1.5 全概率公式与贝叶斯公式

全概率公式和贝叶斯公式主要用于计算比较复杂事件的概率,它们实质上是加法公式和乘法公式的综合运用.

综合运用

加法公式

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

A、B互斥

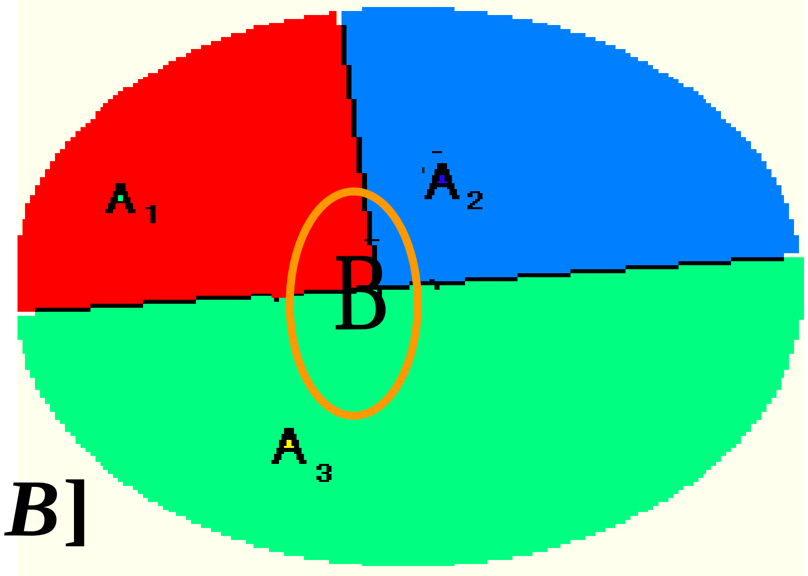
乘法公式

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

$P(A)>0$

*

引例 市场上有甲、乙、丙三家工厂生产的同一品牌产品,三家工厂的市场占有率分别为1/4、1/4、1/2,次品率分别为2%、1%、3%，试求市场上该品牌产品的次品率。



解 设A_1、A_2、A_3分别表示"买到一件甲、乙、丙厂产品"，B="买到一件次品".

$$P(B)=P(\Omega B)=P[(A_1+A_2+A_3)B]$$

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

$$=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$

$$=\frac{1}{4}\times 0.02+\frac{1}{4}\times 0.01+\frac{1}{2}\times 0.03\approx 0.0225$$

*

一、全概率公式

定理：设事件$A_1, A_2, \cdots, A_n$构成完备事件组, 且 $P(A_i)>0\ (i=1, 2, \cdots, n)$, 则对任一事件$B$, 有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B \mid A_i)$$

例1　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。

完备事件组? 事件B?

*

设A_i="飞机被i人击中" (i=0,1,2,3), B="飞机被击落"

依题意, $P(B|A_0)=0, P(B|A_1)=0.2, P(B|A_2)=0.6, P(B|A_3)=1$

为求$P(A_i)$, 设 H_i={飞机被第i人击中} (i=1,2,3)

$$P(A_1)=P(H_1\bar{H}_2\bar{H}_3+\bar{H}_1H_2\bar{H}_3+\bar{H}_1\bar{H}_2H_3)$$

$$=0.4\times0.5\times0.3+0.6\times0.5\times0.3+0.6\times0.5\times0.7=0.36$$

$$P(A_2)=P(H_1H_2\bar{H}_3+H_1\bar{H}_2H_3+\bar{H}_1H_2H_3)$$

$$=0.4\times0.5\times0.3+0.4\times0.5\times0.7+0.6\times0.5\times0.7=0.41$$

$$P(A_3)=P(H_1H_2H_3)=0.4\times0.5\times0.7=0.14$$

于是 $$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$

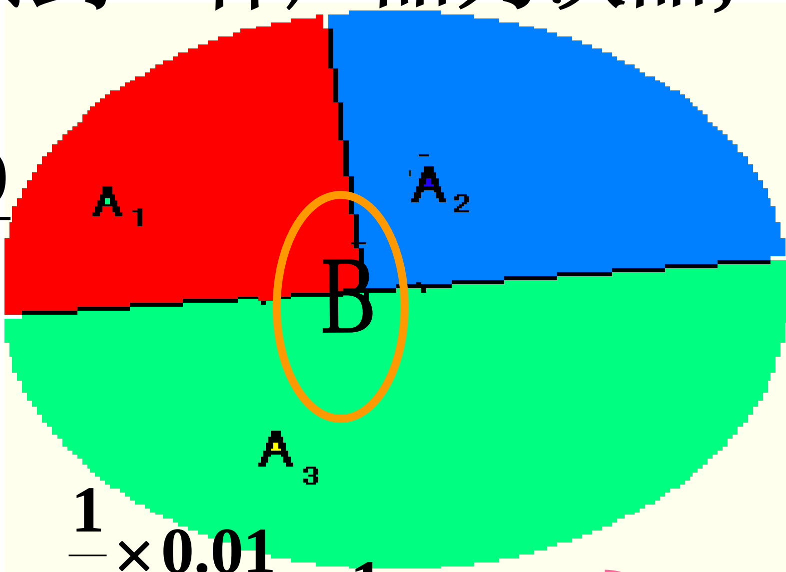
$$=0.36\times0.2+0.41\times0.6+0.14\times1=0.458$$

即飞机被击落的概率为0.458。

*

实际中还有一类问题——已知结果求原因

引例(续)某产品甲、乙、丙厂市场占有率1/4、1/4、1/2,次品率2%、1%、3%,(1)求市场上该品牌产品次品率(0.0225);(2)现买到一件产品为次品,求它由甲厂生产的概率.



解 $$P(A_1 \mid B) = \frac{P(A_1B)}{P(B)} = \frac{P(A_1)P(B \mid A_1)}{0.0225}$$

$$= \frac{\frac{1}{4}\times 0.02}{0.0225} = \frac{2}{9} \approx 0.222$$

$$P(A_2 \mid B) = \frac{P(A_2B)}{P(B)} = \frac{P(A_2)P(B \mid A_2)}{0.0225} = \frac{\frac{1}{4}\times 0.01}{0.0225} = \frac{1}{9} \approx 0.111$$

$$P(A_3 \mid B) = \frac{P(A_3B)}{P(B)} = \frac{P(A_3)P(B \mid A_3)}{0.0225} = \frac{\frac{1}{2}\times 0.03}{0.0225} = \frac{6}{9} \approx 0.667$$

*

二、贝叶斯公式

定理：设事件$A_1, A_2, \cdots, A_n$构成完备事件组,且$P(A_i)>0(i=1, 2, \cdots, n)$，对事件$B$, 若$P(B)>0$，则有

$$P(A_k|B)=\frac{P(A_kB)}{P(B)}=\frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}\quad(k=1, 2, \cdots, n)$$

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.这类问题在实际中更常见.

*

$$P(A_k|B)=\frac{P(A_kB)}{P(B)}=\frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}\quad (k=1,2,\cdots,n)$$

在贝叶斯公式中，$P(A_i)$和$P(A_i|B)$分别称为原因的**先验概率**和**后验概率**.

$P(A_i)$ $(i=1,2,\ldots,n)$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 商店论箱出售玻璃杯,每箱20只,其中每箱含0,1,2只次品的概率分别为0.8, 0.1,0.1,某顾客选中一箱,从中任选4只检查,结果都是好的,便买下了这一箱. 问这一箱不含次品的概率是多少?

完备事件组? 事件B?

解:设A_i=“任取一箱,含i只次品” (i=0,1,2),

B=“从一箱中任取4只检查,结果都是好的”

由已知:$P(A_0)=0.8,\ P(A_1)=0.1,\ P(A_2)=0.1$

$$P(B\mid A_0)=1 \qquad P(B\mid A_1)=\frac{C_{19}^4}{C_{20}^4}=\frac{4}{5} \qquad P(B\mid A_2)=\frac{C_{18}^4}{C_{20}^4}=\frac{12}{19}$$

由Bayes公式:

$$P(A_0\mid B)=\frac{P(A_0)P(B\mid A_0)}{\sum_{i=0}^{2}P(A_i)P(B\mid A_i)}=\frac{0.8\times 1}{0.8\times 1+0.1\times 4/5+0.1\times 12/19}\approx 0.848$$

*

例3 数字通讯中, 信源发0、1两种信号, 其中发0的概率0.55, 发1的概率0.45. 由于信道中存在干扰, 发0的时候, 接收端分别以概率0.9、0.05和0.05接收为0、1和“不清”. 发1的时候, 接收端分别以概率0.85、0.05和0.1接收为1、0和“不清”.现接收端接收到一个“1”的信号, 问发端发的是0的概率是多少?

解：设A=“发射端发射0”, B=“接收端收到1”

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

$$=\frac{0.55\times0.05}{0.55\times0.05+0.45\times0.85}=0.067$$

*

例4　从20000份患有疾病d_1,d_2,d_3的病历卡中统计得到下表数据. 其中：$S=\{S_1,S_2,S_3,S_4\}$.

S_1=食欲不振，S_2=胸痛，S_3=呼吸急促, S_4=发热.

疾病	人数	出现S中一个或几个症状人数
d_1	7750	7500
d_2	5250	4200
d_3	7000	3500

一个具有S中的某些症状的人前来求诊，在没有别的可资依据的诊断手段情况下,诊断该病人患有哪一种疾病较合适?

该问题观察的个数很多,用频率作为概率的近似值.

解:设A_i="患者患有疾病d_i"(i=1,2, 3),

B="患有出现S中的某些症状"

$$P\left(A_1\right)=\frac{7750}{20000}=0.3875, \qquad P\left(B \mid A_1\right)=\frac{7500}{7750}=0.9677,$$

$$P\left(A_2\right)=\frac{5250}{20000}=0.2625, \qquad P\left(B \mid A_2\right)=\frac{4200}{5250}=0.8,$$

$$P\left(A_3\right)=\frac{7000}{20000}=0.35 \qquad P\left(B \mid A_3\right)=\frac{3500}{7500}=0.5.$$

$$P\left(B\right)=P\left(A_1\right)P\left(B \mid A_1\right)+P\left(A_2\right)P\left(B \mid A_2\right)+P\left(A_3\right)P\left(B \mid A_3\right)$$

$$=0.3875\times 0.9677+0.2625\times 0.8+0.35\times 0.5$$

$$=0.76$$

*

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$
$$=0.3875\times0.9677+0.2625\times0.8+0.35\times0.5$$
$$=0.76$$

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{0.3875\times0.9677}{0.76}=0.4934$$

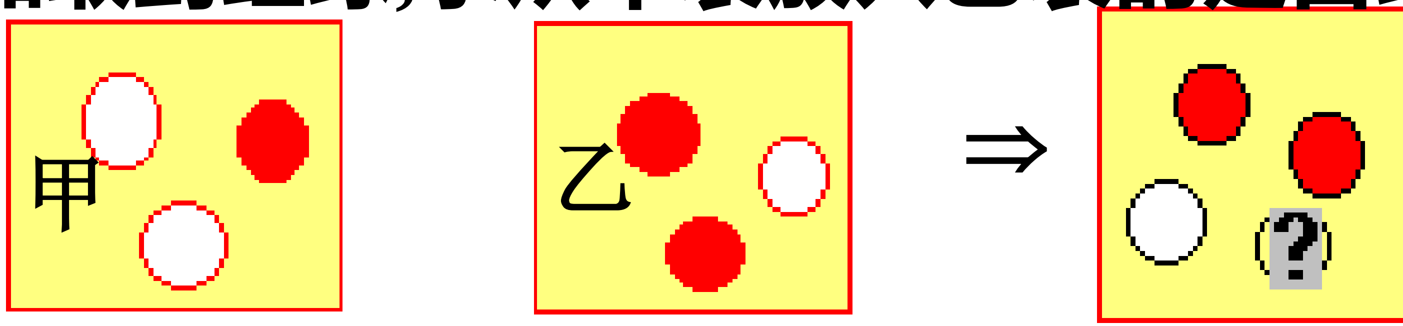
$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{P(B)}=\frac{0.2625\times0.8}{0.76}=0.2763$$

$$P(A_3|B)=\frac{P(A_3)P(B|A_3)}{P(B)}=\frac{0.35\times0.5}{0.76}=0.2303$$

从而推测病人患有疾病d_1较为合理.

*

例5 设甲袋中有2个白球,1个红球,乙袋中有2个红球,1个白球,手感上均无区别.今从甲袋任取一球放入乙袋,搅匀后再从乙袋任取一球,求此球是红球的概率;若已知取到红球,求从甲袋放入乙袋的是白球的概率.



解:设A=“从甲袋放入乙袋的是白球”,

B=“从乙袋中任取一球是红球”

$$P(B)=P(A)P(B\mid A)+P(\overline{A})P(B\mid \overline{A})=\frac{2}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{3}{4}=\frac{7}{12}$$

$$P(A\mid B)=\frac{P(A)P(B\mid A)}{P(B)}=\frac{\frac{2}{3}\times\frac{1}{2}}{\frac{7}{12}}=\frac{4}{7}$$

*

作业：习题1.5

2, 9(2), 13, 16