第一章教学计划

教学内容：

条件概率，乘法公式，全概率公式，贝叶斯公式。

教学目的及目标：

掌握条件概率及相关公式。

教学重点：

条件概率，概率乘法公式，全概率公式。

教学难点：

全概率公式。

§1.3　条件概率

一、条件概率的定义及性质

1、概念及引例

在解决许多概率问题时，往往需要在有某些附加信息(条件)下求事件的概率.

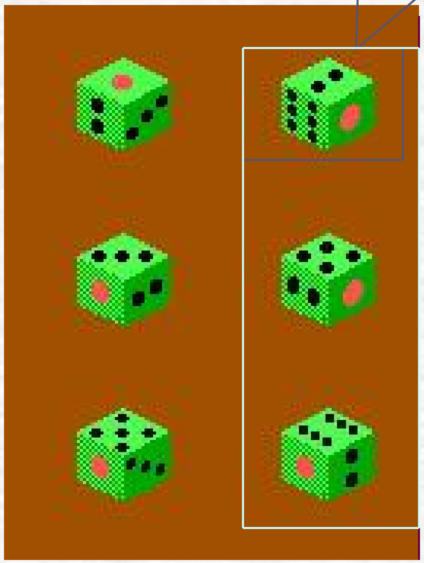
在事件B发生的条件下（隐含P(B)>0)，事件A发生的概率称为A对B的**条件概率**，记作$P(A|B)$.

一般来说，—$P(A|B)\neq P(A)$ 。如下例

例如，掷一颗均匀的骰子，A={掷出2点}，

B={掷出偶数点}， $P(A)$=1/6， $P(A|B)$=?

分析：事件B已经发生，因此，这时试验的所有可能结果构成的集合就是B，



B中共有3个元素，它们的出现是等可能的，其中只有1个在A中，

于是$P(A|B)$= 1/3.

容易看到，这里

$$P(A|B)=\frac{1}{3}=\frac{1/6}{3/6}=\frac{P(AB)}{P(B)}.$$

◁ ◀Back ▷

2．条件概率的定义:

设$(\Omega, \mathscr{F}, P)$为一概率空间，A，B是两事件，且P(B)>0，称

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

为在事件B发生的条件下事件A发生的**条件概率**

注：（1）若P(A)>0，同样可定义

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

（2）条件概率P（•|B）满足概率定义的三条公理。

这表明，条件概率也是一种概率，因此，概率的一切性质都适用于条件概率，例如：

$$P(\Phi|B)=0$$

$$P(\overline{A}|B)=1-P(A|B)$$

$$P(A_1 \cup A_2|B)=P(A_1|B)+P(A_2|B)-P(A_1A_2|B)$$

…………

2. 计算

一般有两种方法：

(1) 定义法：$P(B|A)=P(AB)/P(A)$

(2) 缩小样本空间法： $P(B|A)=\mu_{AB}/\mu_{A}$

例1 掷两颗均匀骰子,已知第一颗掷出6点,问“掷出点数之和不小于10”的概率是多少?

解: 设A={掷出点数之和不小于10}

B={第一颗掷出6点}

应用定义

解法1:

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{3/36}{6/36}=\frac{1}{2}$$

解法2:

$$P(A|B)=\frac{3}{6}=\frac{1}{2}$$

在B发生后的缩减样本空间中计算

Back

条件概率$P(A|B)$与$P(A)$的区别:

$P(A)$与$P(A|B)$的区别在于两者发生的条件不同,它们是两个不同的概念,在数值上一般也不同.

条件概率$P(A|B)$与$P(A)$数值关系:

何时一定有:

$$P(A|B) \geq P(A)?$$

或 $P(A|B) \leq P(A)$?

以及$P(A)=P(A|B)$？

三、概率乘法公式

由条件概率的定义：$P(A\mid B)=\dfrac{P(AB)}{P(B)}$

若已知$P(B)$, $P(A|B)$, 可以反求$P(AB)$:

若$P(B)>0$，则$P(AB)=P(B)P(A|B)$ (2)

将A、B

若$P(A)$

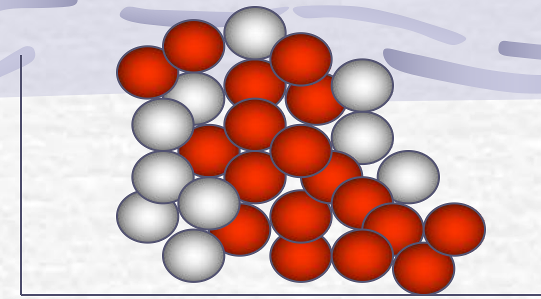
而

(2)和(3)式都称为乘法公式，利用它们可计算两个事件同时发生的概率

故$P(A)>0$，则$P(AB)=P(A)P(B|A)$ (3)

◁ Back ▷

乘法公式应用举例

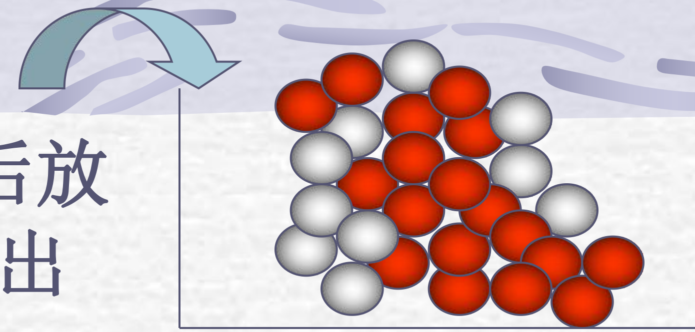


b个白球, r个红球

例4 波里亚罐子模型

一个罐子中包含b个白球和r个红球. 随机地抽取一个球，观看颜色后放回罐中，并且再加进c个与所抽出的球具有相同颜色的球. 这种手续进行四次，试求第一、二次取到白球且第三、四次取到红球的概率.

随机取一个球，观看颜色后放回罐中，并且再加进c个与所抽出的球具有相同颜色的球.

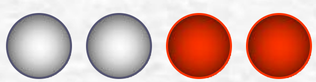


b个白球, r个红球

解：设W_i={第i次取出是白球}, i=1,2,3,4

R_j={第j次取出是红球}，j=1,2,3,4

于是$W_1W_2R_3R_4$表示事件“连续取四个球，第一、第二个是白球，第三、四个是红球.”

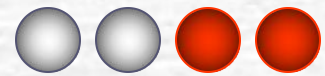


用乘法公式容易求出

$$P(W_1W_2R_3R_4)$$

$$=P(W_1)\mathrm{P}(W_2|W_1)\mathrm{P}(R_3|W_1W_2)\mathrm{P}(R_4|W_1W_2R_3)$$

$$=\frac{b}{b+r}\frac{b+c}{b+r+c}\frac{r}{b+r+2c}\frac{r+c}{b+r+3c}$$



当 c>0 时，由于每次取出球后会增加下一次也取到同色球的概率. 这是一个**传染病模型**. 每次发现一个传染病患者，都会增加再传染的概率.

推广到多个事件的乘法公式:

当$P(A_1A_2)>0$时，有

$$P(A_1A_2A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)$$

一般地,

当$P(A_1A_2\ldots A_{n-1})>0$时，有

$$P(A_1A_2\ldots A_n)$$

$$=P(A_1)P(A_2|A_1)\ldots P(A_n|A_1A_2\ldots A_{n-1})$$

◀Back

注意$P(AB)$与$P(A \mid B)$的区别：

"B发生"：在$P(AB)$中是结果，

在$P(A \mid B)$中是条件！

请看下面的例子

例2 甲、乙两厂共同生产1000个零件，其中300件是乙厂生产的. 而在这300个零件中，有189个是标准件，现从这1000个零件中任取一个，问这个零件是乙厂生产的标准件的概率是多少?

设B={零件是乙厂生产}

A={是标准件}

所求为$P(AB)$.

设B={零件是乙厂生产}

A={是标准件}

所求为$P(AB)$.

若改为“发现它是乙厂生产的,
问它是标准件的概率是多少?”

求的是 $P(A|B)$.

B发生,
在$P(AB)$中作为结果;
在$P(A|B)$中作为条件.

◀Back

*例3 设某种动物由出生算起活到20年以上的概率为0.8，活到25年以上的概率为0.4. 问现年20岁的这种动物，它能活到25岁以上的概率是多少？

解：设A={能活20年以上}，B={能活25年以上}

所求为$P(B|A)$.

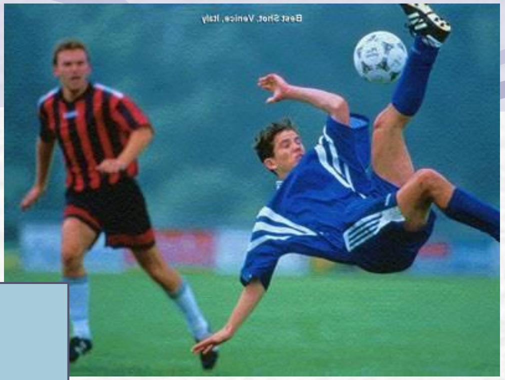
依题意，$P(A)=0.8$, $P(B)=0.4$

$$P(B\mid A)=\frac{P(AB)}{P(A)}=\frac{P(B)}{P(A)}=\frac{0.4}{0.8}=0.5$$

Back

抽签的公平性

5个球迷一张入场券.抽签.



入场券

5张同样的卡片，1张上写有“入场券”，其余4张空白. 将它们放在一起，洗匀，让5个人依次抽取.

用A_i表示“第i个人抽到入场券”，$i=1,2,3,4,5$.

则 $\overline{A_i}$表示“第i个人未抽到入场券”

显然，$P(A_1)=1/5$，$P(\overline{A_1})=4/5$

Back

由于 $A_2 = A_1A_2 + \overline{A}_1A_2 = \overline{A}_1A_2$

因为若第2个人抽到了入场券，第1个人肯定没抽到.

由乘法公式

$$P(A_2) = P(\overline{A}_1)P(A_2 \mid \overline{A}_1)$$

$$= (4/5)(1/4) = 1/5$$

同理，

$$P(A_3) = P(\overline{A}_1\overline{A}_2A_3) = P(\overline{A}_1)P(\overline{A}_2 \mid \overline{A}_1)P(A_3 \mid \overline{A}_1\overline{A}_2)$$

$$= (4/5)(3/4)(1/3) = 1/5$$

继续做下去就会发现，每个人抽到“入场券”的概率都是1/5.

Back

三、全概率公式与贝叶斯公式

综合运用

加法公式

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

A、B互斥

乘法公式

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

$P(A)>0$

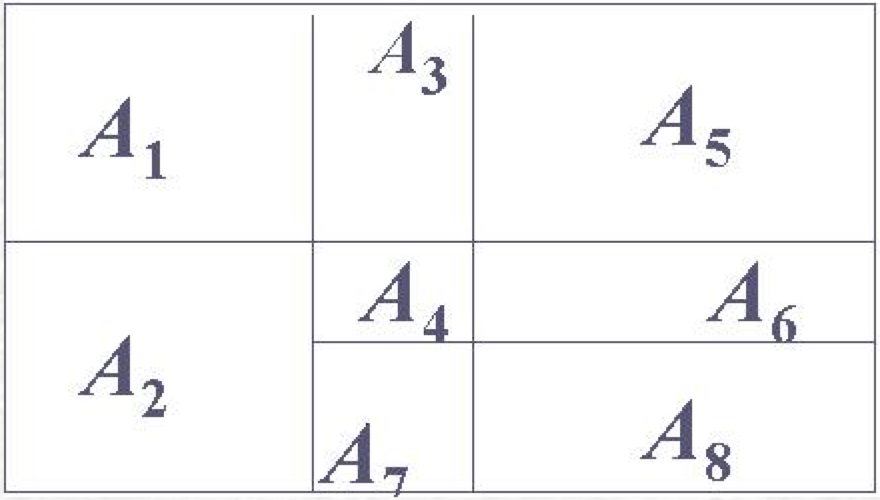
样本空间的划分：设Ω为试验E的样本空间，

$A_1, A_2, \cdots, A_n$为E的一组事件，若

（1） $A_iA_j=\Phi$, $i\neq j$ ， i ， j $=1, 2, \cdots, n$;

（2） $A_1\cup A_2\cup\cdots\cup A_n=\Omega$,

则称A_1，A_2，$\cdots, A_n$为样本空间Ω的一个**划分**。

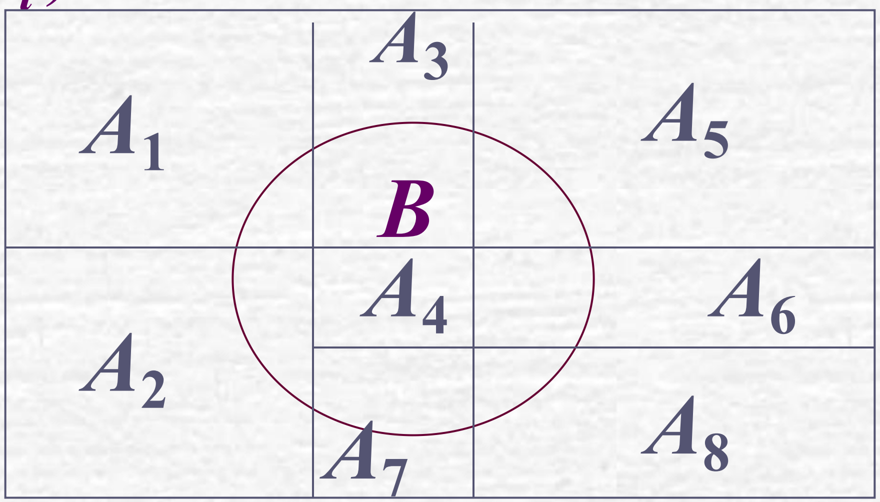


Back

全概率公式:

设$A_1,A_2,\ldots,A_n$是试验E的样本空间Ω的一个划分，且$P(A_i)>0$，$i=1,2,\ldots,n$, B是任一事件, 则

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



Back

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

全概率公式的来由，不难由上式看出:

"全"部概率$P(B)$被分解成了许多部分之和.

它的理论和实用意义在于:

在较复杂情况下直接计算$P(B)$不易,但B总是伴随着某个A_i出现，适当地去构造这一组A_i往往可以简化计算.

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 ($i=1,2,\ldots,n$)，其中，B 由原因 A_i 所引发的概率是

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

每一原因都可能导致 B 发生，故 B 发生的概率是各原因引发 B 的概率的总和。这就是全概率公式.

例　　盒中12个新乒乓球，每次比赛从盒中任取3个球，用后放回。第三次比赛时3个球，取到3个新球的概率。

解：设A:第三次比赛时取到3个新球，

　B_i:第二次比赛时取到i个新球（i=0,1,2,3）。

则B_i构成了第三次比赛取球试验的样本空间的一个划分。

$$P(B_0)=\frac{C_3^3}{C_{12}^3},\quad P(B_1)=\frac{C_9^1\cdot C_3^2}{C_{12}^3},$$

$$P(B_2)=\frac{C_9^2C_3^1}{C_{12}^3},P(B_3)=\frac{C_9^3}{C_{12}^3}$$

$$P(A \mid B_0) = \frac{C_9^3}{C_{12}^3}, \quad P(A \mid B_1) = \frac{C_8^3}{C_{12}^3},$$

$$P(A \mid B_2) = \frac{C_7^3}{C_{12}^3}, P(A \mid B_3) = \frac{C_6^3}{C_{12}^3}$$

于是

$$P(A) = \sum_{i=0}^{3} P(B_i) P(A \mid B_i) = \frac{441}{3025} \approx 0.146.$$

贝叶斯公式:

设$A_1,A_2,\ldots,A_n$是试验E的样本空间Ω的一个划分，且$P(A_i)>0$，$i=1,2,\ldots,n$, B是任一事件且$P(B)>0$, 则

$$P(A_i \mid B) = P(A_i)P(B \mid A_i) \Big/ \sum_{j=1}^{n} P(A_j)P(B \mid A_j)$$

$$i = 1,2,\cdots,n$$

该公式于1763年由英国统计学家贝叶斯(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.

Thomas Bayes
1702 - 1763

Born: 1702 in London, England
Died: 7 April 1763 in Tunbridge Wells, Kent, England

例2：对以往数据分析结果表明，当机器调整得良好时，产品的合格率为90%，而当机器发生某一故障时，其合格率为30%。每天早上机器开动时，机器调整良好的概率为75%，试求已知某日早上第一件产品是合格品时，机器调整良好的概率是多少？

解：设A：“产品合格”，B：“机器调整良好”，

则$P(A|B)=0.9$，$P(A|\overline{B})=0.3$, $P(B)=0.75$，

所求概率为$P(B|A)$。由贝叶斯公式

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

$$=\frac{0.9\times 0.75}{0.9\times 0.75+0.3\times 0.25}=0.9$$

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.

贝叶斯公式 $P(A_i|B)=\dfrac{P(A_i)P(B\mid A_i)}{\sum\limits_{j=1}^{n}P(A_i)P(B\mid A_i)}$

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

贝叶斯公式从数量上刻划了这种变化。

知 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$P(A_i\mid B)$有了新的估计.

例3：假定用甲胎蛋白法诊断肝癌，$P(A|B_1)=0.99$, $P(A|B_2)=0.05$，其中B_1表示“被检验者患有肝癌”，$B_2=\overline{B}_1$，A 表示“被检验者试验反应为阳性”。据调查某地区居民的肝癌发病率$P(B_1)=0.0004$。现若由该地区某居民检验结果呈阳性，问他患肝癌的概率$P(B_1|A)$是多少？

解：$$P(B_1 \mid A)=\frac{P(B_1)P(A \mid B_1)}{P(B_1)P(A \mid B_1)+P(B_2)P(A \mid B_2)}$$

$$=\frac{0.0004 \times 0.99}{0.0004 \times 0.99+0.9996 \times 0.05}=0.00786$$

在实际中，医生常用另一些辅助方法先进行初查，排除大量明显不是肝癌的人，当医生怀疑某人有可能患肝癌时，才建议用甲胎蛋白法检验。这时在被怀疑的对象中，肝癌的发病率已显著提高了，比如说

P（B_1）=0.4，这时再用贝叶斯公式进行计算，可得

$$P(B_1 \mid A)=\frac{0.99 \times 0.4}{0.99 \times 0.4+0.05 \times 0.6}=0.9296$$

这样就大大提高了甲胎蛋白法的准确率了。

练：

一男子到闹市区去，他遇到背后袭击并被抢劫，他断定凶手是个白人。然而，当调查这一案件的法院在相似的条件下多次重复现场情况时，受害者正确识别袭击者种族的次数约占**80%**。求袭击者为白人的概率。

解：记**A:**“袭击者为白人”，**B:** “受害者指认袭击者为白人”，所求概率为**P(A|B)**。由题意

$$P(B \mid A)=0.8, P(\overline{B} \mid \overline{A})=0.8$$

设该地白人比例为**p**，即**P(A)=p**，则

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

$$=\frac{p\times 0.8}{p\times 0.8+(1-p)\times 0.8}=\frac{4p}{1+3p}$$

若p=1/2, 则所求概率为0.8;

若p<1/2, 则所求概率小于0.8;

若p>1/2, 则所求概率大于0.8。