2 条件概率与独立性

2.1 条件概率

2.1.1 条件概率

在解决很多实际的概率问题时, 往往需要考虑在某些附加信息(条件)下的事件发生的概率. 通常记事件A发生的条件下事件B发生的概率为$P(B|A)$, 称为事件A条件下事件B的概率.

例2.1. 投掷一枚骰子后观察点数. 事件B表示观察到3点, 事件A表示观察到奇数点, 求$P(B), P(B|A)$.

解. 整个样本空间$\Omega = \{1, 2, \cdots, 6\}$, 以及事件$B = \{3\}$和事件$A = \{1, 3, 5\}$. 根据古典概型得到

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

□

条件概率的本质: 缩小了有效的样本空间. 下面给出条件概率的形式化定义.

定义2.1. 设A, B为同一空间的两个随机事件且$P(A) > 0$, 称$P(B|A) = P(AB)/P(A)$为事件A发生的条件下事件B发生的概率, 简称**条件概率**.

若事件A发生的条件下事件B发生, 则试验结果必为AB. 由于A发生的条件下考虑B事件的发生, 因此将A看作新的样本空间, 称为缩减的样本空间. 下面给出条件概率的性质:

- 有界性: $0 \leq P(B|A) \leq 1$

  Proof. 由$AB \subset A$, 有$0 \leq P(AB) \leq P(A)$, 根据$P(A) > 0$得到$0 \leq \frac{P(AB)}{P(A)} \leq 1$. □

- 规范性：$P(\Omega|A) = 1$

  Proof. 由$\Omega \cap A = A$得$P(\Omega|A) = \frac{P(A\Omega)}{P(A)} = \frac{P(A)}{P(A)} = 1$. □

- 容斥原理: $P(B_1 \cup B_2|A) = P(B_1|A) + P(B_2|A) - P(B_1B_2|A)$.

  Proof. 根据条件概率的定义有$P(B_1 \cup B_2|A) = P((B_1 \cup B_2) \cap A)/P(A)$, 根据事件的分配率和容斥原理有

  $$P((B_1 \cup B_2) \cap A) = P(AB_1 \cap AB_2) = P(AB_1) + P(AB_2) - P(AB_1AB_2) = P(AB_1B_2),$$

  代入完成证明. □

- 补事件: $P(B|A) = 1 - P(\overline{B}|A)$.

  Proof. 由于事件B和$\overline{B}$互不相容, 有$P(\Omega|A) = 1 = P(B \cup \overline{B}|A) = P(B|A) + P(\overline{B}|A)$. □

- 可列可加性: 设$B_1, B_2, \cdots, B_i, \cdots$是两两互不相容的事件, 则$P(\cup_{i=1}^{\infty} B_i|A) = \sum_{i=1}^{\infty} P(B_i|A)$.

  Proof. 由于$B_1, B_2, \cdots, B_i, \cdots$是两两互不相容的事件, 则$AB_1, AB_2, \cdots, AB_i, \cdots$也是两两互不相容的事件, 根据定义有

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

□

例2.2. 一盒子装有4只不同的产品, 其中3只一等品, 1只二等品. 从盒子中不放回取产品两次, 每次任取一只, 设事件A表示第一次拿到一等品, 事件B表示第二次取到一等品, 求$P(B|A)$.

解. [第一种方法] 将3只一等产品分别标为$\{1, 2, 3\}$, 二等品编号4. 事件(i, j)表示第一,二次分别取到第i, j号产品.
$\Omega = \{(i, j), i \neq j\}, |\Omega| = 4 \times 3 = 12,$
$A = \{(i, j), i \neq j, i \neq 4\}, |A| = 3 \times 3 = 9,$
$B = \{(i, j), i \neq j, j \neq 4\}, |B| = 3 \times 3 = 9,$
$AB = \{(i, j), i \neq j, i \neq 4, j \neq 4\}, |AB| = 3 \times 2 = 6.$
所以$P(B) = \frac{9}{12} = \frac{3}{4}, P(AB) = \frac{6}{12} = \frac{1}{2}, P(B|A) = \frac{P(AB)}{P(A)} = \frac{1/2}{3/4} = \frac{2}{3}$.

[第二种方法: 缩减样本空间] 当第一次取得一等品后剩下2只一等品, 1只2等品, 因此有$P(B|A) = \frac{2}{3}$. □

例2.3. 掷两次均匀的骰子, 已知第一次掷出6点, 问两次掷出点数之和≥ 10的概率.

解. [第一种方法] 首先有样本空间$\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$, 事件$A = \{(i, j) : i = 6, 1 \leq j \leq 6\}$, 事件$B = \{(i, j) : i + j \geq 10\}$, 事件$AB = \{(6, j) : j = 4, 5, 6\}$, 于是有$P(B|A) = 1/2$.

[第一种方法: 缩减空间] 第一次掷6点的情况下, 第二次掷的样本空间$\Omega = \{1, 2, \cdots, 6\}$, 事件$B = \{$两次掷的点数之和 $\geq 10\} = \{4, 5, 6\}$, 所以$P(B|A) = 1/2$. □

2.1.2 乘法公式

根据条件概率公式$P(B|A) = P(AB)/P(A)$可得

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

由此可推广到多个事件的乘法公式: 若$P(A_1A_2 \cdots A_{n-1}) > 0$, 由定义可得

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

例2.4. 一批灯泡100只, 其中10只为次品, 其余为正品. 做不放回抽取, 每次抽取一只, 求第三次才是正品的概率.

解. 设事件A_i ={第i次抽到的正品} $(i=1,2,3)$, 事件B ={第3次才抽到的正品}. 于是有$B=\bar{A}_1\bar{A}_2\ A_3$, 以及

$$P(B)=P(\bar{A}_1\bar{A}_2\ A_3)=P(\bar{A}_1)P(\bar{A}_2|\bar{A}_1)P(A_3|\bar{A}_1\bar{A}_2)=\frac{10}{100}\times\frac{9}{99}\times\frac{90}{98}=\frac{9}{1078}.$$

□

例2.5. *假设有一串钥匙n把, 只有一把能打开门. 任取一把开门, 用后分开, 求第k次打开门的概率.*

解. [第一种方法: 抽签原理] 第k次打开门的事件由前$k-1$次都没有打开门, 以及第k次打开门的事件构成. 第k次打开门的概率为

$$\frac{\binom{n-1}{k-1}}{\binom{n}{k-1}}\times\frac{1}{n-k+1}=\frac{1}{n}.$$

[第二种方法: 乘积公式] 设事件A_i表示第i次不能打开门, 则第k次打开门的事件可以表示为$A_1A_2\cdots A_{k-1}\bar{A}_k$. 于是有

$$\begin{aligned}
&P(A_1A_2\cdots A_{k-1}\overline{A}_k)\\
&= P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_{k-1}|A_1\cdots A_{k-2})P(\overline{A}_k|A_1A_2\cdots A_k)\\
&= \frac{n-1}{n}\times\frac{n-2}{n-1}\times\cdots\times\frac{n-(k-1)}{n-(k-2)}\times\frac{1}{n-(k-1)}=\frac{1}{n}
\end{aligned}$$

□

习题(Matching问题)**.** *有n对夫妻参加一次活动, 所有夫妻被随机两两分成n组, 每组1男1女, 问n对夫妻恰好两两配对的概率.*

2.1.3 全概率公式(Law of total probability)

全概率公式用于复杂的概率计算, 本质上是加法和乘法的综合运用: 对互不相容的事件A,B, $P(A\cup B)=P(A)+P(B)$; 乘法$P(AB)=P(A)P(B|A)$.

定义2.2 (样本空间的划分)**.** *若事件$A_1,A_2,\cdots,A_n$满足: i) 互斥性或互不相容性: $A_i\cap A_j=\emptyset$ $(i\neq j)$; ii) 完备性: $\cup_{i=1}^n A_i=\Omega$, 则称$A_1,A_2,\cdots,A_n$ 为空间Ω的一个划分.*

注1): 若$A_1,A_2,\cdots,A_n$为样本空间的一个划分, 则每次试验, 事件$A_1,A_2,\cdots,A_n$有且仅有一事件发生.

注2): 若$n=2$, 则$A_1=\overline{A}_2$, 即A_1,A_2互为补事件, 对立事件.

定理2.1 (全概率公式)**.** *设$A_1,A_2,\cdots,A_n$为样本空间Ω的一个划分. 对任意事件B, 有*

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

该公式称为全概率公式.

Proof. 根据事件的分配律可得$B = B\cap\Omega = B\cap(\cup_{i=1}^n A_i) = \cup_{i=1}^n BA_i$, 由$A_i\cap A_j = \emptyset$可得$BA_i\cap BA_j = \emptyset$, 即$BA_i$和$BA_j$两两不相容. 由概率的可列可加性得到

$$P(B) = P\left(B\bigcap\left(\bigcup_{i=1}^n A_i\right)\right) = P\left(\bigcup_{i=1}^n BA_i\right) = \sum_{i=1}^n P(BA_i) = \sum_{i=1}^n P(A_i)P(B|A_i).$$

□

[直觉解释]: 将事件B看作某一过程的结果, 将$A_1, A_2, \cdots, A_n$看作该过程的若干原因, 如果

i) 每一原因发生的概率已知, 即$P(A)$已知,

ii) 每一原因对结果B的影响已知, 即$P(B|A_k)$已知,

则$P(B)$可求.

例2.6. *有一种同型号产品由甲乙丙三家工厂生产, 其生产的市场份额分别为30%, 50%, 20%, 三家工厂的次品率分别为2%, 1%, 1%. 求这批产品中任取一件是次品的概率.*

解. 设事件$B =$\{任取一件是次品\}, $A_i =$\{任取一件为第i厂的产品\}, ($i =$甲,乙,丙). 所以

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3) = 1.3\%.$$

□

例2.7 (推迟决定问题)**.** *随机抛n次均匀的硬币, 证明正面向上的次数是偶数(奇数)的概率是1/2.*

Proof. [第一种方法: 全概率公式] 令A表示前$n-1$次抛硬币向上的次数为偶数, B表示前n次抛硬币向上的次数为偶数, 则$\bar{A}$表示前$n-1$次抛硬币向上的次数为奇数. 于是有:

$$P(B) = P(A)P(B|A) + P(\overline{A})P(B|\overline{A}) = \frac{P(A)}{2} + \frac{P(\overline{A})}{2} = \frac{1}{2}.$$

[第二种方法: 直接计算概率] 正面向上的次数是偶数分别是: $0, 2, 4, \ldots, 2k$ $(2k \le n)$, 于是有

$$\sum_{0\le k\le n/2} \binom{n}{2k}\left(\frac{1}{2}\right)^{2k}\left(\frac{1}{2}\right)^{n-2k} = \frac{1}{2^n}\sum_{0\le k\le n/2}\binom{n}{2k} = \frac{1}{2}.$$

[第三种方法: 推迟决定原则(Principle of deferred decision)] 无论前$n-1$次抛正面朝上的次数为奇数或偶数, 结果的奇偶性取决于最后一次, 机会各半. □

2.1.4 贝叶斯公式(Bayes' Law)

实际中还存在另一类问题, 已知结果找原因, 因果互换: 即在观察到事件B已发生的条件下, 寻找导致B发生的概率.

定理2.2 (贝叶斯公式(Bayes' law)). *设$A_1, A_2, \cdots, A_n$为样本空间Ω的一个划分. 若事件$P(B)>0$, 则*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}.$$

Proof. 由全概率公式可知$P(A_i|B) = \frac{P(A_iB)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}$. □

[直觉解释]: 将事件B看作某一过程的结果, 将$A_1, A_2, \cdots, A_n$看作过程中的若干原因, 如果

1) 每一原因发生的概率$P(A_i)$已知;

2) 每一原因A_i对结果B的影响已知, 即$P(B|A_i)$已知,

已知事件B已发生, 求事件B由第i个原因引起的概率$P(A_i|B)$.

特别地, 当$n=2$时, 即事件A和$\bar{A}$为Ω的一个划分, 有

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

例2.8. *已知有3家元件厂, 编号分别为$1, 2, 3$, 它们的次品率分别为$0.02, 0.01, 0.03$, 所占市场份额分别为$0.15, 0.8, 0.05$. 在仓库中随机取一件, 若已知取到的是次品, 求此产品出自三家工厂的概率.*

解. 设事件A_i ={为所需产品来自第i厂} ($i = 1, 2, 3$), 于是有$P(A_1) = 0.15, P(A_2) = 0.8, P(A_3) = 0.05$. 设事件$B$表示取到次品, 有$P(B|A_1) = 0.02$, $P(B|A_1) = 0.01$, $P(B|A_1) = 0.03$. 根据全概率公式有

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3) = 0.0125.$$

根据贝叶斯公式有

$$\begin{aligned} P(A_1|B) &= \frac{P(A_1)P(B|A_1)}{P(B)} = \frac{0.15 \cdot 0.02}{0.0125} = \frac{30}{125} \\ P(A_2|B) &= \frac{P(A_2)P(B|A_2)}{P(B)} = \frac{0.8 \cdot 0.01}{0.0125} = \frac{80}{125} \\ P(A_3|B) &= \frac{P(A_3)P(B|A_3)}{P(B)} = \frac{0.05 \cdot 0.03}{0.0125} = \frac{15}{125} \end{aligned}$$

因此次品出自第2工厂的概率最大. □

习题. *已知事件A为病人被诊断为肝癌, 事件C为病人患有肝癌, $P(A|C) = 0.95$, $P(\overline{A}|\overline{C}) = 0.9$, $P(C) = 0.0004$. 求$P(C|A)$.*

例2.9 (三囚徒问题). *三犯人A, B, C均被判为死刑, 被单独隔离. 法官随机赦免了其中一人, 看守知道谁被赦免但不会说. 犯人A问看守: B和C谁会被判死刑? 看守的策略: 1) 若赦免B, 则说C; 2) 若赦免C, 则说B; 3) 若赦免A, 则以$1/2$的概率说B或C. 看守回答A: 犯人B会被执行死刑. 犯人A兴奋不已, 因为生存的概率为$1/2$. 犯人A将此事告诉犯人C, C同样高兴因为他觉得A的生存几率为$1/3$, 而自己的生存几率为$2/3$. 那么谁错了?*

解. 设事件$A=\{A$被赦免$\}$, 事件$B=\{B$被赦免$\}$, 事件$C=\{C$被赦免$\}$, 由题意可知$P(A)=P(B)=P(C)=1/3$. 设事件D=$\{$看守说B被赦免$\}$, 则有$P(D|A)=1/2$, $P(D|B)=0$, $P(D|C)=1$. 由全概率公式有

$$P(D)=P(A)P(D|A)+P(B)P(D|B)+P(C)P(D|C)=1/2.$$

由贝叶斯公式有

$$P(A|D)=\frac{P(A)P(D|A)}{P(D)}=\frac{1}{3}\quad P(C|D)=\frac{P(C)P(D|C)}{P(D)}=\frac{2}{3}$$

所以A的推断不正确, C的推断正确. □

例2.10 (三门问题). *美国游戏节目: 参赛者看到三扇关闭的门, 其中一门后面是汽车, 两门后面是山羊, 选中有车的门可赢得汽车. 当参赛者选完一扇门但未开启, 节目主持人开启剩下两门的一门露出山羊. 问题: 参赛者是否要换到另一扇关上的门?*

解. 若不换门, 赢的概率为1/3, 若换门, 赢的概率为2/3. □

2.2 独立性

在一般情况下, 由条件概率定义知$P(B|A)=P(AB)/P(A)\neq P(B)$, 事件$A$的发生对事件$B$的发生有影响. 但在很多实际应用中, 一事件的发生对另一事件的发生可能没有任何影响, 即事件的独立性.

定义2.3. *若事件A,B满足$P(AB)=P(A)P(B)$, 则称事件A,B是相互独立的.*

注1): 若事件A和B满足$P(A)P(B)>0$, 则有

$$P(AB)=P(A)P(B)\Leftrightarrow P(B|A)=P(B)\Leftrightarrow P(A|B)=P(A).$$

注2): 不可能事件(概率为0)或必然事件(概率为1)与任何事件都是独立的.

性质2.1. *若事件A,B独立, 则A与$\overline{B}$, $\overline{A}$与B, $\overline{A}$与$\overline{B}$也互相独立.*

Proof. 根据事件差公式$P(A-B)=P(A)-P(AB)$有

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

同理证明$P(\bar{A}B)=P(\bar{A})P(B)$. 利用容斥原理有

$$\begin{aligned}P(\bar{A}\bar{B}) &= 1-P(A\cup B)=1-P(A)-P(B)+P(AB)\\ &= 1-P(A)-P(B)+P(A)P(B)=(1-P(A))(1-P(B))=P(\bar{A})P(\bar{B}),\end{aligned}$$

从而完成证明. □

例2.11. *从一副不含大小王的扑克牌中随机抽取一张扑克, 设事件$A=\{$抽到10$\}$和事件$B=\{$抽到黑色的扑克$\}$. 求事件A与B是否独立?*

解. 不含大小王的扑克共52张, 黑色扑克26张, 4张10, 因此有$P(A) = 4/52 = 1/13$以及$P(B) = \frac{1}{2}$. 进一步有

$$P(AB) = \frac{2}{52} = P(A)P(B),$$

所以事件A和B是独立的. □

在很多真实应用中, 事件的独立性往往会根据实际问题进行判断, 例如:

- 甲乙两人独立进行射击打靶, 由于两人互不影响, 因此甲乙中靶的事件是相互独立的;
- 在机器学习中通常假设训练数据是独立同分布采样;
- 从n件产品中随机抽取2件, 设事件A_i表示第i件是合格品$(i = 1, 2)$. 若有放回抽取, 则事件A_1, A_2是独立. 若不放回抽取, 则不独立.

独立与互斥(互不相容)的关系:

- 事件A和B独立: $P(AB) = P(A)P(B)$, 独立与概率相关, 反映事件的概率属性
- 事件A和B互斥: $AB = \emptyset$, 互斥限于事件之间的运算关系, 与概率无关.

独立与互斥反映事件不同的性质, 无必然联系:

- 事件A和B独立$\not\Rightarrow$事件A和B互斥, 事件A和B互斥$\not\Rightarrow$事件A和B独立(可用例子证明);
- 若事件A 和B满足$P(A)P(B) > 0$, 则有: 事件A和B 独立$\Rightarrow$事件A和B不互斥, 事件A和B互斥$\Rightarrow$事件A和B不独立.

习题. *若事件A, B互斥且$P(A)P(B) > 0$, 下面哪些说法正确:*
i) $P(B|A) > 0$, ii) $P(A|B) = P(A)$, iii) A, B不独立, iv) $P(A|B) = 0$;

若事件A, B独立且$P(A)P(B) > 0$, 下面哪些说法正确:
i) $P(B|A) > 0$, ii) $P(A|B) = P(A)$, iii) $P(A|B) = 0$, iv) $P(AB) = P(A)P(B)$.

定义2.4. *若三个事件A, B, C满足$P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$, $P(BC) = P(B)P(C)$, 且$P(ABC) = P(A)P(B)P(C)$, 称事件A, B, C独立.*

注: 三事件A, B, C的独立性与三事件两两独立有所不同: 三事件A, B, C独立$\Rightarrow$三事件A, B, C两两独立; 反之不一定成立, 还需要满足$P(ABC) = P(A)P(B)P(C)$. 下面定义n个事件的独立性:

定义2.5. *设$A_1, A_2, \cdots, A_n$为n个事件. 若$1 \leq k \leq n$, $1 \leq i_1 \leq i_2 \leq \cdots i_k \leq n$,*

$$P(A_{i_1} \cdots A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}),$$

即$A_1, A_2, \cdots, A_n$中任意k $(k \leq n)$ 件事件独立, 则称$A_1, A_2, \cdots, A_n$相互独立.

例2.12. 三人独立地破译一份密码, 每人单独能破译的概率分别为$1/5, 1/3, 1/4$, 问三人中至少有一人能破译密码的概率.

解. 设事件A_i={第i个人破译密码} $(i = 1, 2, 3)$, 则有$P(A_1) = 1/5$, $P(A_2) = 1/3$, $P(A_3) = 1/4$.

[第一种方法: 容斥原理直接+ 独立性]

$$\begin{aligned} & P(A_1 \cup A_2 \cup A_3) \\ = \ & P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3) \\ = \ & \frac{1}{5} + \frac{1}{4} + \frac{1}{3} - \frac{1}{20} - \frac{1}{15} - \frac{1}{12} + \frac{1}{60} = 0.6 \end{aligned}$$

[第二种方法: 独立性假设]

$$P(A_1 \cup A_2 \cup A_3) = 1 - P(\bar{A}_1\bar{A}_2\cdots\bar{A}_n) = 1 - P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3) = 1 - \frac{4}{5}\cdot\frac{2}{3}\cdot\frac{3}{4} = 0.6.$$

□

设独立事件$A_1, A_2, \cdots, A_n$发生的概率分别为$p_1, p_2, \cdots, p_n$, 则$A_1, A_2, \cdots, A_n$至少有一事件发生的概率为

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - P(\bar{A}_1\bar{A}_2\cdots\bar{A}_n) = 1 - (1 - p_1)(1 - p_2)\cdots(1 - p_n).$$

事件$A_1, A_2, \cdots, A_n$中至少有一事件不发生的概率为

$$P(\bar{A}_1 \cup \bar{A}_2 \cup \cdots \cup \bar{A}_n) = 1 - P(A_1A_2\cdots A_n) = 1 - P_1P_2\cdots P_n.$$

小概率原理: 事件A在一次试验中发生的概率非常小, 但经过多次独立地重复试验, 事件A是必然发生的, 数学上称为小概率原理. 即, 独立事件$A_1, A_2, \cdots, A_n$中每一事件发生的概率$P(A_i) = p$ $(i \in [n])$都非常小, 但随着$n \to \infty$, 有

$$P(A_1A_2\cdots A_n) = 1 - P(\bar{A}_1\bar{A}_2\cdots\bar{A}_n) = 1 - (1 - p)^n \to 1.$$

独立重复多次的小概率事件亦可成立必然事件.

独立事件$A_1, A_2, \cdots, A_n$发生的概率$P(A_i) = p$ $(i \in [n])$, 则n个事件中恰有k个事件发生的概率为

$$\binom{n}{k}p^k(1 - p)^{n-k}.$$

例2.13. 在冷战时期, 美国的导弹精度99%, 苏联的导弹精度60%, 但苏联的导弹数量特别多, 导弹的数量能否弥补精度的不足?

解. 苏联的策略: 每次独立发射n枚导弹. 设事件A_i={第i枚能命中目标}, 则n枚导弹击中目标的概率为

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - (1 - 0.6)^n \geq 0.99 \Rightarrow n \geq 5,$$

因此, 每次独立发射5枚导弹, 目标的概率高于99%. □

事实上, 在上例中, 如果美国的导弹精度为90%, 苏联的导弹精度为70%, 则苏联每次只需独立发射两枚导弹即可达到91%.

例2.14. *一串电路图: A, B, C, D, E, F, G是电路元件, 电路元件各自下方的数字表示正常工作的概率. 若各电路元件之间相互独立. 求电路正常工作的概率.*

解. 若电路正常工作的事件记为W, 则有$W = A \cap B \cap (C \cup D \cup E) \cap (F \cup G) \cap H$. 根据独立性假设有

$$P(W) = P(A)P(B)P(C \cup D \cup E)P(F \cup G)P(H).$$

再根据$P(C \cup D \cup E) = 1 - P(\bar{C})P(\bar{D})P(\bar{E}) = 1 - (0.3)^3 = 0.973$和$P(F \cup G) = 1 - P(\bar{E})P(\bar{G}) = 0.9375$, 可得$P(W) = 0.782$. □

2.3 案例分析

2.3.1 利用独立性计算矩阵的乘法

例2.15. *给定$n \times n$的矩阵A, B, C (n非常大), 验证$AB = C$是否成立?*

针对该问题, 若直接执行矩阵乘法并验算, 时间复杂度为$O(n^3)$; 采用分治法的复杂度为$O(n^{\log_2^7})$, 目前最好$O(n^{2.37})$. 下面利用算法的独立性来判断$AB = C$是否成立.

核心思想: 独立随机产生一个向量$r \in \{0, 1\}^d$, 判断

$$A(Br) = Cr?$$

计算$A(Br)$和Cr的计算复杂度均为$O(n^2)$, 但$A(Br) = Cr$并不能得出$AB = C$; 将这个过程独立重复k, 则可以以很高的概率证明$AB = C$, 这个算法称为Freivalds算法.

Freivalds算法:

Input: A, B, C
Output: Yes/No
For $i = 1 : k$
Generate an $n \times 1$ random $0 - 1$ vector r
Compute $p = A \times (Br) - Cr$
If $p \neq 0$ then Return 'No'.
EndFor
Return 'Yes'.

该算法的计算复杂度为$O(kn^2)$. 现在研究算法的正确性: 若返回'No', 则$AB \neq C$; 若返回'Yes', 并不一定有$AB = C$, 下面计算$AB = C$成立的概率.

设$D = AB - C \neq 0$, 则D必存在某一些不为0的元素, 不妨设为$d_{11} \neq 0$. 在一次循环中, 如果$p = 0$,

表 1: 训练数据集

训练集	特征1	特征2	$\cdots$	特征n	标记
$\mathbf{x}_1$	x_{11}	x_{12}	$\cdots$	x_{1n}	y_1
$\mathbf{x}_2$	x_{21}	x_{22}	$\cdots$	x_{2n}	y_2
$\vdots$					$\vdots$
$\mathbf{x}_m$	x_{m1}	x_{m2}	$\cdots$	x_{mn}	y_m

即$Dr = 0$ $(r = (r_1, r_2, \ldots, r_n)^\top)$, 则

$$r_1 = \frac{-\sum_{j=2}^n d_{1j} r_j}{d_{11}}$$

由于r_1以$1/2$的概率从$\{0,1\}$中随机选择的, 因此上式成立的概率不超过$1/2$, 由于k重循环是独立运行的, 因此k重循环中同时有$p = 0$的概率不超过$1/2^k$, 最后取$k = (\log n)$.

算法的计算复杂度为$O(n^2 \log n)$, 若算法返回'No', 则$AB \neq C$; 若返回'Yes', $P(AB = C) > 1 - 1/n$.

2.3.2 朴素贝叶斯分类器

在机器学习中, 已知给出一个训练数据集$\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_m, y_m)\}$, 如表 1所示. 称$x_i$为示例, y_i为分类标记, 这里不妨假设考虑二类分类问题, 即$y_i \in \{+1, -1\}$, 同时假设所有的特征都是离散值.

如果新来一个示例$\mathbf{x}' = (x'_1, x'_2, \ldots, x'_n)$, 如何给出分类标记$y' = \pm 1$? 这里介绍朴素贝叶斯分类器.

定义2.6 (条件独立性)**.** *对于随机三事件A, B, C, 如果有$P(AB|C) = P(A|C)P(B|C)$成立, 称事件A和B在条件C下是独立的, 简称条件独立.*

朴素贝叶斯分类的基本假设: 在分类标记给定的条件下, 特征之间是独立的, 即

$$P[X_1 = x'_1, X_2 = x'_2, \ldots, X_n = x'_n | y' = +1] = \prod_{i=1}^n P[X_i = x'_i | y' = +1],$$
$$P[X_1 = x'_1, X_2 = x'_2, \ldots, X_n = x'_n | y' = -1] = \prod_{i=1}^n P[X_i = x'_i | y' = -1].$$

基于特征的条件独立性假设, 对于新的示例$x' = (x'_1, x'_2, \ldots, x'_n)$, 根据贝叶斯公式可得:

$$\begin{aligned}
&P(y' = +1|x') \\
&= P(y' = +1|X_1 = x'_1, \ldots, X_n = x'_n) \\
&= \frac{P(y' = +1)P(X_1 = x'_1, \ldots, X_n = x'_n | y' = +1)}{P(y' = +1)P(X_1 = x'_1, \ldots, X_n = x'_n | y' = +1) + P(y' = +1)P(X_1 = x'_1, \ldots, X_n = x'_n | y' = +1)} \\
&= \frac{P(y' = +1)\prod_{i=1}^n P[X_i = x'_i | y' = +1]}{P(y' = +1)\prod_{i=1}^n P[X_i = x'_i | y' = +1] + P(y' = +1)\prod_{i=1}^n P[X_i = x'_i | y' = +1]}
\end{aligned}$$

同理可得

$$P(y' = -1|x') = \frac{P(y' = -1)\prod_{i=1}^n P[X_i = x'_i | y' = -1]}{P(y' = +1)\prod_{i=1}^n P[X_i = x'_i | y' = +1] + P(y' = +1)\prod_{i=1}^n P[X_i = x'_i | y' = +1]}$$

通过训练数据集计算$P(y'=-1)$, $P(y'=+1)$, $P[X_i=x_i'|y'=+1]$以及$P[X_i=x_i'|y'=-1]$.

对于新示例x', 朴素贝叶斯最后分类的结果为:
如果$P(y'=-1|x')\geq P(y'=-1|x')$, 则$y'=-1$, 否则为$y'=+1$.