

第五节 条件概率

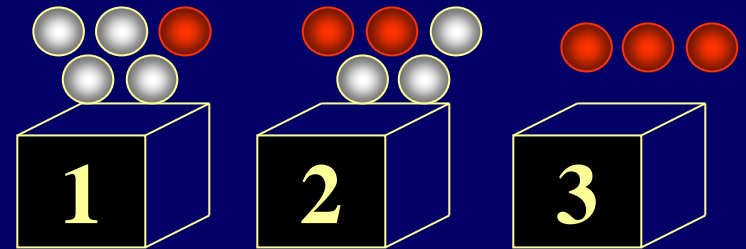
- 全概率公式
- 贝叶斯公式
- 小结 布置作业

三、全概率公式

看一个例子:

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红3白球,3号箱装有3红球.某人从三箱中任取一箱,从中任意摸出一球,求取得红球的概率.

解 记 $A_i = \{\text{球取自} i \text{号箱}\}$,
 $i = 1, 2, 3$;
 $B = \{\text{取得红球}\}$



其中 A_1, A_2, A_3 两两互斥

B 发生总是伴随着 A_1, A_2, A_3 之一同时发生,



即 $B = A_1B + A_2B + A_3B$,

且 A_1B 、 A_2B 、 A_3B 两两互斥

运用加法公式得到

$$P(B) = P(A_1B) + P(A_2B) + P(A_3B)$$

对求和中的每一项运用乘法公式得

$$P(B) = \sum_{i=1}^3 P(A_i)P(B | A_i)$$

代入数据计算得: $P(B) = 8/15$

将此例中所用的方法推广到一般的情形, 就得到在概率计算中常用的全概率公式.



定义 设 S 为随机试验 E 的样本空间, B_1, B_2, \dots, B_n 是 E 的一组事件, 如果满足

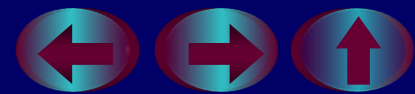
$$(1) B_i B_j = \varnothing \quad (i \neq j)$$

$$(2) B_1 \cup B_2 \cup \dots \cup B_n = S$$

则称 B_1, B_2, \dots, B_n 为**完备事件组**, B_1, B_2, \dots, B_n 为 S 的一个划分.

注意: 若 B_1, B_2, \dots, B_n 为样本空间的一个划分, 则对每次试验, 事件组 B_1, B_2, \dots, B_n 中必有且仅有一个事件发生.

可见, S 的划分是将 S 分割成若干个互斥事件.



定理 1 设试验 E 的样本空间为 S , B_1, B_2, \dots, B_n 为 S 的一个划分, 且 $P(B_i) > 0$ ($i = 1, 2, \dots, n$), 则对样本空间中的任一事件 A , 恒有

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

证明 因为 $A = AS = A(B_1 \cup B_2 \cup \dots \cup B_n)$
 $= AB_1 \cup AB_2 \cup \dots \cup AB_n$

并且 $AB_i \cap AB_j = \varnothing$, ($i \neq j$), 所以

$$\begin{aligned} P(A) &= P(AB_1) + P(AB_2) + \dots + P(AB_n) \\ &= P(B_1)P(A/B_1) + \dots + P(B_n)P(A/B_n) \\ &= \sum_{i=1}^n P(B_i)P(A/B_i) \end{aligned}$$



$$P(A) = \sum_{i=1}^n P(B_i)P(A/B_i) \longrightarrow \text{全概率公式.}$$

全概率公式的基本思想是把一个未知的复杂事件分解为若干个已知的简单事件再求解,而这些简单事件组成一个互不相容事件组,使得某个未知事件 A 与这组互不相容事件中至少一个同时发生,故在应用此全概率公式时,关键是要找到一个合适的 S 的一个划分.

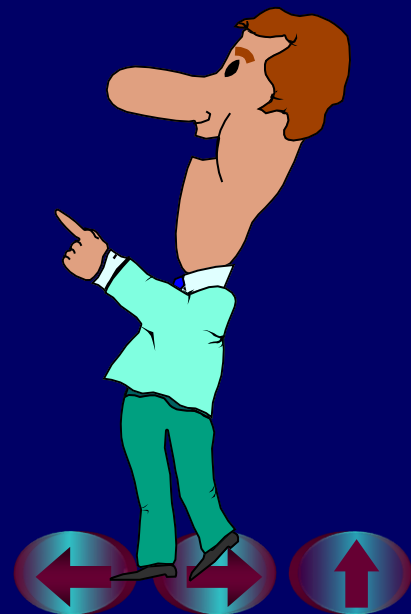


我们还可以从另一个角度去理解 全概率公式.

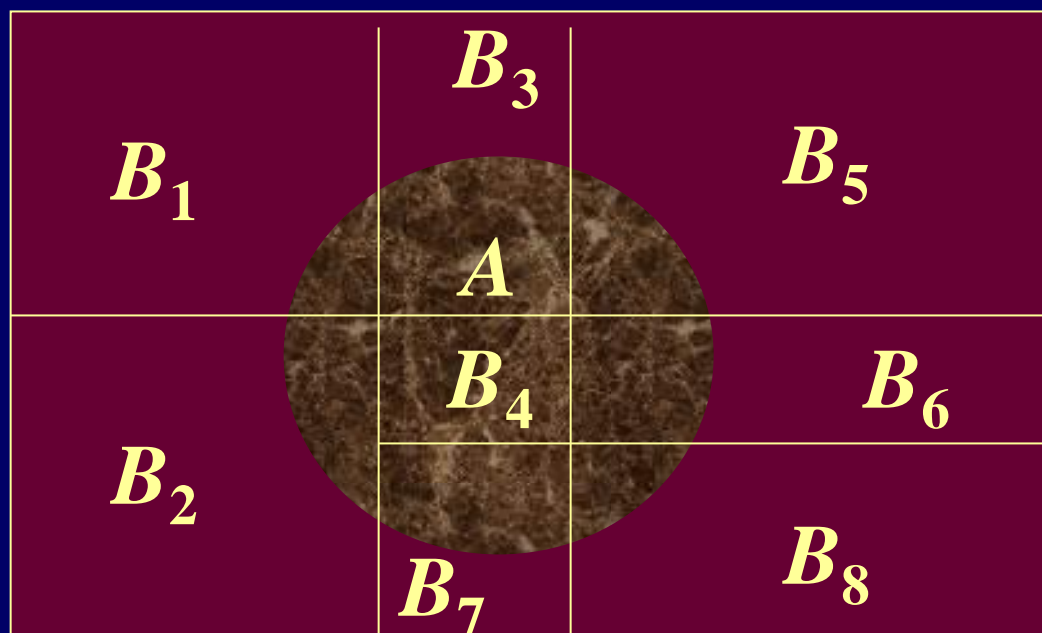
某一事件 A 的发生有各种可能的原因，如果 A 是由原因 $B_i (i=1,2,\dots,n)$ 所引起，则 A 发生的概率是

$$P(AB_i)=P(B_i)P(A |B_i)$$

每一原因都可能导致 A 发生，故 A 发生的概率是各原因引起 A 发生概率的总和，即全概率公式.



由此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 B_i 是原因
 B 是结果

例 甲、乙、丙三人同时对飞机进行射击,三人击中的概率分别为0.4、0.5、0.7. 飞机被一人击中而击落的概率为0.2,被两人击中而击落的概率为0.6,若三人都击中,飞机必定被击落,求飞机被击落的概率.

解

设 $A = \{\text{飞机被击落}\}$
 $B_i = \{\text{飞机被} i \text{人击中}\}, i=1,2,3$

则 $A = B_1A + B_2A + B_3A$

由全概率公式

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)$$



依题意,

$$P(A|B_1) = 0.2,$$

$$P(A|B_2) = 0.6,$$

$$P(A|B_3) = 1$$



为求 $P(B_i)$,

设 $H_i=\{\text{飞机被第}i\text{人击中}\}$, $i=1,2,3$

可求得

$$P(B_1) = P(H_1 \overline{H_2} \overline{H_3} \cup \overline{H_1} H_2 \overline{H_3} \cup \overline{H_1} \overline{H_2} H_3)$$

$$P(B_2) = P(H_1 H_2 \overline{H_3} \cup \overline{H_1} H_2 H_3 \cup H_1 \overline{H_2} H_3)$$

$$P(B_3) = P(H_1 H_2 H_3)$$

将数据代入计算得

$$P(B_1)=0.36; P(B_2)=0.41; P(B_3)=0.14.$$



于是

$$\begin{aligned}P(A) &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) \\&= 0.36 \times 0.2 + 0.41 \times 0.6 + 0.14 \times 1 \\&= 0.458\end{aligned}$$

即飞机被击落的概率为0.458.



四、贝叶斯公式

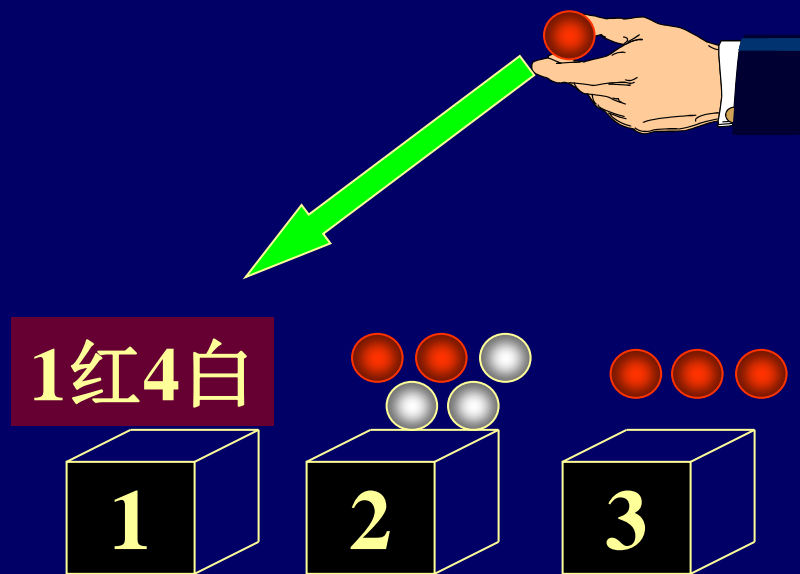
看一个例子：

某人从任一箱中任意摸出一球，发现是红球，求该球是取自1号箱的概率。

或者问：

该球取自哪号箱的可能性最大？

这一类问题是“**已知结果求原因**”。在实际中更为常见，它所求的是条件概率，是已知某结果发生条件下，探求各原因发生可能性大小。

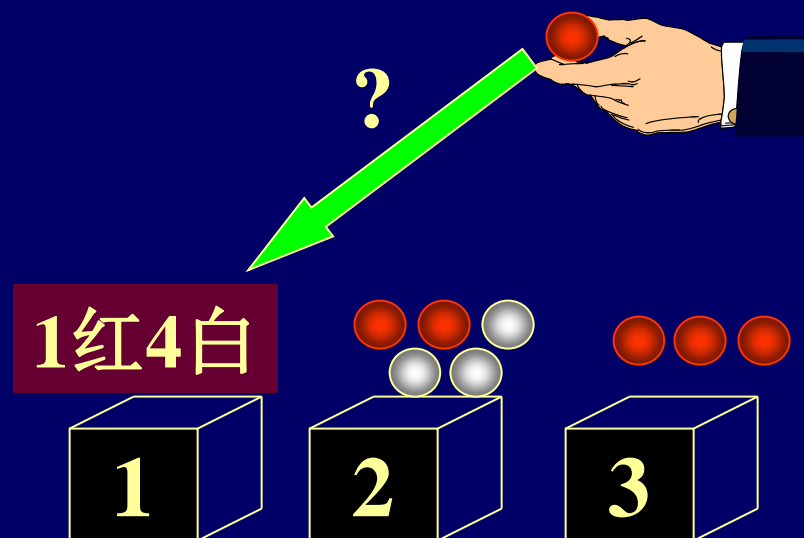


接下来我们介绍为解决这类问题而引出的

贝叶斯公式



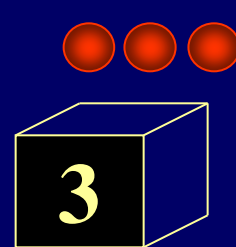
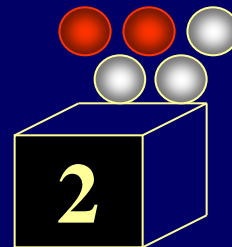
有三个箱子，分别编号为1,2,3，1号箱装有1个红球4个白球，2号箱装有2红球3白球，3号箱装有3红球。某人从三箱中任取一箱，从中任意摸出一球，发现是红球,求该球是取自1号箱的概率。



某人从任一箱中任意摸出一球，发现是红球，求该球是取自1号箱的概率。

记 $A_i = \{\text{球取自} i \text{号箱}\}, i=1,2,3;$
 $B = \{\text{取得红球}\}$

1红4白



求 $P(A_1|B)$

$$P(A_1 | B) = \frac{P(A_1 B)}{P(B)} = \frac{P(A_1)P(B | A_1)}{\sum_{k=1}^3 P(A_k)P(B | A_k)}$$

运用全概率公式
计算 $P(B)$

将这里得到的公式一般化，就得到

贝叶斯公式



定理2(贝叶斯公式) 设 A_1, A_2, \dots, A_n 为样本空间的一个划分, B 为 Ω 中的任一事件, 且 $P(B) > 0$, 则恒有

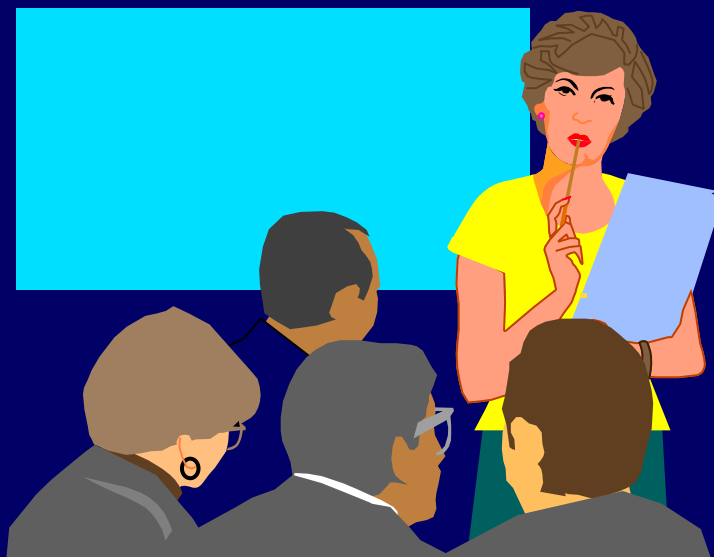
$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^n P(A_j)P(B | A_j)}$$
$$i = 1, 2, \dots, n$$

该公式于1763年由贝叶斯 (Bayes) 给出. 它是在观察到事件 B 已发生的条件下, 寻找导致 B 发生的每个原因的概率.



贝叶斯公式在实际中有很多应用.

■ 它可以帮助人们确定某结果（事件 B ）发生的最可能原因.



例 某一地区患有癌症的人占0.005，患者对一种试验反应是阳性的概率为0.95，正常人对这种试验反应是阳性的概率为0.04，现抽查了一个人，试验反应是阳性，问此人是癌症患者的概率有多大？

求解如下：设 $C=\{\text{抽查的人患有癌症}\}$ ，
 $A=\{\text{试验结果是阳性}\}$ ，

则 \bar{C} 表示“抽查的人不患癌症”。

已知 $P(C)=0.005, P(\bar{C})=0.995$ ，
 $P(A|C)=0.95, P(A|\bar{C})=0.04$

求 $P(C|A)$ 。



由贝叶斯公式，可得

$$P(C | A) = \frac{P(C)P(A | C)}{P(C)P(A | C) + P(\bar{C})P(A | \bar{C})}$$

代入数据计算得 $P(C | A) = 0.1066$

现在来分析一下结果的意义.



1. 这种试验对于诊断一个人是否患有癌症有无意义？
2. 检出阳性是否一定患有癌症？

1. 这种试验对于诊断一个人是否患有癌症有意义.

如果不做试验,抽查一人,他是患者的概率

$$P(C)=0.005$$

患者阳性反应的概率是0.95, 若试验后得阳性反应则根据试验得来的信息, 此人是患者的概率为

$$P(C | A)= 0.1066$$

从0.005增加到0.1066,将近增加约21倍.



2. 即使你检出阳性, 尚可不必过早下结论你有癌症, 这种可能性只有10.66% (平均来说, 1000个人中大约只有107人确患癌症), 此时医生常要通过再试验来确认.

试验结果为阳性, 此人确患癌症的概率为

$$P(C | A)=0.1066$$



■ 在贝叶斯公式中, $P(A_i)$ 和 $P(A_i | B)$ 分别称为原因的验前概率和验后概率.

$P(A_i)$ ($i=1,2,\dots,n$) 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 $P(A_i | B)$ 有了新的估计.

贝叶斯公式从数量上刻划了这种变化



例2 在数字通讯中,由于随机干扰,当发出信号"0"时,收到信号"0","不清","1"的概率分别是0.7,0.2和0.1;当发信号"1"时,收到信号为"1","不清"和"0"的概率分别是0.9,0.1和0,如果整个发报过程中"0"和"1"出现的概率分别是0.6和0.4,当收到"不清"时,试推测原发信号是什么?

解 设 $B = \{\text{发出信号"0"}\}$, 则 $\bar{B} = \{\text{发出信号"1"}\}$
 $A = \{\text{收到信号"不清"}\}$,
则 B 与 \bar{B} 为 $\Omega = \{\text{发出信号"0"或"1"}\}$ 的一个划分.



故收到信号为“不清”而原发信号为“0”的概率为

$$\begin{aligned} P(B | A) &= \frac{P(AB)}{P(A)} = \frac{P(B)P(A | B)}{P(B)P(A | B) + P(\bar{B})P(A | \bar{B})} \\ &= \frac{0.6 \times 0.2}{0.6 \times 0.2 + 0.4 \times 0.1} = 0.75. \end{aligned}$$

而收到信号为“不清”而原发信号为“1”的概率为

$$P(\bar{B} | A) = 1 - P(B | A) = 1 - 0.75 = 0.25.$$

因此,可以推测原发信号很可能(确切地说有75%的可能)是“0”.



例3(肺结核确诊率问题) 假定患肺结核的人通过接受胸部透视,被诊断出的概率为0.95,而未患肺结核的人通过透视,被诊断为有病的概率为0.002,又设某城市成年居民患肺结核的概率为0.1%.现若从该城市居民中随机选出一人来,通过透视被诊断为有肺结核,求这个人确实患有肺结核的概率是多少?

解 设 $A = \{\text{通过胸透诊断有肺结核}\}$,

$B = \{\text{确实患有肺结核}\}$, 则 $\bar{B} = \{\text{未患肺结核}\}$,
故所求概率为

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



五、小结

这一讲我们介绍了

全概率公式

贝叶斯公式

它们是加法公式和乘法公式的综合运用,同学们可通过进一步的练习去掌握它们.



六、 布置作业

习题1-4 (p23) : 4、5、8、11

