

概率论与数理统计

第一章 概率论的基本概念

练习:

两船欲停靠同一个码头, 设两船到达码头的时间内各不相干, 而且到达码头的时间内在一昼夜内是等可能的. 如果两船到达码头后需在码头停留的时间分别是1 小时与2 小时, 试求在一昼夜内, 任一船到达时, 需要等待空出码头的概率.

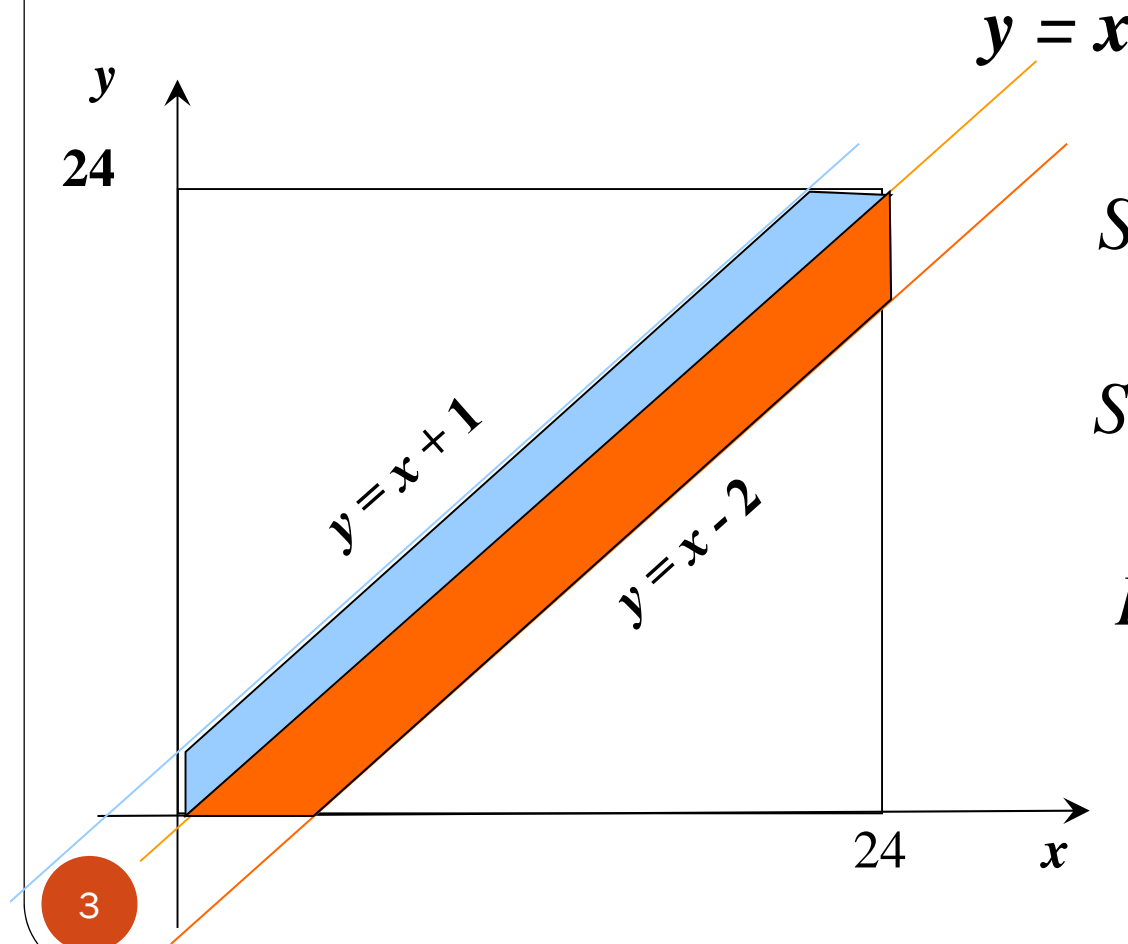
解 设船1 到达码头的瞬时为 x , $0 \leq x < 24$

船2 到达码头的瞬时为 y , $0 \leq y < 24$

设事件 A 表示任一船到达码头时需要等待空出码头

$$S = \{(x, y) \mid 0 \leq x < 24, 0 \leq y < 24\}$$

$$A = \{(x, y) \mid (x, y) \in S, \\ 0 \leq y - x \leq 1, 0 \leq x - y \leq 2\}$$



$$S_S = 24^2$$

$$S_{\bar{A}} = \frac{1}{2} (23^2 + 22^2)$$

$$P(A) = 1 - \frac{S_{\bar{A}}}{S_S} = 0.1207$$

§ 5 条件概率

- ◆ 条件概率
- ◆ 乘法定理
- ◆ 全概率公式与贝叶斯公式

一、条件概率

1. 引例 将一枚硬币抛掷两次,观察其出现正反两面的情况,设事件 A 为“至少有一次为正面”,事件 B 为“两次掷出同一面”.现在来求已知事件 A 已经发生的条件下事件 B 发生的概率.

分析 设 H 为正面, T 为反面. $S = \{ HH, HT, TH, TT \}$.

$$A = \{ HH, HT, TH \}, \quad B = \{ HH, TT \}, \quad P(B) = \frac{2}{4} = \frac{1}{2}.$$

事件 A 已经发生的条件下事件 B 发生的概率,记为

$$P(B|A), \quad \text{则} \quad P(B|A) = \frac{1}{3} = \frac{1/4}{3/4} = \frac{P(AB)}{P(A)} \neq P(B).$$

2. 定义

设 A, B 是两个事件, 且 $P(A) > 0$, 称

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

为在事件 A 发生的条件下事件 B 发生的条件概率.

同理可得
$$P(A|B) = \frac{P(AB)}{P(B)}$$

为事件 B 发生的条件下事件 A 发生的条件概率.

3. 性质

(1) 非负性 : $P(B|A) \geq 0$;

(2) 规范性 : $P(S|B) = 1, P(\emptyset|B) = 0$;

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

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

(5) 可列可加性 : 设 B_1, B_2, \dots 是两两不相容的事件, 则有

$$P\left(\bigcup_{i=1}^{\infty} B_i \middle| A\right) = \sum_{i=1}^{\infty} P(B_i|A).$$

二、乘法定理

设 $P(A) > 0$, 则有 $P(AB) = P(B|A)P(A)$.

设 A, B, C 为事件, 且 $P(AB) > 0$, 则有

$$P(ABC) = P(C|AB)P(B|A)P(A).$$

推广 设 A_1, A_2, \dots, A_n 为 n 个事件, $n \geq 2$,

且 $P(A_1 A_2 \cdots A_{n-1}) > 0$, 则有

$$P(A_1 A_2 \cdots A_n) = P(A_n | A_1 A_2 \cdots A_{n-1}) \times \\ P(A_{n-1} | A_1 A_2 \cdots A_{n-2}) \times \cdots \times P(A_2 | A_1) P(A_1).$$

例1 一盒子装有4只产品,其中有3只一等品、1只二等品.从中取产品两次,每次任取一只,作不放回抽样.设事件 A 为“第一次取到的是一等品”、事件 B 为“第二次取到的是一等品”.试求条件概率 $P(B|A)$.

解 将产品编号,1,2,3为一等品;4号为二等品.

以 (i, j) 表示第一次、第二次分别取到第 i 号、第 j 号产品,则试验的样本空间为

$$S = \{(1,2), (1,3), (1,4), (2,1), (2,3), (2,4), \dots, (4,1), (4,2), (4,3)\},$$

$$A = \{(1,2), (1,3), (1,4), (2,1), (2,3), (2,4), \\ (3,1), (3,2), (3,4)\},$$

$$AB = \{(1,2), (1,3), (2,1), (2,3), (3,1), (3,2)\},$$

由条件概率的公式得

$$\begin{aligned} P(B|A) &= \frac{P(AB)}{P(A)} \\ &= \frac{6/12}{9/12} = \frac{2}{3}. \end{aligned}$$

例2 某种动物由出生算起活**20**岁以上的概率为**0.8**, 活到**25**岁以上的概率为**0.4**, 如果现在有一个**20**岁的这种动物, 问它能活到**25**岁以上的概率是多少?

解 设 A 表示“能活 **20** 岁以上”的事件,
 B 表示“能活 **25** 岁以上”的事件,

则有
$$P(B|A) = \frac{P(AB)}{P(A)}.$$

因为 $P(A) = 0.8$, $P(B) = 0.4$, $P(AB) = P(B)$,

所以
$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{0.4}{0.8} = \frac{1}{2}.$$

抓阄是否与次序有关?

例3 五个阄, 其中两个阄内写着“有”字, 三个阄内不写字, 五人依次抓取, 问各人抓到“有”字阄的概率是否相同?



解 设 A_i 表示“第 i 人抓到有字阄”的事件,

$$i=1,2,3,4,5. \quad \text{则有 } P(A_1) = \frac{2}{5},$$

$$P(A_2) = P(A_2 S) = P(A_2 \cap (A_1 \cup \bar{A}_1))$$

$$= P(A_1 A_2 \cup \overline{A_1} A_2) = P(A_1 A_2) + P(\overline{A_1} A_2)$$

$$= P(A_1)P(A_2|A_1) + P(\overline{A_1})P(A_2|\overline{A_1})$$

$$= \frac{2}{5} \times \frac{1}{4} + \frac{3}{5} \times \frac{2}{4} = \frac{2}{5},$$

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

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

$$\begin{aligned}
&= P(A_1)P(\overline{A_2}|A_1)P(A_3|A_1\overline{A_2}) + P(\overline{A_1})P(A_2|\overline{A_1})P(A_3|\overline{A_1}A_2) \\
&\quad + P(\overline{A_1})P(\overline{A_2}|\overline{A_1})P(A_3|\overline{A_1}\overline{A_2}) \\
&= \frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{2}{5},
\end{aligned}$$

依此类推 $P(A_4) = P(A_5) = \frac{2}{5}$.

故抓阄与次序无关.

摸球试验

例4 设袋中装有 r 只红球、 t 只白球.每次自袋中任取一只球,观察其颜色然后放回,并再放入 a 只与所取出的那只球同色的球,若在袋中连续取球四次,试求第一、二次取到红球且第三、四次取到白球的概率.

解 设 A_i ($i = 1, 2, 3, 4$) 为事件“第 i 次取到红球”

则 $\overline{A_3}$ 、 $\overline{A_4}$ 为事件第三、四次取到白球.

因此所求概率为

$$\begin{aligned} & P(A_1 A_2 \overline{A_3} \overline{A_4}) \\ &= P(\overline{A_4} | A_1 A_2 \overline{A_3}) P(\overline{A_3} | A_1 A_2) P(A_2 | A_1) P(A_1) \\ &= \frac{t+a}{r+t+3a} \cdot \frac{t}{r+t+2a} \cdot \frac{r+a}{r+t+a} \cdot \frac{r}{r+t}. \end{aligned}$$

此模型被波利亚用来作为描述传染病的数学模型.

例5 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 $9/10$.试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”.

因为 $B = \overline{A_1} \overline{A_2} \overline{A_3}$,

$$\begin{aligned} \text{所以 } P(B) &= P(\overline{A_1} \overline{A_2} \overline{A_3}) = P(\overline{A_3} | \overline{A_1} \overline{A_2}) P(\overline{A_2} | \overline{A_1}) P(\overline{A_1}) \\ &= (1 - \frac{9}{10})(1 - \frac{7}{10})(1 - \frac{1}{2}) = \frac{3}{200}. \end{aligned}$$

三、全概率公式与贝叶斯公式

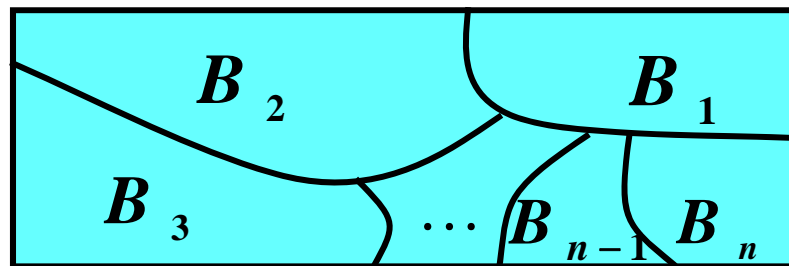
1. 样本空间的划分

定义 设 S 为试验 E 的样本空间, B_1, B_2, \dots, B_n 为 E 的一组事件, 若

$$(i) \quad B_i B_j = \emptyset, i \neq j, i, j = 1, 2, \dots, n;$$

$$(ii) \quad B_1 \cup B_2 \cup \dots \cup B_n = S.$$

则称 B_1, B_2, \dots, B_n 为样本空间 S 的一个划分.



2. 全概率公式

定理 设试验 E 的样本空间为 S , A 为 E 的事件, B_1, B_2, \dots, B_n 为 S 的一个划分, 且 $P(B_i) > 0 (i = 1, 2, \dots, n)$, 则

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_n)P(B_n)$$



全概率公式

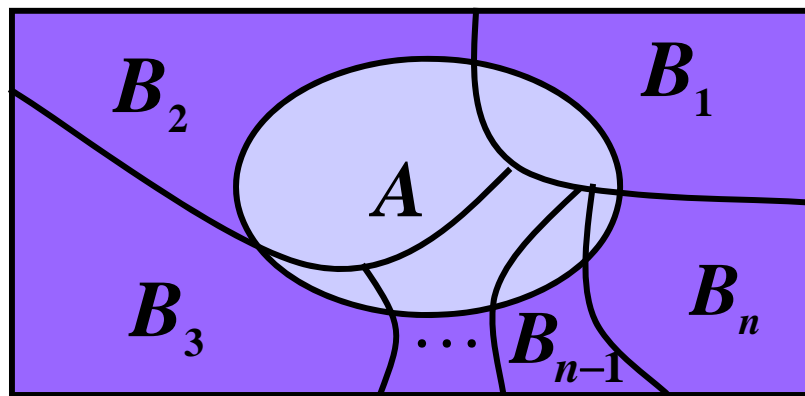
$$\begin{aligned}\text{证明 } A &= AS = A \cap (B_1 \cup B_2 \cup \cdots \cup B_n) \\ &= AB_1 \cup AB_2 \cup \cdots \cup AB_n.\end{aligned}$$

$$\text{由 } B_i B_j = \emptyset \Rightarrow (AB_i)(AB_j) = \emptyset$$

$$\Rightarrow P(A) = P(AB_1) + P(AB_2) + \cdots + P(AB_n)$$

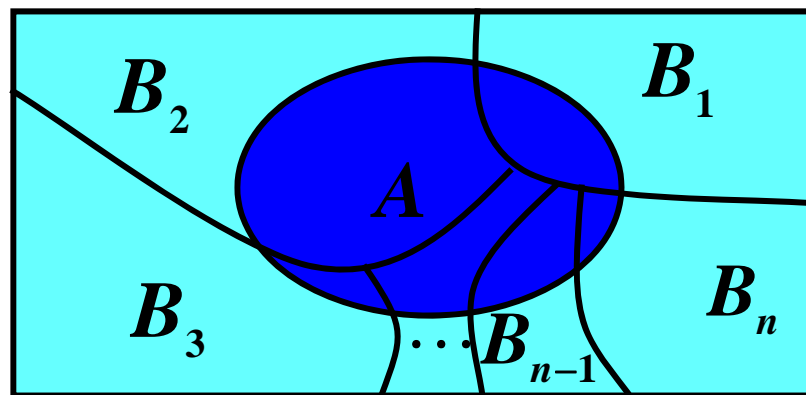
$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n).$$

图示



化整为零
各个击破

说明 全概率公式的主要用处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.



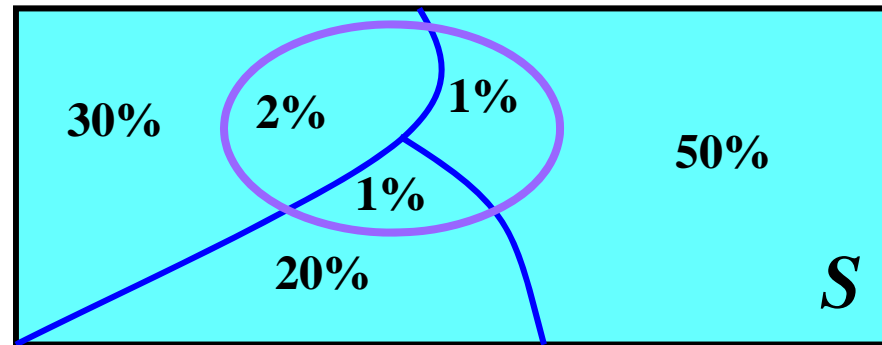
例6 有一批同一型号的产品，已知其中由一厂生产的占 **30%**，二厂生产的占 **50%**，三厂生产的占 **20%**，又知这三个厂的产品次品率分别为 **2%**，**1%**，**1%**，问从这批产品中任取一件是次品的概率是多少？

解 设事件 A 为“任取一件为次品”，

事件 B_i 为“任取一件为 i 厂的产品”， $i = 1, 2, 3$.

$$B_1 \cup B_2 \cup B_3 = S,$$

$$B_i B_j = \emptyset, \quad i, j = 1, 2, 3.$$



由全概率公式得

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3).$$

$$P(B_1) = 0.3, \quad P(B_2) = 0.5, \quad P(B_3) = 0.2,$$

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

$$\begin{aligned} \text{故 } P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\ &= 0.02 \times 0.3 + 0.01 \times 0.5 + 0.01 \times 0.2 = 0.013. \end{aligned}$$

3. 贝叶斯公式

贝叶斯资料

定理 设试验 E 的样本空间为 S . A 为 E 的事件, B_1, B_2, \dots, B_n 为 S 的一个划分, 且 $P(A) > 0, P(B_i) > 0, (i = 1, 2, \dots, n)$, 则

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, \quad i = 1, 2, \dots, n.$$

称此为贝叶斯公式.

证明

$$\begin{aligned} P(B_i|A) &= \frac{P(B_i A)}{P(A)} \\ &= \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, \quad i = 1, 2, \dots, n. \end{aligned}$$

例7 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8 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 **98%**,而当机器发生某种故障时,其合格率为 **55%**.每天早上机器开动时,机器调整良好的概率为 **95%**.试求已知某日早上第一件产品是合格品时,机器调整得良好的概率是多少?

解 设 A 为事件“产品合格”,
 B 为事件“机器调整良好”.
则有

$$P(A|B) = 0.98, \quad P(A|\bar{B}) = 0.55,$$



$$P(B) = 0.95, \quad P(\bar{B}) = 0.05,$$

由贝叶斯公式得所求概率为

$$\begin{aligned} P(B|A) &= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} \\ &= \frac{0.98 \times 0.95}{0.98 \times 0.95 + 0.55 \times 0.05} = 0.97. \end{aligned}$$

即当生产出第一件产品是合格品时,此时机器调整良好的概率为 0.97.

先验概率与后验概率

上题中概率 **0.95** 是由以往的数据分析得到的, 叫做**先验概率**.

而在得到信息之后再重新加以修正的概率 **0.97** 叫做**后验概率**.

例9 根据以往的临床记录,某种诊断癌症的试验具有如下的效果:若以 A 表示事件“试验反应为阳性”,以 C 表示事件“被诊断者患有癌症”,则有 $P(A|C) = 0.95$, $P(\bar{A}|\bar{C}) = 0.95$. 现在对自然人群进行普查,设被试验的人患有癌症的概率为0.005,即 $P(C) = 0.005$, 试求 $P(C|A)$.

解 因为 $P(A|C) = 0.95$,

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

$$P(C) = 0.005, \quad P(\bar{C}) = 0.995,$$



由贝叶斯公式得所求概率为

$$P(C|A) = \frac{P(A|C)P(C)}{P(A|C)P(C) + P(A|\bar{C})P(\bar{C})}$$
$$= 0.087.$$

即平均1000个具有阳性反应的人中大约只有87人患有癌症.

小结

1. 条件概率 $P(B|A) = \frac{P(AB)}{P(A)}$ $\xrightarrow{\text{乘法定理}}$ $P(AB) = P(B|A)P(A)$

\downarrow

全概率公式

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)$$

\downarrow

贝叶斯公式

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, \quad i = 1, 2, \cdots, n$$

2. 条件概率 $P(B|A)$ 与积事件概率 $P(AB)$ 的区别.

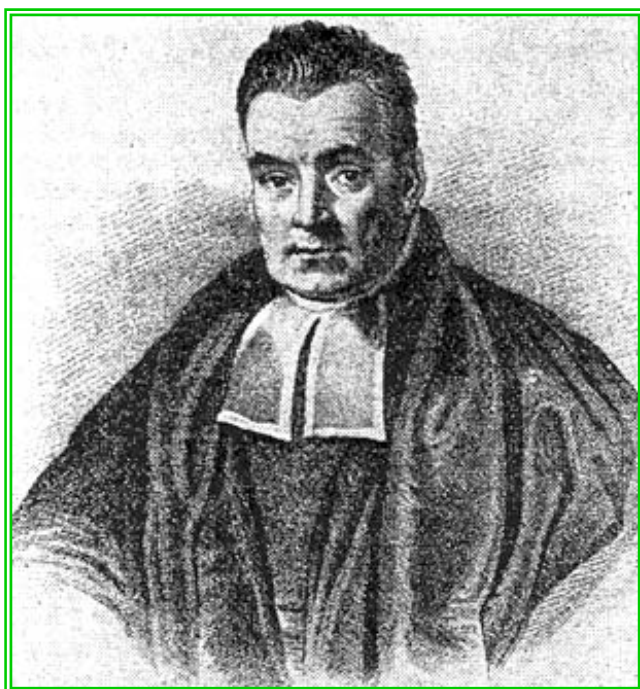
$P(AB)$ 表示在样本空间 S 中, AB 发生的概率, 而 $P(B|A)$ 表示在缩小的样本空间 S_A 中, B 发生的概率. 用古典概率公式, 则

$$P(B|A) = \frac{AB \text{ 中基本事件数}}{S_A \text{ 中基本事件数}},$$

$$P(AB) = \frac{AB \text{ 中基本事件数}}{S \text{ 中基本事件数}},$$

一般来说, $P(B|A)$ 比 $P(AB)$ 大.

贝叶斯资料



Thomas Bayes

**Born: 1702 in London,
England**

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

作业：课后习题 14 、 19 、 24