概率为

$$\mathrm{P}(\text{飞机不出现, 报警}) = \mathrm{P}(A^c \cap B) = \mathrm{P}(A^c)\mathrm{P}(B|A^c) = 0.95 \times 0.10 = 0.095,$$

$$\mathrm{P}(\text{飞机出现, 未报警}) = \mathrm{P}(A \cap B^c) = \mathrm{P}(A)\mathrm{P}(B^c|A) = 0.05 \times 0.01 = 0.000\,5.$$

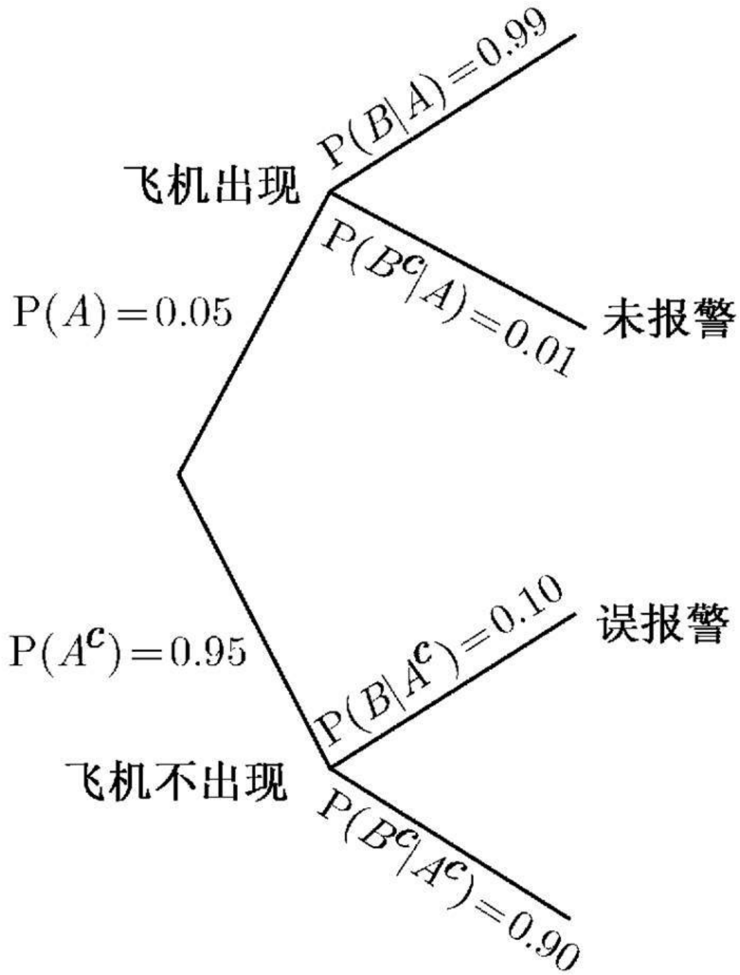

图 1.9　例1.9有关雷达探测的事件的序贯树形图表示

由上例的启示,我们可以利用序贯树形图计算概率,规则如下.

(a) 设立一个序贯树形图,让关心的事件处于图的末端(叶子),由根结点一直到叶子的路径上的每一个结点代表一个事件.而我们所关心的事件的发生是由根结点一直到叶子的一系列事件发生的结果.

(b) 在路径的每个分枝上写上相应的条件概率.

(c) 叶子所代表的事件是相应的分枝上的条件概率的乘积.

数学上可以这样来表示：事件 A 发生的充要条件是一系列事件 $A_1, \cdots, A_n$ 全都发生,即 $A = A_1 \cap A_2 \cap \cdots \cap A_n$. A 发生就是 A_1 发生,接着 A_2 发生等,正如序贯树形图上

n 个结点上的事件顺次发生. A 发生的概率由如下规则给出(也可见图1.10).

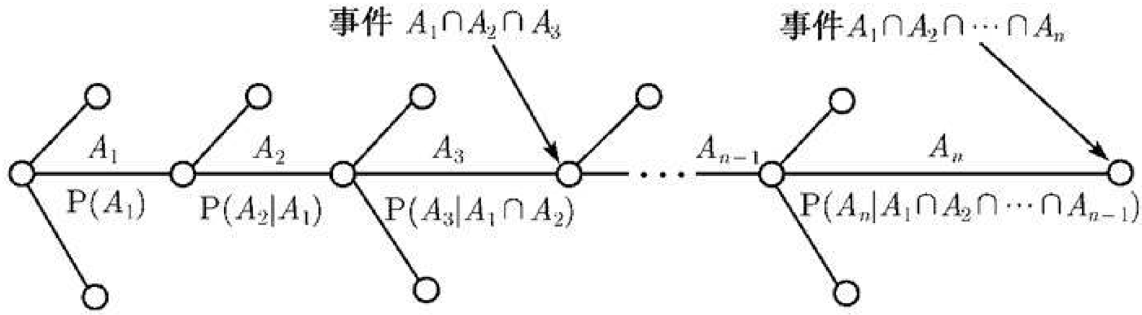

图 1.10　乘法规则的序贯树形图表示. 事件 $A = \cap_{i=1}^{n} A_i$ 用一段路径表示, 或等价地用这一段路径的末端叶子表示, 而路径上的每段树枝表示相应的事件 $A_1, \cdots, A_n$. 在树枝的旁边同时注明相应的条件概率.

路径的末端相应于事件 $A = A_1 \cap A_2 \cap \cdots \cap A_n$, 其概率为由根部到该点的树枝上标示的条件概率的乘积

$$\mathrm{P}(\cap_{i=1}^{n} A_i) = \mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\mathrm{P}(A_3|A_1 \cap A_2) \cdots \mathrm{P}(A_n| \cap_{i=1}^{n-1} A_i).$$

注意, 在图上每一个中间的点也代表一个事件, 例如第 $i+1$ 个结点代表事件 $A_1 \cap A_2 \cap \cdots \cap A_i$. 它们的概率等于相应的条件概率的乘积, 这些乘积因子都已在相应的 树枝下方列明. 例如, 事件 $A_1 \cap A_2 \cap A_3$ 相应于图上的第4个结点, 其概率为

$$\mathrm{P}(A_1 \cap A_2 \cap A_3) = \mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\mathrm{P}(A_3|A_1 \cap A_2)$$

乘法规则

假定所有涉及的条件概率都是正的, 我们有

$$\mathrm{P}(\cap_{i=1}^{n} A_i) = \mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\mathrm{P}(A_3|A_1 \cap A_2) \cdots \mathrm{P}(A_n| \cap_{i=1}^{n-1} A_i).$$

现在我们来证明乘法规则: 由下列恒等式

$$\mathrm{P}(\cap_{i=1}^{n} A_i) = \mathrm{P}(A_1)\frac{\mathrm{P}(A_2 \cap A_1)}{\mathrm{P}(A_1)} \cdot \frac{\mathrm{P}(A_3 \cap A_1 \cap A_2)}{\mathrm{P}(A_1 \cap A_2)} \cdots \frac{\mathrm{P}(\cap_{i=1}^{n} A_i)}{\mathrm{P}(\cap_{i=1}^{n-1} A_i)}.$$

再利用条件概率的定义, 上式右端变成

$$\mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\mathrm{P}(A_3|A_1 \cap A_2) \cdots \mathrm{P}(A_n| \cap_{i=1}^{n-1} A_i).$$

对于两个事件 A_1 和 A_2 的情况, 乘法规则就是条件概率的定义.

例 1.10　从52张扑克牌中连续无放回地抽取3张牌. 我们希望求出3张牌中没有红桃的概率. 假定在抽取的时候, 一堆牌中的每一张牌都是等可能地被抽取的. 根据对称性, 52张牌中任意3张牌的组合被抽取的可能性都是相同的. 一个想法简单但是计算麻烦的方法是: 数清楚不含红桃的3张牌的可能组数, 再除以所有3张牌的可能组数. 现在利用试验的序贯树形图表示法以及乘法规则进行计算(见图1.11).

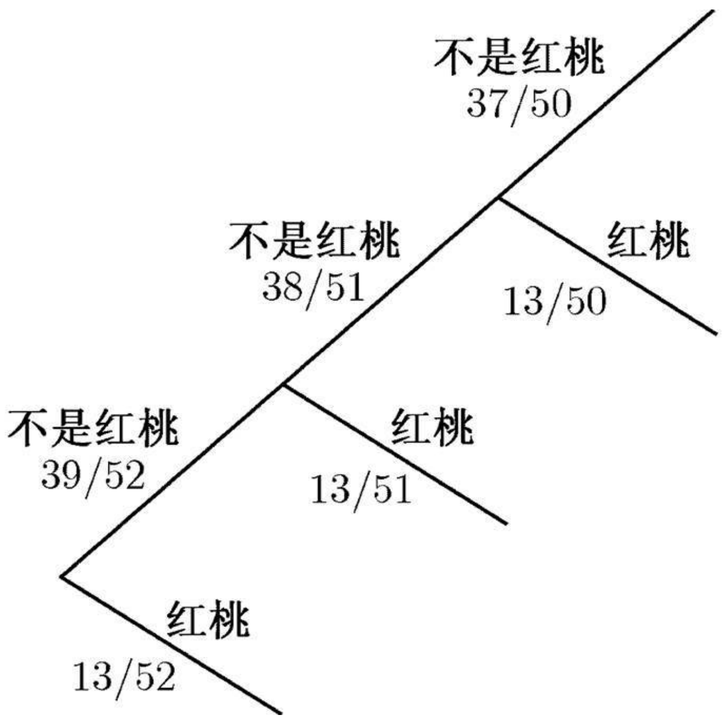

图 1.11　例1.10中抽取3张扑克牌的试验的序贯树形图表示

定义

$$A_i = \{第i张牌不是红桃\}, \quad i = 1, 2, 3.$$

现在利用乘法规则

$$\mathrm{P}(A_1 \cap A_2 \cap A_3) = \mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\mathrm{P}(A_3|A_1 \cap A_2),$$

计算3张牌中没有红桃的概率 $\mathrm{P}(A_1 \cap A_2 \cap A_3)$. 由于52张牌中有39张不是红桃,我们得到

$$\mathrm{P}(A_1) = \frac{39}{52}.$$

由于第一次抽出一张不是红桃,剩下51张牌中有38张不是红桃,因此

$$\mathrm{P}(A_2|A_1) = \frac{38}{51}.$$

最后,由于前面两张不是红桃,剩下50张牌中有37张不是红桃,这样

$$\mathrm{P}(A_3|A_1 \cap A_2) = \frac{37}{50}.$$

这些条件概率列于序贯树形图(图1.11)的相应树枝的上方.现在只需将路径上的(条件)概率相乘,得到

$$P(A_1 \cap A_2 \cap A_3) = \frac{39}{52} \cdot \frac{38}{51} \cdot \frac{37}{50}.$$

注意,由于在序贯树形图上已经标明了许多(条件)概率,其他的一些事件也可以相应地计算.例如

$$P(\text{第一张不是红桃, 第二张牌是红桃}) = \frac{39}{52} \cdot \frac{13}{51}.$$

$$P(\text{第一、第二两张不是红桃, 第三张牌是红桃}) = \frac{39}{52} \cdot \frac{38}{51} \cdot \frac{13}{50}.$$

例 1.11 一个班由4个研究生和12个本科生组成,随机地将这16人分成4个4人组.问每个组分得一个研究生的概率有多大?在这个问题中,什么是随机地分组呢?可以将分组问题看成随机地选位子(不妨将位子 $s_1, \cdots, s_4$ 看成第一组,而将位子 $s_5, \cdots, s_8$ 看成第二组,等等),每个人都有相同的可能性选择16个位子中任意一个位子,当若干个位子被某些学生选定以后,没有选定位子的同学以完全平等的资格去选择剩下的位子.下面基于图1.12所示的序贯树形图,使用乘法规则来计算所需概率.现在设4个研究生的代号为1,2,3,4.考虑事件

$$A_1 = \{\text{学生 1 和 2 分在不同的组}\},$$

$$A_2 = \{\text{学生 1、2 和 3 分在不同的组}\},$$

$$A_3 = \{\text{学生 1、2、3 和 4 分在不同的组}\}.$$

我们所求的概率为 $P(A_3)$.利用乘法规则:

$$P(A_3) = P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2).$$

现在不妨设学生1已经选定了位子,在剩余的15个位子中只有12个位子与学生1分在不同的组内.显然学生2与学生1分在不同组内的可能性为12/15,即

$$P(A_1) = \frac{12}{15}.$$

类似地,当学生1和学生2已经分在2个不同组以后,学生3只有选择剩下14个位子中的8个位子,才能与学生1、2处于不同的组.这说明

$$P(A_2|A_1) = \frac{8}{14}.$$

在学生1、2和3被分派在不同组的条件下,学生4只有在13个位子中选择其中的4个位子之一,才能与他们处于不同的组内.这样

$$P(A_3|A_1 \cap A_2) = \frac{4}{13}.$$

将三个概率相乘,得到所求的概率为

$$\frac{12}{15}\cdot\frac{8}{14}\cdot\frac{4}{13}.$$

反映这种试验的序贯树形图见图1.12.

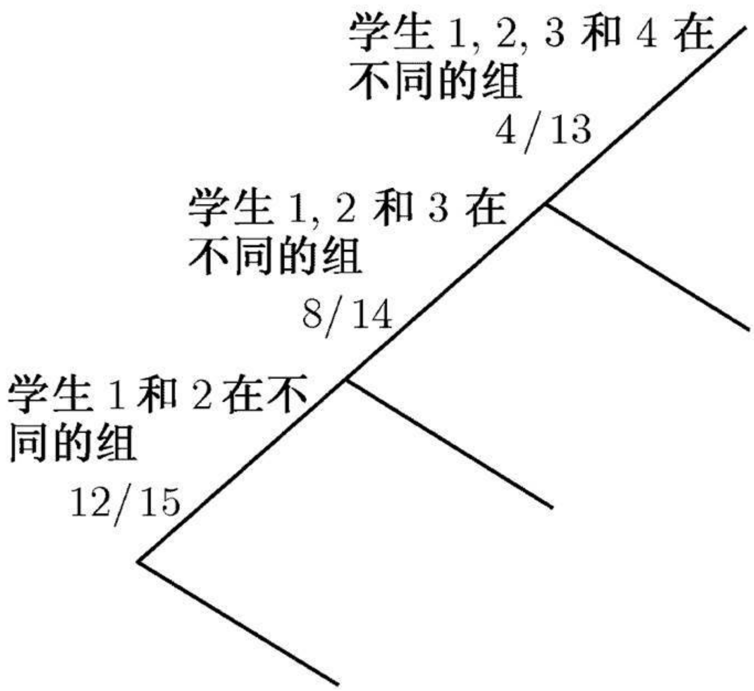

图 1.12　例1.11中学生分组试验的序贯树形图表示

例 1.12(蒙提·霍尔问题,也称三门问题)　这是美国有奖游戏节目中的一个经常出现的智力测验问题.你站在三个封闭的门前,其中一个门后有奖品.当然,奖品在哪一个门后是完全随机的.当你选定一个门以后,你的朋友打开其余两扇门中的一扇空门,显示门后没有奖品.此时你可以有两种选择,保持原来的选择,或改选另一扇没有被打开的门.当你作出最后选择以后,如果打开的门后有奖品,这个奖品就归你.现在有三种策略:

(a) 坚持原来的选择;

(b) 改选另一扇没有被打开的门;

(c) 你首先选择1号门,当你的朋友打开的是2号空门,你不改变主意.当你的朋友打开的是3号空门你改变主意,选择2号门.

最好的策略是什么呢?现在计算在各种策略之下赢得奖品的概率.

在策略 (a) 之下,你的初始选择会决定你的输赢.由于奖品的位置是随机地确定的,你得奖的概率只能是1/3.

在策略（b）之下,如果奖品的位置在你原来指定的门后(概率为1/3),由于你改变了主意,因而失去了获奖的机会.如果奖品的位置不在你原来指定的门后(概率2/3),而你的朋友又将没有奖品的那一扇门打开,当你改变选择的时候,你改变选择后所指定的门后一定有奖品.所以你获奖的概率为2/3.因此(b)比（a）好.

在策略（c）之下,由于提供的信息不够充分,还不能确定你赢得奖品的概率.答案依赖于你的朋友打开空门的方式.现在讨论两种情况.

第一种情况是：当奖品的位置是在1号门后,假定你的朋友总是打开2号空门(当奖品是在2号或3号门后的时候,你的朋友没有选择的余地).现在假定奖品是在1号门后(概率为1/3),你的朋友打开2号门,你不改主意,你得到奖品.当奖品在2号门后面的时候(概率为1/3),你的朋友打开3号空门,你改变主意,你也得到奖品.当奖品在3号门后面的时候(概率为1/3),你的朋友打开2号空门,你不改变主意,你就失去了得奖的机会.这样,你获奖的概率为2/3.说明在这种情况下,策略（c）与策略（b）一样好.

第二种情况是：假定奖品是在1号门后,你的朋友随机地打开2号门或3号门(概率各为1/2).当奖品在1号门后的情况下(概率为1/3),你的朋友打开2号门,此时按你的策略,你不改主意,得到了奖品(概率1/6).但是,如果你的朋友打开的是3号空门,此时你改变了主意,失去了得奖的机会.如果奖品是在2号门后(概率1/3),你的朋友打开3号空门,按你的策略,你改变主意,你就赢得奖品.如果奖品是在3号门后(概率1/3),你的朋友打开2号空门,按你的策略你不改变主意,你就失去奖品.综合起来,在你的朋友这种开门策略之下,你赢得奖品的概率为1/6+1/3=1/2.这时候,策略（c)比策略（b）差.

1.4 全概率定理和贝叶斯准则

本节中我们将讨论条件概率的某些应用.我们首先引入一个计算事件概率的定理.

全概率定理

设 $A_1, A_2, \cdots, A_n$ 是一组互不相容的事件,形成样本空间的一个分割(每一个试验结果必定使得其中一个事件发生).又假定对每一个 i ,$\mathrm{P}(A_i) > 0$.则对于任何事件 B ,下列公式成立

$$\begin{aligned}\mathrm{P}(B) &= \mathrm{P}(A_1 \cap B) + \cdots + \mathrm{P}(A_n \cap B)\\ &= \mathrm{P}(A_1)\mathrm{P}(B|A_1) + \cdots + \mathrm{P}(A_n)\mathrm{P}(B|A_n).\end{aligned}$$

图1.13形象化地展示了全概率定理的内容并给出了证明.直观上,将样本空间分割成若干事件 A_i 的并($A_1, \cdots, A_n$ 形成样本空间的一个分割),然后任意事件 B 的概率等于事件 B 在 A_i 发生的情况下的条件概率的加权平均,而权数刚好等于这些事件 A_i 的无条件概率.这条定理的一个主要应用是计算事件 B 的概率.直接计算事件 B 的概率有点难度,但是若条件概率 $\mathrm{P}(B|A_i)$ 是已知的或是很容易推导计算时,全概率定理就成为计算 $\mathrm{P}(B)$ 的有力工具.应用这条定理的关键是找到合适的分割 $A_1, \cdots, A_n$,而合适的分割又与问题的实际背景有关.

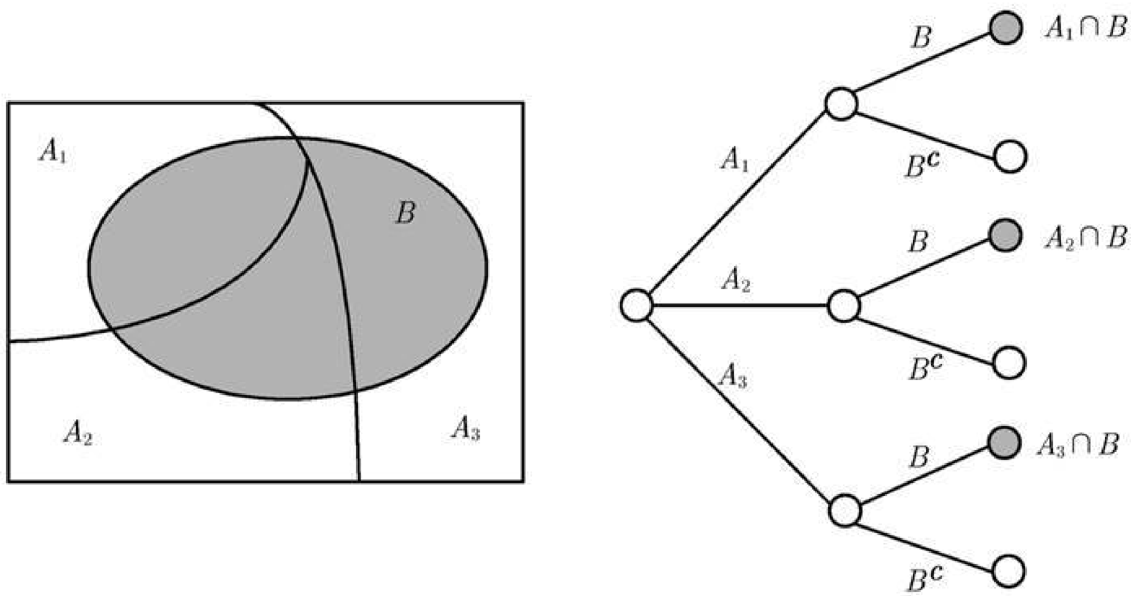

图 1.13 全概率定理的形象化展示和证明.由于事件 $A_1, A_2, \cdots, A_n$ 形成样本空间的一个分割,事件 B 可以分解成不相交的 n 个事件的并,即

$$B = (A_1 \cap B) \cup \cdots \cup (A_n \cap B).$$

利用可加公理,得到

$$\mathrm{P}(B) = \mathrm{P}(A_1 \cap B) + \cdots + \mathrm{P}(A_n \cap B).$$

利用条件概率之定义,我们得到

$$\mathrm{P}(A_i \cap B) = \mathrm{P}(A_i)\mathrm{P}(B|A_i).$$

将上式代入前一式中得到

$$\mathrm{P}(B) = \mathrm{P}(A_1)\mathrm{P}(B|A_1) + \cdots + \mathrm{P}(A_n)\mathrm{P}(B|A_n).$$

我们也可以用等价的序贯树形图来说明全概率定理(右图).叶子 $A_i \cap B$ 的概率等于由根部到叶子上的概率的乘积 $\mathrm{P}(A_i)\mathrm{P}(B|A_i)$.而事件 B 由图上显示的3个叶子组成,将它们的概率相加就得到$\mathrm{P}(B)$

例 1.13　你参加一个棋类比赛,其中50%是一类棋手,你赢他们的概率为0.3;25%是二类棋手,你赢他们的概率是0.4;剩下的是三类棋手,你赢他们的概率是0.5.从他们中间随机地选一位棋手与你比赛,你的胜算有多大?

记 A_i 表示你与 i 类棋手相遇的事件.依题意

$$\mathrm{P}(A_1) = 0.5, \quad \mathrm{P}(A_2) = 0.25, \quad \mathrm{P}(A_3) = 0.25.$$

记 B 为你赢得比赛的事件.我们有

$$\mathrm{P}(B|A_1) = 0.3, \quad \mathrm{P}(B|A_2) = 0.4, \quad \mathrm{P}(B|A_3) = 0.5.$$

这样,利用全概率定理,你在比赛中胜出的概率为

$$\begin{aligned}\mathrm{P}(B) &= \mathrm{P}(A_1)\mathrm{P}(B|A_1) + \mathrm{P}(A_2)\mathrm{P}(B|A_2) + \mathrm{P}(A_3)\mathrm{P}(B|A_3) \\ &= 0.5 \cdot 0.3 + 0.25 \cdot 0.4 + 0.25 \cdot 0.5 \\ &= 0.375.\end{aligned}$$

例 1.14　你抛掷一个均匀的有4个面的骰子.如果得到1或2,你可以再抛掷一次,否则就停止抛掷.你抛掷得到的点数总和至少为4的概率有多大?

记 A_i 为第一次抛掷均匀骰子后得到的点数为 i 的事件.注意,对每一个 i ,$\mathrm{P}(A_i) = 1/4$.记 B 为抛掷得到的点数总和至少为4的事件.在 A_1 发生的条件下,只有第二次抛掷得到3或4,总点数才能至少为4,这样,事件 B 的条件概率为1/2.类似地,如果第一次抛掷时 A_2 发生,只有当第二次抛掷得到2、3或4时,事件 B 才发生,相应的条件概率为3/4.如果第一次抛掷时 A_3 发生,此时不容许抛掷第二次,在这种情况下得到的点数总和在4以下.[5]因此

> [5]如果第一次抛掷时 A_4 发生,虽然不容许第二次抛掷骰子,但是你得到的点数总和已为4.——译者注

$$\mathrm{P}(B|A_1) = \frac{1}{2}, \quad \mathrm{P}(B|A_2) = \frac{3}{4}, \quad \mathrm{P}(B|A_3) = 0, \quad \mathrm{P}(B|A_4) = 1.$$

利用全概率定理,得到

$$\mathrm{P}(B)=\frac{1}{4}\cdot\frac{1}{2}+\frac{1}{4}\cdot\frac{3}{4}+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 1=\frac{9}{16}.$$

在具有序贯特征的试验中,可以多次重复地利用全概率定理进行概率计算.下面是一个例子.

例 1.15 爱丽丝在上一门概率课.在每周周末的时候,她可能跟上课程或跟不上课程.如果她在某一周是跟上课程的,那么她在下周跟上课程的概率为0.8(下周跟不上课程的概率为0.2).然而,如果她在某一周没有跟上课程,那么她在下周跟上课程的概率变为0.4(下周跟不上课程的概率为0.6).现在假定,在第一周上课以前认为她是能够跟上课程的.经过三周的学习,她能够跟上课程的概率有多大?

令 U_i 和 B_i 分别表示经过 i 周学习后跟上和跟不上课程的事件.按照全概率定理, $\mathrm{P}(U_3)$ 可由下式给出

$$\mathrm{P}(U_3)=\mathrm{P}(U_2)\mathrm{P}(U_3|U_2)+\mathrm{P}(B_2)\mathrm{P}(U_3|B_2)=\mathrm{P}(U_2)\cdot 0.8+\mathrm{P}(B_2)\cdot 0.4.$$

对于 $\mathrm{P}(U_2)$ 和 $\mathrm{P}(B_2)$,又可以利用全概率定理

$$\mathrm{P}(U_2)=\mathrm{P}(U_1)\mathrm{P}(U_2|U_1)+\mathrm{P}(B_1)\mathrm{P}(U_2|B_1)=\mathrm{P}(U_1)\cdot 0.8+\mathrm{P}(B_1)\cdot 0.4,$$
$$\mathrm{P}(B_2)=\mathrm{P}(U_1)\mathrm{P}(B_2|U_1)+\mathrm{P}(B_1)\mathrm{P}(B_2|B_1)=\mathrm{P}(U_1)\cdot 0.2+\mathrm{P}(B_1)\cdot 0.6.$$

最后,由于爱丽丝在刚刚开始上课的时候是能够跟上课程的,我们有

$$\mathrm{P}(U_1)=0.8,\quad \mathrm{P}(B_1)=0.2.$$

从前面三个方程式解得

$$\mathrm{P}(U_2)=0.8\cdot 0.8+0.2\cdot 0.4=0.72,$$
$$\mathrm{P}(B_2)=0.8\cdot 0.2+0.2\cdot 0.6=0.28,$$

再利用关于 $\mathrm{P}(U_3)$ 的等式,得到

$$\mathrm{P}(U_3)=0.72\cdot 0.8+0.28\cdot 0.4=0.688.$$

我们也可以为计算 $\mathrm{P}(U_3)$ 构造一个试验的序贯树形图.将随机事件 U_3 进行分解,利用概率论的乘法与加法规则计算 $\mathrm{P}(U_3)$.然而,有时候,基于全概率定理的计算方法更加方便.例如,我们希望计算经过20周的学习以后,爱丽丝能够跟上课程的概率 $\mathrm{P}(U_{20})$.此时,按照序贯树形图进行计算十分烦琐,因为树形图有20层,有 2^{20} 个树叶.另一方面,利用全概率定理,得到递推公式

$$\mathrm{P}(U_{i+1})=\mathrm{P}(U_i)\cdot 0.8+\mathrm{P}(B_i)\cdot 0.4,$$
$$\mathrm{P}(B_{i+1})=\mathrm{P}(U_i)\cdot 0.2+\mathrm{P}(B_i)\cdot 0.6,$$

加上初始条件 $\mathrm{P}(U_1)=0.8$、 $\mathrm{P}(B_1)=0.2$ 后,那么在计算机上计算是十分简便的.

推理和贝叶斯准则

全概率定理是与著名的贝叶斯准则联系在一起的. 贝叶斯准则将形如 $P(A|B)$ 的条件概率与形如 $P(B|A)$ 的条件概率联系起来.

贝叶斯准则

设 $A_1, A_2, \cdots, A_n$ 是一组互不相容的事件, 形成样本空间的一个分割(每一个试验结果必定使得其中一个事件发生). 又假定对每一个 i, $P(A_i) > 0$. 则对于任何事件 B, 只要它满足 $P(B) > 0$, 下列公式成立

$$\begin{aligned}P(A_i|B) &= \frac{P(A_i)P(B|A_i)}{P(B)} \\ &= \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)}.\end{aligned}$$

为证明贝叶斯准则, 只需注意到 $P(A_i)P(B|A_i)$ 与 $P(A_i|B)P(B)$ 是相等的, 因为根据条件概率的定义它们都等于 $P(A_i \cap B)$, 这样得到了第一个等式. 至于第二个等式, 只需对 $P(B)$ 利用全概率公式即可.

贝叶斯准则还可以用来进行**因果推理**. 有许多“原因”可以造成某一“结果”. 现在设我们观察到某一结果, 希望推断造成这个结果出现的“原因”. 现在设事件 $A_1, \cdots, A_n$ 是原因, 而 B 代表由原因引起的结果. $P(B|A_i)$ 表示在因果模型中由“原因” A_i 造成结果 B 出现的概率(见图1.14). 当观察到结果 B 的时候, 我们希望反推结果 B 是由原因 A_i 造成的概率 $P(A_i|B)$. $P(A_i|B)$ 为由于代表新近得到的信息 B 之后 A_i 出现的概率, 称之为**后验概率**, 而原来的 $P(A_i)$ 就称为**先验概率**.

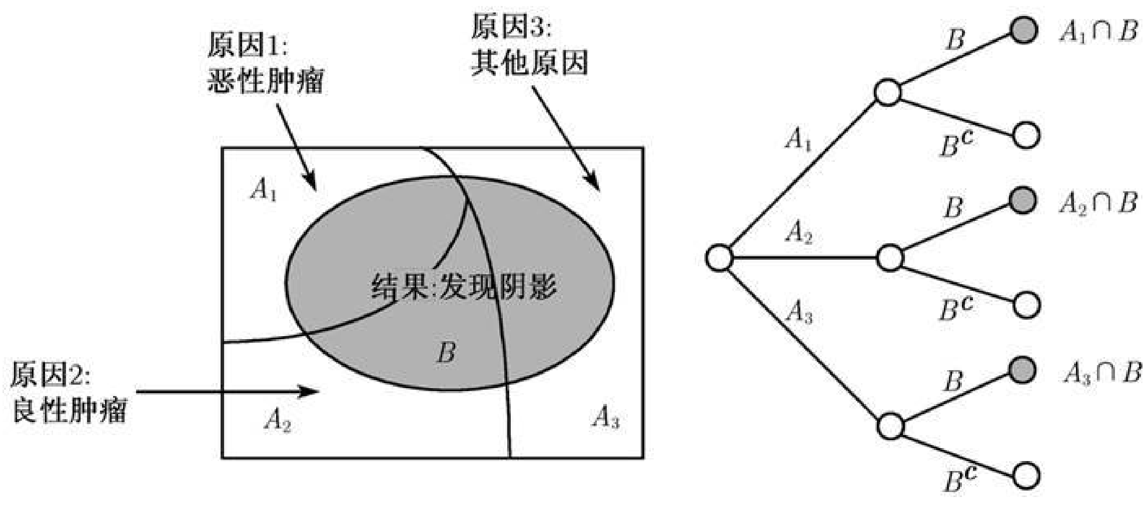

图 1.14　一个蕴涵于贝叶斯准则中的推论的例子. 我们在某病人的X光片中发现一个阴影(事件 B, 代表“结果”). 我们希望对造成这种结果的三个原因进行分析. 这三个原因彼此不相容, 并且造成这个结果的原因一定是三者之一：原因1(事件A_1)是恶性肿瘤, 原因2(事件A_2)是良性肿瘤, 原因3(事件A_3)是肿瘤外的其他原因. 假定我们已经知道 $P(A_i)$ 和 $P(B|A_i), i = 1, 2, 3$. 现在我们已经发现了阴影(事件 B 发生), 利用贝叶斯准则, 这些原因的条件概率为

$$\mathrm{P}(A_i|B)=\frac{\mathrm{P}(A_i)\mathrm{P}(B|A_i)}{\mathrm{P}(A_1)\mathrm{P}(B|A_1)+\mathrm{P}(A_2)\mathrm{P}(B|A_2)+\mathrm{P}(A_3)\mathrm{P}(B|A_3)},\quad i=1,2,3.$$

在右图给出了一个序贯树形图,可用序贯树形图给出条件概率计算的另一种等价的解释.图中第一个深灰的叶子表示恶性肿瘤并出现阴影,其概率为 $\mathrm{P}(A_1\cap B)$,且所有深灰的叶子表示片子中出现阴影,其概率为 $\mathrm{P}(B)$.而由恶性肿瘤造成阴影的条件概率 $\mathrm{P}(A_1|B)$ 是两个概率相除的结果

例 1.16 现在回到雷达探测器的例1.9和图1.9.记

$$A=\{飞机出现\},$$

$$B=\{雷达报警\}.$$

例 1.9 中给出的条件为

$$\mathrm{P}(A)=0.05,\quad \mathrm{P}(B|A)=0.99,\quad \mathrm{P}(B|A^c)=0.1.$$

在贝叶斯准则中令 $A_1=A$ 和 $A_2=A^c$,得到

$$\begin{aligned}\mathrm{P}(飞机出现|雷达报警)&=\mathrm{P}(A|B)\\&=\frac{\mathrm{P}(A)\mathrm{P}(B|A)}{\mathrm{P}(A)\mathrm{P}(B|A)+\mathrm{P}(A^c)\mathrm{P}(B|A^c)}\\&=\frac{0.05\cdot0.99}{0.05\cdot0.99+0.95\cdot0.1}\\&\approx0.342\ 6.\end{aligned}$$

例 1.17 现在回到例1.13的棋类比赛问题.此处 A_i 表示你与 i 类棋手相遇的事件.由例中给出的条件知,

$$\mathrm{P}(A_1)=0.5,\quad \mathrm{P}(A_2)=0.25,\quad \mathrm{P}(A_3)=0.25.$$

记 B 表示你赢得比赛的事件,你胜出的概率为

$$\mathrm{P}(B|A_1)=0.3,\quad \mathrm{P}(B|A_2)=0.4,\quad \mathrm{P}(B|A_3)=0.5.$$

现在假定你已经得胜,问你的对手为一类棋手的概率 $\mathrm{P}(A_1|B)$ 有多大?

利用贝叶斯准则得

$$\begin{aligned}\mathrm{P}(A_1|B)&=\frac{\mathrm{P}(A_1)\mathrm{P}(B|A_1)}{\mathrm{P}(A_1)\mathrm{P}(B|A_1)+\mathrm{P}(A_2)\mathrm{P}(B|A_2)+\mathrm{P}(A_3)\mathrm{P}(B|A_3)}\\&=\frac{0.5\cdot0.3}{0.5\cdot0.3+0.25\cdot0.4+0.25\cdot0.5}\\&=0.4.\end{aligned}$$

例 1.18(假阳性之迷) 设对于某种少见的疾病的检出率为0.95：如果一个被检的人有这种疾病,其检查结果为阳性的概率为0.95;如果该人没有这种疾病,其检查结果为阴性的概率是0.95.现在假定某一人群中患有这种病的概率为0.001,并从这个总体中随机地抽取一个人进行检测,检查结果为阳性.现在问这个人患这种病的概率有多大?

记 A 为这个人有这种疾病,B 为经检验这个人为阳性.利用贝叶斯准则,

$$\begin{aligned}
\mathrm{P}(A|B) &= \frac{\mathrm{P}(A)\mathrm{P}(B|A)}{\mathrm{P}(A)\mathrm{P}(B|A)+\mathrm{P}(A^c)\mathrm{P}(B|A^c)} \\
&= \frac{0.001\cdot 0.95}{0.001\cdot 0.95+0.999\cdot 0.05} \\
&\approx 0.018\,7.
\end{aligned}$$

尽管检验方法非常精确,一个经检测为阳性的人仍然不大可能真正患有这种疾病(患有该疾病的概率小于2%).根据《经济学人》（*The Economist*）1999年2月20日的报道,在一家美国著名的大医院中80%的受访者不知道这类问题的正确答案,大部分人回答这个经检测为阳性的人患病的概率为0.95!

1.5 独立性

1.4节中我们引入了条件概率 $P(A|B)$ 的概念. 这个条件概率刻画了事件 B 的发生给事件 A 带来的信息. 一个有趣且重要的特殊情况是事件 B 的发生并没有给事件 A 带来新的信息, 它没有改变事件 A 发生的概率, 即

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

在上述等式成立的情况下, 我们称事件 A 是**独立**于事件 B 的. 注意, 由条件概率的定义可知 $P(A|B) = P(A \cap B)/P(B)$, 上式等价于

$$P(A \cap B) = P(A)P(B).$$

我们将后者作为事件 A 和事件 B 相互独立的正式定义, 其原因是后者包括了 $P(B) = 0$ 的情况, 而当 $P(B) = 0$ 的时候, $P(A|B)$ 是没有定义的. 在这个关系中 A 和 B 具有对称的地位. 因此 A 独立于 B 蕴涵着 B 独立于 A. 这样我们可以称 A 和 B 是相互独立的, 或 A 和 B 是**相互独立的事件**.

人们容易从直观判定独立性. 例如, 若它们分别是在两个不同的并且没有相互作用的物理过程的控制下发生的事件, 我们就可以判定它们相互独立. 另一方面, 事件之间的独立性不能直观地从样本空间中的事件看出来. 通常认为, 若两个事件互不相容, 就可以判定它们相互独立, 事实上, 恰巧相反, 若事件 A 和事件 B 互不相容, 并且 $P(A) > 0$ 和 $P(B) > 0$ 成立, 则它们永远不会相互独立, 因为 $A \cap B = \varnothing$, 从而 $P(A \cap B) = 0 \neq P(A)P(B)$. 例如, A 和 A^c在$P(A) \in (0,1)$ 的情况下是不独立的(除非 $P(A) = 0$, 或 $P(A) = 1$), 这是因为 A 发生可以确切地告诉你 A^c一定不会发生, A 的发生与否的确会给事件 A^c 的发生与否带来信息.

例 1.19 考虑连续两次抛掷一个具有4个面的对称的骰子, 假定16种可能的试验结果是等概率的, 每个试验结果的概率为1/16.

(a) 事件

$$A_i = \{第一次抛掷后得\ j\}, \quad B_j = \{第二次抛掷后得\ j\}$$

是否相互独立?我们有

$$\mathrm{P}(A_i \cap B_j) = \mathrm{P}(\text{两次抛掷的结果是}(i,j)) = \frac{1}{16},$$

$$\mathrm{P}(A_i) = \frac{A_i\text{中的试验结果数}}{\text{总的试验结果数}} = \frac{4}{16},$$

$$\mathrm{P}(B_j) = \frac{B_j\text{中的试验结果数}}{\text{总的试验结果数}} = \frac{4}{16}.$$

由于 $\mathrm{P}(A_i \cap B_j) = \mathrm{P}(A_i)\mathrm{P}(B_j)$, 可知 A_i 与 B_j 是相互独立的. 在两次抛掷骰子的试验中, 离散的均匀概率律(等概率模型)蕴涵着两次抛掷的独立性.

(b) 事件

$$A = \{\text{第一次抛掷后得 } 1\}, \quad B = \{\text{两次抛掷的总和为 } 5\}$$

是否相互独立?这个问题的答案不是很明显. 我们有

$$\mathrm{P}(A \cap B) = \mathrm{P}(\text{两次抛掷的结果为}(1,4)) = \frac{1}{16},$$

$$\mathrm{P}(A) = \frac{\text{事件}A\text{中所含的试验结果数}}{\text{所有可能的结果数}} = \frac{4}{16}.$$

事件B由试验结果$(1,4)$、$(2,3)$、$(3,2)$和$(4,1)$组成, 因此

$$\mathrm{P}(B) = \frac{\text{事件}B\text{中所含的试验结果数}}{\text{所有可能的结果数}} = \frac{4}{16}.$$

这样, $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$, 即 A 和 B 相互独立.

(c) 事件

$$A = \{\text{两次抛掷的最大数为 } 2\}, \quad B_j = \{\text{两次抛掷的最小数为 } 2\}$$

是否相互独立?直观上看这两个事件是不独立的, 因为两次抛掷的最小数蕴涵着两次抛掷的最大数的信息. 例如, 如果最小数为2, 最大数不可能为1. 现在用定义证明它们不独立. 我们有

$$\mathrm{P}(A\cap B)=\mathrm{P}(\text{两次抛掷的结果为}(2,2))=\frac{1}{16},$$

同时

$$\mathrm{P}(A)=\frac{A\text{中的试验结果数}}{\text{总的试验结果数}}=\frac{3}{16},$$

$$\mathrm{P}(B)=\frac{B\text{中的试验结果数}}{\text{总的试验结果数}}=\frac{5}{16},$$

得到 $\mathrm{P}(A)\mathrm{P}(B)=15/(16)^2$. $\mathrm{P}(A\cap B)\neq\mathrm{P}(A)\mathrm{P}(B)$, 故它们并不独立.

最后, 我们要指出, 若事件 A 和事件 B 相互独立, 那么 B 发生, 不会对 A 的发生与否提供任何信息. 同样, 凭直观想象, B 不发生, 也不会对 A 的概率提供任何信息. 事实上, 我们可以证明, 若 A 和 B 相互独立, 则 A 和 B^c 也相互独立(见本章后的习题).

1.5.1 条件独立

前面已经提到, 在给定某事件的条件下, 诸事件的条件概率形成符合要求的概率律. 因此我们可以讨论在条件概率律下的独立性. 特别地, 在给定 C 之下, 若事件 A 和事件 B 满足

$$\mathrm{P}(A\cap B|C)=\mathrm{P}(A|C)\mathrm{P}(B|C),$$

则称 A 和 B 在给定 C 之下**条件独立**. 为了导出条件独立的另一个特征, 利用条件概率的定义和乘法规则, 得到

$$\begin{aligned}\mathrm{P}(A\cap B|C)&=\frac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(C)}\\&=\frac{\mathrm{P}(C)\mathrm{P}(B|C)\mathrm{P}(A|B\cap C)}{\mathrm{P}(C)}\\&=\mathrm{P}(B|C)\mathrm{P}(A|B\cap C).\end{aligned}$$

比较前面两组等式的最右端, 只要 $\mathrm{P}(B|C)\neq 0$, 那么 $\mathrm{P}(B|C)$ 这个因子就可以消掉, 得到

$$\mathrm{P}(A|B\cap C)=\mathrm{P}(A|C),$$

这是条件独立的另一个等价定义(要求 $\mathrm{P}(B|C)\neq 0$). 这个等式说明在给定 C 发生的条件之下, 进一步假定 B 也发生, 并不影响事件 A 的条件概率.

有意思的是, A 和 B 两个事件相互独立并不包含条件独立, 反过来也是如此. 下面请看两个例子.

例 1.20 考虑抛掷两枚均匀的硬币. 这个试验的4种可能结果都是等可能的. 令

$$H_1=\{第一枚硬币正面向上\},$$
$$H_2=\{第二枚硬币正面向上\},$$
$$D=\{两枚硬币的试验结果不同\}.$$

事件 H_1 和事件 H_2 是相互独立的. 但是

$$\mathrm{P}(H_1|D)=\frac{1}{2},\quad \mathrm{P}(H_2|D)=\frac{1}{2},\quad \mathrm{P}(H_1\cap H_2|D)=0,$$

这样, $\mathrm{P}(H_1\cap H_2|D)\neq \mathrm{P}(H_1|D)\mathrm{P}(H_2|D)$, 从而 H_1和H_2 并不条件独立.

这个例子可以推广. 对于任何概率模型, 记 A 和 B 是相互独立的事件, C 是一个满足条件 $\mathrm{P}(C)>0$、$\mathrm{P}(A|C)>0$ 和 $\mathrm{P}(B|C)>0$ 的事件, 并且 $A\cap B\cap C$ 为空集. 这样, 由于 $\mathrm{P}(A\cap B|C)=0$ 和 $\mathrm{P}(A|C)\mathrm{P}(B|C)>0$, A 和 B 不可能条件独立(给定C).

例 1.21 有两枚硬币, 一枚蓝的, 一枚红的. 在抛掷硬币之前, 先按1/2的概率随机地选定一枚硬币, 然后进行连续两次独立地抛掷硬币的试验. 硬币是不均匀的. 蓝的硬币在抛掷的时候以0.99的概率正面向上. 而红的那一枚硬币在抛掷的时候以0.01的概率正面向上.

记 B 为选定蓝色的硬币的事件, H_i 为第 i 次抛掷时出现正面向上. 当选定硬币以后, 由于我们抛掷硬币的时候, 两次抛掷的结果不会互相影响, H_1 和 H_2 是相互独立的事件. 这样

$$\mathrm{P}(H_1\cap H_2|B)=\mathrm{P}(H_1|B)\mathrm{P}(H_2|B)=0.99\cdot 0.99.$$

另一方面, H_1 和 H_2 并不独立. 直观上, 当我们知道第一次抛掷的结果是正面向上, 我们就想到这是一枚蓝色的硬币, 此时可以预料到第二次抛掷硬币的结果也是正面向上.[6]数学上, 可如下证明. 利用全概率定理, 我们得到

> [6]因此两次抛掷的结果是不独立的. ——译者注

$$\mathrm{P}(H_1)=\mathrm{P}(B)\mathrm{P}(H_1|B)+\mathrm{P}(B^c)\mathrm{P}(H_1|B^c)=\frac{1}{2}\cdot 0.99+\frac{1}{2}\cdot 0.01=\frac{1}{2},$$

由对称性可知 $\mathrm{P}(H_2)=1/2$. 但是对于 $H_1\cap H_2$, 利用全概率定理得到

$$\begin{aligned}\mathrm{P}(H_1\cap H_2)&=\mathrm{P}(B)\mathrm{P}(H_1\cap H_2|B)+\mathrm{P}(B^c)\mathrm{P}(H_1\cap H_2|B^c)\\&=\frac{1}{2}\cdot 0.99\cdot 0.99+\frac{1}{2}\cdot 0.01\cdot 0.01\approx\frac{1}{2}.\end{aligned}$$

这样 $\mathrm{P}(H_1\cap H_2)\neq \mathrm{P}(H_1)\mathrm{P}(H_2)$, 即 H_1和H_2 是相互依赖的, 即使在给定 B 的条件下是相互独立的.

现在把关于独立性的结论总结一下.

独立性

- 两个事件 A 和 B 称为相互独立的,如果它们满足

  $\mathrm{P}(A \cap B)=\mathrm{P}(A)\mathrm{P}(B)$.

  若 B 还满足 $\mathrm{P}(B)>0$,则独立性等价于

  $\mathrm{P}(A|B)=\mathrm{P}(A)$.

- 若 A 与 B 相互独立,则 A 与 B^c 也相互独立.

- 设事件 C 满足 $\mathrm{P}(C)>0$,两个事件 A 和 B 称为在给定 C 的条件下条件独立,如果它们满足

  $\mathrm{P}(A \cap B|C)=\mathrm{P}(A|C)\mathrm{P}(B|C)$.

  若进一步假定 $\mathrm{P}(B \cap C)>0$,则 A 和 B 在给定 C 的条件下的条件独立性与下面的条件是等价的

  $\mathrm{P}(A|B \cap C)=\mathrm{P}(A|C)$.

- 独立性并不蕴涵条件独立性,反之亦然.

1.5.2 一组事件的独立性

两个事件的相互独立性的概念能够推广到多个事件的相互独立性.

几个事件的相互独立性的定义

设 $A_1,\cdots,A_n$ 为 n 个事件. }若它们满足

$$\mathrm{P}\left(\bigcap_{i\in S} A_i\right)=\prod_{i\in S}\mathrm{P}(A_i)\text{对}\{1,2,\cdots,n\}\text{的任意子集}S\text{成立},$$

则称 $A_1,\cdots,A_n$ 为相互独立的事件.

关于事件 A_1,A_2,A_3,独立性条件归结为下列4个条件:

$$\begin{aligned}
\mathrm{P}(A_1\cap A_2)&=\mathrm{P}(A_1)\mathrm{P}(A_2),\\
\mathrm{P}(A_1\cap A_3)&=\mathrm{P}(A_1)\mathrm{P}(A_3),\\
\mathrm{P}(A_2\cap A_3)&=\mathrm{P}(A_2)\mathrm{P}(A_3),\\
\mathrm{P}(A_1\cap A_2\cap A_3)&=\mathrm{P}(A_1)\mathrm{P}(A_2)\mathrm{P}(A_3).
\end{aligned}$$

前面3个等式说明任意两个事件是相互独立的,这种性质称为**两两独立**.但是第4个条件也非常重要,它并不是前面3个等式的推论.反过来,第4个条件也不包含前3个条件.下面两个例子说明了这些事实.

例 1.22(两两独立并不包含独立) 设试验是抛掷两枚均匀的硬币.考虑下列事件:

$$H_1 = \{\text{第一次扔得正面}\},$$
$$H_2 = \{\text{第二次扔得正面}\},$$
$$D = \{\text{两次扔得的结果不相同}\}.$$

由定义可知 H_1 和 H_2 是相互独立的.现在证明 H_1 和 D 也是相互独立的.注意到

$$\mathrm{P}(D|H_1) = \frac{\mathrm{P}(H_1 \cap D)}{\mathrm{P}(H_1)} = \frac{1/4}{1/2} = \frac{1}{2} = \mathrm{P}(D),$$

可知 D 与 H_1 是相互独立的.D 与 H_2 的相互独立性可以类似地证明.另一方面,由

$$\mathrm{P}(H_1 \cap H_2 \cap D) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \mathrm{P}(H_1)\mathrm{P}(H_2)\mathrm{P}(D),$$

可知三个事件是不独立的.

例 1.23(等式 $\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)$ 不包含独立) 设试验是抛掷两个均匀的骰子(正六面体):

$$A = \{\text{第一次扔得 1、2 或 3}\},$$
$$B = \{\text{第一次扔得 3、4 或 5}\},$$
$$C = \{\text{两次扔得的点数之和为 9}\}.$$

我们有

$$\mathrm{P}(A \cap B) = \frac{1}{6} \neq \frac{1}{2} \cdot \frac{1}{2} = \mathrm{P}(A)\mathrm{P}(B),$$
$$\mathrm{P}(A \cap C) = \frac{1}{36} \neq \frac{1}{2} \cdot \frac{4}{36} = \mathrm{P}(A)\mathrm{P}(C),$$
$$\mathrm{P}(B \cap C) = \frac{1}{12} \neq \frac{1}{2} \cdot \frac{4}{36} = \mathrm{P}(B)\mathrm{P}(C).$$

这样3个事件是不独立的,并且任何一对事件也不相互独立的.但是下面的等式是成立的

$$\mathrm{P}(A \cap B \cap C) = \frac{1}{36} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{4}{36} = \mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C).$$