§1.4 全概率公式和贝叶斯公式

全概率公式(The Law of Total Probability)和贝叶斯公式(Bayes’ Theorem)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

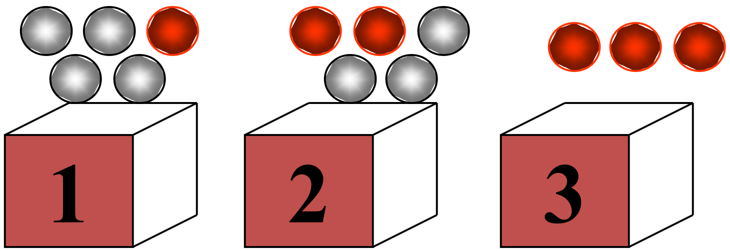
加法公式	乘法公式
$P(A+B)=P(A)+P(B)$	$P(AB)= P(A)P(B\|A)$
A、B互斥	$P(A)>0$

一. 全概率公式The Law of Total Probability

例1 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={球取自i号箱}，

B ={取得红球}



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

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

对每一项运用乘法公式

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

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.

外文书上的推导

$$p(X)=\sum_{Y}p(X,Y)=\sum_{Y}p(X|Y)p(Y)$$

Sum rule
加法公式

Product rule
乘法公式

全概率公式

设S为随机试验的样本空间，$A_1,A_2,\ldots,A_n$是两两互斥的事件，且有$P(A_i)>0$，$i=1,2,\ldots,n$，

$\bigcup_{i=1}^{n} A_i = S$， 则对任一事件$B$，有

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

称满足上述条件的$A_1,A_2,\ldots,A_n$为完备事件组.

证明 $A_1,\ A_2,\ \cdots,\ A_n$ 两两互不相容，

得 $A_1B,\ A_2B,\ \cdots,\ A_nB$ 也两两互不相容；

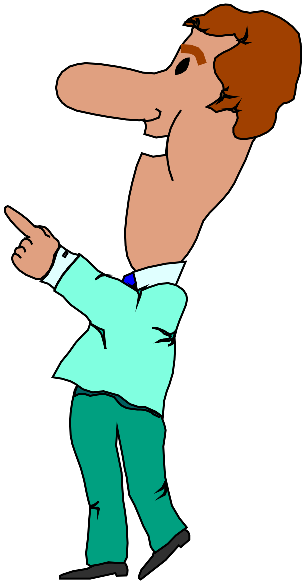
$$P(A_iBA_jB) = P(A_iBA_j) = P(\emptyset) = 0$$

我们还可以从另一个角度去理解 **全概率公式**

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

$$P(BA_i)=P(A_i)P(B\,|A_i)$$

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



全概率公式的关键：

数学模型

完备事件组

例2.甲、乙、丙三个车间生产同一种产品，其产量分别占总数的25%,35%,40%，次品率分别为5%,4%,2%，从这批产品中任取一件，求它是次品的概率.

A_1，A_2，A_3 分别表示产品由甲、乙、丙车间生产

B表示产品为次品

完备事件组

全概率公式

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

$$= 0.0345$$

例3 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.

完备事件组是什么？　　甲、乙、丙？

设B={飞机被击落}

A_i={飞机被i个人击中}, $i = 0, 1, 2, 3$

由全概率公式

$$P(B)= P(A_0)P(B|A_0)+ P(A_1)P(B|A_1)+ P(A_2)P(B|A_2)+ P(A_3)P(B|A_3)$$

由全概率公式

$$P(B)=P(A_0)P(B\,|A_0)+P(A_1)P(B\,|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B\,|A_3)$$

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_j={飞机被第j人击中}, j=甲,乙,丙

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

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

$$P(A_3)=P(H_1H_2H_3)$$

加法公式　　独立性

$P(A_1)=0.36$; $P(A_2)=0.41$; $P(A_3)=0.14$.

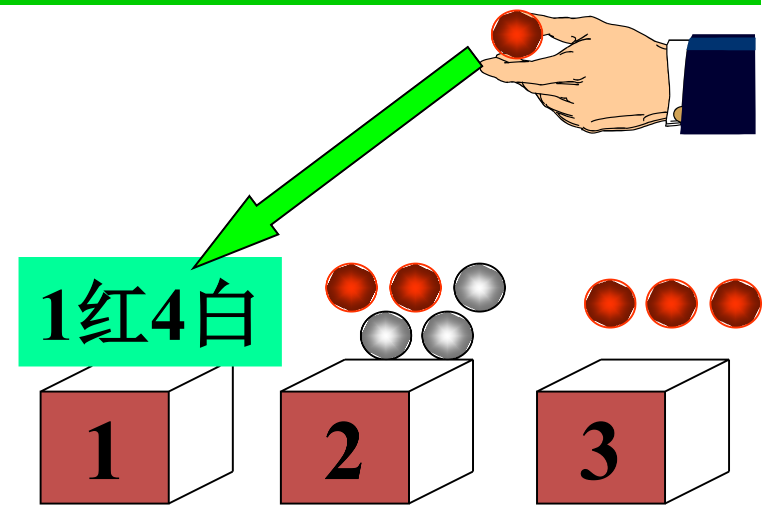
$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$

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

即飞机被击落的概率为0.458.

实际中还有下面一类问题“已知结果求原因”

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.

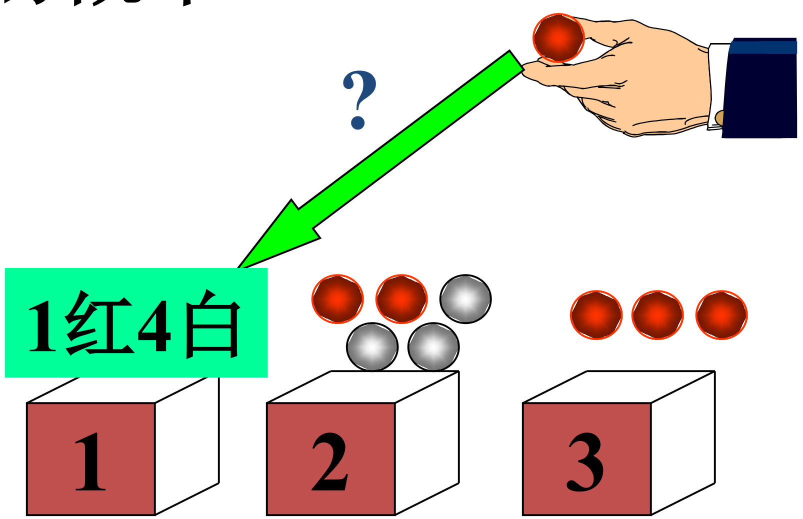


或者问:

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，求各原因发生可能性大小.

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．



记 A_i={球取自i号箱}, i=1,2,3; B ={取得红球}

求$P(A_1|B)$.

条件概率公式

乘法公式

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

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

二. 贝叶斯公式Bayes' Theorem

设$A_1, A_2, \ldots, A_n$是完备事件组，则对任一事件B，有

先验概率
Prior Prob.

相似度
Likelihood

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

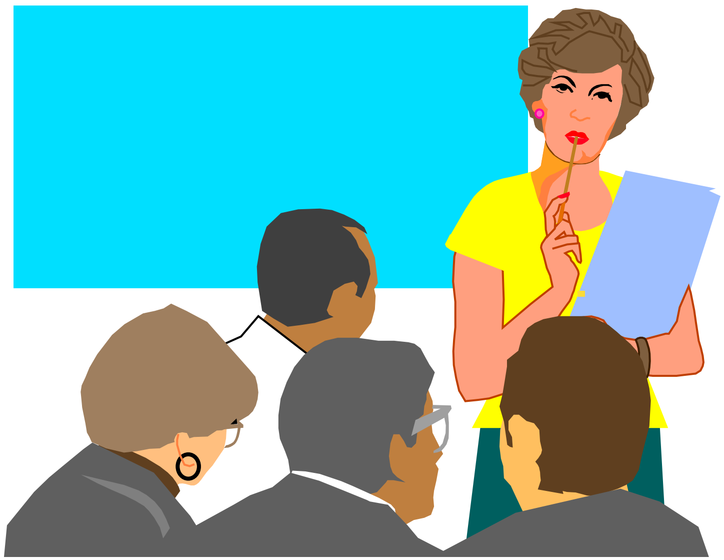
后验概率
Posterior Prob.

全概率
Sum over Space hypotheses

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A_i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**某结果发生的最可能原因**.



Bayes公式的实际意义

假定 $A_1, A_2, \cdots, A_n$ 为导致试验结果的“原因”

称 $P(A_i)$ $(i=1,2,\cdots,n)$ 为先验概率

先验概率反映了各种“原因”发生的可能性大小（在试验前是知道的，以往的经验得到）

若试验产生“结果”事件 B，则要探讨事件发生的“原因”

$$P(A_i \mid B) \quad (i=1,2,\cdots,n)$$

称 $P(A_i|B)$为后验概率

后验概率反映了试验后对各种“原因”发生的可能性大小的推断

① 后验概率可以通过 Bayes 公式进行计算

后验转先验

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

② Bayes 方法广泛应用于网络、分类、诊断、估计、检验、判别、推理等方面

Bayes公式的重要意义在于利用人们掌握的先验知识来推断后验概率

例如：肺癌计算机自动辅助诊断系统

假定B为“肺癌”。A_1为“抽烟”，A_2为“喝酒”，A_3为“唱歌”，A_4为“生活在雾霾严重的城市”。要确定那种习惯是“肺癌”的主要杀手？

应用统计方法确定先验概率 $P(A_i)$: 比如A_3表示人们“抽烟”的概率，可以进行抽样统计的方法获得。随机抽取一大帮人，“抽烟”的频率。

可以应用统计方法获得$P(B|A_i)$，如$P(B|A_3)$表示“唱歌”多的人得肺癌B的概率。随机抽取的一大帮人中，“唱歌”的人中得肺癌的概率

$$P(B \mid A_i) \quad (i = 1, 2, \cdots, n)$$

应用 Bayes 公式计算机可计算出后验概率

$$P(A_i \mid B) \quad (i = 1, 2, \cdots, n)$$

例2. 甲、乙、丙三个车间生产同一种产品，其产量分别占总数的25%, 35%, 40%，次品率分别为5%, 4%, 2%，今随机地从中任取一件，发现是次品，问这产品由哪个车间生产的可能性较大？

$$P(B)=0.0345$$

$$P(A_1\mid B)=\frac{P(A_1)P(B\mid A_1)}{P(B)}\approx 0.362$$

贝叶斯公式

$$P(A_2\mid B)=\frac{P(A_2)P(B\mid A_2)}{P(B)}\approx 0.406$$

$$P(A_3\mid B)=\frac{P(A_3)P(B\mid A_3)}{P(B)}\approx 0.232$$

该产品由乙车间生产的可能性最大。

例4 用甲胎蛋白检测法(AFP)诊断肝病，已知确实患肝病者被诊断为肝病的概率为0.95，未患肝病者被误诊为肝病的概率为0.02，假设人群中肝病的发病率为0.0004，现在有一个人被诊断为患有肝病，求此人确实为肝病患者的概率。

设 A={肝病患者}，B={被诊断为患有肝病}，

由贝叶斯公式，

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

$$=\frac{0.0004\times0.95}{0.0004\times0.95+(1-0.0004)\times0.02}\approx0.0187$$

总结

1. 全概率公式
2. 贝叶斯公式

Summary

1.1 随机事件定义,关系,运算

定义：样本空间；基本、复合、必然、不可能事件
关系：包含、相等、并、交、差、互斥、对立
运算：吸收、重余、幂等、差化积、交换、结合、分配、反演

1.2 随机事件的概率

古典定义
统计定义
公理化定义

1.3概率的性质与运算

性质：非负、规范、可列可加性
运算：加法、逆事件、减法、广义加减、乘法
条件概率、事件独立性、伯努利概型、二项公式

1.4 全概率公式和贝叶斯公式

全概率公式
贝叶斯公式