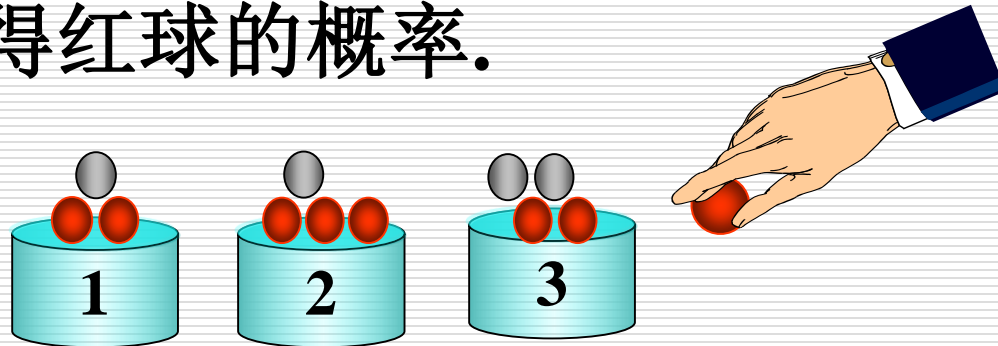


全概率公式和贝叶斯公式

1.5.1 全概率公式

引例：有三个罐子,1号装有 2 红 1 黑球, 2号装有 3 红 1 黑球, 3号装有 2 红 2 黑球. 某人从中**随机取一罐**, 在从中**任意取出一球**, 求取得红球的概率.



如何求取得红球的概率？ ？ ？

1.5.1 全概率公式

定理1.2 设试验 E 的样本空间为 Ω , A_1, A_2, \dots, A_n 为 E 的一组事件, 且满足:

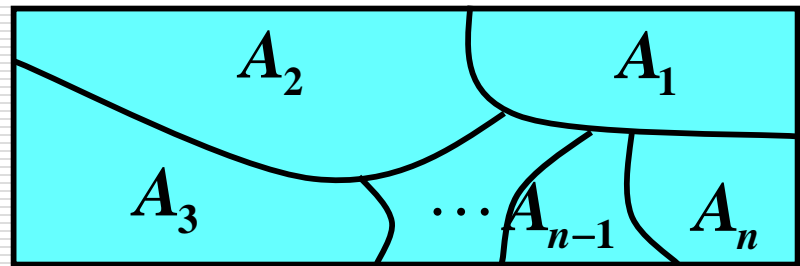
(1) A_1, A_2, \dots, A_n 两两互不相容, $P(A_i) > 0, i = 1, 2, \dots, n$;

(2) $\bigcup_{i=1}^n A_i = \Omega$

则对任一事件 B , 有 $P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$ (1.7)

(1.7)称为**全概率公式**.

称满足(1)和(2)的 A_1, A_2, \dots, A_n 为**完备事件组**或**样本空间的一个划分**.



证明

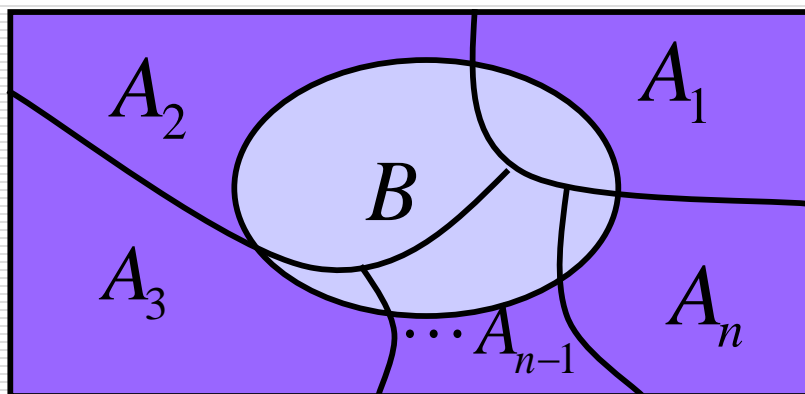
$$\begin{aligned} B &= B\Omega = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) \\ &= BA_1 \cup BA_2 \cup \cdots \cup BA_n. \end{aligned}$$

由 $A_i A_j = \emptyset \Rightarrow (BA_i)(BA_j) = \emptyset$

$$\Rightarrow P(B) = P(BA_1) + P(BA_2) + \cdots + P(BA_n)$$

$$\begin{aligned} \Rightarrow P(B) &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots \\ &+ P(A_n)P(B|A_n) \end{aligned}$$

图示



类似于曲
边梯形求
面积化整
为零各个
击破

全概率公式的使用

我们把事件B看作某一过程的结果，

把 $A_1, A_2, \dots, A_n \dots$ 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)

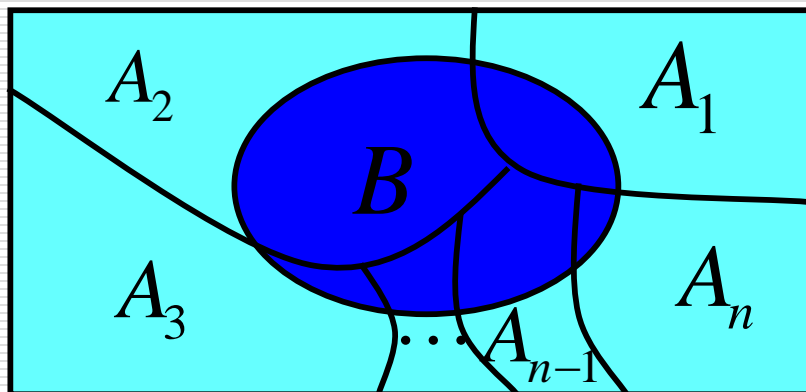
而且每一原因对结果的影响程度已知，

(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。

(即求 $P(B)$)

说明 事实上全概率公式的主要用途在于它可以将一个复杂事件的概率计算问题,分解为若干个简单事件的概率计算问题,最后应用概率的可加性求出最终结果.



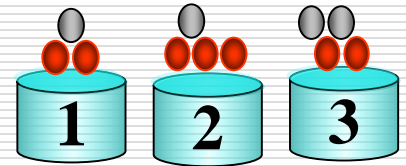
1.5.1 全概率公式

再看引例 1

有三个罐子,1号装有 2 红 1 黑球, 2号装有 3 红 1 黑球, 3号装有 2 红 2 黑球. 某人从中随机取一罐, 再从中任意取出一球, 求取得红球的概率.

解 记 $A_i = \{ \text{取到的是 } i \text{ 号罐} \} \ i=1, 2, 3; B = \{ \text{取得红球} \}$

A_1, A_2, A_3 的发生都会导致 B 发生,
 A_1, A_2, A_3 构成完备事件组.



由全概率公式得 $P(B) = \sum_{i=1}^3 P(A_i)P(B|A_i)$

依题意: $P(B|A_1)=2/3, \quad P(B|A_2)=3/4, \quad P(B|A_3)=1/2,$

$$P(A_i)=1/3, i=1, 2, 3$$

代入数据计算得: $P(B) \approx 0.639.$

1.5.1 全概率公式

【例1.15】 假设有3箱同种型号零件，里面分别装有50件、30件、40件，而且一等品分别有20件、12件和24件，现在任取一箱，从中不放回地先后取出两个零件，试求：

(1)先取出的零件是一等品的概率；

(2)两次取出的零件均为一等品的概率.

解：设 A_i = “任取的一箱为第 i 箱零件”， $i = 1, 2, 3$,

B_j = “第 j 次取到的是一等品”， $j = 1, 2$.

由题意知 A_1 、 A_2 和 A_3 构成完备事件组，

且 $P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$

1.5.1 全概率公式

$$(1) \quad P(B_1 | A_1) = \frac{20}{50} = 0.4, \quad P(B_1 | A_2) = \frac{12}{30} = 0.4$$

$$P(B_1 | A_3) = \frac{24}{40} = 0.6$$

由全概率公式得

$$\begin{aligned} P(B_1) &= \sum_{i=1}^3 P(A_i)P(B_1|A_i) \\ &= \frac{1}{3}(0.4 + 0.4 + 0.6) = 0.467. \end{aligned}$$

1.5.1 全概率公式

(2) 因为

$$P(B_1B_2 | A_1) = \frac{C_{20}^2}{C_{50}^2} = 0.1551$$

$$P(B_1B_2 | A_2) = \frac{C_{12}^2}{C_{30}^2} = 0.1517$$

$$P(B_1B_2 | A_3) = \frac{C_{24}^2}{C_{40}^2} = 0.3538$$

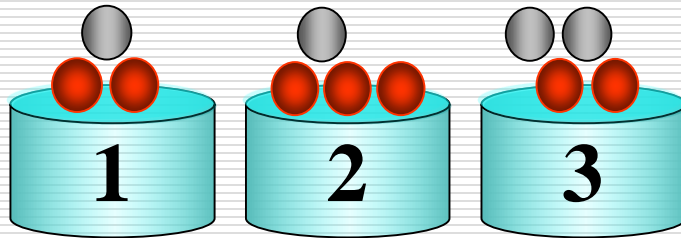
由全概率公式得

$$\begin{aligned} P(B_1B_2) &= \sum_{i=1}^3 P(A_i)P(B_1B_2|A_i) \\ &= \frac{1}{3}(0.1551 + 0.1517 + 0.3538) = 0.22 \end{aligned}$$

1.5.1 全概率公式

引例 2：

某人从任一罐中任意摸出一球，发现是红球，求该球是取自 1 号罐的概率。



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。

下面就介绍为解决这类问题而引出的公式：

Bayes(贝叶斯)公式

1.5 全概率公式和贝叶斯公式

1.5.2 贝叶斯公式

定理1.3 设试验 E 的样本空间为 Ω ， B 为 E 的事件， A_1, A_2, \dots, A_n 为完备事件组，且 $P(B) > 0$ ， $P(A_i) > 0, i = 1, 2, \dots, n$ ，则

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}, i = 1, 2, \dots, n \quad (1.8)$$

(1.8)式称为**贝叶斯公式**.

1.5.2 贝叶斯公式

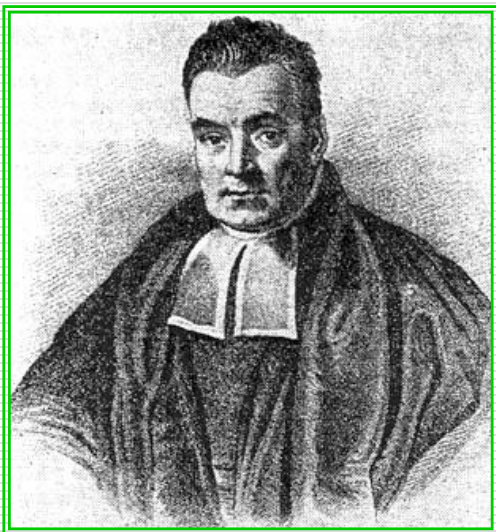
证明 由条件概率公式、乘法公式及全概率公式知:

$$\begin{aligned} P(B|A_i) &= \frac{P(BA_i)}{P(B)} \\ &= \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}, \quad i = 1, 2, \dots, n. \end{aligned}$$

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.

1.5.2 贝叶斯公式

贝叶斯公式是英国哲学家Bayes于1763首先提出的，经过多年的发展和完善，由这一公式的思想已经发展成为一整套统计推断方法，即“Bayes方法”，这一方法在计算机诊断、模式识别、基因组成、蛋白质结构等很多方面都有应用。



Thomas Bayes

**Born: 1702 in London,
England**

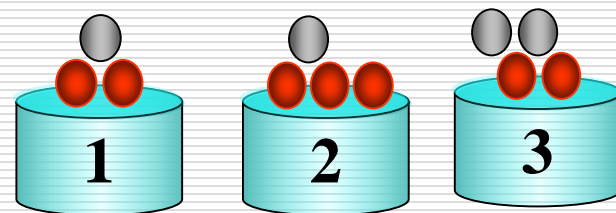
**Died: 17 Apr. 1761 in
Tunbridge Wells, Kent,
England**

1.5.2 贝叶斯公式

再看引例 2

某人从任一罐中任意摸出一球，发现是红球，求该球是取自 1 号罐的概率。

解 记 $A_i = \{ \text{取到第 } i \text{ 号罐} \} \quad i=1, 2, 3;$
 $B = \{ \text{取得红球} \}$



A_1, A_2, A_3 是完备事件组。

由贝叶斯公式得
$$P(A_1 | B) = \frac{P(B | A_1)P(A_1)}{\sum_{i=1}^3 P(A_i)P(B | A_i)}$$

其中 $P(B | A_1) = 2/3, \quad P(B | A_2) = 3/4,$
 $P(B | A_3) = 1/2, \quad P(A_i) = 1/3, \quad i=1, 2, 3.$

代入数据计算得: $P(A_1 | B) \approx 0.348$

1.5.2 贝叶斯公式

特别有：

设事件 A 、 B 为试验 E 的两事件，由于 A 和 \bar{A} 是一个完备事件组，若 $P(A) > 0$ ， $P(\bar{A}) > 0$ ， $P(B) > 0$ ，贝叶斯公式的一种常用简单形式为

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

在应用全概率公式与贝叶斯公式时，有两个问题需要弄清楚：

1、如何确定完备事件组

一般，可从下列两个方面来寻找完备事件组：

当事件的发生与相继两个试验有关时，从**第一试验入手**寻找完备事件组；

当事件的发生是由诸多**两两互斥**的原因而引起的，可以这些“原因”为完备事件组。

2、如何区分是用全概率公式还是用贝叶斯公式

“**由因求果**”用全概率公式，“**执果求因**”用贝叶斯公式。

例 某电子设备制造厂所用的元件是由三家元件制造厂提供的.根据以往的记录有以下的数据:

元件制造厂	次品率	提供元件的份额
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05

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) 在仓库中随机地取一只元件,求它是次品的概率;

(2) 在仓库中随机地取一只元件,若已知取到的是次品,为分析此次品出自何厂,需求出此次品由三家工厂生产的概率分别是多少.试求这些概率.

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

则 B_1, B_2, B_3 是样本空间 S 的一个划分,

且 $P(B_1) = 0.15$, $P(B_2) = 0.80$, $P(B_3) = 0.05$,

$$P(A|B_1) = 0.02, \quad P(A|B_2) = 0.01, \quad P(A|B_3) = 0.03.$$

(1) 由全概率公式得

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\ &= 0.0125. \end{aligned}$$

(2) 由贝叶斯公式得

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{0.02 \times 0.15}{0.0125} = 0.24.$$

$$P(B_2|A) = \frac{P(A|B_2)P(B_2)}{P(A)} = 0.64,$$

$$P(B_3|A) = \frac{P(A|B_3)P(B_3)}{P(A)} = 0.12.$$

故这只次品来自第 2 家工厂的可能性最大.

【例4】 设在12只乒乓球中有9只新球和3只旧球, 第一次比赛取出3只, 用后放回去; 第二次比赛又取出3只, 求第二次取到的3只球中有2只为新球的概率.

【解】 这里有两个相继“试验”：“第一次取出3只”和

“第二次取出3只”. 因此, 可根据“第一次试验”的各种情

形确定完备事件组. 第一次取出3只球有4种情况: 没有新球、有一只新球、有两只新球和全是新球, 分别用事件表示为:

$$B_0, B_1, B_2, B_3$$

设A为事件: “第二次取出2新1旧”, 则由古典概率计

算公式[超几何分布]得:

$$P(B_0) = \frac{C_3^3}{C_{12}^3} = \frac{1}{220} \quad [\text{从9新3旧中取3旧}]$$

$$P(B_1) = \frac{C_9^1 C_3^2}{C_{12}^3} = \frac{27}{220} \quad [\text{从9新3旧中取1新2旧}]$$

$$P(B_2) = \frac{C_9^2 C_3^1}{C_{12}^3} = \frac{108}{220} \quad [\text{从9新3旧中取2新1旧}]$$

$$P(B_3) = \frac{C_9^3}{C_{12}^3} = \frac{84}{220} \quad [\text{从9新3旧中取3新}]$$

注意：第二次取球时12只球的新旧组成是随第一次取出的3球组成的变化而变化,易得:

$$P(A | B_0) = \frac{C_9^2 C_3^1}{C_{12}^3} = \frac{108}{220}$$

[从9新3旧中取2新1旧]

$$P(A | B_1) = \frac{C_8^2 C_4^1}{C_{12}^3} = \frac{112}{220}$$

[从8新4旧中取2新1旧]

$$P(A | B_2) = \frac{C_7^2 C_5^1}{C_{12}^3} = \frac{105}{220}$$

[从7新5旧中取2新1旧]

$$P(A | B_3) = \frac{C_6^2 C_6^1}{C_{12}^3} = \frac{90}{220}$$

[从6新6旧中取2新1旧]

由全概率公式得:

$$P(A) = \sum_{k=0}^3 P(B_k) P(A | B_k)$$

由全概率公式得:

$$P(A) = \sum_{k=0}^3 P(B_k)P(A|B_k)$$

$$= \frac{1}{220} \cdot \frac{108}{220} + \frac{27}{220} \cdot \frac{112}{220} + \frac{108}{220} \cdot \frac{105}{220} + \frac{84}{220} \cdot \frac{90}{220}$$

$$= \frac{22032}{48400} = \frac{1377}{3025} \quad \blacksquare$$

$$\approx 0.45520661157$$

1.5.2 贝叶斯公式

【例1.16】 玻璃杯成箱出售，每箱20只，假设各箱含0，1，2只残次品的概率分别是0.8，0.1和0.1，某顾客欲购一箱玻璃杯，在购买时，售货员随即取出一箱，顾客开箱随机地查看四只，若无残次品，则买下该箱玻璃杯，否则退回，试求：

- (1) 顾客买下该箱的概率 α ;
- (2) 在顾客买下的一箱中，确实没有残次品的概率 β .

1.5.2 贝叶斯公式

解：设 B = “顾客买下该箱玻璃杯”，

A_i = “抽到的一箱中有 i 件残次品”， $i = 0, 1, 2$.

(1) 事件 B 在下面三种情况下均会发生：抽到的一箱中没有残次品、有1件残次品或有2件次品。

显然 A_0, A_1, A_2 是完备事件组， $P(A_0) = 0.8, P(A_1) = 0.1, P(A_2) = 0.1$

由题意知 $P(B|A_0) = \frac{C_{20}^4}{C_{20}^4} = 1, P(B|A_1) = \frac{C_{19}^4}{C_{20}^4} = \frac{4}{5}, P(B|A_2) = \frac{C_{18}^4}{C_{20}^4} = \frac{12}{19}$

由全概率公式得

$$P(B) = P(A_0)P(B|A_0) + P(A_1)P(B|A_1) + P(A_2)P(B|A_2) = 0.94$$

1.5.2 贝叶斯公式

(2) 由贝叶斯公式

$$\beta = P(A_0|B) = \frac{P(A_0)P(B|A_0)}{P(B)}$$

$$= \frac{0.8 \times 1}{0.94} = 0.85$$

1.5.2 贝叶斯公式

【例1.17】 根据以往的记录，某种诊断肝炎的试验有如下效果：对肝炎病人的试验呈阳性的概率为0.95；非肝炎病人的试验呈阴性的概率为0.95。对自然人群进行普查的结果为：有千分之五的人患有肝炎。现有某人做此试验结果为阳性，问此人确有肝炎的概率为多少？

1.5.2 贝叶斯公式

解：设 A = “某人确有肝炎”，

B = “某人做此试验结果为阳性”；

由已知条件有 $P(B|A) = 0.95$

$$P(\bar{B}|\bar{A}) = 0.95, \quad P(A) = 0.005$$

从而 $P(\bar{A}) = 1 - P(A) = 0.995$

$$P(B|\bar{A}) = 1 - P(\bar{B}|\bar{A}) = 0.05$$

由贝叶斯公式，有 $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\bar{A})P(B|\bar{A})}$

$$= \frac{0.005 \times 0.95}{0.005 \times 0.95 + 0.995 \times 0.05} = 0.087$$

1.5.2 贝叶斯公式

本题的结果表明, 虽然 $P(B|A) = 0.95$,
 $P(\bar{B}|\bar{A}) = 0.95$ 这两个概率都很高. 但是, 即试验
阳性的人有肝炎的概率只有8.7%.

$$P(B|A) \quad P(A|B)$$

■ 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。

$P(A_i)$ ($i=1,2,\dots,n$) 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

伊索寓言“孩子与狼”讲的是一个小孩每天到山上放羊，山里有狼出没。第一天，他在山上喊“狼来了！狼来了！”，山下的村民闻声便去打狼，发现狼没有来；第二天仍是如此；第三天，狼真的来了，可无论小孩怎么喊叫，也没有人来救他，因为前两次他说了谎，人们不再相信他了。

现在用贝叶斯公式来分析此寓言中村民对这个小孩的可信程度是如何下降的。

首先记事件A为“小孩说谎”，记事件B为“小孩可信”。不妨设村民过去对这个小孩的印象为

$$P(B)=0.8, P(\bar{B})=0.2$$

我们现在用贝叶斯公式来求 $P(B|A)$ ，亦及这个小孩说了一次谎后，村民对他的可信程度的改变。

在贝叶斯公式中我们要用到 $P(A|B)$ 和 $P(A|\bar{B})$ ，这两个概率的含义是：前者为“可信”(B)的孩子“说谎”(A)的可能性，后者为“不可信”的孩子“说谎”的可能性。设：

$$P(A|B)=0.1, P(A|\bar{B})=0.5$$

第一次村民上山打狼，发现狼没有来，即小孩说了谎（A）。村民根据这个信息，对小孩的可信程度改变为（用贝叶斯公式）

$$\begin{aligned}P(B|A) &= \frac{P(B)P(A|B)}{P(B)P(A|B) + P(\bar{B})P(A|\bar{B})} \\&= \frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.2 \times 0.5} = 0.444\end{aligned}$$

这表明村民上了一次当后，对这个小孩的可信程度由原来的0.8调整为0.444，也就是

$$P(B) = 0.444, P(\bar{B}) = 0.556$$

在此基础上，我们再用一次贝叶斯公式计算 $P(B|A)$ 亦即这个小孩第二次说谎后，村民对他的可信程度改变为：

$$P(B|A) = \frac{0.444 \times 0.1}{0.444 \times 0.1 + 0.556 \times 0.5} = 0.138$$

这表明村民们经过两次上当，对这个小孩的可信程度已经从0.8下降到0.138，如此低的可信程度，村民听到第三次呼叫怎么再会上山打狼呢？

课堂练习

有一台用来检验产品质量的仪器，已知一只次品经检验被认为是次品的概率为0.99，而一只正品经检验被认为是次品的概率0.005，已知产品的次品率为4%，若一产品经检验被认为是次品，求它确为次品的概率。

解 设 $A = \{\text{产品经检验被认为是次品}\}$
 $B = \{\text{产品确为次品}\}$

1.5.2 贝叶斯公式

由题设知

$$P(B) = 0.04, \quad P(\bar{B}) = 0.96,$$

$$P(A | B) = 0.99, \quad P(A | \bar{B}) = 0.005,$$

由贝叶斯公式，所求概率为

$$\begin{aligned} P(B | A) &= \frac{P(A | B)P(B)}{P(A | B)P(B) + P(A | \bar{B})P(\bar{B})} \\ &= \frac{0.99 \times 0.04}{0.99 \times 0.04 + 0.005 \times 0.96} = 0.892 \end{aligned}$$
