第一章 概率论的基本概念

§3 条件概率

§3条件概率

目录索引

- <u>条件概率</u>
- 二 乘法定理
- 三 全概率公式和贝叶斯公式

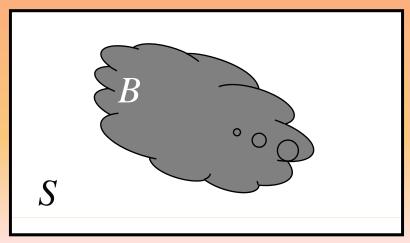


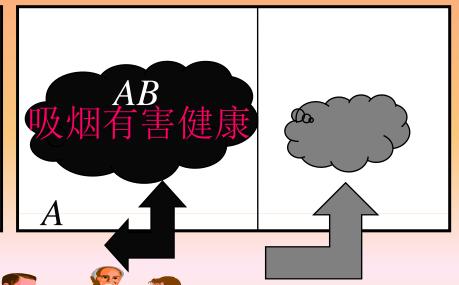
第一章 概率论的基本概念

一条件概率

§3条件概率

条件概率是概率论中一个重要而实用的概念。 它所考虑的是事件 A 已经发生的条件下事件 B 发生的概率。









返回主目录

条件概率

§3条件概率

设A、B是某随机试验中的两个事件,且P(A)>0则称事件B在"事件A已发生"这一附加条件下的

概率为在事件A已发生的条件下事件B的条件概率

,简称为B在A之下的条件概率,记为P(B|A)

例1 盒中有4个外形相同的球,它们的标号分别为1、2、3、4,每次从盒中取出一球,有放回地取两次.

则该试验的所有可能的结果为

(1, 1) (1, 2) (1, 3) (1, 4)

(2, 1) (2, 2) (2, 3) (2, 4)

(3,1) (3,2) (3,3) (3,4)

(4, 1) (4, 2) (4, 3) (4, 4)

其中(i, j)表示第一次取i号球,第二次取j号球



设A={第一次取出球的标号为2}

B={ 取出的两球标号之和为 4 }

则事件B所含的样本点为

(1,3) (2,2) (3,1)

因此事件B的概率为:

$$P(B) = \frac{3}{16}$$

若我们考虑在事件A发生的条件下,事件B发生的概率并记此概率为: P(B|A)

由于已知事件A已经发生,则该试验的所有可能结果为



$$(2, 1)$$
 $(2, 2)$ $(2, 3)$ $(2, 4)$

这时,事件B是在事件A已经发生的条件下的概率 ,因此这时所求的概率为

$$P(B|A) = \frac{1}{4}$$

注: 由例1可以看出,事件在"条件A已发生这附加条件的概率与不附加这个条件的概率是不同的.

因此,有必要引入下面的定义:



设A、B是某随机试验中的两个事件,且

§3条件概率

则 $P(B|A) = \frac{P(AB)}{P(A)}$

称为在事件A已发生的条件下事件B的条件概率, 简称为B在A之下的条件概率。

在例 1 中,我们已求得
$$P(B) = \frac{3}{16}$$
, $P(B|A) = \frac{1}{4}$ 还可求得 $P(A) = \frac{4}{16}$, $P(AB) = \frac{1}{16}$ 故有 $P(B|A) = \frac{P(AB)}{P(A)}$

条件概率的性质:

§3条件概率

- 1 非负性:对任意事件B,有 $P(B|A) \ge 0$
- 2 规范性: P(S|A)=1;
- 3 可列可加性:如果随机事件 B_1 , B_2 ,…, B_n , 而 两 可不相容,则

$$P\left(\bigcup_{n=1}^{\infty} B_n \middle| A\right) = \sum_{n=1}^{\infty} P(B_n \middle| A)$$

第一章 概率论的基本概念

§3条件概率

例 2已知某家庭有3个小孩,且至少有一个是女孩,求该家庭至少有一个男孩的概率.

解: 设 A={ 3个小孩至少有一个女孩 } B={ 3个小孩至少有一个男孩 }

則
$$P(A)=1-P(\overline{A})=1-\frac{1}{8}=\frac{7}{8}$$
 $P(AB)=1-\frac{1}{8}=\frac{6}{8}=\frac{6}{8}$
所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{8}{7}=\frac{6}{7}$

两个事件的乘法公式

§3条件概率

由条件概率的计算公式

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

我们得

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

这就是两个事件的乘法公式.

多个事件的乘法公式

§3条件概率

设 A_1 , A_2 , …, A_n 为n个随机事件,且

$$P(A_1 A_2 \cdots A_{n-1}) > 0$$

则有

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2 | A_1)$$

$$P(A_3 | A_1 A_2) \cdots P(A_n | A_1 A_2 \cdots A_{n-1})$$

这就是n个事件的乘法公式.

例 3 袋中有一个白球与一个黑球,现每次从中取出一球,若取出白球,则除把白球放回外再加进一个白球,直至取出黑球为止.求取了n次都未取出黑球的概率.

解:

设
$$B = \{ \mathbf{取} \mid n \text{ 次都未取出黑球} \}$$

$$A_i = \{ \hat{\mathbf{x}} \mid i \text{ 次取出白球} \} \quad (i = 1, 2, \dots, n)$$
则 $B = A_1 A_2 \cdots A_n$

由乘法公式, 我们有



第一章 概率论的基本概念

§3条件概率

$$P(B) = P(A_1 \ A_2 \cdots A_n)$$

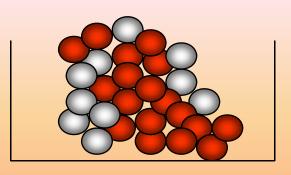
$$= P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2 \cdots A_{n-1})$$

$$= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \cdots \cdot \frac{n}{n+1}$$

$$= \frac{1}{n+1}$$

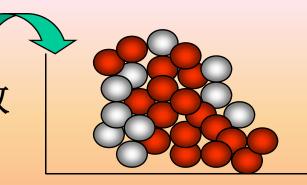
乘法公式应用举例

(波里亚罐子模型)



t个白球,r个红球

一个罐子中包含t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a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a个与所抽出的球具有相同颜色的球.



t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 $P(W_1W_2R_3R_4)$

 $=P(W_1)P(W_2|W_1)P(R_3|W_1W_2)P(R_4|W_1W_2R_3)$

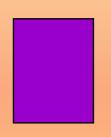
$$=\frac{t}{t+r}\frac{t+a}{t+r+a}\frac{r}{t+r+2a}\frac{r+a}{t+r+3a}$$

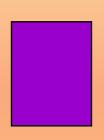
当 *a*>0 时,由于每次取出球后会增加下一次也取到同色球的概率.这是一个传染病模型.每次发现一个传染病患者,都会增加再传染的概率.

一场精彩的足球赛将要举行,

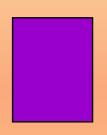
5个球迷好不容易才搞到一张入场券

大家都想去,只好用抽签的方法来解决





入场 券



5张同样的卡片,只有一张上写有"入场券",其余的什么也没写.将它们放在一起,洗匀,让5个人依次抽取.

"先抽的人当然要比后抽的人抽到的机会大."

后抽比先抽的确实吃亏吗?

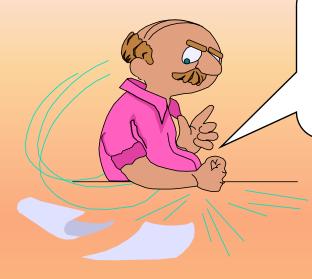
我们用 A_i 表示"第i个人抽到入场券" i=1,2,3,4,5.

则 Ā表示"第i个人未抽到入场券"

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

也就是说,

第1个人抽到入场券的概率是1/5.



"大家不必争先恐后,你们一个一个 按次序来,谁抽到'入场券'的机会都 一样大."

到底谁说的对呢?让我们用概率论的知识来计算一下,每个人抽到"入场券"的概率到底有多大?

"先抽的人当然要比后抽的人抽到的机会大。"

由于
$$A_2 = \overline{A_1}A_2$$

由乘法公式

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

因为若第2个人抽到 了入场券,第1个人 肯定没抽到.

也就是要想第2个人抽到入场券,必须第1 个人未抽到,

计算得:

$$P(A_2) = (4/5)(1/4) = 1/5$$

同理,第3个人要抽到"入场券",必须 第1、第2个人都没有抽到. 因此

$$P(A_3) = P(\overline{A_1}\overline{A_2}A_3) = P(\overline{A_1})P(\overline{A_2} \mid \overline{A_1})P(A_3 \mid \overline{A_1}\overline{A_2})$$

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

继续做下去就会发现,每个人抽到"入场券"的概率都是1/5.

这就是有关抽签顺序问题的正确解答。也就是说, 抽签不必争先恐后。

例 4 设某光学仪器厂制造的透镜,第一次落下时打破的概率为 1/2 ,若第一次落下未打破,第二次落下打破的概率为 7/10 ,若前两次落下未打破,第三次落下打破的概率为 9/10 。求透镜落下三次而未打破的概率。

解:以 A_i (i=1,2,3)表示事件"透镜第i次落下打破",以B表示事件"透镜落下三次而未打破",有:

$$P(B) = P(\overline{A}_1 \overline{A}_2 \overline{A}_3) = P(\overline{A}_3 | \overline{A}_1 \overline{A}_2) P(\overline{A}_2 | \overline{A}_1) P(\overline{A}_1)$$

$$= (1 - \frac{9}{10})(1 - \frac{7}{10})(1 - \frac{1}{2}) = \frac{3}{200}.$$

注意P(AB)与 $P(A \mid B)$ 的区别!

请看下面的例子

例5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 $}$ 300个 $A={$ 是标准件 $}$ 乙厂生产

111个
非标准件189个是
标准件

甲生产 700 个 设 $B={$ 零件是乙厂生产 $}$

 $A=\{是标准件\}$

300个

111个 非标准件 **189个**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?" 甲生产 700 个

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.

例6 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

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

条件概率P(A|B)与P(A)的区别

每一个随机试验都是在一定条件下进行的,设A是随机试验的一个事件,则P(A)是在该试验条件下事件A发生的可能性大小.

而条件概率P(A|B)是在原条件下又添加"B发生"这个条件时A发生的可能性大小。虽然P(A|B)仍是概率,但样本空间已经改变.

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

条件概率P(A|B)是在原条件下又添加"B发生"这个条件时A发生的可能性大小.那么,是否一定有:

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

或 $P(A|B) \leq P(A)$?

请思考!!

条件概率总结

• 条件概率在原样本空间当中是否是某个事件的概率?

答:不是,它是一个整体概念。由于样本空间已发生改变,所以它不再是原样本空间当中的事件发生的概率。

如何判断是否是条件概率?

- 1.明确给出了在A发生的前提或条件下,求事件B的概率
- 2.经过分析已知事件A发生,求事件B的概率
- 3.后验概率:事情已经发生了,让你找原因

条件概率如何获得?

- 1.从原题当中直接给出
- 2.使用公式,或者叫定义法求出
- 3.改变样本空间,用其它方法求出。

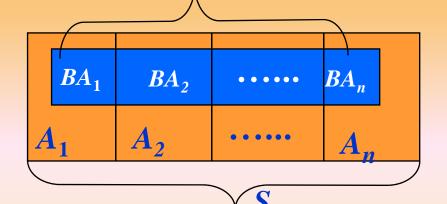
全概率公式和贝叶斯公式

定义 设S为试验E的样本空间, $A_1, A_2, \cdots A_n$ 为E的一组事件。若满足

(1)
$$A_i A_j = \emptyset, i \neq j, i, j = 1, 2, \dots, n;$$

$$(2) A_1 \cup A_2 \cup \cdots \cup A_n = S. B = BA_1 \cup BA_2 \cdots \cup BA_n$$

则称为 $A_1, A_2, \cdots A_n$ 样本空间 S 的一个划分。





全概率公式:

§3条件概率

设随机事件 A_1 , A_2 , …, A_n … 以及 B 满足:

$$(1)$$
. A_1 , A_2 , …, A_n … 两两互不相容;

$$(2). \bigcup_{n=1}^{\infty} A_n = S \quad \coprod \quad B \subset \bigcup_{n=1}^{\infty} A_n ;$$

(3).
$$P(A_n) > 0 \ (n = 1, 2, \dots)$$

则有
$$P(B) = \sum_{n=1}^{\infty} P(A_n) P(B|A_n)$$

全概率公式的证明

§3条件概率

由条件:

$$B \subset \bigcup_{n=1}^{\infty} A_n$$

得

$$B = \bigcup_{n=1}^{\infty} (A_n B)$$

 $B = BA \cup BA \cup BA$ $BA_1 \quad BA_2 \quad \cdots \quad BA_n$ $A_1 \quad A_2 \quad \cdots \quad A_n$

而且由 A_1 , A_2 , …, A_n … 两两互不相容,

得 A_1B , A_2B , …, A_nB … 也两两互不相容;

全概率公式的证明(续)

§3条件概率

所以由概率的可列可加性,得

$$P(B) = \sum_{n=1}^{\infty} P(A_n B)$$

再由条件 $P(A_n) > 0$ ($n = 1, 2, \dots$),得

$$P(A_n B) = P(A_n)P(B|A_n)$$

代入公式(1),得

$$P(B) = \sum_{n=1}^{\infty} P(A_n B) = \sum_{n=1}^{\infty} P(A_n) P(B|A_n)$$



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

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

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.

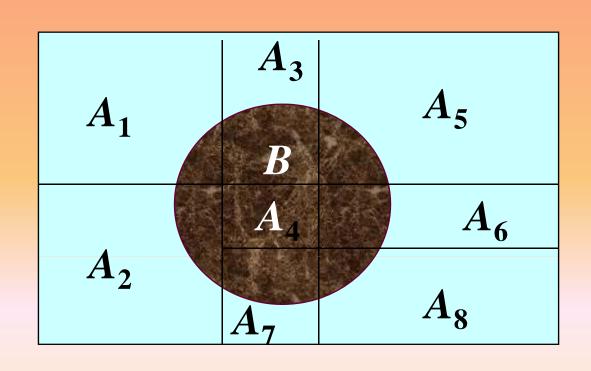
某一事件B的发生有各种可能的原因 (i=1,2,...,n),如果B是由原因 A_i 所引起,则 B发生的概率是

$$P(BA_i)=P(A_i)P(B \mid A_i)$$

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.



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 A_i 是原因 B是结果

全概率公式的使用

§3条件概率

我们把事件B看作某一过程的结果,

把 A_1 , A_2 , …, A_n …看作该过程的若干个原因,

根据历史资料,每一原因发生的概率已知,

(即
$$P(A_n)$$
已知)

而且每一原因对结果的影响程度已知,

(即
$$P(B|A_n)$$
已知)

则我们可用全概率公式计算结果发生的概率.



§3条件概率

例7 某小组有20名射手,其中一、二、三、四级射手人数分别为2、6、9、3. 又若选一、二、三、四级射手参加比赛,则在比赛中射中目标的概率分别为0.85、0.64、0.45、0.32,今随机选一人参加比赛,试求该小组在比赛中射中目标的概率.

解:

设 $B = \{$ 该小组在比赛中射中目标 $\}$ $A_i = \{$ 选i级射手参加比赛 $\}(i=1, 2, 3 4)$ 由全概率公式,有



$$P(B) = \sum_{n=1}^{4} P(A_n) P(B|A_n)$$

$$= \frac{2}{20} \times 0.85 + \frac{6}{20} \times 0.64 + \frac{9}{20} \times 0.45 + \frac{3}{20} \times 0.32$$

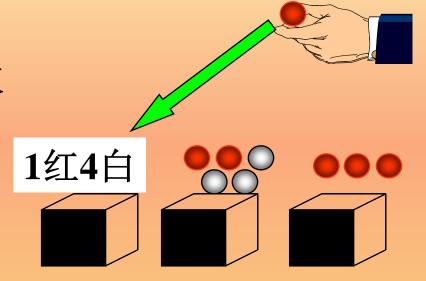
$$= 0.5275$$

实际中还有下面一类问题,是"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.

接下来我们介绍为解决这类问题而引出的

贝叶斯公式

Bayes 公式

§3条件概率

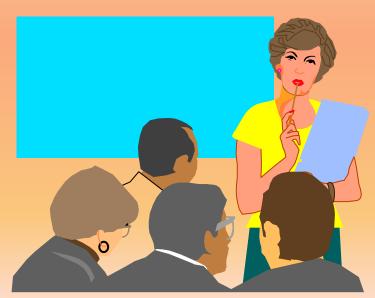
设随机事件 A_1 , A_2 , …, A_n … 以及B 满足 (1). A_1 , A_2 , …, A_n … 两两互不相容; (2). $\bigcup_{n=0}^{\infty} A_n = S$ 且 $B \subset \bigcup_{n=0}^{\infty} A_n$;

$$(2)$$
. $\bigcup_{n=1}^{n=1} A_n = S$ 出. $B \subset \bigcup_{n=1}^{n=1} A_n$; (3) . $P(A_n) > 0$ $(n=1, 2, \cdots)$

$$P(A_n \mid B) = \frac{P(A_n B)}{P(B)} = \frac{P(B \mid A_n)P(A_n)}{\infty}, \quad n = 1, 2, \cdots, \infty$$

$$\sum_{j=1}^{\infty} P(B \mid A_j)P(A_j)$$
公 返回主目录

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.



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,所以也被称为逆概公式。

Bayes公式的使用

§3条件概率

我们把事件B看作某一过程的结果,

把 A_1 , A_2 , …, A_n …看作该过程的若干个原因,

根据历史资料,每一原因发生的概率已知,

(即
$$P(A_n)$$
已知)

而且每一原因对结果的影响程度已知,

(即
$$P(B|A_n)$$
已知)

如果已知事件B已经发生,要求此时是由第i个原 因引起的概率,则用Bayes公式

(即求
$$P(A_i|B)$$
)



例 8 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 $C=\{$ 抽查的人患有癌症 $\}$, $A=\{$ 试验结果是阳性 $\}$,

则 \overline{C} 表示"抽查的人不患癌症" 已知 $P(C)=0.005, P(\overline{C})=0.995,$ $P(A|C)=0.95, P(A|\overline{C})=0.04$



求P(C|A).

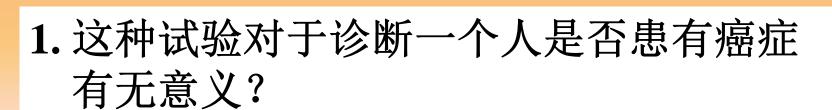
由贝叶斯公式,可得

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

代入数据计算得:

 $P(C \mid A) = 0.1066$

现在来分析一下结果的意义.



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无意义?

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 $P(C \mid A) = 0.1066$

从0.005增加到0.1066,将近增加约21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 $P(C \mid A)=0.1066$

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.

例 9

§3条件概率

袋中有10个黑球,5个白球.现掷一枚均匀的骰子, 掷出几点就从袋中取出几个球. 若已知取出的球 全是白球, 求掷出3点的概率.

解:

设: B={ 取出的球全是白球 }

$$A_i = \{$$
掷出 i 点 $\}$ $(i=1, 2, \dots, 6)$

则由Bayes公式,得

$$P(A_{3}|B) = \frac{P(A_{3})P(B|A_{3})}{\sum_{i=1}^{6} P(A_{i})P(B|A_{i})}$$



例9(续)

=0.04835

$$= \frac{\frac{1}{6} \times \frac{C_5^3}{C_{15}^3}}{\sum_{i=1}^5 \frac{1}{6} \times \frac{C_5^i}{C_{15}^i} + \frac{1}{6} \times 0}$$

§3条件概率

例 10 某电子设备制造厂所用的晶体管是由三家元件厂提供的。根据以往的记录有以下的数据。

元件制造厂	次品率	提供晶体管的份额
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05
		$egin{array}{cccccccccccccccccccccccccccccccccccc$

例10 (续)

§3条件概率

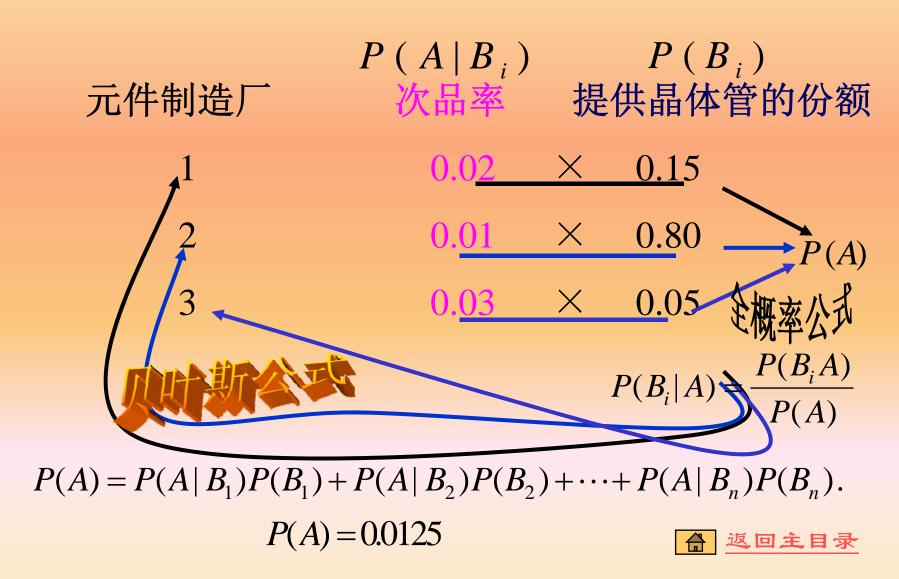
设这三家工厂的产品在仓库中是均匀混合的,且无区别的标志。

- (1) 在仓库中随机的取一只晶体管, 求它是次品的概率。
- (2) 在仓库中随机的取一只晶体管,若已知取到的是次品试分析此次品出自那家工厂的可能性最大。

解: 设A表示"取到的是一只次品", B_i (i=1,2,3)表示"取到的产品是由第i家工厂提供的",

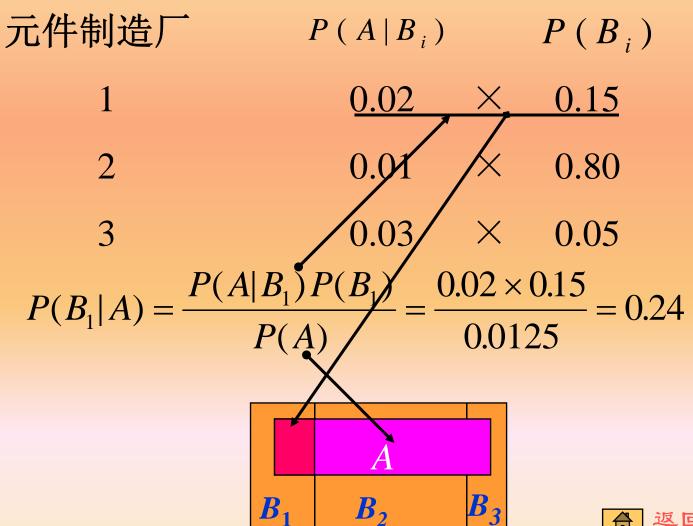


例10 (续)



第一章 概率论的基本概念

例10(续)



例10(续)

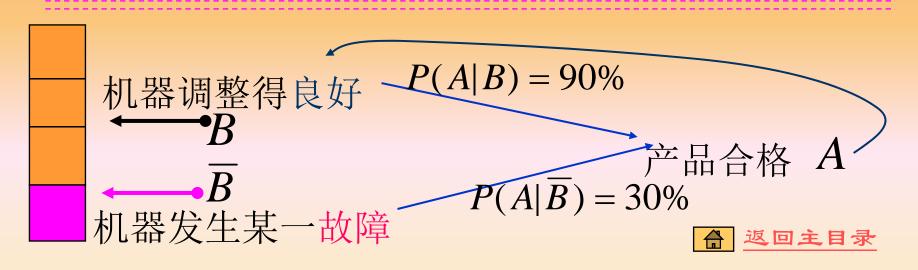
$$P(B_1 \mid A) = \frac{3}{12.5} = 24\%$$
,

$$P(B_2 \mid A) = \frac{8}{12.5} = 64\%$$
,

$$P(B_3 \mid A) = \frac{1.5}{12.5} = 12\%$$
.

§3条件概率

例 11 对以往的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%。已知某天早上第一件产品是合格品,试求机器调整得良好的概率是多少?



第一章 概率论的基本概念

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

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

下面我们再回过头来看一下贝叶斯公式



贝叶斯公式
$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_i)P(B \mid A_i)}$$

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

 $P(A_i)(i=1,2,...,n)$ 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。

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.值得一提的是,后来的学者依据贝叶斯公式的思想发展了一整套统计推断方法,叫作"贝叶斯统计".可见贝叶斯公式的影响。