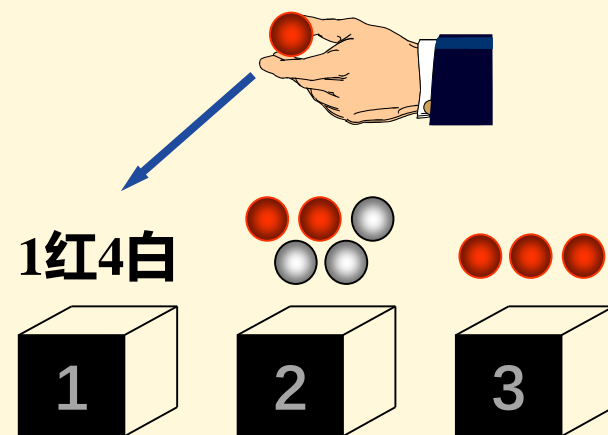
所以由全概率公式得:

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)=(1/2)\times P(B|A_1)+(1/2)\times P(B|A_2)$$

则调查结束后, $P(B)$ 可以用频率代替, $P(B|A_1)$ 记为同性恋的比例 p , $P(B|A_2)$ 是可以统计出来的, 记为 p_1 .

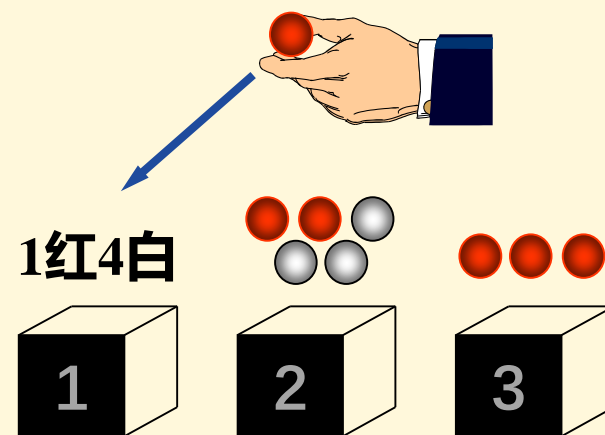
可求得: $p=2P(B)-p_1$.

例8.有三个箱子，分别编号为1, 2, 3, 1号箱装有1个红球4个白球，2号箱装有2个红3个白球，3号箱装有3个红。某人从三箱中任取一箱，从中任意摸出一球，发现是红球，求该球是取自1号箱的概率。



这一类问题在实际中更为常见，它所求的是条件概率，即已知结果发生的条件下，求某原因发生可能性的大小。

例8.有三个箱子，分别编号为1, 2, 3, 1号箱装有1个红球4个白球，2号箱装有2个红3个白球，3号箱装有3个红。某人从三箱中任取一箱，从中任意摸出一球，发现是红球，求该球是取自1号箱的概率。



解：记 $A_i = \{\text{摸到的球取自} i \text{号箱}\}$, $i=1,2,3$; $B = \{\text{取到的球为红球}\}$

$$P(A_1 | B) = \frac{P(A_1 B)}{P(B)} = \frac{P(A_1)P(B | A_1)}{\sum_{k=1}^3 P(A_k)P(B | A_k)} = \frac{\frac{1}{3} \times \frac{1}{5}}{\frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{2}{5} + \frac{1}{3} \times 1} = \frac{1}{8}$$



Thomas Bayes
(约1701-1761)

设 S 为试验 E 的样本空间, B 为 E 的事件, A_1, \dots, A_n 为 S 的一个划分, 且 $P(A_i) > 0, i=1, \dots, n, P(B) > 0$, 则有

$$P(A_j | B) = \frac{P(A_j)P(B | A_j)}{\sum_{i=1}^n P(A_i)P(B | A_i)}, \quad j = 1, 2, \dots, n$$

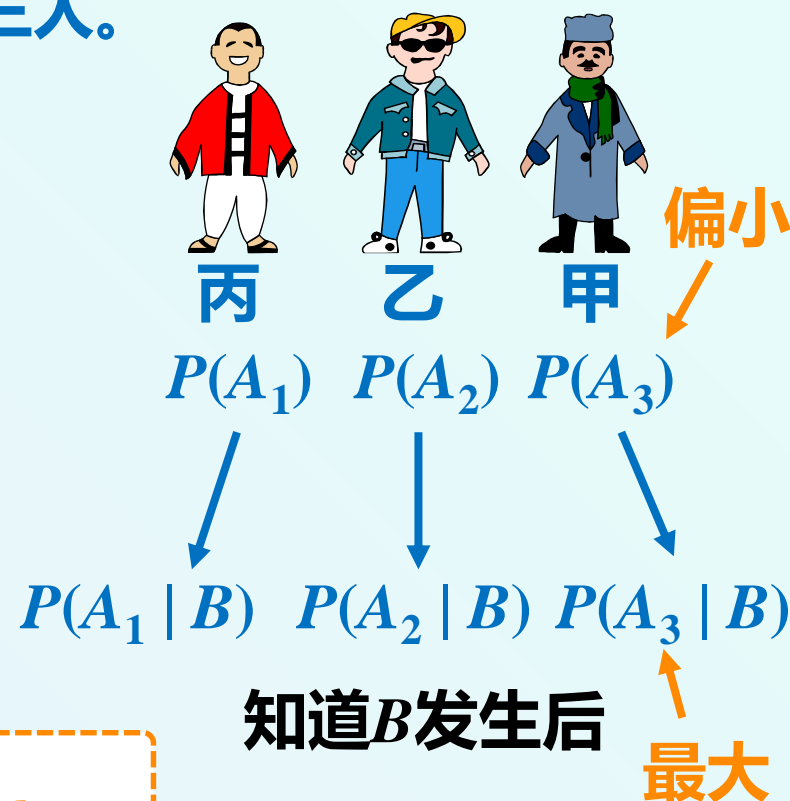
贝叶斯公式在实际中有很多应用, 它可以帮助人们确定引起某结果(事件 B)发生的最可能原因。

例如: 某地发生了一个案件, 怀疑对象有甲、乙、丙三人。

在不了解案情细节(事件 B)之前, 侦破人员根据过去的前科, 对他们作案的可能性有一个估计。

但在知道案情细节后, 这个估计就有了变化。

原来认为作案可能性较小的某甲, 现在可能变成了重点嫌疑犯。



$P(A_j)$ 和 $P(A_j | B)$ 分别称为原因的先验概率和后验概率。

$P(A_j)$ 是在没有进一步信息(不知道事件 B 是否发生)的情况下, 人们对诸事件发生可能性大小的认识。

当有了新的信息(知道 B 发生), 人们对诸事件发生可能性大小 $P(A_j|B)$ 有了新的估计。

贝叶斯公式 从数量上刻划了这种变化。

人们依据贝叶斯公式的思想发展了一整套的统计推断方法, 叫做贝叶斯统计。
可见其在统计上的作用。

例9.对以往数据分析结果表明,当机器调整良好时,产品的合格品率为98%。当机器发生某种故障时,产品的合格品率为55%。根据以往经验,每天早上机器开动时,机器调整良好的概率为95%。这一日早上生产了一件产品,发现是合格品,问该日机器调整良好的概率是多少?

解: 记 $B = \{\text{该日早上生产了一件合格品}\}$

$A_1 = \{\text{该日机器调整良好}\}$, $A_2 = \{\text{该日机器发生某种故障}\}$

由Bayes公式可得

$$P(A_1 | B) = \frac{P(A_1)P(B | A_1)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2)} = \frac{0.95 \times 0.98}{0.95 \times 0.98 + 0.05 \times 0.55} = 0.97$$

$P(A_1)$ 先验概率

$P(A_1|B)$ 后验概率

例10. 据调查某地区居民患某种癌症的概率为0.0004。现用一方法检查该种疾病，若呈阴性，表明不患病；若呈阳性，表明患病。由于技术和操作不完善等原因，有病者未必检出阳性，无病者也有可能呈阳性反应。根据经验，已知有病者检出阳性的概率为0.99，无病者错检为阳性的概率为0.05。现设某人已检出阳性，问他患病的概率是多少？

解：设事件 B 表示一居民检验出阳性，事件 A 表示一居民患该种疾病。

则： $P(A)=0.0004$ $P(B|A)=0.99$ $P(B|\bar{A})=0.05$ $P(\bar{A})=0.9996$

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\bar{A})P(B|\bar{A})} = \frac{0.99 \times 0.0004}{0.99 \times 0.0004 + 0.05 \times 0.9996} = 0.00786$$

先验概率

后验概率

1. 这种试验对于诊断一个人是否患有癌症有无意义?

分析：如果不做试验，抽查一人，他患癌症的概率(先验): $P(A)=0.0004$

若试验后得阳性反应，则根据试验得来的信息，此人患癌症的概率(后验)为:
 $P(A|B)=0.00786$

从0.0004增加到0.00786，将近增加约20倍。

说明这种试验对于诊断一个人是否患有癌症有意义。

2. 检出阳性是否一定患有癌症?

试验结果为阳性，此人确患癌症的概率为 $P(A|B) = 0.00786$

即使你检出阳性，尚可不必过早下结论有癌症，这种可能性只有0.786%(平均来说，10000人中大约只有78人确患癌症)，此时医生常要通过再试验来确认。

进一步降低错检的概率是提高检验精度的关键，但在实际中由于技术和操作等原因，这是很困难的。实际中，常常采用复查的方法来降低错误率。此时先验概率变为 $P(A) = 0.00786$ ，再利用贝叶斯公式计算得：

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\bar{A})P(B|\bar{A})} = \frac{0.99 \times 0.00786}{0.99 \times 0.00786 + 0.05 \times 0.99214} = 0.1356$$



作业：37, 43, 46, 47, 49, 53, 56

第 4 讲

谢谢观看