

目录

第一部分 条件概率: $P(A|条件B) = \frac{P(A \cap B)}{P(条件B)} = \frac{\text{事件}A,B\text{同时发生的概率}}{\text{事件}B\text{发生的概率}}$ 2

1 “条件概率”的意思 2

2 条件概率的性质 4

2.1 性质: $P(A|条件B) \geq 0$ 4

2.2 性质: $P(\Omega|条件B) = 1$ 4

2.3 性质: $P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B) - P(A_1 A_2 | B)$ 4

2.4 性质: $P(A | B) = 1 - P(\overline{A} | B)$ 4

2.5 性质: 可列可加性: 若 $A_1, A_2, \dots A_n, \dots$ 是“互不相容”的事件, 则有: $P(\sum_{i=1}^{\infty} A_i | B) = \sum_{i=1}^{\infty} P(A_i | B) \leftarrow$ 即: “和的概率”, 等于 “概率的和” 4

3 “条件概率”的乘法公式: $P(前后) = P(后) \cdot P(前|后) = P(前) \cdot P(后|前)$ 4

4 传染病模型 8

5 全概率公式: $P(B) = \underbrace{P(A_1) \cdot P(B|A_1)} + \underbrace{P(A_2) \cdot P(B|A_2)} + \dots + \underbrace{P(A_n) \cdot P(B|A_n)}$ 9

6 贝叶斯公式 Bayes’ theorem : $P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^n [P(A_i) \cdot P(B|A_i)]} = \frac{P(A_k B)}{P(B)}$ 14

6.1 先验概率 (从经验来推后果) & 后验概率 (更新迭代经验) 14

6.2 贝叶斯公式: 从 “果”, 来推是 “某因” 的可能性大小 14

文件名

第一部分 条件概率:

$$P(A|\text{条件}B) = \frac{P(A \cap B)}{P(\text{条件}B)} = \frac{\text{事件}A,B\text{同时发生的概率}}{\text{事件}B\text{发生的概率}}$$

1 “条件概率”的意思

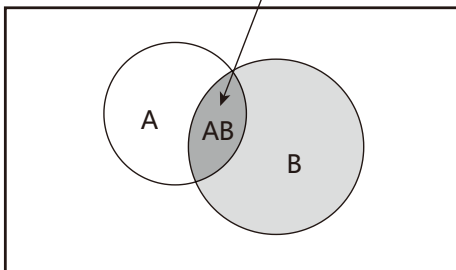
条件概率是: 有 A, B 两个事件, 和样本空间 Ω . 其中 $P(B) > 0$, 则, 在 B 已经发生的条件下, A 发生的概率, 就叫做 A 对 B 的“条件概率”. 记作: $P(A|\text{条件}B)$, 读作“在 B 发生的条件下, A 发生的概率”.

$$\text{即, 条件概率公式是: } P(A|\text{condition}B) = \frac{\overbrace{\text{在 B 发生条件下,A 发生的样本点数}}^{\text{这个分子即: AB 同时发生了}}}{\text{B 里面有多少个样本点}} = \frac{n_{AB}}{n_B}$$

$$\text{还可写成: } P(A|\text{condition}B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{n_{AB}}{n}}{\frac{n_B}{n}} = \frac{n_{AB}}{n_B}$$

这块交集, 就是在B发生的前提下, A发生的概率

$$\text{即: } P(A|\text{条件}B) = \frac{P(AB)}{P(B)}$$



如上图所示, 注意: 概率是个比值, 所以你光有分子那块的交集值, 是没用的, 它还需要与另一个数 (分母) 去比.

上面公式中, $P(AB)$ (\leftarrow 也可写作 $P(A \cap B)$) 的计算公式是什么呢?

- 如果事件 A, 和事件 B 是相互独立的, 则 $P(AB) = P(A) \cdot P(B)$
- 如果事件 A, 和事件 B 不相互独立, 则只能用“条件概率”公式, 来求 $P(AB)$, 即:

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

注意: “条件概率”, 和 “分步骤法” 的区别:

- 分步骤法 (用乘法): 前后每一步骤的事件是相互独立的, 彼此没有条件关系. 比如, 第一步你结婚, 第二步我结婚. 我们这两件事发生的概率互不影响.

- 条件概率 (里面也有用到乘法): 前面的事件, 有可能会 (但并不一定) 影响到后面事件的发生概率. 即前后事件之间并不互相独立.

会影响的例子: 比如一共有 100 个上岸机会, 则第一步你上岸的成功概率, 会影响到第二步我上岸的成功概率. (你若成功, 留给我的名额数量就会更少.)

彼此独立的例子: 比如在你回国的条件下, 我出门的概率. 两者发生的概率毫无关系. 你回不回国, 跟我会出不出门没半毛钱关系.

例

有 6 个球, 各有编号. 我们先定义下这些事件:

- B: 取到偶数编号的球
- A_1 : 取到 1 号球
- A_2 : 取到 2 号球
- A_5 : 取到大于 4 号的球

则:

取到 1 号球的概率

$$- \overbrace{P(A_1)}^{\text{1号球选1}} = \frac{\overbrace{C_1^1}^{\text{1号球选1}}}{\underbrace{C_6^1}_{\text{全 6 选1}}} = \frac{1}{6} = 0.166667$$

偶数编号的球里面, 取不到奇数编号的球

$$- P(A_1|B) = \frac{\text{在 B 条件里面, 取到 } A_1 \text{ (即 1 号球)}}{\text{B: 取到偶数编号的球}} = \frac{\overbrace{0}^{\text{偶数编号的球里面, 取不到奇数编号的球}}}{\underbrace{C_3^1}_{\text{3个偶数球里面取 1 个}}} = 0$$

1个编号 2 的球里面, 取 1 个

$$- P(A_2|B) = \frac{\overbrace{C_1^1}^{\text{1个编号 2 的球里面, 取 1 个}}}{\underbrace{C_3^1}_{\text{3个偶数球里面取 1 个}}} = \frac{1}{3}$$

5,6号与偶数的交集, 只有 6 号一个球

$$- P(A_5|B) = \frac{\text{在 B 条件里面, 取到大于 4 号的球}}{\text{B: 取到偶数编号的球}} = \frac{\overbrace{1}^{\text{5,6号与偶数的交集, 只有 6 号一个球}}}{3}$$

例

某地 10 万人, 其中平均有 40 人患甲病; 有 34 人患乙病; 两种病都患的有 32 人. 则, 我们设:

→ 事件 A : 是一个人得甲病

→ 事件 B : 是一个人得乙病

即有:

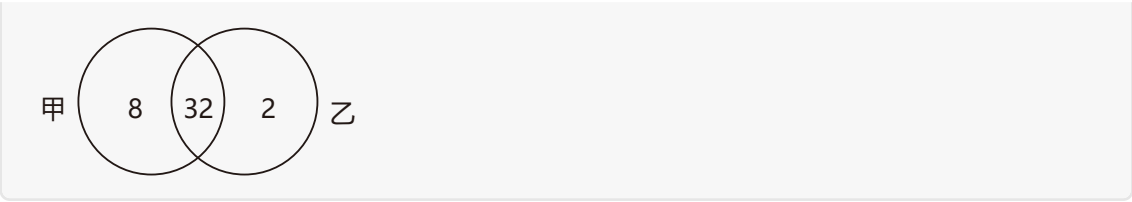
$$P(A) = \frac{40}{100000} = 0.0004$$
$$P(B) = \frac{34}{100000} = 0.00034$$
$$P(AB) = \frac{32}{100000} = 0.00032$$

根据公式: $P(AB) = P(A) \cdot \underbrace{P(B|A)}_{= \frac{P(AB)}{P(A)}} = P(A) \cdot \frac{P(AB)}{P(A)}$,

$$\leftarrow \text{即 } P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{32}{100000}}{\frac{40}{100000}} = 0.8$$

根据公式: $P(AB) = P(B) \cdot P(A|B) = P(B) \cdot \frac{P(AB)}{P(B)}$

$$\leftarrow \text{即 } P(A|B) = \frac{P(AB)}{P(B)} = \frac{\frac{32}{100000}}{\frac{34}{100000}} = 0.941176$$

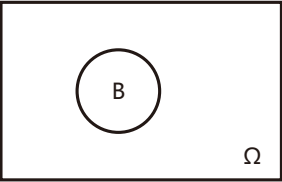


2 条件概率的性质

2.1 性质: $P(A|条件B) \geq 0$

2.2 性质: $P(\Omega|条件B) = 1$

$P(\Omega | B)=1$



2.3 性质: $P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B) - P(A_1 A_2 | B)$

2.4 性质: $P(A | B) = 1 - P(\bar{A} | B)$

2.5 性质: 可列可加性: 若 $A_1, A_2, \dots, A_n, \dots$ 是“互不相容”的事件, 则有: $P(\sum_{i=1}^\infty A_i | B) = \sum_{i=1}^\infty P(A_i | B) \leftarrow$ 即: “和的概率”, 等于 “概率的和”

3 “条件概率” 的乘法公式:

$$P(前后) = P(后) \cdot P(前|后) = P(前) \cdot P(后|前)$$

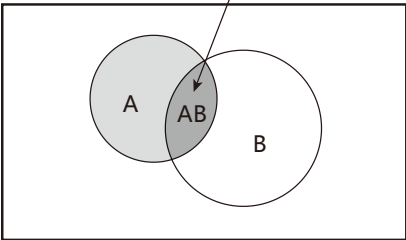
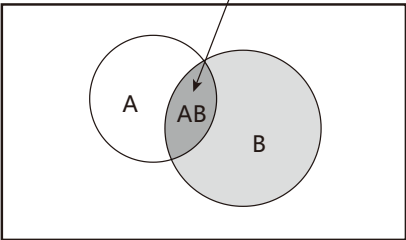
推导过程:

因为
$$\begin{cases} P(A|B) = \frac{P(AB)}{P(B)} \rightarrow \text{即 } P(AB) = \underbrace{P(B) \cdot P(A|B)}_{\text{乘法, 是交集}\cap\text{的概念}} \quad \textcircled{1} \\ P(B|A) = \frac{P(AB)}{P(A)} \rightarrow \text{即 } P(AB) = \underbrace{P(A) \cdot P(B|A)}_{\text{乘法, 是交集}\cap\text{的概念}} \quad \textcircled{2} \end{cases}$$

①和②, 就是“乘法公式”. 即 $P(AB) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$
其中 $P(A)>0, P(B)>0$

$P(AB) = P(B) * P(A | 条件B)$

$P(AB) = P(A) * P(B | 条件A)$



同理, 多个事件的乘法公式就是:

$$\rightarrow P(ABC) = \underbrace{P(A)} \cdot \underbrace{P(B|A)} \cdot \underbrace{P(C|BA)}$$

↑ 上面 “从右往左” 看, 就是按 A,B,C 的顺序

→ $P(A_1 A_2 \cdots A_n) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_2 A_1) \cdot \dots \cdot P(\underbrace{A_n}_{\text{即最后一个数字n, 写在前面.}} \mid \underbrace{A_{n-1} \dots A_2 A_1}_{\substack{\text{后面的条件,} \\ \text{最后一个数字是:} \\ \text{n少一位}}})$

↑ 上面“从右往左”看, 就是按 A_1, A_2, \dots, A_n 的顺序

例

有 100 件产品, 次品率 = 10%, 即有 10 件次品. 做不放回抽样, 问: 第 3 次才取到合格品的概率是?

我们先令:

- A_1 表示第 1 次取, 就取到了合格品
- A_2 表示第 2 次取, 取到了合格品
- A_3 表示第 3 次取, 取到了合格品

那么第 3 次才取到合格品, 就是:

$$\begin{aligned} P(\overline{A_1} \overline{A_2} A_3) &= \underbrace{P(\overline{A_1})}_{\text{分步骤来做}} \cdot \underbrace{P(\overline{A_2} | \overline{A_1})}_{\substack{\text{第1次取, 取到次品} \\ \text{在第1次取到次品的条件下,} \\ \text{第2次取到也是次品}}} \cdot \underbrace{P(A_3 | \overline{A_2} \overline{A_1})}_{\substack{\text{在第1, 2次取到次品的条件下,} \\ \text{第3次取到正品}}} \\ &= \frac{\overbrace{C_{10}^1}^{\text{第一步, 10次品里面取1}}}{C_{100}^1} \cdot \frac{\overbrace{C_9^1}^{\substack{\text{第二步, 因为第一步} \\ \text{已经拿走1件次品,} \\ \text{这第二步就只能在9件次品} \\ \text{里面取1件次品了}}}}{C_{99}^1} \cdot \frac{\overbrace{C_{90}^1}^{\substack{\text{既然第三步要拿到正品,} \\ \text{我们就从90件正品里面取1了}}}}{C_{98}^1} \\ &= \frac{10}{100} \cdot \frac{9}{99} \cdot \frac{90}{98} = 0.00834879 \end{aligned}$$

例

某产品:

- 甲公司占 60% 市场份额, 且其产品合格率是 90%
- 乙公司占 40% 市场份额, 且其产品合格率是 80%

我们先定义下这些事件:

- J : 表示产品是甲的
- \bar{J} : 表示产品是乙的
- Q (qualified): 表示产品是“合格”的
- \bar{Q} : 表示产品是“不合格”的

问, 你买一个产品, 是甲公司的, 并且是合格的概率是?

$$P(JQ) = \underbrace{P(J)}_{=0.6} \cdot \underbrace{P(Q | J)}_{\text{甲的合格率}=0.9} = 0.54$$

问, 你买一个产品, 是乙公司的, 并且是合格的概率是?

$$P(\bar{J}Q) = \underbrace{P(\bar{J})}_{=0.4} \cdot \underbrace{P(Q | \bar{J})}_{\text{乙条件下的合格率}=0.8} = 0.32$$

例

抽签, 共 10 签, 其中有 4 个为“成功上岸”的好签. 甲乙丙三人, 按顺序依次去抽, 不放回.

我们先设定事件:

- A: 表示甲抽到“成功”
- B: 表示乙抽到“成功”
- C: 表示丙抽到“成功”

问, (1) 甲抽到“成功”的概率? $P(A) = \frac{C_{4\text{好签}}^1}{C_{10\text{签}}^1} = \frac{4}{10} = 0.4$

(2) 甲乙都抽到“成功”的概率?

$$\begin{aligned} P(AB) &= \underbrace{P(A)}_{\text{第 1 步: 甲先成功}} \cdot \underbrace{P(B|A)}_{\text{第 2 步: 在甲成功的前提下, 乙再成功}} \\ &= \frac{\overbrace{C_{4\text{好签}}^1}^{\text{甲先抽掉一张好签}}}{C_{10\text{签}}^1} \cdot \frac{\overbrace{C_{4\text{好签}-1}^1}^{\text{乙就只能从剩下的 3 张好签中来抽了}}}{C_{10\text{签}-1}^1} = \frac{4}{10} \cdot \frac{3}{9} = 0.133333 \end{aligned}$$

(3) 甲失败, 乙成功的概率?

$$\begin{aligned} P(\overline{A}B) &= \underbrace{P(\overline{A})}_{\text{第 1 步: 甲先失败}} \cdot \underbrace{P(B|\overline{A})}_{\text{第 2 步: 在甲失败的前提下, 乙再成功}} \\ &= \frac{\overbrace{C_{6\text{坏签}}^1}^{\text{甲先从共 6 张坏签中取 1}}}{C_{10\text{签}}^1} \cdot \frac{\overbrace{C_{4\text{好签}}^1}^{\text{乙从共 4 张好签中取 1}}}{C_{10\text{签}-1}^1} = \frac{6}{10} \cdot \frac{4}{9} = 0.266667 \end{aligned}$$

(4) 甲乙丙都抽到“成功”的概率?

$$\begin{aligned} P(ABC) &= \underbrace{P(A)}_{\text{第 1 步: 甲先成功}} \cdot \underbrace{P(B|A)}_{\text{第 2 步: 在甲成功的前提下, 乙再成功}} \cdot \underbrace{P(C|BA)}_{\text{第 3 步: 在甲乙都成功的前提下, 丙再成功}} \\ &= \frac{C_{4\text{好签}}^1}{C_{10\text{签}}^1} \cdot \frac{C_{\text{还剩 3 好签}}^1}{C_{\text{还剩 9 签}}^1} \cdot \frac{C_{\text{还剩 2 好签}}^1}{C_{\text{还剩 8 签}}^1} \\ &= \frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = 0.0333333 \end{aligned}$$

例

你的新邻居, 是个 is very shy and withdrawn, invariably helpful but with very little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, ant a passion for detail.

你觉得他更可能是 – 图书管理员, 还是农民?

大多数人会认为他是图书馆管理员. 其实, 这种判断是”非理性”的. 问题的关键在于 – 你没有把这两种职业的”人数比例”考虑进去.

→ 在美国,农民与图书馆管理员,两者的数量之比是20:1.

根据这个比例,我们可以假设他们人数的概率分别为:

$$\begin{cases} P(\text{农民职业}) = \frac{20}{100} \\ P(\text{图书管理员职业}) = \frac{1}{100} \end{cases}$$

→ 假如你听到”彬彬有礼”这类描述,你的直觉是:

40%的图书馆管理员符合这个描述,而只有10%的农民符合这个描述.

如果这是你的估计,那就意味着:

$$\begin{cases} \text{在100\% 的纯粹农民群体中,彬彬有礼者的占比概率是:} \\ \qquad \qquad \qquad P(\text{彬彬有礼} | \text{农民职业}) = 0.1 \\ \text{在100\% 的纯粹图书管理员群体中,彬彬有礼者的占比概率是:} \\ \qquad \qquad \qquad P(\text{彬彬有礼} | \text{图书管理员职业}) = 0.4 \end{cases}$$

→ 在既有农民,又有图书管理员的混合群体中:

$$\begin{cases} \text{在属于农民职业的前提下,是彬彬有礼者的概率是:} \\ P(\text{农民职业} \cap \text{彬彬有礼}) = \underbrace{P(\text{农民职业})}_{\text{第一步,是农民的概率}} \cdot \underbrace{P(\text{彬彬有礼} | \text{农民职业})}_{\text{第二步,是在农民前提下,彬彬有礼者的概率}} \\ \qquad \qquad \qquad = 0.2 \cdot 0.1 = 0.02 \\ \text{在属于图书管理员职业的前提下,是彬彬有礼者的概率是:} \\ P(\text{图书管理员职业} \cap \text{彬彬有礼}) = P(\text{图书管理员职业}) \cdot P(\text{彬彬有礼} | \text{图书管理员职业}) \\ \qquad \qquad \qquad = 0.01 \cdot 0.4 = 0.004 \end{cases}$$

→ $P(\text{在彬彬有礼者中,取1人为图书管理员的概率})$

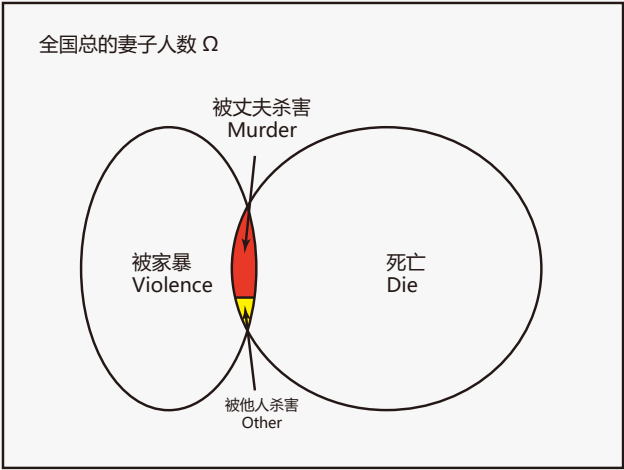
$$\begin{aligned} &= \frac{P(\text{图书管理员职业} \cap \text{彬彬有礼})}{P(\text{农民职业} \cap \text{彬彬有礼}) + P(\text{图书管理员职业} \cap \text{彬彬有礼})} \\ &= \frac{0.004}{0.02 + 0.004} = 0.166667 \end{aligned}$$

所以,即使你认为“符合这个描述的人是一个图书馆管理员的可能性,是一个农民的 4 倍”,也抵不过农民的数量很多.

例

辛普森杀妻案,原告证明辛普森常常家暴前妻. 他们认为,长期家暴说明辛普森有杀妻的动机. 被告律师则举出数据反驳说,美国有 400 万被家暴的妻子,但只有 1432 名被丈夫杀害,这个概率只有 $\frac{1432}{400万}$ = 比 1/2500 还低. 所以家暴证明不了辛普森谋杀. 被告想表达的是: 在“家暴”这个事件前提条件下,丈夫谋杀妻子的概率不高. 即 $P(\text{丈夫家暴} \cap \text{丈夫杀妻}) =$ 概率值很低.

你怎么看? 事实上,被告举出的概率,不适用于这个案子上. 因为本案的妻子已经死亡,“妻子已死”也变成了一个已经存在的前提条件. 所以现在我们要看的概率就是: $P(\text{丈夫家暴} \cap \text{妻子已死亡} \cap \text{是丈夫杀妻})=?$ 即: 在“被家暴”且“死亡”的妻子数量里面 (这里就有两个前提条件了,而不是仅一个前提条件),有多少是被丈夫杀害的?



即:

- 辛普森律师一方的概率公式是: $\frac{\text{红色}}{\text{Violence}} < \frac{1}{2500}$
- 妻子一方律师的概率公式是: $\frac{\text{红色}}{\text{红色} + \text{黄色}} = 93\%$

根据美国 1992 年发布的数据推算: 每 10 万个被家暴的妇女中, 有 43 个会被谋杀. 其中 40 个是被丈夫谋杀, 其他 3 个是被丈夫以外的人谋杀. 那么, 条件概率就是:

$$P(A|B) = \frac{P(AB)}{P(B)}$$
$$P(\text{丈夫杀} | \text{家暴} \cap \text{妻死}) = \frac{P(\text{家暴} \cap \text{妻死} \cap \text{丈夫杀})}{P(\text{家暴} \cap \text{妻死})} = \frac{\frac{40}{100000}}{\frac{43}{100000}} = 0.930233$$

你仔细体会一下两者的不同:

- 辛普森方, 是说: 在所有“之后活着和死去”的被家暴的妻子里, 被丈夫杀了的可能性是多大. 即 $\frac{\text{丈夫杀害}}{\text{条件: 1. 被家暴}}$
- 妻子方, 是说: 在所有“死去”的被家暴的妻子里, 被丈夫杀了的可能性是多大? 即 $\frac{\text{丈夫杀害}}{\text{条件: 1. 被家暴 \& 2. 死亡}}$

不过, 即使概率高达 93%, 也不能绝对证明辛普森杀了妻子. 因为“条件概率”只表示统计意义上的“相关性”, 并不代表“因果关系”. 即只说明: 家暴和谋杀妻子之间有很强相关性。

4 传染病模型

例

有红球 a 个, 黑球 b 个. 你从中取出一个球, 看到其颜色后, 把它放回, 并同时再放入 c 个与你看到的颜色相同的球. 问: 连续 3 次都是取出红球的概率?

先设定事件:

- A_1 : 表示你第 1 次, 取出的是红球
- A_2 : 表示你第 1 次, 取出的是红球
- A_3 : 表示你第 3 次, 取出的是红球

$$P(\underbrace{A_1 A_2 A_3}_{\text{连续3次取到红球}}) = \underbrace{P(A_1)}_{\text{第一步: 第一次就取到红球}} \cdot \underbrace{P(A_2 | A_1)}_{\text{第二步: 在第一次取到红球的前提下, 第二次也取到红球}} \cdot \underbrace{P(A_3 | A_2 A_1)}_{\text{第三步: 再前两次都取到红球的前提下, 第三次也取到红球}}$$

$$= \frac{C_{a\text{红}}^1}{C_{a\text{红} + b\text{黑}}^1} \cdot \frac{C_{a\text{红} + c\text{红}}^1}{C_{a\text{红} + b\text{黑} + c\text{红}}^1} \cdot \frac{C_{a\text{红} + 2c\text{红}}^1}{C_{a\text{红} + b\text{黑} + 2c\text{红}}^1}$$

在第一次取到红球后,
又放入c个红球.
所以第二次取时,
红球就一共有a + c个

为什么总数没有减1个?
因为是“放回抽样”.
总数不变

上面可以看出:

- 当 c 红 = 0 时, 就是正常的 “放回抽样”.
- 当 c 红 = -1 时, 就是 “不放回抽样”. 即把之前步骤中取到的球, 拿走了, 不放回总体中.
- 当 c 红 >0 时, 就是本例的 “传染病模型”.

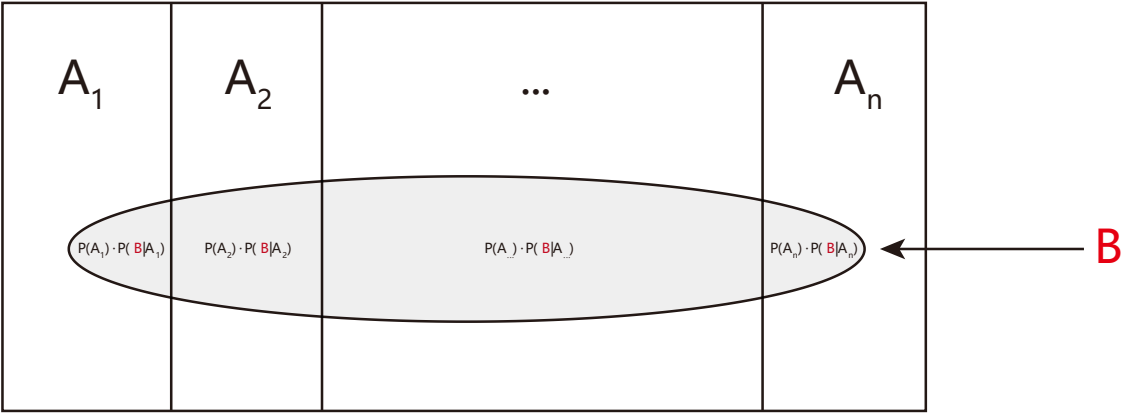
5 全概率公式: $P(B) =$

$$\underbrace{P(A_1) \cdot P(B|A_1)} + \underbrace{P(A_2) \cdot P(B|A_2)} + \dots + \underbrace{P(A_n) \cdot P(B|A_n)}$$

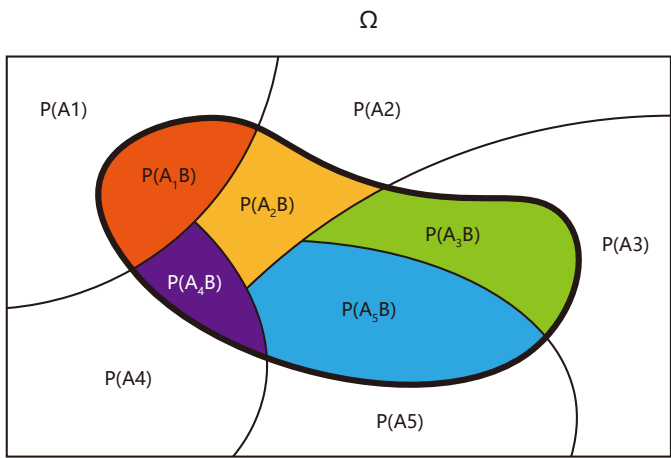
全概率公式 Total Probability Theorem:
如果 A_1, A_2, \dots, A_n 构成一个 “完备事件组”, 即: (1) 这些事件两两互不相容, (2) 其 “和”(或 “并集”) 为全集 Ω , (3) $P(A_i) > 0$.

则有:
$$\sum_{i=1}^n [P(A_i) \cdot P(B|A_i)] = P(B)$$

即有:
$$P(B) = \underbrace{P(A_1) \cdot P(B|A_1)} + \underbrace{P(A_2) \cdot P(B|A_2)} + \dots + \underbrace{P(A_n) \cdot P(B|A_n)}$$



Total Probability Theorem
全概率公式



上图, 粗线部分围起来的整块有彩色区域, 就是 B.
B 的概率, 就等于 = 每一个彩色块的概率, 加总起来.

比如第 1 块, 橙色的概率, 就是 A1 和 B 的交集, 即 $= P(A_1 \cap B)$
 $P(B)$ = 所有 5 块彩色的概率加起来. 即得到下图中的“全概率公式”.

全概率公式:

$$\begin{aligned} P(B) &= \text{第 1 块的概率} + \text{第 2 块的概率} + \dots + \text{第 } n \text{ 块的概率} \\ &= P(A_1B) + P(A_2B) + \dots + P(A_nB) \\ &= P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \dots + P(A_n) \cdot P(B|A_n) \\ &= \sum_{i=1}^n [P(A_i) \cdot P(B|A_i)] \end{aligned}$$

并有: $P(B) = \underbrace{P(A) \cdot P(B|A)}_{\text{A中,B的概率的具体值}} + \underbrace{P(\overline{A}) \cdot P(B|\overline{A})}$

注意: 上式中, $P(B|A)$ 这块只是个比例而已. 即 B 在 A 中的比例. 即 $\frac{B}{A}$. 但单纯的比例是没用的. 比如, alice 说她的收入只有 bob 的 1/10, 但 1/10 依然没有告诉你 alice 的收入到底是多少? 所以, 比例值还需要乘上一个基数. 这个“基数”就是 bob 本身的收入, 比如是 10000 元, 你才能知道 alice 的收入是 $10000 \cdot \frac{1}{10} = 1000$ 元.

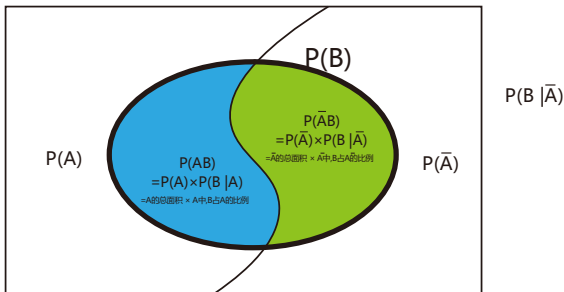
同理, 本处的公式, $P(B|A)$ 这个比例, 还要乘上“ $P(A)$ 本身的值”作为基数, 我们才能最终知道 $P(AB)$ 的具体值到底是多少.

事实上, $P(B|A)$ 就是 B 占 A 的比例. 即 $\frac{B}{A}$.
而 $P(A) \cdot P(B|A)$ 就是 AB 的交集面积占整个全集 Ω 的比例, 即 $\frac{A \cap B}{\Omega}$

如果我们把全集分为两部分: A 和 \overline{A} , 则, B 的部分, 就是:

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

如下图:



5 全概率公式:
$$P(B) = \underbrace{P(A_1) \cdot P(B|A_1)} + \underbrace{P(A_2) \cdot P(B|A_2)} + \dots + \underbrace{P(A_N) \cdot P(B|A_N)} \quad 11$$

比如, 全集 Ω (亚欧板块) 被分成两部分: 一块是 A (亚洲), 另一块是 \overline{A} (欧洲). 全集中有子集 B (俄罗斯), 被 A 和 \overline{A} 各自分割了一部分土地. 那么:

$$\underbrace{P(B)}_{\text{俄罗斯占亚欧板块的面积比例}} = \underbrace{P(A)}_{\text{亚洲占亚欧板块的比例}} \cdot \underbrace{P(B|A)}_{\text{亚洲中的俄罗斯部分, 占亚洲的比例}} + \underbrace{P(\overline{A})}_{\text{欧洲占亚欧板块的比例}} \cdot \underbrace{P(B|\overline{A})}_{\text{欧洲中的俄罗斯部分, 占欧洲的比例}}$$

例

一个工厂, 有 4 条生产线, 情况如下:

	生产线 1	生产线 2	生产线 3	生产线 4
产量	15%	20%	30%	35%
不合格率	0.05	0.04	0.03	0.02

问: 从该工厂的产品中, 任取一件, 是“不合格品”的概率?

我们先设定事件:

- A_1 : 表示是生产线 1 中的产品
- A_2 : 表示是生产线 2 中的产品
- A_3 : 表示是生产线 3 中的产品
- A_4 : 表示是生产线 4 中的产品
- B : 表示是次品

那么, 你任取一件为不合格的概率, 不就是整个工厂总的不合格概率么?! 即 $=P(B)$

$$\begin{aligned} P(B) &= \underbrace{\underbrace{P(A_1)}_{\text{产品属于生产线 1 的概率}} \cdot \underbrace{P(B|A_1)}_{\text{生产线 1 中的次品率}}}_{\text{第 1 条生产线中(的条件下), 不合格品的概率}} + P(A_2) \cdot P(B|A_2) + P(A_3) \cdot P(B|A_3) + P(A_4) \cdot P(B|A_4) \\ &= (15\% \cdot 0.05) + (20\% \cdot 0.04) + (30\% \cdot 0.03) + (35\% \cdot 0.02) \\ &= 0.0315 \end{aligned}$$

例

有 10 台机器人, 3 台是次品. 已经卖出去了 2 台 (是正品还是次品未知).

问: 再取 1 台, 是正品的概率?

首先, 我们定义事件:

- B_{00} : B(bad), 表示前两次取, 都是次品 (用 0 表示)
- B_{10} : 表示前两次取, 是一正 (用 1 表示), 一次 (用 0 表示). 至于顺序是“正, 次”还是“次, 正”, 都行
- B_{11} : 表示前两次取, 都是正品
- G_{xx3} : G(good), 表示第三次取, 是正品

那么, 第 3 次取到正品 $P(G_{xx3})$ 的情况, 就有这 3 种可能性:

- (第 1 次取到) 次, (第 2 次取到) 次, (第 4 次取到) 正.

$$\text{即 } \rightarrow = \underbrace{P(B_{00})}_{\text{前两次取到次品}} \cdot \underbrace{P(G_{xx3} | B_{00})}_{\text{在前两次取到次品的条件下, 第 3 次取到正品}}$$

- 次, 正, 正. 即 $\rightarrow = P(B_{10}) \cdot P(G_{xx3} | B_{10})$
- 正, 正, 正. 即 $\rightarrow = P(B_{11}) \cdot P(G_{xx3} | B_{11})$

上面这三种可能性并存, 就是 “和”(并集) 的概念. 用加法:

$$\begin{aligned} G_{xx3} &= \underbrace{P(B_{00}) \cdot P(G_{xx3} | B_{00})}_{\text{第1种情况: 2次, 1正}} + \underbrace{P(B_{10}) \cdot P(G_{xx3} | B_{10})}_{\text{第2种情况: 1次1正, 再正}} + \underbrace{P(B_{11}) \cdot P(G_{xx3} | B_{11})}_{\text{第3种情况: 前2正, 再1正}} \\ &= \underbrace{\frac{C_{\text{总}10}^2 \cdot C_{\text{总}10-2}^1}{C_{\text{总}10}^2 \cdot C_{\text{总}10-2}^1}}_{\text{第1种情况: 2次, 1正}} + \underbrace{\frac{C_{\text{总}7\text{正}}^1 \cdot C_{\text{总}3\text{次}}^1}{C_{\text{总}10}^2}}_{\text{第2种情况: 1次1正, 再1正}} + \underbrace{\frac{C_{\text{总}7\text{正}}^2 \cdot C_{\text{还剩总}5\text{正}}^1}{C_{\text{总}10}^2 \cdot C_{\text{还剩总}5\text{正}}^1}}_{\text{第3种情况: 前2正, 再1正}} \\ &= \frac{C_3^2 \cdot C_7^1}{C_{10}^2 \cdot C_8^1} + \frac{C_7^1 \cdot C_3^1}{C_{10}^2 \cdot C_8^1} + \frac{C_7^2 \cdot C_5^1}{C_{10}^2 \cdot C_8^1} = 0.7 \end{aligned}$$

注意: 上面的第2种情况, 先取1正1次, 或1次1正的情况, 就是:

$$\begin{aligned} &= \underbrace{\frac{C_7^1}{C_{10}^1} \cdot \frac{C_3^1}{C_9^1}}_{\text{先取1正, 再取1次}} + \underbrace{\frac{C_3^1}{C_{10}^1} \cdot \frac{C_7^1}{C_9^1}}_{\text{先取1次, 再取1正}} = 0.466667 \\ &= \frac{C_7^1 \cdot C_3^1}{C_{10}^2} = 0.466667 \end{aligned}$$

总10个里面取2个,
其中1个从正品里面取,
1个从次品里面取

一件事是“分步骤”来做的, 用乘法

		第 1 次取到	第 2 次取到	第 3 次取到
“或”的情况 用加法	情况1	次	次	正
	情况2	次	正	正
	情况3	正	正	正

例

有 10 件产品, 其中次品的数量, 有三种可能性: 0 件 /1 件 /2 件, 即这三种可能性中的每一种, 发生的概率是 1/3.

同时, 检验时也存在 “误检” 情况:

	\rightarrow 被检验成 \rightarrow		的概率是
正品		次品	0.02
正品	\rightarrow	正品	0.98
次品	\rightarrow	正品	0.05
次品		次品	0.96

问: 这批产品能通过检验 (即事件 S_2) 的可能性是多少? 即本题要求 $P(S_2)=?$

这要分两种情况来讨论 (“和” 的概念, 用加法):

- 1. 正品被误检 (成 “假”) 时的情况
- 2. 次品被误检 (成 “真”) 时的情况

我们先定义各种事件:

- B_0 : B(bad). 表示总的 10 件产品中, 存在 0 件次品. 该事件的概率, 题目已经告诉我们: $P(B_0) = \frac{1}{3}$
- B_1 : 表示总的 10 件产品中, 存在 1 件次品. $P(B_1) = \frac{1}{3}$
- B_2 : 表示总的 10 件产品中, 存在 2 件次品. $P(B_2) = \frac{1}{3}$
- S_1 : S(sample. (v.) 抽样检验; 取样; 采样) 表示任意抽检一次, 抽到了正品. (但这里还有个问题不清晰, 就是说这个正品, 到底是它本身就是“正品”; 还是说只是抽验认为它是“正品”?)
- $\overline{S_1}$: 表示任意抽检一次, 抽到了次品.
- S_2 : 表示再次检验, 并“通过验证” (注意: 有误检率存在. 所以通过检验的, 未必是“正品”; 反之亦然).

本题要求的 $P(S_2)$, 实际上就是: “无论第一次抽, 认为是正是次; 在第二次检验时, 都认为是正品” 的东西. 即: $P(S_2) = \underbrace{P(S_1) \cdot P(S_2 | S_1)}_{\text{第一次抽为正品, 第二次检验为正}} + \underbrace{P(\overline{S_1}) \cdot P(S_2 | \overline{S_1})}_{\text{第一次抽为次品, 第二次检验为正}}$

那么我们先考算 $P(S_1)$ 和 $P(\overline{S_1})$.

→ $P(S_1)$: 是在具体“次品”数量未知的情况下, 抽 1 次就得到“正品”的概率.

$$\begin{aligned}
 P(S_1) &= \underbrace{P(B_0) \cdot P(S_1 | B_0)}_{\text{在总数中有 0 次品的条件下, 抽 1 次得到正品的概率}} + \underbrace{P(B_1) \cdot P(S_1 | B_1)}_{\text{总数中含有 1 次品, 抽 1 次取到正}} + \underbrace{P(B_2) \cdot P(S_1 | B_2)}_{\text{总数中含有 2 次品, 抽 1 次取到正}} \\
 &= \underbrace{\frac{1}{3} \cdot \frac{C_{\text{总 10 正}}^1}{C_{\text{总 10}}^1}}_{\text{总 10 中含有 0 次品}} + \underbrace{\frac{1}{3} \cdot \frac{C_{\text{总 9 正}}^1}{C_{\text{总 10}}^1}}_{\text{总 10 中含有 1 次品}} + \underbrace{\frac{1}{3} \cdot \frac{C_{\text{总 8 正}}^1}{C_{\text{总 10}}^1}}_{\text{总 10 中含有 2 次品}} \\
 &= 0.9
 \end{aligned}$$

所以: $P(\overline{S_1}) = 1 - P(S_1) = 1 - 0.9 = 0.1$

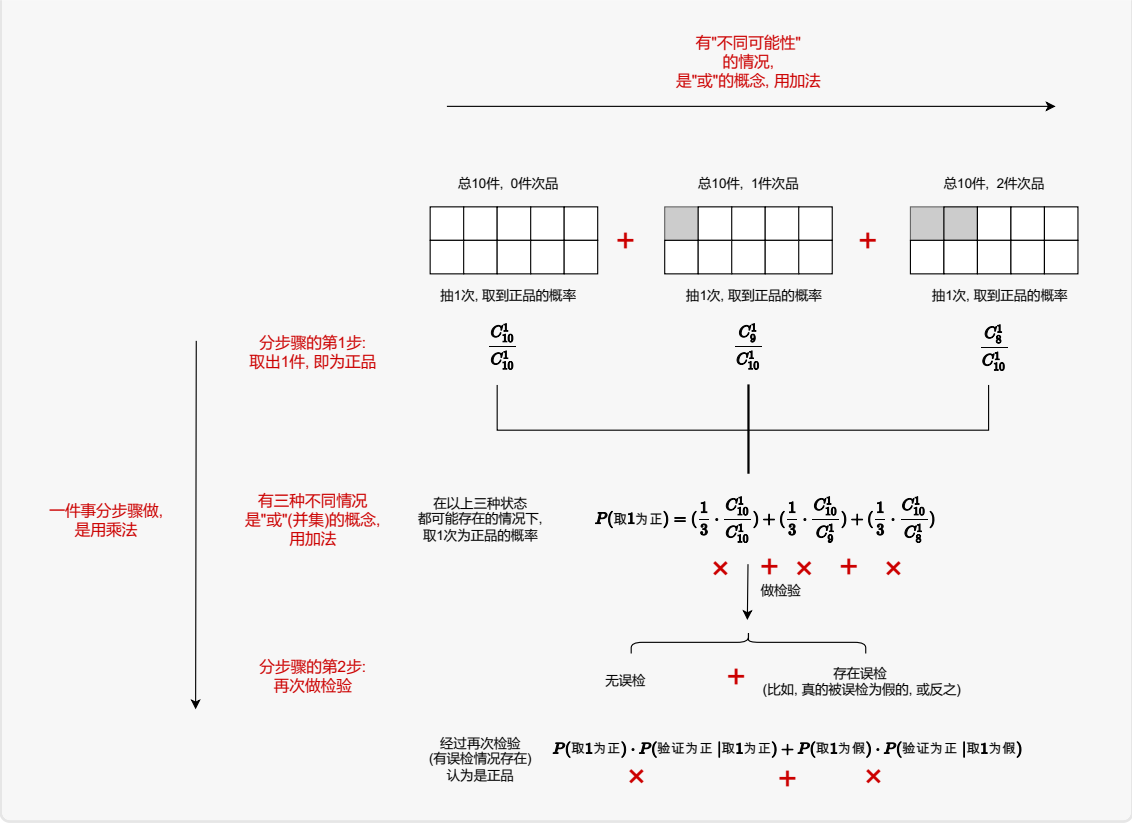
于是, 我们就能得到:

$$\begin{aligned}
 P(S_2) &= \underbrace{\overbrace{P(S_1)}^{=0.9} \cdot \overbrace{P(S_2 | S_1)}^{\text{从上面的表格中可知, 正品被检验为正品, 概率为0.98}}}_{\text{第一次抽为正品, 第二次检验为正}} + \underbrace{\overbrace{P(\overline{S_1})}^{=0.1} \cdot \overbrace{P(S_2 | \overline{S_1})}^{\text{次品被检验为正品, 概率是0.05}}}_{\text{第一次抽为次品, 第二次检验为正}} \\
 &= (0.9 \cdot 0.98) + (0.1 \cdot 0.05) = 0.887
 \end{aligned}$$

6 贝叶斯公式 BAYES' THEOREM:

$$P(A_K|B) = \frac{P(A_K) \cdot P(B|A_K)}{\sum_{I=1}^N [P(A_I) \cdot P(B|A_I)]} = \frac{P(A_K B)}{P(B)}$$

14



6
贝叶斯公式
BAYES' THEOREM:

$$P(A_K|B) = \frac{P(A_K) \cdot P(B|A_K)}{\sum_{I=1}^N [P(A_I) \cdot P(B|A_I)]} = \frac{P(A_K B)}{P(B)}$$

15

“条件概率”的公式是 $P(B|A) = \frac{P(AB)}{P(A)}$ ①

可进一步变换为： $P(AB) = P(A) \cdot P(B|A)$

改变A和B的顺序，就得到： $\overbrace{P(BA)}^{P(AB)}$ = $P(B) \cdot P(A|B)$ ②

$P(AB)$
是AB同时发生的概率，
其实就 = $P(BA)$

把公式②代入公式①，就得到：
$$P(B|A) = \frac{\overbrace{P(AB)}^{P(BA) = P(B) \cdot P(A|B)}}{P(A)}$$

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

即： $P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

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

③ ← 这个就是“贝叶斯公式”

上面“贝叶斯公式”的意思就是说：“在现象 B 出现的条件下，事件 A 发生的概率”（即 P(A|B)），就等于“事件 A 发生的概率（即 P(A)）”，乘以“事件 A 发生条件下，事件 B 出现的概率”（即 P(B|A)），再除以“事件 B 出现的概率”（即 P(B)）。

推导 2：

有条件概率： $P(B|A_i) = \frac{P(A_i \cap B)}{P(A_i)}$ ①

即： $P(A_i B) = P(A_i) \cdot P(B|A_i)$ ②

然后，对①式，我们把A_i和B 倒一倒，让B作为条件，即有：

$$P(A_i|B) = \frac{\overbrace{P(A_i B)}^{\text{分子即式②的内容}}}{\underbrace{P(B)}_{\text{分母即“全概率公式”}}} \leftarrow \text{分母的 } P(B) = \sum_{i=1}^n [P(A_i) \cdot P(B|A_i)]$$

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

$$= \frac{\overbrace{P(A_i) \cdot P(B|A_i)}^{i \text{ 可取 } 1 \text{ 到 } n \text{ 中的任何索引值，即任何一条路径}}}{\sum_{i=1}^n [P(A_i) \cdot P(B|A_i)]} = \frac{\overleftarrow{\text{条件概率}}}{\overleftarrow{\text{全概率公式}}} \quad \text{③} \leftarrow \text{这个就是“贝叶斯公式”}$$

推导 3：

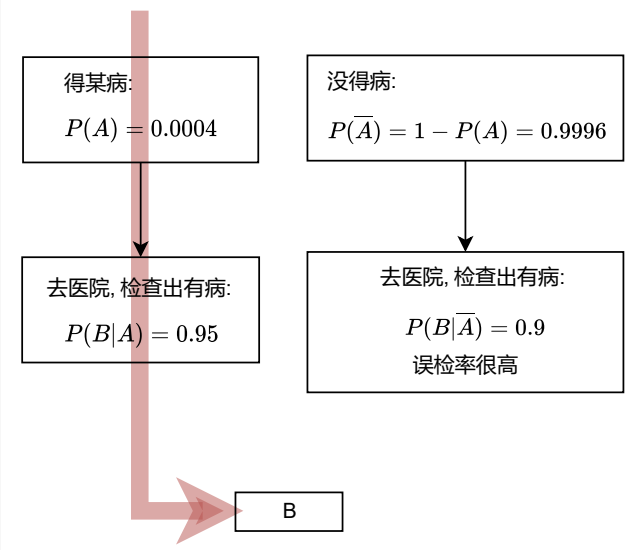
$$P(A_K|B) = \frac{P(A_K) \cdot P(B|A_K)}{\sum_{I=1}^N [P(A_I) \cdot P(B|A_I)]} = \frac{P(A_K B)}{P(B)} \quad 16$$

则: 从 B 结果, 来倒推推断“它是由第 A_k 个原因造成的”的可能性占多大的概率, 就由这个公式 (贝叶斯公式) 给出:

$$P(\text{原因}A_k \mid \text{结果}B)$$
$$= \frac{P(\text{原因}A_k \cap \text{结果}B)}{P(\text{结果}B)}$$
$$= \frac{P(\text{原因}A_k) \cdot P(\text{结果}B \mid \text{原因}A_k)}{\sum_{i=1}^n [P(\text{原因}A_i) \cdot P(\text{结果}B \mid \text{原因}A_i)]}$$
$$= \frac{A_k \text{这个人, 他的责任量}}{\text{所有相关的人(所有的}A\text{), 责任加起来的总和量}}$$

例

得某病 ← 事件 A
医院检查出有病 ← 事件 B

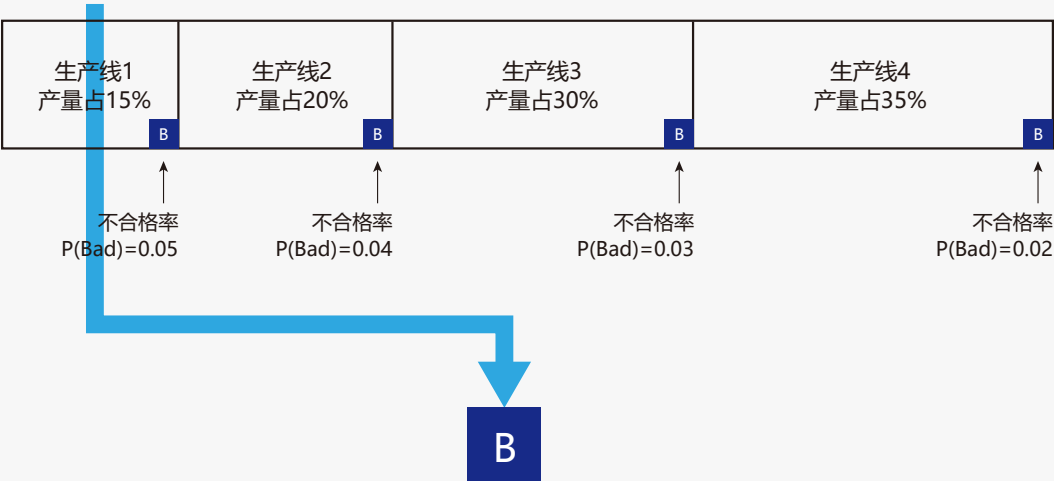


在某人检测出有病 (B) 的前提下, 问 “其是真实患病 (A)” 的概率是多少? (即走的是上面红色路径的概率). 即求: $P(A|B) = ?$

$$P(A|B) = \frac{P(AB)}{P(B)}$$
$$= \text{或用贝叶斯公式} \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^n [P(A_i) \cdot P(B|A_i)]}$$
$$= \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})}$$
$$= \frac{0.0004 \cdot 0.95}{(0.0004 \cdot 0.95) + (0.9996 \cdot 0.9)} = 0.000422213$$

- 所以, 你可以看出:
- 全概率公式, 是从 “原因” 来推 “结果的可能性是多少”.
 - 贝叶斯公式, 是从 “结果” 来倒推其 “是从哪一种原因得来的” 的可能性. 即 $P(\text{原因}_i \mid \text{结果})$

例



一个工厂, 有 4 条生产线, 情况如下:

	生产线 1	生产线 2	生产线 3	生产线 4
产量	15%	20%	30%	35%
不合格率	0.05	0.04	0.03	0.02

我们设:

- A_1 : 代表生产线 1 的产品
- A_2 : 代表生产线 2 的产品
- A_3 : 代表生产线 3 的产品
- A_4 : 代表生产线 4 的产品
- B (bad): 代表不合格品.

则, 这个工厂的总的合格率:

即 全概率公式

$$P(B) = \underbrace{P(A_1) \cdot P(B|A_1)}_{\text{第 1 条生产线的产量占比, 乘以第 1 条生产线的合格率}} + \dots + P(A_4) \cdot P(B|A_4)$$
$$= (0.15 \cdot 0.05) + (0.2 \cdot 0.04) + (0.3 \cdot 0.03) + (0.35 \cdot 0.02) = 0.0315$$

全概率公式的图解

设有样本空间 S 及其中某一事件为 B , 按如下方式构造一完备事件组 $\{A_i\} (i=1, 2, \dots, 10)$, 注意到事件 A_i 发生的条件下事件 B 发生的概率为 $P(BA_i)=P(A_i)P(B|A_i)$, 则

$$P(B)=P(A_1)P(B|A_1)+\dots+P(A_{10})P(B|A_{10})$$

现在问: 在发生了 B 结果的前提下, 它是由 A_1 原因 (路径) 引起的可能性, 是多少? ← 这就是要用“贝叶斯公式”了. 从“结果”, 来问“某原因”导致的可能性是多少.

A₁生产线产量占0.15

A₁生产线的不合格率=0.05

$$P(A_1|B) = \frac{P(A_1 B)}{P(B)} = \frac{\overbrace{P(A_1)} \cdot \overbrace{P(B|A_1)}}{\underbrace{P(B)}} = \frac{0.15 \cdot 0.05}{0.0315} = 0.238095$$

这个工厂总的合格率=0.0315

同理, 其他原因 (其他生产线带来的次品) 的可能性是:

$$P(A_2|B) = \frac{P(A_2 B)}{P(B)} = \frac{P(A_2) \cdot P(B|A_2)}{P(B)} = \frac{0.2 \cdot 0.04}{0.0315} = 0.253968$$

$$P(A_3|B) = \frac{P(A_3 B)}{P(B)} = \frac{P(A_3) \cdot P(B|A_3)}{P(B)} = \frac{0.3 \cdot 0.03}{0.0315} = 0.285714$$

$$P(A_4|B) = \frac{P(A_4 B)}{P(B)} = \frac{P(A_4) \cdot P(B|A_4)}{P(B)} = \frac{0.35 \cdot 0.02}{0.0315} = 0.222222$$

所以, 是第 3 条生产线造成的原因的可能性最大, 因为其概率值最高.

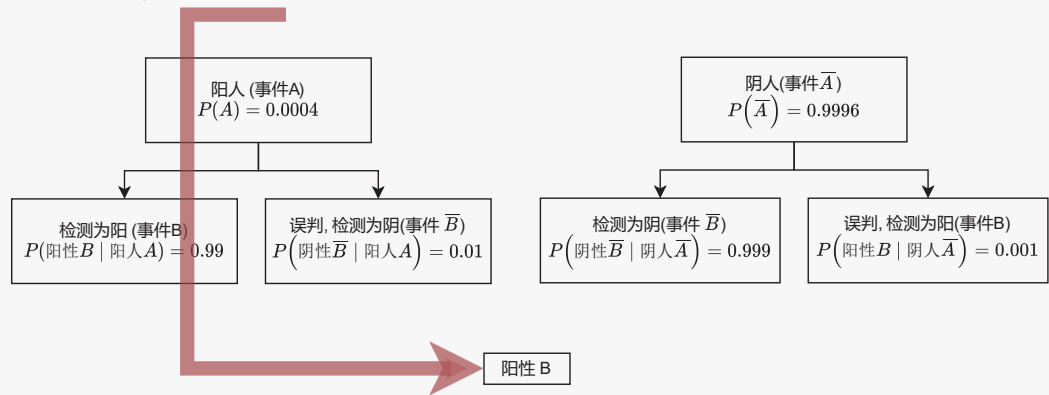
例

某病, 发病率是 0.0004. 有种检测方法, 但存在误诊情况.

我们先设定事件:

- 事件 A: 表示某人真有病, 即阳性
- 事件 \bar{A} : 表示某人无病, 阴性
- 事件 B: 表示检测认为该人阳性.

该检测方法, 准确度如下:



问: 一个人被检测为“阳”, 那么他真的是“得病” 的概率是? 即问: $P(\text{阳人}A | \text{阳性}B = ?)$

$$P(\text{前后})=P(\text{前})\cdot P(\text{后} | \text{前})$$
$$P(A|B) = \frac{\overbrace{P(AB)}}{\underbrace{P(B)}} = \frac{P(A) \cdot P(B|A)}{P(B)}$$

任何一个人被检测为阳的概率

$$= \frac{\overbrace{P(A)}^{=0.0004} \cdot \overbrace{P(B|A)}^{=0.99}}{\overbrace{P(A)}^{=0.0004} \cdot \overbrace{P(B|A)}^{=0.99} + \overbrace{P(\bar{A})}^{=0.9996} \cdot \overbrace{P(B|\bar{A})}^{=0.001}}$$

无误诊, 将患者(A)确认为阳性(B) 有误诊, 将阴人(\bar{A})误诊为阳性

$$= \frac{0.0004 \cdot 0.99}{0.0004 \cdot 0.99 + 0.9996 \cdot 0.001} = 0.283749$$

6 贝叶斯公式 BAYES' THEOREM:

$$P(A_K|B) = \frac{P(A_K) \cdot P(B|A_K)}{\sum_{I=1}^N [P(A_I) \cdot P(B|A_I)]} = \frac{P(A_K B)}{P(B)}$$

20



概率告诉我们: 要相信长期中的期望. $0.99^{365} = 0.025518$, 而 $1.01^{365} = 37.7834$.
篮球领域有一句名言——“训练时, 用正确姿势投丢的球, 比用错误姿势投进的球, 更有价值.”
站在当下, 未来任何事都只是一个概率. 所谓坚持, 所谓努力, 其实就是寻找一个大概率成功的方向, 然后相信系统, 相信长期主义. 当然, 你得坚持活着. 等到长期的到来.

但行为经济学家发现, 人们在决策过程中, 往往并不遵循“贝叶斯规律”, 而是给予最近发生的事件和最新的经验, 以更多的权重值, 更看重近期的事件. 面对复杂问题, 人们往往会走捷径, 依据可能性, 而非概率来做决策. 这种对经典模型的系统性偏离, 称为“偏差”. 因此, 投资者在决策判断时, 并非绝对理性.
但长期以来, 由于缺乏有力的“能结合人类决策中的理性和感性因素”的替代工具, 经济学家不得不在分析中坚持“贝叶斯法则”.