互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

第 3 章　条件概率和独立性

3.1　引言

本章我们将介绍条件概率，这是概率论中最重要的概念之一．这个概念的重要性是双重的．首先，我们在计算某些事件的概率时，同时具有某些关于该试验的附加信息，此时概率应该是条件概率．其次，即使没有附加信息可用，也可以利用条件概率的方法计算某些事件的概率，而这种方法常可以使计算非常简单．

3.2　条件概率

同时掷两枚骰子，假设 36 种结果都是等可能发生的，因此每种结果发生的概率为 1/36．进一步假设已知第一枚骰子的点数为 3．在这些条件下，两枚骰子点数之和为 8 的概率是多大？为了计算这个概率，推理如下：给定第一枚骰子的点数为 3，那么掷两枚骰子共有 6 种可能结果：(3，1)，(3，2)，(3，3)，(3，4)，(3，5)，(3，6)．因为每个结果发生的概率都一样，那么，这 6 种结果也应该是等可能的．即在给定第一枚骰子为 3 的情况下，6 种结果(3，1)，…，(3，6)中每一个结果发生的(条件)概率应该是 1/6，而样本空间中其他 30 个点的(条件)概率应该是 0．这样，在第一枚骰子的点数为 3 的条件下，两枚骰子点数之和为 8 的概率应该是 1/6．

如果我们让 E 和 F 分别表示“两枚骰子点数之和为 8”和“第一枚骰子点数为 3”，利用上述方法，计算得到的概率称为假定 F 发生的情况下 E 发生的条件概率，记为 $P(E|F)$．

用如下方式可以推导出一个对于所有 E 和 F 都适用的计算 $P(E|F)$ 的一般公式：如果 F 发生了，那么为了 E 发生，其结果必然是既属于 E 也属于 F 的一点，即这个结果必然属于 EF．既然已知 F 已经发生，F 成了新的样本空间，因此 E 发生的(条件)概率必然等于 EF 发生的概率与 F 发生的概率之比．也就是说，我们有如下定义． 56

定义　如果 $P(F)>0$，那么

$$P(E|F)=P(EF)/P(F) \tag{2.1}$$

例 2a　乔伊 80％肯定他把失踪的钥匙放在了他外套两个口袋中的一个里．他 40％确定放在左口袋，40％确定放在右口袋．如果检查了左口袋发现没有找到钥匙，那么钥匙在右口袋的条件概率是多少？

解　如果令 L 表示“钥匙在左口袋”这个事件，令 R 表示“钥匙在右口袋”这个事件，则所求概率为

$$P(R|L^c)=P(RL^c)/P(L^c)=P(R)/[1-P(L)]=2/3 \qquad ■$$

如果一个有限样本空间 S 中每个结果都是等可能的，那么这些加上条件的事件在子集 $F\subset S$ 中，F 中的所有结果都是等可能的．在这些例子中，我们用 F 这样的样本空间可以很容易地计算出形如 $P(E|F)$ 的条件概率．的确，在处理简化的样本空间时，我们可以用

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

更简单、更容易理解的方法．下面两个例子可以阐明这一点．

例 2b　抛掷一枚硬币两次，假定样本空间 $S=\{(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T}),(\mathrm{T},\mathrm{H}),(\mathrm{T},\mathrm{T})\}$中 4 个样本点是等可能发生的，求给定以下事件后两枚硬币都是正面朝上的条件概率：(a) 第一枚正面朝上；(b) 至少有一枚正面朝上．

解　令 $B=\{(\mathrm{H},\mathrm{H})\}$表示事件"两枚硬币都是正面朝上"；令 $F=\{(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T})\}$表示事件"第一枚硬币正面朝上"；令 $A=\{(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T}),(\mathrm{T},\mathrm{H})\}$表示事件"至少有一枚硬币正面朝上"．那么(a)中所求概率为

57

$$P(B|F)=P(BF)/P(F)=\frac{P(\{(\mathrm{H},\mathrm{H})\})}{P(\{(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T})\})}=\frac{1/4}{2/4}=1/2$$

对于(b)，有

$$P(B|A)=P(BA)/P(A)=\frac{P(\{(\mathrm{H},\mathrm{H})\})}{P(\{(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T}),(\mathrm{T},\mathrm{H})\})}=\frac{1/4}{3/4}=1/3$$

因此，已知第一枚硬币正面朝上的条件下，两枚硬币都是正面朝上的条件概率为 1/2，而已知至少有一枚硬币正面朝上的条件下，两枚硬币都是正面朝上的概率为 1/3．很多学生对后者感到吃惊．他们推断，给定至少有一枚正面朝上这个事件后就有两种结果：两枚都正面朝上，或者只有一枚正面朝上．他们犯的错误是把这两种结果看成等可能的了．最初有 4 个等可能的结果，因为"至少一枚正面朝上"表明"结果不是(T，T)"，所以只剩 3 个结果：(H，H)，(H，T)，(T，H)．而这三种结果都是等可能的，(H，H)只是其中一个结果，因此其条件概率为 1/3 是很自然的了．■

例 2c　桥牌游戏里，52 张牌平均发给东、西、南、北四家．如果南和北一共有 8 张黑桃，那么东有剩下 5 张黑桃里的 3 张的概率是多大？

解　或许计算这个概率的最简单的方法是缩减样本空间．即已知南北 26 张牌中共有 8 张黑桃，那么还剩下 26 张牌，其中正好 5 张是黑桃，将要分给东西两家．由于每种分法都是等可能的，因此，东家 13 张牌中正好有 3 张黑桃的条件概率是

$$\frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}}\approx 0.339$$
■

在式(2.1)两边同时乘以 $P(F)$，可以得到

$$P(EF)=P(F)P(E|F) \tag{2.2}$$

换言之，公式(2.2)说明了 E 和 F 同时发生的概率，等于 F 发生的概率乘以在 F 发生的条件下 E 发生的条件概率．公式(2.2)在计算事件的交的概率时非常有用．

例 2d　在选修法语课还是化学课这件事上，席琳犹豫不决．她估计如果选修法语课，则有 1/2 的概率获得"A"等成绩，而如果选修化学课，则有 2/3 的概率获得"A"等成绩．如果席琳通过掷硬币来做决定，那么她将选修化学课，并获得"A"等成绩的概率是多大？

解　如果用 C 表示"席琳选修化学课"，用 A 表示"她获得了'A'等成绩"，那么所求概

58 率为 $P(CA)$，利用公式(2.2)计算如下：

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P(CA)=P(C)P(A|C)=\frac{1}{2}\times\frac{2}{3}=\frac{1}{3}$$ ■

例 2e 假设坛子中有 8 个红球与 4 个白球，我们无放回地取出两个球.

(a) 假定每次取球时，坛子中各个球被取中的可能性是一样的，那么取出的两个球都为红球的概率是多少?

(b) 现在假定球的质量不同，每个红球的质量为 r，每个白球的质量为 w，假设每次抽到一给定球的概率是其质量除以当时坛子中球的总质量，则两次取出的都为红球的概率是多少?

解 (a) 令 R_1 和 R_2 分别表示第一次与第二次取出红球的事件. 若取出的第一个球是红球，那么坛子中剩下 7 个红球和 4 个白球，因此，

$$P(R_2|R_1)=\frac{7}{11}$$

由于 $P(R_1)$显然等于 8/12，因此所求概率为

$$P(R_1R_2)=P(R_1)P(R_2|R_1)=\left(\frac{2}{3}\right)\left(\frac{7}{11}\right)=\frac{14}{33}$$

当然，这个概率也可按下式计算得出：

$$P(R_1R_2)=\frac{\binom{8}{2}}{\binom{12}{2}}$$

(b) 我们再次令 R_i 表示事件“第 i 次取出的是红球”，并利用公式

$$P(R_1R_2)=P(R_1)P(R_2|R_1)$$

现在，对红球进行标记，记 $B_i(i=1,\cdots,8)$表示第一次取出的是第 i 个红球. 那么，

$$P(R_1)=P\left(\bigcup_{i=1}^{8}B_i\right)=\sum_{i=1}^{8}P(B_i)=8\frac{r}{8r+4w}$$

进而，假定第一次取出的球是红色的，那么坛子里有 7 个红球和 4 个白球. 同理可得，

$$P(R_2|R_1)=\frac{7r}{7r+4w}$$

那么，两个都是红球的概率是

$$P(R_1R_2)=\frac{8r}{8r+4w}\frac{7r}{7r+4w}$$ ■

公式(2.2)的推广有时也称为乘法规则，它提供了任意个事件交的概率的计算方法.

乘法规则

$$P(E_1E_2E_3\cdots E_n)=P(E_1)P(E_2|E_1)P(E_3|E_1E_2)\cdots P(E_n|E_1\cdots E_{n-1})$$

59

为了证明乘法规则，只需对等式右边利用条件概率的定义，可得

$$P(E_1)\frac{P(E_1E_2)}{P(E_1)}\frac{P(E_1E_2E_3)}{P(E_1E_2)}\cdots\frac{P(E_1E_2\cdots E_n)}{P(E_1E_2\cdots E_{n-1})}=P(E_1E_2\cdots E_n)$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

例 2f　在第 2 章例 5m 的配对问题上，我们知道 P_N，即 N 个人从 N 个帽子中随意挑选并且没有配对的概率为

$$P_N = \sum_{i=0}^{N}(-1)^i/i!$$

那么 N 个人中刚好有 k 个人配对成功的概率是多少呢？

解　让我们先考虑固定 k 个人配对成功的情况，即确定这 k 个人全配对成功而其他人都没有配对的概率．设 E 表示这 k 个人都拿到自己帽子的事件，G 表示剩下的 $N-k$ 个人都没有拿到自己帽子的事件，我们有

$$P(EG) = P(E)P(G|E)$$

现在，令 $F_i(i=1,\cdots,k)$表示第 i 个成员刚好配对成功，那么，

$$\begin{aligned}P(E) &= P(F_1F_2\cdots F_k) = P(F_1)P(F_2|F_1)P(F_3|F_1F_2)\cdots P(F_k|F_1\cdots F_{k-1}) \\ &= \frac{1}{N}\frac{1}{N-1}\frac{1}{N-2}\cdots\frac{1}{N-k+1} = \frac{(N-k)!}{N!}\end{aligned}$$

假设这 k 个人都配对好了正确的帽子，剩下的 $N-k$ 个人随机地从另外 $N-k$ 顶帽子中选择．那么 $N-k$ 人中没有一个人配对成功的概率等于这 $N-k$ 个人在这 $N-k$ 顶帽子选择时没有配对．因此

$$P(G|E) = P_{N-k} = \sum_{i=0}^{N-k}(-1)^i/i!$$

表明只有 k 个人配对成功的概率为

$$P(EG) = \frac{(N-k)!}{N!}P_{N-k}$$

由于上面的结果对于 N 个人中的任意 k 个人适用，因此

$$P(\text{刚好 } k \text{ 个配对成功}) = P_{N-k}/k! \approx e^{-1}/k! \qquad \text{当 } N \text{ 很大时}$$ ■

60 我们现在可以运用乘法规则得到第 2 章例 5h(b)的第二种解法．

例 2g　一副 52 张牌随机地分成 4 堆，每堆 13 张．计算每一堆正好有一张 A 的概率．

解　定义事件 $E_i(i=1,2,3,4)$如下：

$E_1 = \{$黑桃 A 在任何一堆里$\}$

$E_2 = \{$黑桃 A 和红桃 A 在不同的堆里$\}$

$E_3 = \{$黑桃 A,红桃 A 和方块 A 在不同的堆里$\}$

$E_4 = \{$4 张 A 在不同的堆里$\}$

所求概率为 $P(E_1E_2E_3E_4)$，利用乘法规则，

$$P(E_1E_2E_3E_4) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3)$$

由于 E_1 为样本空间 S，

$$P(E_1) = 1$$

为确定 $P(E_2|E_1)$，考虑包含黑桃 A 的那一堆．因为红桃 A 可以在黑桃 A 这一堆，也可分在其余 3 堆中．红桃 A 分在其余 3 堆的可能性为

$$P(E_2|E_1) = 1 - \frac{12}{51} = \frac{39}{51}$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

方块 A 可以分在黑桃 A 或红桃 A 这些堆里，也可以分在其余的两堆里．分到其余两堆的可能性为

$$P(E_3 \mid E_1E_2) = 1 - \frac{24}{50} = \frac{26}{50}$$

梅花 A 可以分在黑桃 A、红桃 A 或方块 A 这些堆里，或者分在另外一堆里．方块 A 在另外一堆的概率为

$$P(E_4 \mid E_1E_2E_3) = 1 - \frac{36}{49} = \frac{13}{49}$$

因此，我们可以得到所求的每堆恰好有 1 张 A 的概率为

$$P(E_1E_2E_3E_4) = \frac{39 \times 26 \times 13}{51 \times 50 \times 49} \approx 0.105$$

即大概有 10.5%的机会每堆牌中恰好有一个 A(习题 13 将利用乘法规则给出另一种解法)． ■

注释 $P(E|F)$的定义与概率的频率解释是一致的．为此，假设进行了 n 次独立重复试验(n 相当大)．若我们只考虑事件 F 发生的那些试验，此时 $P(E|F)$近似地等于事件 E 发生的相对频率．由于概率 $P(F)$是事件 F 发生的频率的极限，在 n 次独立重复试验中，事件 F 会近似地发生 $nP(F)$次．类似地，事件 EF 会近似地发生 $nP(EF)$次．这样，在 F 发生的近 $nP(F)$次试验中，事件 E 也发生的相对频率近似地等于

$$\frac{nP(EF)}{nP(F)} = \frac{P(EF)}{P(F)}$$ 61

当 n 越来越大时，其相对频率趋于 $P(EF)/P(F)$．这个值也就是 $P(E|F)$的频率定义．

3.3 贝叶斯公式

设 E 和 F 为两个事件，可以将 E 表示为

$$E = EF \cup EF^c$$

这样，E 中的结果，要么同时属于 E 和 F，要么只属于 E 但不属于 F(见图 3-1)．显然，EF 和 EF^c 是互不相容的，因此，根据公理 3，我们有

$$\boxed{\begin{aligned} P(E) &= P(EF) + P(EF^c) = P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)[1 - P(F)] \end{aligned}} \tag{3.1}$$

公式(3.1)说明事件 E 发生的概率等于在 F 发生的条件下 E 的条件概率与在 F 不发生的条件下 E 发生的条件概率的加权平均，其中加在每个条件概率上的权重就是作为条件的事件发生的概率．这是一个极其有用的公式，它使得我们能够通过以第二个事件发生与否作为条件来计算第一个事件的概率．也就是说，在许多问题中，直接计算某个事件的概率很困难，但是一旦知道第二个事件发生与否，就容易计算了．我们接下来用一些例子阐述这点．

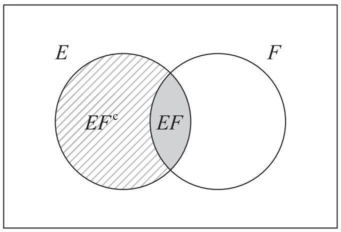


图 3-1 $E = EF \cup EF^c$．EF = 阴影区域；EF^c = 条纹区域

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

例 3a 第 1 部分　保险公司认为人可以分为两类，一类易出事故，另一类则不易出事故. 统计表明，一个易出事故者在一年内发生事故的概率为 0.4，而对不易出事故者来说，这个概率则减少到 0.2，若假定第一类人占人口的比例为 30%，现有一个新人来投保，那么该人在购买保单后一年内将出事故的概率有多大？

解　以这个投保客户是不是易出事故者作为条件，我们将得到所求概率. 令 A_1 表示
62 “投保客户一年内将出事故”这一事件，而以 A 表示“投保人为容易出事故者”这一事件，则所求概率 $P(A_1)$ 为

$$P(A_1)=P(A_1|A)P(A)+P(A_1|A^c)P(A^c)=0.4\times0.3+0.2\times0.7=0.26$$ ■

例 3a 第 2 部分　假设一个新投保人在购买保单后一年内出了事故，那么他是易出事故者的概率是多大？

解　所求概率为 $P(A|A_1)$，可从下式计算得到：

$$P(A|A_1)=\frac{P(AA_1)}{P(A_1)}=\frac{P(A)P(A_1|A)}{P(A_1)}=\frac{0.3\times0.4}{0.26}=\frac{6}{13}$$ ■

例 3b　考虑一副 52 张扑克牌的如下玩法：将洗好的一副牌都牌面朝下放好，一次翻开一张，玩家只有一次机会可以猜接下来翻开的一张是否是黑桃 A，如果是，那么玩家获胜；如果不是，那么玩家输. 另外，如果直到最后一张还没有翻开，而此前没有出现过黑桃“A”，且玩家也没有猜过，那么玩家也获胜. 什么是好的策略？什么是差的策略？

解　其答案是：任何一种策略，获胜的概率都是 1/52. 为了说明这点，我们将用归纳的方法证明一个更强的结论：对于 n 张牌，其中有一张牌为黑桃“A”，那么不管采取何种策略，获胜的概率都是 $1/n$. 这点显然对 $n=1$ 是正确的. 假设对 $n-1$ 张牌，该结论也成立. 现在考虑 n 张牌，对于任一给定策略，令 p 表示按该策略第一次就猜牌的概率. 如果第一次就猜牌，那么获胜的概率为 $1/n$. 另外，如果按策略第一次不猜牌，那么获胜的概率就是第一张牌不是黑桃 A 的概率 $(n-1)/n$，乘以在第一张牌不是黑桃“A”的条件下，获胜的条件概率. 而此条件概率就等于含一张黑桃 A 的 $n-1$ 张牌的玩牌游戏中获胜的概率，利用归纳假设，该条件概率为 $1/(n-1)$，因此，按策略第一次不猜牌的条件下，获胜的概率为

$$\frac{n-1}{n}\,\frac{1}{n-1}=\frac{1}{n}$$

因此，令 G 表示“第一次就猜牌”这一事件，我们可得

$$\begin{aligned}P(\{\text{获胜}\})&=P(\{\text{获胜}|G\})P(G)+P(\{\text{获胜}|G^c\})(1-P(G))\\&=\frac{1}{n}\times p+\frac{1}{n}\times(1-p)=\frac{1}{n}\end{aligned}$$ ■

例 3c　在回答一道多项选择题时，学生或者知道正确答案，或者就猜一个. 令 p 表
63 示他知道正确答案的概率，则 $1-p$ 表示猜的概率. 假定学生猜中正确答案的概率为 $1/m$，此处 m 就是多项选择题的可选择答案数. 在已知他回答正确的条件下，该学生知道正确答案的概率是多少？

解　令 C 和 K 分别表示事件“该学生回答正确”和“该学生知道正确答案”. 这样

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P(K|C)=\frac{P(KC)}{P(C)}=\frac{P(C|K)P(K)}{P(C|K)P(K)+P(C|K^c)P(K^c)}$$

$$=\frac{p}{p+(1/m)(1-p)}=\frac{mp}{1+(m-1)p}$$

例如，如果 $m=5$，$p=1/2$，那么在已知该学生回答正确的条件下，他知道正确答案的条件概率为 5/6. ■

例 3d 一项血液化验有 95%的把握诊断某种疾病，但是，这项化验用于健康人也会有 1%的“伪阳性”结果(即如果一个健康人接受这项化验，则化验结果误诊此人患该疾病的概率为 0.01). 如果该疾病的患者事实上仅占总人口的 0.5%，若某人化验结果为阳性，则此人确实患该疾病的概率是多少？

解 令 D 表示“接受化验的这个人患该疾病”这一事件，令 E 表示“其化验结果为阳性”这一事件，所求概率为

$$P(D|E)=\frac{P(DE)}{P(E)}=\frac{P(E|D)P(D)}{P(E|D)P(D)+P(E|D^c)P(D^c)}$$

$$=\frac{0.95\times 0.005}{0.95\times 0.005+0.01\times 0.995}=\frac{95}{294}\approx 0.323$$

因此，在验血结果为阳性的人当中，真正患该病的人只有 32%. 对于这一结果，许多学生感到非常吃惊(因为验血似乎是个好办法，他们总认为这个数值应该高得多). 因此，有必要给出第二个解法. 与前一个解法比较，第二个解法尽管不严格，但却更直观.

由于事实上患该疾病的人占总人口的比例为 0.5%，平均地算，接受化验的每 200 个人中应有 1 个患者，而这项化验只能保证疾病的患者被诊断为患病的概率为 0.95，因此，平均来说，每 200 个接受化验者能保证有 0.95 个人被诊断出，并且此人真的患病. 但另一方面(平均来说)，在其余 199 个健康人中，这项化验会错误地诊断出 199×0.01 个人患该病，因此，每当诊断出 0.95 个病人时(平均地说)总有 199×0.01 个健康人被误诊为患病. 64
于是，当验血结果确定某人患该病时，正确诊断所占比例为

$$\frac{0.95}{0.95+199\times 0.01}=\frac{95}{294}\approx 0.323$$ ■

利用公式(3.1)还可以根据附加信息对某事件的概率进行修正. 下面的例子就是它的一个应用.

例 3e 假设某药剂师考虑如下诊断方案：如果我至少有 80%的可能确定病人确实有此病，那么我会建议手术；而如果我并不确定，那么我会推荐做进一步的检查，该检查是昂贵的，有时也是痛苦的. 现在，开始我仅仅有 60%的把握认为琼斯患有此病，因此我推荐做了检查项目 A，该检查对于确有此病的患者给出阳性结果，而对健康人却不会给出阳性结果. 经检查琼斯的结果是阳性后，正当我建议手术时，琼斯给了我另一个信息，他患有糖尿病. 这个信息使问题复杂化，尽管它并不影响我一开始认为他患有此病的 60%的把握，但是却影响了检查项目 A 的效果. 因为虽然该检查项目对健康人不给出阳性，但是对于患有糖尿病却不患有这种疾病的人来说，有 30%的可能给出阳性结果. 那么我现在该如何做？是做进一步检查，还是立即手术？

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

解　为了决定是否建议手术，医生首先要计算在检查项目 A 结果为阳性的情况下，琼斯患该病的概率．令 D 表示“琼斯患此病”这一事件，E 表示“检查项目 A 结果为阳性”这一事件，那么所求条件概率 $P(D|E)$为

$$P(D|E)=\frac{P(DE)}{P(E)}=\frac{P(E|D)P(D)}{P(E|D)P(D)+P(E|D^c)P(D^c)}=\frac{0.6\times 1}{1\times 0.6+0.3\times 0.4}=0.833$$

注意，我们以琼斯是否患有此病为条件计算了项目 A 结果为阳性的概率，并且利用了如下事实：因为琼斯患有糖尿病，已知在其不患上述疾病的条件下项目 A 结果为阳性的条件概率 $P(E|D^c)$等于 0.3，因此，医生现在有超过 80%的把握确定琼斯患有此病，所以应该建议手术． ■

例 3f　在某刑事调查过程中，调查员有 60%的把握认为嫌疑人确实犯有此罪．假定现在得到了一份新的证据，表明罪犯有某个身体特征(左撇子、光头或者棕色头发等)，如果有 20%的人有这种特征，那么在嫌疑犯具有这种特征的条件下，检查官认为他确实犯此罪

65 的把握有多大?

解　令 G 表示“嫌疑犯确实犯此罪”这一事件，而 C 表示“他具有罪犯的该身体特征”，那么有

$$P(G|C)=\frac{P(GC)}{P(C)}=\frac{P(C|G)P(G)}{P(C|G)P(G)+P(C|G^c)P(G^c)}=\frac{1\times 0.6}{1\times 0.6+0.2\times 0.4}\approx 0.882$$

其中我们假定了嫌疑犯事实上没犯罪却有该身体特征的概率为 0.2，也就是具有这个特征的人口的比例． ■

例 3g　1965 年 5 月，在 Buenos Aires 举行的世界桥牌锦标赛上，英国一对著名的桥牌手 Terrence Reese 和 Boris Schapiro 被指控作弊，说是他们用手指作暗号暗示他们红桃的张数．Reese 和 Schapiro 都否认这项指控．事后，英国桥牌协会举行了一个听证会．听证会按法律的程序进行，既包含控方，也包含辩方．双方都有目击证人．在接下来的调查过程中，控方检查了 Reese 和 Schapiro 打的几手牌，并声称他们的打法与通过作弊已知了红桃的张数的打法是吻合的．针对这个观点，辩方律师指出，他们的打法也同样与标准打法一致．然而，控方指出，既然他们的打法与其作弊的假设是一致的，那么就应该是支持这种假设．你如何理解控方的理由?

解　此问题基本上是新的证据(在此例中，牌的打法)如何影响某个特定假设成立的概率的问题．现在，令 H 表示“某个特定假设(如 Reese 和 Schapiro 确实作弊)”，而 E 表示“新的证据”，那么

$$P(H|E)=\frac{P(HE)}{P(E)}=\frac{P(E|H)P(H)}{P(E|H)P(H)+P(E|H^c)[1-P(H)]}\tag{3.2}$$

其中 $P(H)$为在新证据展示之前我们对假设成立的可能性的估值．新证据支持假设成立，如果它使得假设成立的可能性增大，即 $P(H|E)\geqslant P(H)$．利用公式(3.2)，此式等价于

$$P(E|H)\geqslant P(E|H)P(H)+P(E|H^c)[1-P(H)]$$

或等价于

66
$$P(E|H)\geqslant P(E|H^c)$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

即当假设成立时，新证据发生的概率大于假设不成立时发生的概率，就认为新证据支持假设. 事实上，已知新的证据发生的条件下，假设成立的新概率和初始概率的关系可从下式看出：

$$P(H|E)=\frac{P(H)}{P(H)+[1-P(H)]\dfrac{P(E|H^c)}{P(E|H)}}$$

因此，在所考虑的问题中，牌手的打法可以认为是支持假设成立的，除非在他们作弊的条件下这种打法的可能性大于他们不作弊的条件下这种打法的可能性. 而控方并没有作此声明. 因此，他们关于新证据是支持作弊的假设的这一断言是无效的. ■

例 3h 双胞胎可能是同卵双生或者异卵双生. 同卵双生也叫单卵双生，是由一个受精卵分裂成为两个完全一样的部分发育而来的. 因此，同卵双胞胎含有相同的基因. 异卵双生又叫二卵双生，是由两个受精卵植入子宫发育而来的. 异卵双胞胎像不同时间出生的兄弟姐妹一样，基因多少是有些一样的. 为了知道双胞胎中同卵双生所占的比例，加利福尼亚州洛杉矶市的科学家已经指派一名统计学家来研究这个问题. 这位统计学家首先要求市里每一家医院对双胞胎做记录，同时对是否是同卵双生做标记. 然而医院告诉他判断一个新生儿是否是同卵双生并不是一件简单的事，这关系到父母是否愿意自费给孩子做这项复杂而又昂贵的 DNA 检验. 经过一番思考之后，统计学家只让医院提供标记着双胞胎是否是相同性别的所有双胞胎数据列表. 当数据表明约有 64%的双胞胎是性别相同时，统计学家就宣称约有 28%的双胞胎是同卵双生. 他是如何得出这个结论的？

解 统计学家推断同卵双生双胞胎性别总是相同的. 又因为异卵双生双胞胎就相当于普通的兄弟姐妹，所以性别相同的概率也有 1/2. 令 I 表示“同卵双生双胞胎”事件，令 SS 表示“双胞胎性别相同”事件，他在双胞胎是否为同卵双生的条件下计算概率 $P(SS)$，得到

$$P(SS)=P(SS|I)P(I)+P(SS|I^c)P(I^c)$$

或者

$$P(SS)=1\times P(I)+\frac{1}{2}\times[1-P(I)]=\frac{1}{2}+\frac{1}{2}P(I)$$

由 $P(SS)\approx 0.64$ 可以得出 $P(I)\approx 0.28$. ■

当发现新的证据时，假设成立的概率之变化可以表示为假设的“优势比”之变化，其中优势比的概念定义如下 . 67

定义 事件 A 的优势比定义为

$$\frac{P(A)}{P(A^c)}=\frac{P(A)}{1-P(A)}$$

即事件 A 的优势比告诉我们该事件发生的可能性是不发生的可能性的倍数. 举例来说，如果 $P(A)=2/3$，那么 $P(A)=2P(A^c)$，因此，事件 A 的优势比等于 2. 如果某事件的优势比等于 α，那么通常称支持假设成立的优势比为“$\alpha:1$”.

现在考虑假设 H 以概率 $P(H)$成立，如果我们发现了新的证据 E，那么在 E 成立的条件下，H 成立和 H 不成立的条件概率分别为

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P(H|E)=\frac{P(E|H)P(H)}{P(E)}\qquad P(H^c|E)=\frac{P(E|H^c)P(H^c)}{P(E)}$$

因此，引进证据 E 后，假设 H 的新的优势比为

$$\frac{P(H|E)}{P(H^c|E)}=\frac{P(H)}{P(H^c)}\frac{P(E|H)}{P(E|H^c)}\tag{3.3}$$

即 H 的新的优势比值等于它原来的优势比的值乘以新的证据在 H 和 H^c 之下的条件概率比值. 公式(3.3)也验证了例 3f 的结论，因为如果在 H 成立的条件下新证据的概率大于在 H^c 成立的条件下新证据的概率，H 的优势比值是递增的. 反之，是递减的.

例 3i 一个坛子里装了两枚 A 型硬币和一枚 B 型硬币. 当抛 A 型币时，正面向上的概率是 1/4. 当抛 B 型硬币时，正面向上的概率是 3/4. 随机从坛子里取一枚硬币掷，假定掷出的结果是正面向上，则所掷的是 A 型硬币的概率是多少？

解 令 A 表示“掷的是 A 型硬币”这个事件，$B=A^c$ 表示“掷的是 B 型硬币”这个事件. 令 head 表示硬币正面向上这个事件，我们要求的是 $P(A|\text{head})$. 从式(3.3)可以看出，

$$\frac{P(A|\text{head})}{P(A^c|\text{head})}=\frac{P(A)}{P(B)}\frac{P(\text{head}|A)}{P(\text{head}|B)}=\frac{2/3}{1/3}\frac{1/4}{3/4}=2/3$$

所以优势比是 2/3∶1，或者，等价地，掷的是 A 型硬币的概率为 2/5. ■

公式(3.1)可推广如下：假定 F_1，F_2，…，F_n 是互不相容的事件，且

68

$$\bigcup_{i=1}^{n}F_i=S$$

换言之，这些事件中必有一件发生. 记

$$E=\bigcup_{i=1}^{n}EF_i$$

又由于事实上 $EF_i(i=1,\cdots,n)$是互不相容的，我们可以得到如下公式：

$$P(E)=\sum_{i=1}^{n}P(EF_i)=\sum_{i=1}^{n}P(E|F_i)P(F_i)\tag{3.4}$$

因此，公式(3.4)被称为全概率公式，对于事件 F_1，F_2，…，F_n，其中一个或者仅有一个发生，我们可以用这个公式通过 F_i 中一个发生的条件概率来计算 $P(E)$. 公式(3.4)叙述了 $P(E)$等于 $P(E|F_i)$的加权平均，每项的权为事件 F_i 发生的概率.

例 3j 在第 2 章例 5j 中，我们考虑了对于随机洗好的牌，跟在第一个 A 后的特定牌的概率，我们给出了一种推理方法，求出这个概率是 1/52. 这里有个基于条件概率的概率论证：令 E 表示跟在第一个 A 后的特定牌(如 x)这个事件. 为计算 $P(E)$，我们忽略牌 x 和桌上另 51 张牌的相对顺序. 令 $\mathbf{O}$ 表示由

$$P(E)=\sum_{\mathbf{O}}P(E|\mathbf{O})P(\mathbf{O})$$

给出的顺序. 现在，给定 $\mathbf{O}$，这里有 52 种可能的顺序对应第 i 次拿到牌 $x(i=1,\cdots,52)$，但是因为所有 52! 种可能的顺序出现的概率相等. 因为在这些可能之中，牌 x 会跟在第一个 A 后面，所以有 $P(E|\mathbf{O})=1/52$，这表示 $P(E)=1/52$. ■

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

再次，令 F_1，…，F_n 表示一组互不相容且穷举的事件(意思是恰好有这些事件中的一个必须发生). 现在假设 E 发生了(新的证据)，我们想要计算 F_j 发生的概率. 利用公式(3.4)，我们有如下命题.

命题 3.1

$$P(F_j|E)=\frac{P(EF_j)}{P(E)}=\frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n}P(E|F_i)P(F_i)} \tag{3.5}$$ 69

公式(3.5)称为贝叶斯公式，是根据英国哲学家托马斯·贝叶斯命名的. 如果我们把事件 F_j 设想为关于某个问题的各个可能的“假设条件”，那么，贝叶斯公式可以这样理解：它告诉我们，在试验之前对这些假设条件所作的判断[即 $P(F_j)$]，可以如何根据试验的结果来进行修正.

例 3k 一架飞机失踪了，推测它等可能地坠落在 3 个区域. 令 $1-\beta_i(i=1, 2, 3)$表示飞机事实上坠落在第 i 个区域，且被发现的概率(β_i 称为忽略概率，因为它表示忽略飞机的概率，通常由该区域的地理和环境条件决定). 已知对区域 1 的搜索没有发现飞机，求在此条件下，飞机坠落在第 $i(i=1, 2, 3)$个区域的条件概率.

解 令 $R_i(i=1, 2, 3)$表示“飞机坠落在第 i 个区域”这一事件，令 E 表示“对第 1 个区域的搜索没有发现飞机”这一事件，利用贝叶斯公式可得

$$P(R_1|E)=\frac{P(ER_1)}{P(E)}=\frac{P(E|R_1)P(R_1)}{\sum_{i=1}^{3}P(E|R_i)P(R_i)}=\frac{\beta_1\times\frac{1}{3}}{\beta_1\times\frac{1}{3}+1\times\frac{1}{3}+1\times\frac{1}{3}}=\frac{\beta_1}{\beta_1+2}$$

对于 $j=2, 3$，有

$$P(R_j|E)=\frac{P(E|R_j)P(R_j)}{P(E)}=\frac{1\times\frac{1}{3}}{\beta_1\times\frac{1}{3}+\frac{1}{3}+\frac{1}{3}}=\frac{1}{\beta_1+2}\qquad j=2,3$$

注意，当搜索了第 1 个区域没有发现飞机时，飞机坠落在第 $j(j\neq1)$个区域的更新(即条件)概率会增大，而坠落在第 1 个区域的概率会减小. 这是一个常识问题：因为既然在第 1 个区域没有发现飞机，当然飞机坠落在该区域的概率会减少，而坠落在其他区域的概率会增大. 而且飞机坠落在第 1 个区域的条件概率是忽略概率 β_1 的递增函数. 当 β_1 增加时，增大了飞机坠落在第 1 个区域的条件概率. 类似地，$P(R_j|E)(j\neq1)$是 β_1 的递减函数. ■ 70

下一个例子经常被学过概率而又不讲道德的学生们用来从概率知识较少的朋友那里赢钱.

例 3l 假设有 3 张形状完全相同但颜色不同的卡片，第一张两面全是红色，第二张两面全是黑色，而第三张是一面红一面黑. 将这 3 张卡片放在帽子里混合后，随机地取出 1 张放在地上. 如果取出的卡片朝上的一面是红色的，那么另一面为黑色的概率是多少?

解 令 RR，BB，RB 分别表示取出的卡片是“两面红”、“两面黑”以及“一面红一面黑”这三个事件. 再令 R 表示取出的卡片“朝上一面是红色”这一事件. 我们可按如下方式

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

得到所求概率：

$$P(RB|R)=\frac{P(RB\cap R)}{P(R)}=\frac{P(R|RB)P(RB)}{P(R|RR)P(RR)+P(R|RB)P(RB)+P(R|BB)P(BB)}$$

$$=\frac{\frac{1}{2}\times\frac{1}{3}}{1\times\frac{1}{3}+\frac{1}{2}\times\frac{1}{3}+0\times\frac{1}{3}}=\frac{1}{3}$$

因此，答案是 1/3. 但是有些学生猜另一面是黑色的概率为 1/2，他们错误地推理给定正面为红色时这张牌有两种可能，两面全红，或者一面红一面黑，这两种可能性是一样的. 事实上，你可以把三张牌的 6 个面记为(R_1, R_2)，(B_1, B_2)，(R_3, B_3)，其中(R_1, R_2)表示全红的那张牌的两个面，(B_1, B_2)表示两面均为黑色的那张牌的两个面，(R_3, B_3)表示一红一黑的那张牌的两个面. 即 6 个面是以相等的概率出现的. 只有 R_3 出现时，背面是黑色的. 而 R_1，R_2 出现时，背面均为红色的. 因此，当正面为红色时，背面为黑色的概率为 1/3 而不是 1/2. ■

例 3m 镇上新搬来一对夫妇，已知他们有两个孩子. 假设某天遇到该母亲带着一个孩子在散步. 如果这个孩子是女孩，那么她的两个孩子都是女孩的概率是多少？

解 首先，定义如下事件：

G_1：第一个(也即最大的)孩子为女孩；

G_2：第二个孩子为女孩；

G：被看到跟母亲一起散步的为女孩.

而且令 B_1，B_2，B 表示类似上述的事件，其中"女孩"替换为"男孩". 这样，所求概率 $P(G_1G_2|G)$可以表示如下：

71

$$P(G_1G_2|G)=\frac{P(G_1G_2G)}{P(G)}=\frac{P(G_1G_2)}{P(G)}$$

并且，

$$\begin{aligned}P(G)&=P(G|G_1G_2)P(G_1G_2)+P(G|G_1B_2)P(G_1B_2)\\&\quad+P(G|B_1G_2)P(B_1G_2)+P(G|B_1B_2)P(B_1B_2)\\&=P(G_1G_2)+P(G|G_1B_2)P(G_1B_2)+P(G|B_1G_2)P(B_1G_2)\end{aligned}$$

其中最后一个等式用到了结论 $P(G|G_1G_2)=1$ 以及 $P(G|B_1B_2)=0$. 如果我们做出常规假设，即 4 种可能结果 G_1，G_2，B_1，B_2 都是等可能的，那么有

$$P(G_1G_2|G)=\frac{1/4}{1/4+P(G|G_1B_2)/4+P(G|B_1G_2)/4}=\frac{1}{1+P(G|G_1B_2)+P(G|B_1G_2)}$$

因此，答案依赖于在已知事件 G_1B_2 条件下，碰到母亲带着女孩的条件概率，以及在已知事件 G_2B_1 条件下，碰到母亲带着女孩的条件概率. 举例来说，我们假定母亲对于性别没有倾向，母亲带着大孩子一起散步的概率为 p，那么有

$$P(G|G_1B_2)=p=1-P(G|B_1G_2)$$

这表明

$$P(G_1G_2|G)=\frac{1}{2}$$

HZ BOOKS 华章图书

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

另一方面，如果我们假定两个孩子性别不同，母亲带着女孩一起散步的概率为 q，且与孩子出生的次序是独立的，那么有

$$P(G|G_1B_2)=P(G|B_1G_2)=q$$

这意味着

$$P(G_1G_2|G)=\frac{1}{1+2q}$$

举例来说，如果取 $q=1$，即母亲总是选择女孩一起去散步，那么两个都是女孩的条件概率为 1/3. 这点同例 2b 是一致的，因为事件“看见母亲带着一个女孩”与事件“至少有一个女孩”是等价的.

因此，综上所述，此问题是无法求解的. 事实上，即使假设孩子的性别是等可能的，我们仍需要额外的假设才能解决问题. 因为该试验的样本空间包含了如下形式的向量：(s_1, s_2, i)，其中 s_1 表示大孩子的性别，s_2 表示小孩子的性别，i 表示被碰到的孩子的出生次序. 因此，为确定样本空间的点的概率，只知道孩子的性别的概率是不够的，还需要知道母亲所带孩子的出生次序的条件概率(给定大小孩子的性别). ■ 72

例 3n 储物箱里有 3 种不同的一次性手电. 第一种手电使用超过 100 小时的概率为 0.7，而第二种和第三种手电相应的概率分别只有 0.4 和 0.3. 假设箱子里的手电，20%为第一种，30%为第二种，50%第三种.

(a) 随机挑一个手电，能使用 100 小时以上的概率是多少？

(b) 已知手电使用超过 100 小时，那么它是第 $j(j=1, 2, 3)$种手电的条件概率是多少？

解 (a) 令 A 表示“挑出的手电能使用 100 小时以上”这一事件，令 $F_j(j=1, 2, 3)$表示挑出了第 j 种手电. 为了计算 $P(A)$，以挑出手电的种类为条件，可得

$$\begin{aligned}P(A)&=P(A|F_1)P(F_1)+P(A|F_2)P(F_2)+P(A|F_3)P(F_3)\\&=0.7\times0.2+0.4\times0.3+0.3\times0.5=0.41\end{aligned}$$

因此，随机挑选的手电能使用 100 小时以上的概率为 0.41.

(b) 利用贝叶斯公式得到所求概率：

$$P(F_j|A)=\frac{P(AF_j)}{P(A)}=\frac{P(A|F_j)P(F_j)}{0.41}$$

因此，

$$P(F_1|A)=0.7\times0.2/0.41=14/41$$
$$P(F_2|A)=0.4\times0.3/0.41=12/41$$
$$P(F_3|A)=0.3\times0.5/0.41=15/41$$

例如，对第一种手电，尽管被选中的初始概率只有 0.2，但是得到手电使用超过 100 小时的信息后，这个概率提升到 14/41≈0.341. ■

例 3o 某罪犯在犯罪现场留下了一些 DNA，法医研究后注意到能够辨认的只有 5 对，而且每个无罪的人，与这 5 对相匹配的概率为 10^{-5}，律师认为罪犯就是该城镇 1 000 000 个居民之一. 在过去 10 年内，该城镇有 10 000 人刑满释放. 他们的 DNA 资料都记录在案，

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

在检查这些 DNA 文档之前，律师认为这 10 000 个有犯罪前科的人犯此罪的概率为 α，而其余 990 000 个居民中的每个人犯此罪的概率为 β，其中 $\alpha=c\beta$.（即他认为最近 10 年内释放的有犯罪前科的人作案的可能性是其他人的 c 倍.）将 DNA 分析结果同这 10 000 个有犯罪前科的人的数据文档对比后，发现只有 AJ 琼斯的 DNA 符合. 假设律师关于 α 和 β 之间的关
73 系是准确的，AJ 作案的可能性有多大？

解　首先，注意，概率之和必等于 1，我们有

$$1=10\,000\alpha+990\,000\beta=(10\,000c+990\,000)\beta$$

因此，

$$\beta=\frac{1}{10\,000c+990\,000}\qquad \alpha=\frac{c}{10\,000c+990\,000}$$

现在，令 G 表示事件“AJ 为作案者”，令 M 表示事件“AJ 是这 10 000 人中唯一与现场 DNA 相匹配的人”. 那么

$$P(G|M)=\frac{P(GM)}{P(M)}=\frac{P(G)P(M|G)}{P(M|G)P(G)+P(M|G^c)P(G^c)}$$

另一方面，如果 AJ 为作案者，此时其他的 9999 人都不是作案者. 这 10 000 人中 AJ 是唯一 DNA 匹配者的概率为

$$P(M|G)=(1-10^{-5})^{9999}$$

如果 AJ 不是作案者，若要他是唯一匹配者，他的 DNA 必须匹配（这以概率 10^{-5} 发生），而所有其他 9999 个有前科的人都必定不是作案者，且都不是匹配者. 现在给定 AJ 不是作案者，我们确定其他有前科的人不是作案者的条件概率. 令 $C=\{$除 AJ 外，其他有前科的人不是作案者$\}$，则

$$P(C|G^c)=\frac{P(CG^c)}{P(G^c)}=\frac{1-10\,000\alpha}{1-\alpha}$$

同样，在除 AJ 外，其他有前科的人不是作案者的前提下，这 9999 人与现场 DNA 相匹配的概率为 $(1-10^{-5})^{9999}$，所以

$$P(M|G^c)=10^{-5}\left(\frac{1-10\,000\alpha}{1-\alpha}\right)(1-10^{-5})^{9999}$$

因为 $P(G)=\alpha$，由上面公式可得

$$P(G|M)=\frac{\alpha}{\alpha+10^{-5}(1-10\,000\alpha)}=\frac{1}{0.9+\frac{10^{-5}}{\alpha}}$$

因此，如果律师最初认为任一有犯罪前科的人作案的可能性是没有犯罪前科的人的 100 倍（即 $c=100$），那么 $\alpha=1/19\,900$，且

$$P(G|M)=\frac{1}{1.099}\approx 0.9099$$

如果律师最初认为 $c=10$，那么 $\alpha=1/109\,000$，且

74 $$P(G|M)=\frac{1}{1.99}\approx 0.5025$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

如果律师最初认为任一有犯罪前科的人作案的可能性与镇里其他人是相同的($c=1$)，那么$\alpha=10^{-6}$，且

$$P(G|M)=\frac{1}{10.9}\approx 0.0917$$

因此，概率变化范围是从大概 9%(此时律师最初假设所有人作案的概率都一样)到 91%(此时他认为每个有犯罪前科的人作案的概率是其他没有犯罪前科的人的 100 倍). ■

3.4 独立事件

本章前面的例子表明：在已知 F 发生的条件下 E 发生的条件概率 $P(E|F)$一般来说不等于 E 发生的非条件概率 $P(E)$. 也就是说，知道了 F 已发生通常会改变 E 的发生机会. 但在一些特殊情形下，$P(E|F)$确实等于 $P(E)$，此时我们称 E 和 F 独立. 即如果已知 F 的发生并不影响 E 发生的概率，那么 E 和 F 就是独立的.

因为 $P(E|F)=P(EF)/P(F)$，所以如果

$$P(EF)=P(E)P(F) \tag{4.1}$$

那么 E 和 F 独立. 由于公式(4.1)关于 E 和 F 是对称的，这就表明只要 E 和 F 独立，那么 F 和 E 也独立. 因此，我们有以下定义.

定义 *对于两个事件 E 和 F，若公式(4.1)成立，则称它们是独立的(independent). 若两个事件 E 和 F 不独立，则称它们是相依的(dependent)，或相互不独立.*

例 4a 从一副洗好的 52 张扑克牌里随机抽取一张牌. 令 E 表示事件"抽取的牌为一张 A"，令 F 表示事件"抽取的牌为一张黑桃"，那么 E 和 F 就是独立的. 因为 $P(EF)=1/52$，而 $P(E)=4/52$ 且 $P(F)=13/52$. ■

例 4b 掷两枚硬币，假设全部 4 个结果出现的可能性是一样的. 令 E 表示事件"第一枚硬币正面朝上"，令 F 表示事件"第二枚硬币反面朝上"，那么 E 和 F 是独立的，因为 $P(EF)=P\{(\mathrm{H},\mathrm{T})\}=\frac{1}{4}$，而 $P(E)=P(\{(\mathrm{H},\mathrm{H}),(\mathrm{H},\mathrm{T})\})=\frac{1}{2}$，且 $P(F)=P(\{(\mathrm{H},\mathrm{T}),(\mathrm{T},\mathrm{T})\})=\frac{1}{2}$. ■

例 4c 掷两枚均匀的骰子，令 E_1 表示事件"骰子点数之和为 6"，令 F 表示事件"第一枚骰子点数为 4"，那么

$$P(E_1F)=P(\{(4,2)\})=\frac{1}{36}$$

而

$$P(E_1)P(F)=\frac{5}{36}\times\frac{1}{6}=\frac{5}{216}$$

75

因此，E_1 和 F 不独立. 直觉上，这个原因是显而易见的，因为如果我们关注掷出点数和为 6(两枚骰子)，那么在第一枚骰子为 4(或者 1，2，3，4，5 中任一个)，我们都还乐观，因为此时仍有机会得到和为 6. 另一方面，如果第一枚骰子为 6，那么我们就不乐观了，因为没有任何机会得到点数和为 6 了. 也就是说，得到和为 6 的概率依赖于第一枚骰子的结果，

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

因此，E_1 和 F 不可能独立.

现在，令 E_2 表示事件"骰子数之和为 7"，那么 E_2 是否和 F 独立？答案是肯定的，因为

$$P(E_2F) = P(\{(4,3)\}) = \frac{1}{36}$$

而

$$P(E_2)P(F) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

我们留给读者去证明这个直觉上的结论：为什么事件"骰子点数之和为 7"同第一枚骰子的点数是独立的. ■

例 4d　如果令 E 表示事件"下届总统是共和党人"，令 F 表示事件"未来一年将会有一次大地震"，那么大多数人会认为 E 和 F 是独立的. 然而，如果另外一个事件 G 是"选举之后两年内会经历经济衰退"，那么对于 E 和 G 是否独立，却存在着长期争论. ■

下面我们证明如果 E 独立于 F，那么 E 也独立于 F^c.

命题 4.1　如果 E 和 F 独立，那么 E 和 F^c 也独立.

证明　假定 E 和 F 独立. 由于 $E=EF\cup EF^c$，且 EF 和 EF^c 显然是互不相容的，所以有

$$P(E) = P(EF) + P(EF^c) = P(E)P(F) + P(EF^c)$$

或者，等价地

$$P(EF^c) = P(E)[1-P(F)] = P(E)P(F^c)$$

命题得以证明. □

因此，如果 E 和 F 是独立的，那么无论得知 F 发生的信息，还是得知 F 不发生的信息，E 发生的概率都是不变的.

现在假设 E 既和 F 独立，又和 G 独立，那么 E 是否一定和 FG 独立？有点不可思议，答案是否定的，下例说明了这一点.

例 4e　掷两枚均匀的骰子. 令 E 表示事件"骰子点数之和为 7"，令 F 表示事件"第一枚骰子点数为 4"以及 G 表示事件"第二枚骰子点数为 3". 从例 4c 可以得知，E 和 F 是独立的，同样的推理表明 E 和 G 也是独立的. 但是，很明显 E 和 FG 是不独立的，因为
76 $P(E|FG)=1$. ■

从例 4e 还可以得出关于三个事件 E，F 和 G 的独立性的合适的定义，三个事件的独立性比要求所有 $\binom{3}{2}$ 对事件的独立更强，我们给出如下定义.

定义　三个事件 E，F 和 G 称为独立的，如果

$$\begin{aligned}P(EFG) &= P(E)P(F)P(G)\\P(EF) &= P(E)P(F)\\P(EG) &= P(E)P(G)\\P(FG) &= P(F)P(G)\end{aligned}$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

注意，如果 E，F 和 G 是独立的，那么 E 与 F 和 G 的任意组合事件都是独立的. 比如，E 和 $F\cup G$ 就是独立的，因为

$$\begin{aligned}P[E(F\cup G)] &= P(EF\cup EG) = P(EF)+P(EG)-P(EFG)\\ &= P(E)P(F)+P(E)P(G)-P(E)P(FG)\\ &= P(E)[P(F)+P(G)-P(FG)] = P(E)P(F\cup G)\end{aligned}$$

当然，还可以把独立性的定义推广到三个以上事件. 事件 E_1，E_2，…，E_n 称为独立的，如果对这些事件的任意子集 $E_{1'}$，$E_{2'}$，…，$E_{r'}$，$r'\leqslant n$，都有

$$P(E_{1'}E_{2'}\cdots E_{r'}) = P(E_{1'})P(E_{2'})\cdots P(E_{r'})$$

最后，我们来定义无限个事件的独立性，如果无限个事件的任意有限个子集都是独立的，则称这无限个事件是独立的.

有时会遇到这种情况，所考虑的概率试验由一系列子试验组成. 例如，连续抛掷一枚硬币这个试验，就可以把每掷一次看作一个子试验. 在许多场合下，假定任一组子试验的结果不影响其他子试验的结果是合理的. 如果真是这样，我们称这些子试验是独立的. 更确切地说，如果任意的事件序列 E_1，E_2，…，E_n，…是独立的，则称这一系列子试验是独立的，这里，事件 E_i 完全由第 i 次子试验的结果所决定.

如果各个子试验彼此相同，即各子试验有相同的(子)样本空间及相同的事件概率函数，那么就称这些试验为重复试验.

例 4f 进行一个独立的无穷序列的重复试验，每次试验成功的概率为 p，失败的概率为 $1-p$，试求如下概率：

(a) 前 n 次试验中至少成功 1 次；(b) 前 n 次试验中恰好成功 k 次；(c) 所有试验结果都成功. 77

解 为计算前 n 次试验中至少成功 1 次的概率，最简单的方法是先求它的对立事件的概率，它的对立事件就是“前 n 次试验全失败了”. 若以 E_i 表示第 i 次试验失败这一事件，则由独立性可得，前 n 次试验全失败的概率是

$$P(E_1E_2\cdots E_n) = P(E_1)P(E_2)\cdots P(E_n) = (1-p)^n$$

因此，(a) 的答案就是 $1-(1-p)^n$.

为计算(b)的答案，考虑任一个由 k 个成功、$n-k$ 个失败组成的前 n 个结果的特定序列. 由独立性知，每个这样的序列发生的概率为 $p^k(1-p)^{n-k}$. 由于共有 $\binom{n}{k}$ 个这样的序列(由 k 个成功与 $n-k$ 个失败组成的排列种数为 $n!/[k!(n-k)!]$)，故(b)中的所求概率为

$$P\{\text{恰有 } k \text{ 次成功}\} = \binom{n}{k}p^k(1-p)^{n-k}$$

为计算(c)的答案，我们由(a)注意到，前 n 次试验全成功的概率为

$$P(E_1^{c}E_2^{c}\cdots E_n^{c}) = p^n$$

因此，运用概率的连续性属性(2.6 节)，我们可得所求概率为

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P\left(\bigcap_{i=1}^{\infty} E_i^c\right) = P\left(\lim_{n\to\infty}\bigcap_{i=1}^{n} E_i^c\right) = \lim_{n\to\infty} P\left(\bigcap_{i=1}^{n} E_i^c\right) = \lim_{n\to\infty} p^n = \begin{cases} 0 & \text{如果 } p<1 \\ 1 & \text{如果 } p=1 \end{cases}$$ ■

例 4g　由 n 个元件组成的系统称为并联的，如果至少有一个元件工作正常，那么整个系统都工作正常(见图 3-2)．对于这样的系统，如果元件 i 工作正常的概率为 p_i，$i=1$，…，n，并且各元件的工作状态相互独立．那么整个系统工作正常的概率是多大？

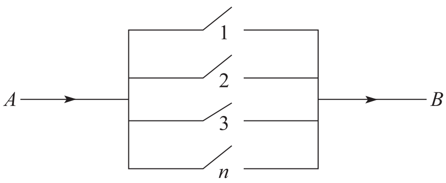


图 3-2　并联系统：只要有一个开关是通的，A 与 B 之间就是通的

解　令 A_i 表示事件“元件 i 工作正常”，那么

$$\begin{aligned} P(\{\text{系统工作正常}\}) &= 1 - P(\{\text{系统工作不正常}\}) \\ &= 1 - P(\{\text{所有元件工作不正常}\}) \\ &= 1 - P\left(\bigcap_i A_i^c\right) = 1 - \prod_{i=1}^{n}(1-p_i) \qquad \text{利用独立性} \end{aligned}$$

78 ■

例 4h　进行独立重复试验，每次试验为掷两枚均匀的骰子，每次试验的结果是两枚骰子点数之和，那么“和为 5”出现在“和为 7”之前的概率是多少？

解　令 E_n 表示事件“前 $n-1$ 次试验中，结果 5 和 7 都不出现，而第 n 次试验出现 5”，那么所求概率为

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$$

因为任一次试验中，$P(\{\text{和为 }5\})=4/36$，且 $P(\{\text{和为 }7\})=6/36$，这样，利用试验的独立性可以得到

$$P(E_n) = \left(1-\frac{10}{36}\right)^{n-1} \times \frac{4}{36}$$

因此，

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \frac{1}{9}\sum_{n=1}^{\infty}\left(\frac{13}{18}\right)^{n-1} = \frac{1}{9} \times \frac{1}{1-13/18} = \frac{2}{5}$$

该结果还可以利用条件概率得到．令 E 表示事件“和为 5 出现在和为 7 之前”，那么以首次试验结果为条件也可得到所求概率，方法如下：令 F 表示事件“第一次试验中骰子的点数之和为 5”，G 表示事件“第一次试验中骰子的点数之和为 7”，H 表示事件“第一次试验中骰子的点数之和既不是 5 也不是 7”．利用全概率公式，有

$$P(E) = P(E|F)P(F) + P(E|G)P(G) + P(E|H)P(H)$$

然而，

$$P(E|F) = 1 \qquad P(E|G) = 0 \qquad P(E|H) = P(E)$$

前两个等式是显而易见的，第三个等式是因为：第一次结果既不是 5，又不是 7，那么这种
79 情况下又相当于试验重新开始了．也就是说，试验者将要连续扔两枚骰子直到两枚骰子的点数之和为 5 或者 7．另一方面，每次试验都是独立的，因此，第一次试验的结果不会对接下来的试验有影响．因为 $P(F)=4/36$，$P(G)=6/36$，$P(H)=26/36$，所以

HZ BOOKS 华章图书

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P(E)=\frac{1}{9}+P(E)\frac{13}{18}$$

或者

$$P(E)=\frac{2}{5}$$

读者可能会发现答案很直观，因为既然 5 出现的概率为 4/36 而 7 出现的概率为 6/36，那么直观地看，5 出现在 7 之前的比例为 4∶6，因此概率就为 4/10，事实上的确如此.

这说明了如果 E 和 F 是一次试验中的两个互不相容事件，那么在连续试验时，事件 E 在事件 F 之前发生的概率为

$$\frac{P(E)}{P(E)+P(F)}$$

■

例 4i 有 n 种类型的优惠券，某人收集到第 i 种优惠券的概率为 p_i，$\sum_{i=1}^{n}p_i=1$. 假定各种券的收集是相互独立的. 假设这个人收集了 k 张优惠券，令 A_i 表示事件“其中至少有一张第 i 种优惠券”，对于 $i\neq j$，计算

(a) $P(A_i)$；(b) $P(A_i\cup A_j)$；(c) $P(A_i|A_j)$

解

$$P(A_i)=1-P(A_i^c)=1-P(\{\text{没有第 } i \text{ 种优惠券}\})=1-(1-p_i)^k$$

上面利用了每种优惠券的收集都是独立的，不是第 i 种优惠券的概率为 $1-p_i$. 类似地

$$\begin{aligned}P(A_i\cup A_j)&=1-P\big((A_i\cup A_j)^c\big)\\&=1-P(\{\text{既没有第 } i \text{ 种优惠券，也没有第 } j \text{ 种优惠券}\})\\&=1-(1-p_i-p_j)^k\end{aligned}$$

此处利用了每种优惠券的收集都是独立的，既不是第 i 种优惠券也不是第 j 种优惠券的概率为 $1-p_i-p_j$.

为了计算 $P(A_i|A_j)$，我们利用等式

$$P(A_i\cup A_j)=P(A_i)+P(A_j)-P(A_iA_j)$$

其中，利用(a)和(b)，可以得到

$$\begin{aligned}P(A_iA_j)&=1-(1-p_i)^k+1-(1-p_j)^k-[1-(1-p_i-p_j)^k]\\&=1-(1-p_i)^k-(1-p_j)^k+(1-p_i-p_j)^k\end{aligned}$$

这样，

$$P(A_i|A_j)=\frac{P(A_iA_j)}{P(A_j)}=\frac{1-(1-p_i)^k-(1-p_j)^k+(1-p_i-p_j)^k}{1-(1-p_j)^k}$$

■

80

下面这个例子陈述了一个在概率论历史中地位很高的问题，即著名的*点数问题*(problem of the points). 具体来说，该问题是这样的：两个赌徒下了注后，就按照某种方式进行赌博，规定得胜者赢得所有赌注. 但在谁也没得胜之前，赌博因故中止了. 此时每人都获得了一些“得分”，那么这些赌本该如何分呢？

这个问题是 1654 年 de Méré 爵士向法国数学家帕斯卡提出的，爵士当时是一个职业赌徒. 在攻克这一难题的过程中，帕斯卡提出了这样一个重要的思想：赌徒赢得赌本的比例

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

取决于如果比赛继续进行下去，他们各自能取胜的概率．帕斯卡解决了一些特殊情形，更重要的是，他开始与法国著名的数学家费马建立了通信联系，讨论该问题．他们之间通信的结果，不仅完全解决了点数问题，而且还为解决很多其他机会游戏问题搭好了框架．有些人把他们建立通信联系的这一天看作概率论的生日．他们之间的著名的通信往来，对于激发欧洲数学家对概率论的兴趣也起了重要的作用．因为当时一流的数学家都认识帕斯卡和费马．例如，在他们建立联系后不久，荷兰的年轻数学家惠更斯也来到了巴黎，和他们一起讨论这些问题和解法，而且，人们对这个新领域的兴趣和积极性也迅速高涨起来．

例 4j 点数问题　假设在独立重复试验中，每次成功的概率为 p，失败的概率为 $1-p$．在 m 次失败之前已有 n 次成功的概率是多大？设想 A 和 B 进行这样的赌博：当试验成功时，A 得 1 分，试验失败时，B 得 1 分，如果 A 先得到 n 分，那么 A 获胜，如果 B 先得到 m 分，那么 B 获胜，所求概率就是 A 获胜的概率．

解　以下将给出两种解答．第一种是帕斯卡给出的，第二种是费马给出的．

令 $P_{n,m}$ 表示在 m 次失败之前已经出现了 n 次成功，以第一次的结果为条件，可得

$$P_{n,m} = pP_{n-1,m} + (1-p)P_{n,m-1}, \qquad n \geqslant 1, m \geqslant 1$$

(为什么？给出理由.)利用很明显的边界条件 $P_{n,0}=0$，$P_{0,m}=1$，该等式能解出 $P_{n,m}$．求解 $P_{n,m}$ 的细节非常枯燥，现在来看看费马的解答．

费马论证了，要使得 n 次成功出现在 m 次失败之前，那么在前 $m+n-1$ 次试验中，至少有 n 次成功．(即使在第 $m+n-1$ 次试验之前，赌博就结束了，我们仍可以假设剩下的试验继续进行.)事实上，如果在前 $m+n-1$ 次试验中至少有 n 次成功，那么至多有 $m-1$

81 次失败，因此 n 次成功必然出现在 m 次失败之前．另一方面，如果 $m+n-1$ 次试验中不超过 n 次成功，那么至少有 m 次失败，因此，n 次成功不会出现在 m 次失败之前．

因此，如例 4f 所示，$m+n-1$ 次试验中恰好有 k 次成功的概率为

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

我们可得所求的 n 次成功出现在 m 次失败之前的概率为

$$P_{n,m} = \sum_{k=n}^{m+n-1} \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}$$ ■

下一个例子将告诉我们，当确定一个人赢得一场比赛的概率时，假设比赛在胜负决出后依然继续时更容易．

例 4k 发球和接球的游戏协议　考虑一个要在两个选手 A 和 B 之间发球和接球的游戏，比如排球、羽毛球或壁球．这个比赛是由有顺序的接球组成，而每个接球由一个选手发球开始，然后持续到一个选手赢．赢得一球的人得一分，当两人中有一人赢得 n 分时，比赛结束．假定 A 开始发球，A 赢球的概率为 p_A，B 赢球的概率为 $q_A=1-p_A$．当 B 先发球时，A 赢球的概率为 p_B，B 赢球的概率为 $q_B=1-p_B$．A 开始发球．同时存在两种可能的协议：一个为得分发球，即上一个球得分的人将发下一次球；一个为交替发球，即两个人交替发球，这样的话，没有连续的两个球是由一个人发的．例如，当 $n=3$ 时，在得分发球的协议下，如果 A 赢得第一球，B 赢得第二球，A 赢得接下来两球，发球者会是 A，B，

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

A，A. 在交替发球协议下，发球者分别是 A，B，A，B，A，…，直到决出胜负比赛结束. 如果你是 A，请问你会选择哪个协议？

解 令人惊奇的是，A 不管选哪一个协议，其结果都是一样的. 为证明这一点，我们不妨假设比赛一直持续到两人的总分为 $2n-1$ 为止. 第一个赢得 n 分的人则是在这场总分为 $2n-1$ 分的比赛中赢得至少 n 分的人. 刚开始，我们要注意，如果使用交替协议，那么选手 A 要发 n 次球，B 要发 $n-1$ 次球. 现在，考虑得分发球协议，同样假设比赛总分持续到 $2n-1$ 分. 由于当赢的人已经决出时，不管是谁在发多余的球，都已无意义，所以我们假设当赢的人已经决出时，输的人发剩下的球. 这里我们要注意现在这个修改过的协议不会改变答案. 下面两个情形说明这一点.

情形 1：A 赢得比赛.

由于 A 先发球，A 的第二次发球紧随 A 的第一次赢球，A 的第三次发球紧随 A 的第二次赢球，这样，A 的第 n 次发球会紧随 A 的第 $n-1$ 次赢球，但是这会是 A 的最后一次发球. 之所以有这样的结果，是因为或者 A 赢球，即 A 已赢得第 n 分，或者 A 输球，即 B 开始不停地发球直至 A 赢得 n 分. 这样我们可以得到，在修正的协议里，A 赢得比赛意味着 A 总共发了 n 次球. 82

情形 2：B 赢得比赛.

由于 A 先发球，B 的第一个发球是紧随着 B 的第一次得分. B 的第二次发球紧随着 B 的第二次得分，这样，B 的第 $n-1$ 次发球紧随着 B 的第 $n-1$ 次得分，但是这会是 B 的最后一次发球，因为当 B 得到这一分时 B 已经赢得比赛，当 B 输了这一分时，则 A 不停地发球直至 B 赢得第 n 分. 这样我们可以看到，B 赢得比赛时，B 将会发 $n-1$ 次球. 这样，由于总分是 $2n-1$，所以 A 将发 n 次球.

这样在这两个协议下，A 总会发 n 次球，而 B 发 $n-1$ 次球. 由于 A 赢得比赛的概率只和是谁发球有关，那么在这两个协议之下，A 赢得比赛的概率是一样的. ■

接下来两个例子和赌博有关，我们可以从中得到重要的结论. ㊀

例 4l 假设开始有 r 个选手，第 i 个选手有 n_i 枚游戏币，$n_i>0$，$i=1$，…，r. 每局选择两个选手来玩游戏，赢的人要从输的人手中获得一枚游戏币. 每个选手当其手中的游戏币输完时将被淘汰. 游戏持续到只剩一个选手，他手中应有所有的游戏币 $n\equiv\sum_{i=1}^{r}n_i$，假设每局游戏是独立的，而且每局游戏两个选手赢的概率是相等的. 求出第 i 个选手赢得比赛的概率.

解 开始，我们假设这里有 n 个选手，而每个人手中只有一枚游戏币. 考虑第 i 个选手，他每局有相同的概率输或赢一枚币，而每局之间是相互独立的. 另外，他会一直参加比赛，直到她手中的游戏币为 0 或 n. 由于这对于 n 个选手都是一样的，那么每个人有相同的概率赢得比赛，这表示每个人有 $1/n$ 的概率赢得比赛. 现在我们假设这 n 个人被分成了 r 个组，第 i 组有 $n_i(i=1，…，r)$个选手. 那么赢的人在第 i 组的概率是 n_i/n，但是由于

㊀ 本节其余内容可以作为选学内容.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

(a) 第 i 组开始有 n_i 枚游戏币，$i=1$，…，r，而且

(b) 每一次游戏由不同组的人玩，赢的概率相等，结果是赢的队多一枚游戏币，输的
83 队少一枚游戏币.

我们可以很容易地看到胜利者来自第 i 组的概率正好就是我们所求的概率，因此 $P_i=n_i/n$. 有趣的是，之前的讨论表明游戏结果和每一局中两个选手的选取并无关系.

在赌徒破产问题中，只有两个赌徒，但是他们赢分的概率不相等. ■

例 4m 赌徒破产问题　两个赌徒，就连续抛掷一枚硬币的结果进行打赌. 对于每一次抛掷的结果，如果是正面朝上，B 将支付给 A 一元，如果是反面朝上，A 将付给 B 一元. 一直这样下去，直到某一方钱输光. 假定连续抛掷硬币是独立的，且每次的结果正面朝上的概率为 p，假定开始时 A 有 i 元，B 有 $N-i$ 元，那么 A 最后能赢得所有钱的概率是多大？

解　令 E 表示事件"开始时 A 有 i 元，B 有 $N-i$ 元，而 A 最后赢得所有钱"，显然，它和 A 最初的钱数有关，记 $P_i=P(E)$. 以第一次掷硬币的结果为条件，令 H 表示事件"第一次抛掷结果为正面朝上"，则

$$P_i = P(E) = P(E|H)P(H) + P(E|H^c)P(H^c) = pP(E|H) + (1-p)P(E|H^c)$$

现在，假定第一次抛掷结果为正面朝上，第一次打赌结束后的状态是：A 有了 $i+1$ 元，而 B 有 $N-(i+1)$元. 因为随后的抛掷都同前面独立并且正面朝上的概率都为 p，因此，从该时刻开始，A 赢得所有钱的概率等同于这种情形：一开始 A 有 $i+1$ 元而 B 有 $N-(i+1)$元. 因此

$$P(E|H) = P_{i+1}$$

类似地，

$$P(E|H^c) = P_{i-1}$$

这样，令 $q=1-p$，可以得到

$$P_i = pP_{i+1} + qP_{i-1} \qquad i = 1,2,\cdots,N-1 \tag{4.2}$$

利用明显的边界条件 $P_0=0$ 和 $P_N=1$，就可以求解方程(4.2). 由于 $p+q=1$，上式等价于

$$pP_i + qP_i = pP_{i+1} + qP_{i-1}$$

或者

84
$$P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1}) \qquad i = 1,2,\cdots,N-1 \tag{4.3}$$

因此，由 $P_0=0$，利用方程(4.3)可得

$$P_2 - P_1 = \frac{q}{p}(P_1 - P_0) = \frac{q}{p}P_1$$

$$P_3 - P_2 = \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1$$

$$\vdots$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$
\begin{aligned}
P_i - P_{i-1} &= \frac{q}{p}(P_{i-1} - P_{i-2}) = \left(\frac{q}{p}\right)^{i-1} P_1 \\
&\vdots \\
P_N - P_{N-1} &= \frac{q}{p}(P_{N-1} - P_{N-2}) = \left(\frac{q}{p}\right)^{N-1} P_1
\end{aligned}
\tag{4.4}
$$

将方程组(4.4)中前 $i-1$ 个等式累加，可以得到

$$
P_i - P_1 = P_1\left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{i-1}\right]
$$

或者

$$
P_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)} P_1 & \text{如果}\dfrac{q}{p} \neq 1 \\ iP_1 & \text{如果}\dfrac{q}{p} = 1 \end{cases}
$$

再利用事实 $P_N=1$，可得

$$
P_1 = \begin{cases} \dfrac{1-(q/p)}{1-(q/p)^N} & \text{如果 } p \neq \dfrac{1}{2} \\ \dfrac{1}{N} & \text{如果 } p = \dfrac{1}{2} \end{cases}
$$

因此

$$
P_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)^N} & \text{如果 } p \neq \dfrac{1}{2} \\ \dfrac{i}{N} & \text{如果 } p = \dfrac{1}{2} \end{cases}
\tag{4.5}
$$

令 Q_i 表示开始时 A 有 i 元，B 有 $N-i$ 元，最后 B 赢得所有钱的概率，那么，由对称性，只需将 p 替换为 q，i 替换为 $N-i$，就有

$$
Q_i = \begin{cases} \dfrac{1-(q/p)^{N-i}}{1-(q/p)^N} & \text{如果 } q \neq \dfrac{1}{2} \\ \dfrac{N-i}{N} & \text{如果 } q = \dfrac{1}{2} \end{cases}
$$

85

而且，因为 $q=1/2$ 等价于 $p=1/2$，因此当 $q\neq 1/2$ 时，有

$$
\begin{aligned}
P_i + Q_i &= \frac{1-(q/p)^i}{1-(q/p)^N} + \frac{1-(p/q)^{N-i}}{1-(p/q)^N} = \frac{p^N - p^N(q/p)^i}{p^N - q^N} + \frac{q^N - q^N(p/q)^{N-i}}{q^N - p^N} \\
&= \frac{p^N - p^{N-i}q^i - q^N + q^i p^{N-i}}{p^N - q^N} = 1
\end{aligned}
$$

当 $p=q=1/2$ 时结论仍成立，所以

$$
P_i + Q_i = 1
$$

用语言来描述，就是 A 和 B 中某一人将赢得所有钱的概率为 1；或者说，A 的钱总在 1 与 $N-1$ 之间而赌博无休止地进行下去的概率为 0(读者必须注意，这场赌博有三个可能

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

结果，而不是两个，即或者 A 胜，或者 B 胜，或者谁也不胜．但我们刚才证明了最后一种结果的概率为 0.）

现在从数值上阐释上述结论．若开始时 A 有 5 元，而 B 有 10 元，则当 $p=1/2$ 时，A 得胜的概率为 1/3，而当 $p=0.6$ 时，A 得胜的概率猛增为

$$\frac{1-\left(\frac{2}{3}\right)^5}{1-\left(\frac{2}{3}\right)^{15}} \approx 0.87$$

赌徒破产问题还有一种特殊情形，称为赌博持续时间(duration of play)问题．这个问题是 1657 年法国数学家费马向荷兰数学家克里斯第安・惠更斯提出的，后来被惠更斯解决．设 A 和 B 每人有 12 枚硬币，他们以抛掷 3 个骰子的方法赌这些钱：若点数为 11(谁掷骰子都可以)，则 A 给 B 一枚硬币，如果点数为 14，则 B 给 A 一枚硬币．谁先赢得所有硬币谁就获胜．因为 $P(\{$点数为 11$\})=27/216$ 及 $P(\{$点数为 14$\})=15/216$，由例 4h 可以看出，对 A 而言，这正是 $p=15/42$，$i=12$，$N=24$ 情形下的赌徒破产问题．一般的赌徒破产问题由数学家詹姆斯・伯努利解决，其结果发表于 1713 年(他去世后的第 8 年).

作为赌徒破产问题的一个应用，考虑如下的药品试验问题，假设为治疗某种疾病，正在研制两种新药．新药 i 的治愈率为 $P_i(i=1, 2)$，即每个接受药品 i 治疗的病人被治愈的概率是 P_i．然而，P_i 是未知的．我们希望知道 $P_1>P_2$ 或 $P_2>P_1$．试验是成对、有序地进行的．对于各对病人，其中一人施以药品 1，另一人用药品 2．当其中一种药的治愈人数超过另一种药的治愈人数的一定数量时，试验就停止．令

$$X_j=\begin{cases}1 & \text{第 } j \text{ 对病人中用药品 1 者治愈了疾病}\\ 0 & \text{其他}\end{cases}$$

$$Y_j=\begin{cases}1 & \text{第 } j \text{ 对病人中用药品 2 者治愈了疾病}\\ 0 & \text{其他}\end{cases}$$

86

设 M 是事先确定的正整数，试验在第 N 次时停止，其中 N 是使下列两个等式之一第一次成立时的那个 n 的值：

$$X_1+\cdots+X_n-(Y_1+\cdots+Y_n)=M$$

或

$$X_1+\cdots+X_n-(Y_1+\cdots+Y_n)=-M$$

若 n 使第一个等式成立，就下结论 $P_1>P_2$；若 n 使第二个等式成立，则下结论 $P_1<P_2$.

为了确定这个方法的好坏，我们需要知道利用这个方法导致错误结论的概率．即在假设 $P_1>P_2$ 是一组药的治愈率的真值的条件下，作出 $P_2>P_1$ 的错误决定的概率．为确定这个概率，注意每次试验的结果：当药品 1 有效，而药品 2 无效时，这个累计差会增加 1，相应的概率为 $P_1(1-P_2)$．但是当药品 2 有效而药品 1 无效时，这个累计差会减少 1，其相应的概率为 $P_2(1-P_1)$．当两种药的疗效相同时，这个累计差保持不变，其相应概率为 $P_1P_2+(1-P_1)(1-P_2)$．如果我们只考虑累计差变化的那些试验，则累计差以概率

$$P=P(\{\text{累计差增加 1}\mid\text{累计差增加 1 或减少 1}\})=\frac{P_1(1-P_2)}{P_1(1-P_2)+P_2(1-P_1)}$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

增加 1，以概率

$$1-P=\frac{P_2(1-P_1)}{P_1(1-P_2)+P_2(1-P_1)}$$

减少 1．因此，作出判断 $P_2>P_1$ 的概率等于在赌徒累计赢 M 元钱之前累计输 M 元钱的概率(赌徒每次赢的概率为 P，每次输赢为 1 元钱)．在公式(4.5)中，令 $i=M$，$N=2M$，这个概率为

$$P(\{下结论\ P_2>P_1\})=1-\frac{1-\left(\frac{1-P}{P}\right)^M}{1-\left(\frac{1-P}{P}\right)^{2M}}=1-\frac{1}{1+\left(\frac{1-P}{P}\right)^M}=\frac{1}{1+\gamma^M}$$

其中 $\gamma=\frac{P}{1-P}=\frac{P_1(1-P_2)}{P_2(1-P_1)}$．例如，$P_1=0.6$，$P_2=0.4$，则当 $M=5$ 时，得出错误结论的概率为 0.017，当 $M=10$ 时，这个概率降到 0.0003． ■

下面我们介绍一种解决非概率问题的概率方法．设有一个集合，我们希望知道在这个集合中是否至少有一个元素具有某种特征．例如，在一群鸡中，我们希望知道是否有鸡患禽流感．我们的方法是随机地从该集合中取一个元素，抽取的方法是使得这个集合中的每个元素被抽取到的概率都大于 0．现在考虑抽到的元素不具有该特征的概率．若这个概率 87
为 1，则在这个集合中没有元素有这种特征．若这个概率小于 1，则在这个集合中至少有一个元素具有这种特征．

本节最后一个例子具体阐明了这种方法．

例 4n 首先给出 n 个顶点的完全图的定义：在平面上有 n 个点(称为顶点)，用 $\binom{n}{2}$ 条线段(称为边)将每对点联结起来．例如图 3-3 给出了 3 个顶点的完全图．假设 n 个顶点的完全图的每条边染成了红色或蓝色．一个有趣的问题是：对于给定的整数 k，是否存在一个染色方法，使得该图上任意给定的 k 个顶点，其相应的 $\binom{k}{2}$ 条边不是同一种颜色．通过概率论证可知，如果 n 不是太大，答案是肯定的．

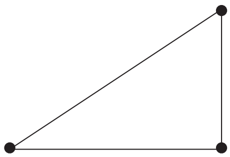


图 3-3　3 个顶点的完全图

论证如下：假设每条边独立等可能地被染成红色或蓝色．即每条边为红色的概率为 1/2．将这 $\binom{n}{k}$ 个 k 个顶点所组成的子集编号，定义事件 E_i，$i=1,\cdots,\binom{n}{k}$ 如下：

$$E_i=\{第\ i\ 个\ k\ 顶点集合的所有边颜色相同\}$$

这样，由于 k 顶点集合的 $\binom{k}{2}$ 条边的每条都等可能地染成红色或蓝色，因此，它们颜色相同的概率为

$$P(E_i)=2\times\left(\frac{1}{2}\right)^{k(k-1)/2}$$

又因为

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P\Big(\bigcup_i E_i\Big) \leqslant \sum_i P(E_i) \qquad \text{布尔不等式}$$

所以我们可得“至少存在一个 k 顶点集合，其所有边颜色相同”的概率 $P\Big(\bigcup_i E_i\Big)$ 满足

88

$$P\Big(\bigcup_i E_i\Big) \leqslant \binom{n}{k}\Big(\frac{1}{2}\Big)^{k(k-1)/2-1}$$

因此，如果

$$\binom{n}{k}\Big(\frac{1}{2}\Big)^{k(k-1)/2-1} < 1$$

或者，等价地

$$\binom{n}{k} < 2^{k(k-1)/2-1}$$

则 $\binom{n}{k}$ 个 k 顶点集合里，“至少存在一个 k 顶点集合，其所有的边颜色都相同”的概率小于 1. 因此，在前述 n 和 k 的条件下，“没有一个 k 顶点集合，其所有边颜色相同”的概率为正数，这意味着至少存在一种染色方法，使得对任意 k 顶点集合，其所有的边染色不全相同. ■

注释

(a) 上面的论证列出了关于 n 和 k 的条件，在这样条件下，存在一种染色方法满足所要求的性质即可，但它并没有告诉我们如何染色，使得所染的颜色满足所要求的命题(当然，可以随机地染色，然后检查所染的颜色是否满足所要求的性质，若不成，再重复一次，直到成功为止).

(b) 将概率引进那些纯粹是确定问题的方法称为概率化方法(probabilistic-method[⊖]). 此方法的其他例子在理论习题 24 以及第 7 章的例 2t 和例 2u 中给出.

3.5 $P(\cdot \mid F)$是概率

条件概率满足普通概率的所有性质，命题 5.1 证明了条件概率 $P(E|F)$满足概率的三条公理.

命题 5.1

(a) $0 \leqslant P(E|F) \leqslant 1$.

(b) $P(S|F)=1$.

(c) 若 $E_i(i=1, 2, \cdots)$为互不相容的事件序列，则

$$P\Big(\bigcup_{i=1}^{\infty} E_i \,|\, F\Big) = \sum_{i=1}^{\infty} P(E_i|F)$$

证明 为了证明(a)，我们必须证明 $0 \leqslant P(EF)/P(F) \leqslant 1$. 不等式左边是显然的，而不等式的右边成立是因为 $EF \subset F$ 成立意味着 $P(EF) \leqslant P(F)$. (b)成立是因为

⊖ 参见 N. Alon, J. Spencer, and P. Erdos, *The Probabilistic Method* (New York: John Wiley & Sons, Inc., 1992).

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P(S|F)=\frac{P(SF)}{P(F)}=\frac{P(F)}{P(F)}=1$$ 89

(c)成立是因为

$$P\left(\bigcup_{i=1}^{\infty} E_i \,|\, F\right)=\frac{P\left(\left(\bigcup_{i=1}^{\infty} E_i\right)F\right)}{P(F)}$$

$$=\frac{P\left(\bigcup_{i=1}^{\infty} E_iF\right)}{P(F)} \qquad \text{因为}\left(\bigcup_{i=1}^{\infty} E_i\right)F=\bigcup_{i=1}^{\infty} E_iF$$

$$=\frac{\sum_{i=1}^{\infty} P(E_iF)}{P(F)}=\sum_{i=1}^{\infty} P(E_i|F)$$

其中，倒数第二个等式成立是因为 $E_iE_j=\varnothing$，这意味着 $E_iFE_jF=\varnothing$. □

如果我们定义 $Q(E)=P(E|F)$，那么根据命题5.1，$Q(E)$可认为是关于 S 中事件的概率函数. 因此，前面证明的关于概率的命题它都满足. 例如，我们有

$$Q(E_1 \cup E_2)=Q(E_1)+Q(E_2)-Q(E_1E_2)$$

或者等价地，

$$P(E_1 \cup E_2|F)=P(E_1|F)+P(E_2|F)-P(E_1E_2|F)$$

而且，如果我们定义条件概率 $Q(E_1|E_2)$为 $Q(E_1|E_2)=Q(E_1E_2)/Q(E_2)$，那么根据式(3.1)可得

$$Q(E_1)=Q(E_1|E_2)Q(E_2)+Q(E_1|E_2^c)Q(E_2^c) \tag{5.1}$$

由于

$$Q(E_1|E_2)=\frac{Q(E_1E_2)}{Q(E_2)}=\frac{P(E_1E_2|F)}{P(E_2|F)}=\frac{P(E_1E_2F)/P(F)}{P(E_2F)/P(F)}=P(E_1|E_2F)$$

所以式(5.1)等价于

$$P(E_1|F)=P(E_1|E_2F)P(E_2|F)+P(E_1|E_2^cF)P(E_2^c|F)$$

例5a 考虑例3a，保险公司认为人可以分为两种不同的类，一类易出事故，另一类不易出事故. 在任意给定的一年内，易出事故者将发生事故的概率为0.4，而对不易出事故者来说，此概率为0.2. 若已知某新保险客户在第一年已经出过一次事故，问他在保险有效的第二年又出一次事故的条件概率是多大？ 90

解 如果令 A 表示“该保险客户是易出事故者”这一事件，而 $A_i(i=1,2)$表示“他在第 i 年出一次事故”. 那么，以他是不是易出事故者为条件，可以算出所求概率 $P(A_2|A_1)$ 如下：

$$P(A_2|A_1)=P(A_2|AA_1)P(A|A_1)+P(A_2|A^cA_1)P(A^c|A_1)$$

而

$$P(A|A_1)=\frac{P(A_1A)}{P(A_1)}=\frac{P(A_1|A)P(A)}{P(A_1)}$$

但是，例3a中已经假设 $P(A)=3/10$，且算出了 $P(A_1)=0.26$，因此

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$P(A|A_1)=\frac{0.4\times0.3}{0.26}=\frac{6}{13}$$

从而

$$P(A^c|A_1)=1-P(A|A_1)=\frac{7}{13}$$

因为 $P(A_2|AA_1)=0.4$，$P(A_2|A^cA_1)=0.2$，所以

$$P(A_2|A_1)=0.4\times\frac{6}{13}+0.2\times\frac{7}{13}\approx0.29$$ ■

例 5b　一只母猩猩生了一只幼猩猩，但是，却不能断定两只公猩猩究竟哪一只是父亲．在进行基因分析之前，有迹象表明第一只公猩猩为父亲的概率为 p，第二只为父亲的概率为 $1-p$．从这三只猩猩身上获得的 DNA 表明，对于一个特殊的基因组，母猩猩具有基因对(A，A)，第一只公猩猩具有基因对(a，a)，而第二只公猩猩具有基因对(A，a)．如果 DNA 检验表明幼猩猩具有基因对(A，a)，那么第一只公猩猩是父亲的概率是多少？

解　令所有概率都是以事件“母猩猩有基因对(A，A)，第一只公猩猩具有基因对(a，a)，第二只公猩猩具有基因对(A，a)”为条件的条件概率．又令 $M_i(i=1，2)$表示第 i 只公猩猩为父亲这一事件．令 $B_{A,a}$ 表示幼猩猩具有基因对(A，a)这一事件，那么如下可得到 $P(M_1|B_{A,a})$：

$$\begin{aligned}P(M_1|B_{A,a})&=\frac{P(M_1B_{A,a})}{P(B_{A,a})}=\frac{P(B_{A,a}|M_1)P(M_1)}{P(B_{A,a}|M_1)P(M_1)+P(B_{A,a}|M_2)P(M_2)}\\&=\frac{1\times p}{1\times p+1/2\times(1-p)}=\frac{2p}{1+p}\end{aligned}$$

因为当 $p<1$ 时 $2p/(1+p)>p$，所以幼猩猩的基因对为(A，a)这一信息增加了第一只公猩
91 猩为父亲的概率．这点很直观，因为相对于 M_2 来说，M_1 成立的条件下，幼猩猩的基因对为(A，a)的可能性更大(各自的条件概率分别为 1 和 1/2)． ■

下面的例子研究游程理论中的一个问题．

例 5c　设有一个独立重复试验序列，每次试验成功的概率为 p，失败的概率为 $q=1-p$．计算长度为 n 的成功游程先于长度为 m 的失败游程出现的概率．

解　记 E 为事件“长度为 n 的成功游程先于长度为 m 的失败游程出现”．以第一次试验结果为条件，得到

$$P(E)=pP(E|H)+qP(E|H^c) \tag{5.2}$$

其中 H 表示第一次试验成功．

现在假定第一次试验成功，为了使得长度为 n 的成功游程先出现，我们希望接着 $n-1$ 次试验均成功．现在令 F 表示事件“第 2 次到第 n 次试验均成功”．以 F 为条件，我们得到

$$P(E|H)=P(E|FH)P(F|H)+P(E|F^cH)P(F^c|H) \tag{5.3}$$

一方面，显然，$P(E|FH)=1$．另一方面，若 F^cH 发生，则说明第一次试验成功，但在后面的 $n-1$ 次试验中，至少有一次试验失败．然而，当失败发生时，前面的成功已经失去作用，试验相当于从失败开始，因此

$$P(E|F^cH)=P(E|H^c)$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

由试验的相互独立性，可知 F 和 H 是独立的，即 $P(F|H)=P(F)=p^{n-1}$，因此，由式(5.3)可知，

$$P(E|H) = p^{n-1} + (1-p^{n-1})P(E|H^c) \tag{5.4}$$

用类似的方法可以得到 $P(E|H^c)$的表达式．即，令 G 表示事件“第 2 次试验直到第 m 次试验均失败”．于是

$$P(E|H^c) = P(E|GH^c)P(G|H^c) + P(E|G^cH^c)P(G^c|H^c) \tag{5.5}$$

GH^c 表示前 m 次试验均失败，故 $P(E|GH^c)=0$．当 G^cH^c 发生时，则说明第一次试验失败，但在后面的 $m-1$ 次试验中，至少有一次试验成功．因此，这次成功使得所有以前的失败已经失去作用，相当于从成功开始，即

$$P(E|G^cH^c) = P(E|H)$$

又因为 $P(G^c|H^c)=P(G^c)=1-q^{m-1}$，所以由式(5.5)可得

$$P(E|H^c) = (1-q^{m-1})P(E|H) \tag{5.6}$$

求解方程(5.4)和方程(5.6)可得

$$P(E|H) = \frac{p^{n-1}}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}}$$

$$P(E|H^c) = \frac{(1-q^{m-1})p^{n-1}}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}}$$ 92

最后得到

$$\begin{aligned} P(E) &= pP(E|H) + qP(E|H^c) = \frac{p^n + qp^{n-1}(1-q^{m-1})}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}} \\ &= \frac{p^{n-1}(1-q^m)}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}} \end{aligned} \tag{5.7}$$

有趣的是，注意到问题的对称性，可知长度为 m 的失败游程先于长度为 n 的成功游程出现的概率也可用式(5.7)进行计算，不过公式中 m 和 n 对换，p 和 q 对换．

$$\begin{aligned} &P(\{\text{长度为 } m \text{ 的失败游程先于长度为 } n \text{ 的成功游程出现}\}) \\ &= \frac{q^{m-1}(1-p^n)}{q^{m-1}+p^{n-1}-q^{m-1}p^{n-1}} \end{aligned} \tag{5.8}$$

因为式(5.7)和式(5.8)之和为 1，所以长度为 n 的成功游程和长度为 m 的失败游程终有一个会出现．

作为式(5.7)的一个具体例子，我们注意，在掷一枚均匀的硬币试验中，长度为 2 的正面游程先于长度为 3 的反面游程出现的概率为 7/10．长度为 2 的正面游程先于长度为 4 的反面游程出现的概率为 5/6． ■

在接下来的例子中我们再次研究配对问题，这次运用条件概率解答问题．

例 5d 在一次聚会上，n 个人摘下他们的帽子，然后把这些帽子混合在一起，每人再随机选择一顶帽子．如某个人选中了他自己的帽子，我们就说出现了一个配对．求以下事件的概率．

(a) 没有配对；(b) 恰有 k 个配对．

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

解　(a)令 E 表示“没有配对”这一事件，它显然与 n 有关，因此可记 $P_n=P(E)$. 以第一个人是否选中自己的帽子(分别记为 M 和 M^c)为条件，有

$$P_n=P(E)=P(E|M)P(M)+P(E|M^c)P(M^c)$$

显然 $P(E|M)=0$，因此

$$P_n=P(E|M^c)\frac{n-1}{n} \tag{5.9}$$

现在，$P(E|M^c)$是在已知 $n-1$ 个人中有一个特殊的人(此人的帽子已被第一人选走)必定选不到自己帽子的条件下，这 $n-1$ 个人选 $n-1$ 顶帽子没有配对的概率. 此处有两种互不相容的选取方式：该特殊的人没有选中第一人的帽子且其余的人中也没有配对；该特殊的人选中了第一人的帽子，且其余的人中也没有配对. 前者的概率正是 P_{n-1}，此时可把该特殊的人理解成第一人，而后者的概率为$[1/(n-1)]P_{n-2}$，我们有

93 $$P(E|M^c)=P_{n-1}+\frac{1}{n-1}P_{n-2}$$

于是由式(5.9)可得

$$P_n=\frac{n-1}{n}P_{n-1}+\frac{1}{n}P_{n-2}$$

或等价地，

$$P_n-P_{n-1}=-\frac{1}{n}(P_{n-1}-P_{n-2}) \tag{5.10}$$

但是，由于 P_n 表示 n 个人在他们的帽子中任选一顶没有配对的概率，我们有

$$P_1=0\qquad P_2=\frac{1}{2}$$

从而由式(5.10)可得

$$P_3-P_2=-\frac{P_2-P_1}{3}=-\frac{1}{3!}\quad 或\quad P_3=\frac{1}{2!}-\frac{1}{3!}$$

$$P_4-P_3=-\frac{P_3-P_2}{4}=\frac{1}{4!}\quad 或\quad P_4=\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}$$

一般地，我们有

$$P_n=\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}-\cdots+\frac{(-1)^n}{n!}$$

(b)为了计算正好有 k 个配对的概率，先考虑固定的某 k 个人，他们且只有他们选中自己的帽子的概率为

$$\frac{1}{n}\,\frac{1}{n-1}\cdots\frac{1}{n-(k-1)}P_{n-k}=\frac{(n-k)!}{n!}P_{n-k}$$

其中 P_{n-k}是已知 k 个人选中自己的帽子，其余 $n-k$ 个人在他们自己的 $n-k$ 顶帽子中选取而没有配对的条件概率，再因这 k 个人有$\binom{n}{k}$种选法，故正好有 k 个配对的概率为

$$\frac{P_{n-k}}{k!}=\frac{\dfrac{1}{2!}-\dfrac{1}{3!}+\cdots+\dfrac{(-1)^{n-k}}{(n-k)!}}{k!}$$

■

HZ BOOKS 华章图书

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

概率论中的一个重要概念是事件的条件独立性. 如果已知 F 发生的条件下，E_1 发生的概率不因 E_2 是否发生而改变，则称事件 E_1 和 E_2 关于给定的事件 F 是条件独立的(conditionally independent). 更确切地说，如果

$$P(E_1 \mid E_2 F) = P(E_1 \mid F) \tag{5.11}$$

或等价地，

$$P(E_1 E_2 \mid F) = P(E_1 \mid F) P(E_2 \mid F) \tag{5.12}$$

则称 E_1 与 E_2 关于 F 是条件独立的.

条件独立的概念容易推广到两个以上事件的情形，我们把它留作习题.

读者应该注意到，条件独立性的概念在例 5a 中已经用过了，在那里，我们假定：在已知保险客户是否为易出事故者的情况下，“他在第 $i(i=1,2,\cdots)$年，出一次事故”这些事件是条件独立的. 下一个例题，有时也称为拉普拉斯继承准则，进一步解释条件独立性的概念. 94

例 5e **拉普拉斯继承准则** 一个盒子中有 $k+1$ 枚不均匀的硬币，抛掷第 i 枚硬币时，其正面朝上的概率为 i/k，$i=0,1,\cdots,k$. 从盒子中随机取出一枚硬币，并重复地抛掷，如果前 n 次抛掷结果都为正面朝上，那么第 $n+1$ 次结果仍为正面朝上的概率是多少？

解 令 $C_i(i=0,1,\cdots,k)$表示开始取出的是第 i 枚硬币这一事件，F_n 表示前 n 次结果都为正面朝上，H 表示第 $n+1$ 次抛掷正面朝上. 所求概率 $P(H \mid F_n)$为：

$$P(H \mid F_n) = \sum_{i=0}^{k} P(H \mid F_n C_i) P(C_i \mid F_n)$$

现在，已知取出的是第 i 枚硬币，假设各次抛掷的结果条件独立是合理的，每次出现正面朝上的概率为 i/k. 于是有

$$P(H \mid F_n C_i) = P(H \mid C_i) = \frac{i}{k}$$

而且，

$$P(C_i \mid F_n) = \frac{P(C_i F_n)}{P(F_n)} = \frac{P(F_n \mid C_i) P(C_i)}{\sum_{j=0}^{k} P(F_n \mid C_j) P(C_j)} = \frac{(i/k)^n [1/(k+1)]}{\sum_{j=0}^{k} (j/k)^n [1/(k+1)]}$$

因此

$$P(H \mid F_n) = \frac{\sum_{i=0}^{k} (i/k)^{n+1}}{\sum_{j=0}^{k} (j/k)^n}$$

但当 k 很大时，可利用积分近似

$$\frac{1}{k} \sum_{i=0}^{k} \left(\frac{i}{k}\right)^{n+1} \approx \int_0^1 x^{n+1} \mathrm{d}x = \frac{1}{n+2}$$

$$\frac{1}{k} \sum_{j=0}^{k} \left(\frac{j}{k}\right)^{n} \approx \int_0^1 x^{n} \mathrm{d}x = \frac{1}{n+1}$$

故对很大的 k 有

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

95

$$P(H|F_n) \approx \frac{n+1}{n+2}$$ ■

例 5f 序贯地补充信息　假设有 n 个互不相容且穷举的假设，其初始概率[有时也称为先验(prior)概率]为 $P(H_i)$，$\sum_{i=1}^{n} P(H_i)=1$．现在，如果得知"事件 E 发生"，那么 H_i 成立的条件概率为[有时称为 H_i 的后验(posterior)概率]：

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_j P(E|H_j)P(H_j)} \tag{5.13}$$

假设现在我们首先知道 E_1 发生，然后 E_2 发生．那么，如果仅仅得知第一条信息，则 H_i 为真假设的条件概率为

$$P(H_i|E_1) = \frac{P(E_1|H_i)P(H_i)}{P(E_1)} = \frac{P(E_1|H_i)P(H_i)}{\sum_j P(E_1|H_j)P(H_j)}$$

而如果得知两条信息，则 H_i 为真假设的条件概率 $P(H_i|E_1E_2)$可以如下计算：

$$P(H_i|E_1E_2) = \frac{P(E_1E_2|H_i)P(H_i)}{\sum_j P(E_1E_2|H_j)P(H_j)}$$

然而，或许有人疑惑，可否这样计算 $P(H_i|E_1E_2)$：利用式(5.13)的右边，将 E 替换为 E_2，将 $P(H_j)$替换为 $P(H_j|E_1)$，$j=1$，…，n．即，将 $P(H_j|E_1)(j\geqslant 1)$作为先验概率，将 E_2 作为新近得到的信息，然后利用式(5.13)来计算后验概率？

解　上述方法是合理的，条件是：对每一个 $j=1$，…，n，在给定 H_j 的条件下，事件 E_1 和 E_2 是条件独立的．如果真是这样，那么

$$P(E_1E_2|H_j) = P(E_2|H_j)P(E_1|H_j) \qquad j=1,\cdots,n$$

因此，

$$\begin{aligned} P(H_i|E_1E_2) &= \frac{P(E_2|H_i)P(E_1|H_i)P(H_i)}{P(E_1E_2)} = \frac{P(E_2|H_i)P(E_1H_i)}{P(E_1E_2)} \\ &= \frac{P(E_2|H_i)P(H_i|E_1)P(E_1)}{P(E_1E_2)} = \frac{P(E_2|H_i)P(H_i|E_1)}{Q(1,2)} \end{aligned}$$

其中 $Q(1,2)=P(E_1E_2)/P(E_1)$．因为上式对所有 i 都成立，所以我们将上式对 i 求和得到

$$1 = \sum_{i=1}^{n} P(H_i|E_1E_2) = \sum_{i=1}^{n} \frac{P(E_2|H_i)P(H_i|E_1)}{Q(1,2)}$$

它说明

96

$$Q(1,2) = \sum_{i=1}^{n} P(E_2|H_i)P(H_i|E_1)$$

这样可得结果

$$P(H_i|E_1E_2) = \frac{P(E_2|H_i)P(H_i|E_1)}{\sum_{i=1}^{n} P(E_2|H_i)P(H_i|E_1)}$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

例如，假设有两枚硬币，选择一枚抛掷. 令 H_i 表示选中了第 $i(i=1, 2)$枚硬币且假设选中第 i 枚硬币后抛掷，正面朝上的概率为 p_i，$i=1, 2$. 令 E_j 表示"对于选中的硬币的第 j 次的抛掷结果". 抛掷以后，即 E_1 发生以后，只需将 $P(H_i)$进行修正，得到 $P(H_i|E_1)$. 若以后还有新的试验结果 E_2，此时只需将 $P(H_i|E_1)$进行修正，得到 $P(H_i|E_1E_2)$，将 $P(H_i|E_1E_2)$重新写成 $P(H_i)$，即忘掉它的历史. 每次得到新的试验结果，只需将修正了的 $P(H_i)$再次进行修正即可，而不必考虑修正的历史. ■

小结

对于任意事件 E 和 F，已知 F 发生的条件下，E 发生的条件概率记为 $P(E|F)$，定义为

$$P(E|F)=\frac{P(EF)}{P(F)}$$

恒等式

$$P(E_1E_2\cdots E_n)=P(E_1)P(E_2|E_1)\cdots P(E_n|E_1\cdots E_{n-1})$$

称为概率的**乘法规则**.

一个有用的恒等式是

$$P(E)=P(E|F)P(F)+P(E|F^c)P(F^c)$$

它可用于通过以 F 是否发生为条件计算 $P(E)$.

$P(H)/P(H^c)$称为事件 H 的**优势比**. 恒等式

$$\frac{P(H|E)}{P(H^c|E)}=\frac{P(H)P(E|H)}{P(H^c)P(E|H^c)}$$

说明当得到一个新的证据 E 后，H 的优势比等于原来的优势比值乘以当 H 成立时新证据发生的概率与 H 不成立时新证据发生的概率的比值.

令 $F_i(i=1, \cdots, n)$为互不相容的事件列，且它们的并为整个样本空间，恒等式

$$P(F_j|E)=\frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n}P(E|F_i)P(F_i)}$$

称为**贝叶斯公式**. 如果事件 $F_i(i=1, \cdots, n)$为一组假设，那么贝叶斯公式说明了如何计算当新证据成立时，这些假设成立的条件概率.

贝叶斯公式的分母用了下面的公式：

$$P(E)=\sum_{i=1}^{n}P(E|F_i)P(F_i)$$

我们称之为**全概率公式**.

如果 $P(EF)=P(E)P(F)$，那么我们称事件 E 和 F 是**独立的**. 该等式等价于 $P(E|F)=P(E)$或 $P(F|E)=P(F)$. 因此，如果知道 E 和 F 其中之一的发生并不影响另一个发生的概率，那么 E 和 F 独立.

事件 $E_1, \cdots, E_n$ 称为独立的，如果对任何子集 $E_{i_1}, \cdots, E_{i_r}$，有

$$P(E_{i_1}\cdots E_{i_r})=P(E_{i_1})\cdots P(E_{i_r})$$

对于任一给定事件 F，$P(E|F)$可以认为是样本空间中事件 E 的概率函数.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

习题

3.1 掷两枚均匀的骰子，求已知两枚骰子点数不同的条件下，至少有一枚点数为 6 的条件概率.

3.2 掷两枚均匀骰子，求给定两枚骰子点数之和为 i 的条件下，第一枚点数为 6 的条件概率. 计算 i 取值从 2 到 12 的所有情况.

3.3 利用公式(2.1)，计算在一手桥牌里，南北两家持有总共 8 张黑桃的条件下，东家持有 3 张黑桃的条件概率.

3.4 掷两枚均匀的骰子，在点数之和为 $i(i=2, 3, \cdots, 12)$ 的条件下，至少有一枚点数为 6 的条件概率是多少？

97

3.5 坛子里有 6 个白球和 9 个黑球. 如果无放回随机抽取 4 个球，那么前两个是白球且后两个是黑球的概率是多少？

3.6 坛子里有 12 个球，其中 8 个白球. 从中有放回(无放回)地抽取 4 个球，那么已知抽取的球中正好有 3 个白球的条件下，第一个球和第三个球是白球的条件概率(有放回和无放回情形下分别计算)是多少？

3.7 如果国王来自有两个孩子的家庭，那么另一个孩子是他姐妹的概率是多大？

3.8 某夫妇有两个孩子，已知老大是女孩的条件下两个孩子都是女孩的条件概率是多大？

3.9 假设有 3 个坛子，坛子 A 有 2 个白球和 4 个红球，坛子 B 有 8 个白球和 4 个红球，坛子 C 有 1 个白球和 3 个红球. 如果从每个坛子各取一个球，那么在正好取了两个白球的条件下，从坛子 A 里取的是白球的条件概率是多大？

3.10 随机无放回地从一副 52 张牌里抽取 3 张，已知第二张和第三张都是黑桃，求第一张是黑桃的条件概率.

3.11 随机无放回地从一副 52 张牌里抽取 2 张，令 B 表示“两张都是 A”，令 A_s 表示“抽中了黑桃 A”，令 A 表示“至少抽中了一张 A”，计算(a) $P(B|A_s)$；(b) $P(B|A)$.

3.12 某大学毕业生计划今夏参加前三场精算师考试. 他将在 6 月份参加第一场考试. 若通过了，则在 7 月份参加第二场. 而若又通过了，则参加 8 月份的第三场. 如果在某场考试失败了，则不允许参加剩下的考试. 他通过首场考试的概率为 0.9；如果通过了首场考试，则他通过第二场考试的条件概率为 0.8；如果通过了前两场，那么他通过第三场的条件概率是 0.7.

(a) 他通过全部三场考试的概率是多大？

(b) 已知他没有通过全部三场考试，那么她在第二场考试失败的条件概率是多大？

3.13 考虑一副 52 张牌(有 4 张 A)随机平均分给 4 家，每家 13 张，我们对每家都有一张 A 的概率 p 感兴趣，令 E_i 表示“第 i 家恰好有 1 张 A”，利用乘法规则计算 $p=P(E_1E_2E_3E_4)$.

3.14 坛子里最初有 5 个白球和 7 个黑球. 每次取出一个球，记下它的颜色后放回坛子，同时再放入相同颜色的 2 个球. 计算如下概率：

(a) 前两个球是黑色的，后两个球是白色的；(b) 前 4 个球中恰好有 2 个是黑色的.

3.15 吸烟的孕妇宫外孕的概率是不吸烟孕妇的两倍. 如果有 32%的孕妇吸烟，那么有百分之几的宫外孕孕妇吸烟？

3.16 98%的婴儿分娩是安全的. 然而有 15%的分娩是剖腹产. 当采用剖腹产时，婴儿的生存概率为 96%. 如果随机选择一个采用非剖腹产的孕妇，那么其婴儿的生存概率是多少？

3.17 某个社区，36%的家庭有一条狗，22%的家庭既有一条狗，又有一只猫，另外，30%的家庭有一只猫.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

(a) 随机选择一个家庭，为既有猫又有狗的概率是多少？

(b) 随机选择一个家庭，已知该家庭有猫的条件下，还有一条狗的条件概率是多少？

3.18 某城市中，46%的人是无党派人士，30%的人属于自由党，24%的人属于保守党. 在最近一次地方选举中，35%的无党派人士、62%的自由党员和58%的保守党员参与了选举. 随机选择一位选民，假定他参与了地方选举，求以下概率：

(a) 他是无党派人士；(b) 他是自由党党员；(c) 他是保守党党员；(d) 参与地方选举的选民比例.

3.19 参加过某个“戒烟班”的人，有48%的女性和37%的男性在结束后一年内坚持没有吸烟，这些人参加了年末的庆功会. 如果一开始，班里有62%的男性，问：

(a) 参加庆功会的女性所占百分比是多少？

(b) 参加庆功会的人数占全班的百分比是多少？

3.20 某大学里，52%的学生为女生，5%的学生为计算机科学专业，2%的学生为计算机科学专业的女生. 如果随机挑选一名学生，求以下条件概率： 98

(a) 在已知该学生主修计算机科学的条件下，该生为女生的条件概率；

(b) 已知该生为女生的条件下，该生主修计算机科学的条件概率.

3.21 总共有500对职业夫妇参与了关于年薪的调查，调查结果见表3-1.

表 3-1

妻子	丈夫	
	低于25 000美元	高于25 000美元
低于25 000美元	212	198
高于25 000美元	36	54

例如，其中有36对夫妇，妻子年薪超过25 000美元，而丈夫年薪低于25 000美元. 如果随机挑选一对夫妇，求以下概率：

(a) 丈夫年薪低于25 000美元；

(b) 丈夫年薪高于25 000美元的条件下，妻子年薪超过25 000美元的条件概率；

(c) 丈夫年薪低于25 000美元的条件下，妻子年薪超过25 000美元的条件概率.

3.22 分别掷一枚红色、蓝色和黄色骰子(都是6面)，我们感兴趣的是蓝骰子点数小于黄骰子点数，而黄骰子点数小于红骰子点数的概率. 即令 B，Y，R 分别表示蓝骰子、黄骰子和红骰子的点数，我们感兴趣的是 $P(B<Y<R)$.

(a) 没有两个骰子点数一样的概率是多大？

(b) 已知没有两个骰子点数一样的情况下，$B<Y<R$ 的概率是多大？

(c) 求 $P(B<Y<R)$.

3.23 坛子Ⅰ有2个白球和4个红球，而坛子Ⅱ有1个白球和1个红球，随机从坛子Ⅰ中取一个球放入坛子Ⅱ，然后随机从坛子Ⅱ中取一个球，

(a) 从坛子Ⅱ中取出的球是白球的概率是多大？

(b) 已知从坛子Ⅱ中取出的是白球，问从第一个坛子中取出的放入第Ⅱ个坛子的球是白球的条件概率是多大？

3.24 在一个坛子里放入两个球，假设在放入之前，每个球分别以概率为1/2涂成黑色，以概率为1/2涂成金色. 假设两个球的涂色是相互独立的.

(a) 已知金色的颜料已经用过(即至少有一个球涂成了金色)，计算两个球都涂成金色的概率.

(b) 假设坛子倒了，一个球掉了出来，是金色，那么其中两个球都是金色的概率是多大？并解释.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

3.25 以下方法用于估计 100 000 人的城镇里的 50 岁以上的人口的数量："当你在街上散步时，数一数你碰到的 50 岁以上的人数，再算出它们占你遇到的人的百分数，这样做几天后，用 100 000 去乘得到的百分数就是所求的估值."对这个方法作出你的评价.

提示：设这个城市中 50 岁以上的人所占比例为 p，另外，令 α_1 表示一个 50 岁以下的人在街上的时间所占的比例，α_2 表示一个 50 岁以上的人的相应比值，这个方法估计的是什么量？什么时候这个估计值近似等于 p？

3.26 假设有 5%的男性和 0.25%的女性为色盲，并假定男性和女性的数量相等. 随机选择一个色盲的人，他是男性的概率是多大？如果男性的数量是女性的两倍呢？

3.27 一公司所有员工都开车上班. 公司希望估计出每个车内员工的平均数. 下面提供的方法中，哪一个是正确的？并给出解释.

(a) 随机地找 n 个员工，问他们所乘的车内有多少员工，求出其平均值；

(b) 随机地选 n 辆车，数一数车内的员工人数，然后求平均值.

3.28 假设一副 52 张牌洗好后扣在桌子上，每次翻开一张，直到出现第一张 A. 已知第一张 A 出现在第 20 张翻牌，问接下来的牌是以下牌的条件概率是多大，

(a) 黑桃 A？(b) 梅花 2？

3.29 盒子里有 15 个网球，其中 9 个球还没用过. 随机地抽取 3 个，用它们练球，之后放回盒子. 随后，又随机地从中再抽取 3 个，求其中没有一个球被用过的概率是多大？

99 **3.30** 两个盒子，一个里面有黑白弹子各一个，另一个有 2 个黑弹子和 1 个白弹子. 随机挑选一个盒子，再随机从中取出一个弹子，问取出的是黑弹子的概率是多大？已知弹子是白色的条件下，选中的盒子是第一个盒子的条件概率是多大？

3.31 阿奎娜夫人接受了一次癌症活体组织检查. 但她不想让检查结果影响她周末的情绪. 若她告诉医生，只有好消息时才打电话通知结果. 这样，当医生不打电话时，她仍然可下结论：她的结果是不好的. 学过概率论的阿奎娜夫人要求医生，首先掷一枚硬币，若硬币为正面朝上，那么当检验结果是好消息时，医生就及时通知她，当检验结果是坏消息时，就不通知她. 若硬币为反面朝上时，医生就不必打电话. 这样，即使医生不打电话，也不意味着"一定是坏消息". 令 α 表示检查结果为癌症的概率，令 β 表示医生不打电话的条件下，检验结果为癌症的条件概率.

(a) α 和 β 哪个大？(b) 求出 β 和 α 之间的关系，验证(a)中的结论.

3.32 某家庭有 j 个孩子的概率为 p_j，其中 $p_1=0.1$，$p_2=0.25$，$p_3=0.35$，$p_4=0.3$. 随机从该家庭挑选一个孩子，已知这孩子为该家庭里最大的孩子，求该家庭有以下数量孩子的条件概率：

(a) 仅有一个孩子. (b) 有 4 个孩子.

如果随机挑选的一个孩子是该家庭里最小的孩子，重做(a)和(b).

3.33 在雨天的时候，乔伊有 0.3 的概率迟到；平时，他有 0.1 的概率迟到. 而明天下雨的概率是 0.7.

(a) 求出乔伊明天不迟到的概率.

(b) 在乔伊没有迟到的条件下，下雨的条件概率是多少？

3.34 在例 3f 中，假设新的证据服从不同可能的解释. 事实上，证据有 90%的可能造假. 请问，在这个例子中，嫌疑犯有多大的概率是有罪的？

3.35 有 0.6 的概率是妈妈把礼物藏起来了，有 0.4 的概率是爸爸把礼物藏起来了. 当妈妈藏时，她有 70%的可能藏楼上，30%的可能藏楼下. 爸爸藏楼上和楼下的概率相等.

(a) 礼物在楼上的概率是多少？

(b) 假设礼物在楼上，请问有多大的可能是爸爸藏起来了？

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

3.36 商店 A，B，C 各有 50 名、75 名和 100 名员工，其中分别有 50%，60%和 70%是女性. 我们假定每个员工的辞职是等可能的，而且没有性别差异. 若有个员工辞职了，而且是女性，则她在 C 店工作的概率是多少？

3.37 (a) 某赌徒口袋里有一枚均匀硬币，还有一枚两面都为正面的硬币. 他随机从中取出一枚并掷之，发现是正面朝上，问掷的是均匀硬币的概率是多大？

(b) 假设他将那枚硬币再掷一次，发现还是正面朝上. 那么它是均匀硬币的概率是多少？

(c) 假设他第三次掷那枚硬币，发现是反面朝上. 那么它是均匀硬币的概率是多少？

3.38 坛子 A 中有 5 个白球和 7 个黑球. 坛子 B 中有 3 个白球和 12 个黑球. 我们投掷一枚均匀的硬币，如果结果为正面朝上，则从 A 中取一个球，而如果结果为反面朝上，则从 B 中取一个球. 假设取出的是白球，问硬币是反面朝上的概率是多大？

3.39 在例 3a 中，已知投保人第一年内没发生事故，问第二年发生事故的条件概率是多大？

3.40 坛子里有 5 个白球和 7 个红球. 考虑按如下方式取出 3 个球：每一次取出一个，并记下其颜色，然后把它放回坛子，同时再放进一个相同颜色的球. 求取出的 3 个球满足下列条件的概率：

(a) 没有白球；(b) 只有 1 个白球；(c) 是 3 个白球；(d) 刚好有 2 个白球.

3.41 一副洗好的牌分成两份，各 26 张. 从其中一份里取出 1 张，发现是 A. 然后将它放入另一份里，洗好后，从中取出一张. 计算这张牌是 A 的概率.

提示：以取出的是否是放进去的那一张为条件.

3.42 全美有 12%的家庭在加州. 全美总共有 1.3%的家庭总收入超过 250 000 美元每年，而全加州有 3.3%的家庭年总收入超过 250 000 美元.

(a) 求非加州地区家庭中年收入超过 250 000 美元的家庭比例.

(b) 假设随机地抽取年收入超过 250 000 美元的一个美国家庭，请问该家庭是加州地区的概率是多少？ 100

3.43 盒子里有 3 枚硬币，第一枚两面都为正面，第二枚为均匀的硬币，第三枚是不均匀的，它出现正面朝上的概率为 75%. 如果从中随机挑一枚硬币来掷发现是正面朝上，那么它是两面都为正面的概率是多少？

3.44 监狱看守通知三个囚犯，在他们中要随机选择一个处决，而把另两个释放. 囚犯 A 请求看守秘密地告诉他，另外两个囚犯中谁将获得自由，A 声称："因为我已经知道他们两人中至少有一个获得自由，所以你泄漏这点消息是无妨的."但是看守拒绝回答这个问题，他对 A 说："如果你知道了你的同伙中谁将获释，那么，你自己被处决的概率将由 1/3 增加到 1/2，因为你就成了剩下的两个囚犯中的一个了."对于看守的上述理由，你怎么评价？

3.45 假定有 10 枚硬币，掷第 i 枚硬币正面朝上的概率为 $i/10$，$i=1$，2，…，10. 如果先随机选择一枚硬币来掷，结果为正面，那么它是第 5 枚硬币的概率是多少？

3.46 在一个给定年份，投保的男司机索赔的概率为 p_m，而投保的女司机索赔的概率为 p_f，其中 $p_f\neq p_m$. 男司机的比例为 α，$0<\alpha<1$. 随机挑选一名司机，令 A_i 表示"该司机第 i 年索赔"这一事件，证明

$$P(A_2\mid A_1)>P(A_1)$$

给出上述不等式的直观解释.

3.47 坛子里有 5 个白球和 10 个黑球. 若掷一枚均匀的骰子，掷出几点就从坛子中取几个球，则取出的球都是白球的概率是多少？在取出的球都是白球的条件下，掷出的骰子点数为 3 的条件概率是多少？

3.48 两个外形一样的橱柜都有 2 个抽屉. A 橱柜每个抽屉里有一枚银币，B 橱柜有一个抽屉里有一枚银币，另一抽屉有一枚金币. 随机挑选一个橱柜，打开其中一个抽屉，发现是一枚银币，求另一个

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

抽屉里也是银币的概率.

3.49 前列腺癌是男性中比较常见的一种癌，作为男性是否患有前列腺癌的指标，医生经常进行一项检查，测量仅由前列腺分泌的PSA(prostate specific antigen)蛋白质水平. 尽管PSA水平能表明癌症，但这种检查是出了名地靠不住. 事实上，一个未患前列腺癌的男性其PSA水平偏高的概率为0.135，而他确实有癌症的情况下，此概率增至0.268. 如果一个医生用其他方法诊断该男士有70%的可能患有前列腺癌，给定下列条件下，他患有前列腺癌的条件概率是多大？

(a) 检查指标为高水平. (b) 检查指标不为高水平.

若假设医生最初有30%的把握认为他患有前列腺癌，重做以上问题.

3.50 假设某保险公司把被保险人分成如下三类："低风险的"、"一般的"、"高风险的". 该保险公司的资料表明，对于上述三种人而言，在一年期内卷入一次事故的概率依次为0.05，0.15和0.30. 如果"低风险的"被保险人占人口的20%，"一般的"占50%，"高风险的"占30%. 试问在固定的一年中出事故的人口占多大比例？如果某被保险人在2012年没出事故，他是"低风险的"概率是多大？是"一般的"概率是多大？

3.51 某员工为找新工作，请其领导写封推荐信. 她估计如果得到了强有力的推荐，那么有80%的可能找到工作；如果得到了一般的推荐，那么有40%的可能找到工作；如果得到比较弱的推荐，那么只有10%的可能找到工作. 而且她估计她得到强有力的推荐、一般的推荐和较弱的推荐的概率分别为0.7，0.2，0.1.

(a) 她有多大的把握认为她会找到新工作？

(b) 已知她找到了新工作，她得到了强有力的推荐、一般的推荐以及较弱的推荐的条件概率各是多少？

(c) 已知她没找到新工作，她得到了强有力的推荐、一般的推荐以及较弱的推荐的条件概率各是多少？

3.52 一个高中学生非常焦急地等待大学录取通知书. 她估计在她被录取(未被录取)的条件下，在下周各天内收到信件的概率见表3-2.

表 3-2

日期	P(收到信件 | 被录取)	P(收到信件 | 未被录取)
周一	0.15	0.05
周二	0.20	0.10
周三	0.25	0.10
周四	0.15	0.15
周五	0.10	0. 20

101 同时，她估计自己被录取的概率为0.6.

(a) 星期一收到信件的概率是多大？

(b) 星期一没收到信件的条件下，星期二收到信件的概率是多大？

(c) 若星期三以前未收到信件，她被录取的概率有多大？

(d) 如果她在星期四收到信件，她被录取的条件概率有多大？

(e) 本周她没收到信件，问她被录取的条件概率有多大？

3.53 一个并联系统只要其中有一个元件正常工作就运行正常. 考虑一个有n个元件的并联系统，假设每个元件独立地正常工作的概率为1/2，计算已知系统工作运行的条件下，元件1正常工作的条件概率.

3.54 在下列(a)到(e)描述的情形中，如果你必须建立一个关于事件E和事件F的数学模型，哪种情形你可认为它们是独立的？试解释原因.

(a) E表示某个女商人是蓝眼睛，而F表示她的秘书也是蓝眼睛.

(b) E表示某教授有辆汽车，F表示他的名字出现在电话簿里.

(c) E表示某男人身高低于6英尺，而F表示其体重超过200磅.

(d) E表示某妇女生活在美国，而F表示她生活在西半球.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

(e) E 表示明天会下雨，而 F 表示后天还会下雨.

3.55 某班已知有 4 个一年级男生，6 个一年级女生，还有 6 个二年级男生. 如果从中随机选择一个学生，且其性别和年级是独立的，那么此班有多少个二年级女生？

3.56 假设你一直在收集优惠券，并假设一共有 m 种优惠券. 若每次收集优惠券时，得到第 i 种优惠券的概率为 p_i，$i=1$，…，m. 假设你正收集第 n 张优惠券，那么它是一张新类型(以前未曾收集到)的概率有多大？

提示：以这张优惠券的种类为条件.

3.57 一个关于股价变化的简化模型为：每天股价上涨一个单位的概率为 p，下跌一个单位的概率为 $1-p$，不同天里的变化认为是独立的.

(a) 2 天后股价仍维持在初始水平的概率是多大？

(b) 3 天后股价上涨一个单位的概率是多大？

(c) 已知三天后股价上涨了一个单位，问第一天股价上涨的条件概率是多大？

3.58 我们希望模拟掷一枚均匀硬币的试验，但是我们只有一枚不一定均匀的硬币，其正面朝上的概率为未知的 p(p 不一定为 1/2). 考虑下面步骤：

(1)掷一次硬币；(2)再掷一次硬币；

(3)如果上面两次结果一样，回到第 1 步；(4)最后一次掷出的结果，作为试验结果.

(a) 证明这样得到的结果，正面朝上或朝下的概率是一样的.

(b) 如果我们简化了步骤，将硬币掷到出现两次不同时为止，将最后一次掷硬币的结果作为试验结果. 这样做可以吗？

3.59 独立抛掷一枚硬币，每次正面朝上的概率为 p，前四次结果如下的概率是多大，

(a) H，H，H，H？(b) T，H，H，H？

(c) 在连续掷硬币过程中，T，H，H，H 出现在 H，H，H，H 之前的概率是多大？

提示：对于(c)，在什么条件下 H，H，H，H 先发生？

3.60 人眼的颜色由一对基因决定. 如果都为蓝色基因，那么他眼睛为蓝色，如果都为棕色基因，那么眼睛为棕色. 如果一个是蓝色基因，一个是棕色基因，那么他眼睛为棕色.（我们称棕色基因比蓝色基因占优势.）一个新生儿独立地从其父母处各得到一个遗传基因，父亲的基因对中的任一基因以相等的概率遗传给他的孩子. 母亲的情况也是一样的. 假设史密斯及其父母眼睛都为棕色，但其姐姐眼睛为蓝色.

(a) 史密斯拥有蓝色基因的概率是多大？

(b) 假设史密斯夫人眼睛为蓝色. 他们第一个孩子的眼睛为蓝色的概率是多大？

(c) 如果他们第一个孩子的眼睛为棕色，它们第二个孩子眼睛为棕色的概率是多大？ 102

3.61 有关白化病的基因记为 A 和 a，只有从其父母都遗传了基因 a 的人才会得白化病. 有基因对 A，a 的人在外表上是正常的，但是因为他能将其基因遗传给下一代，因此称为携带者. 现假设一对正常的夫妇有两个小孩，其中有一个为白化病. 假设另一个未得白化病的孩子将来与一个白化病携带者结婚.

(a) 他们的第一个孩子得白化病的概率是多大？

(b) 已知他们的第一个孩子未得白化病的条件下，第二个孩子得白化病的条件概率是多大？

3.62 芭芭拉和黛安娜出去射击. 假设芭芭拉每次射击击中木鸭子的概率为 p_1，而黛安娜每次射击击中木鸭子的概率为 p_2. 假设她们同时射击同一只木鸭子. 如果木鸭子翻了(表示射中了)，求以下事件的概率.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

(a) 两人都射中了. (b) 芭芭拉射中了.

其中作了怎样的独立性假设?

3.63 A 和 B 被卷入一场决斗. 决斗的规则是：拿起自己的枪，并同时向对方射击. 如果有一人或两人都被射中，那么决斗结束. 如果两人都射空，那么重复过程. 假设每次射击结果都是独立的，且 A 射中 B 的概率为 p_A，B 射中 A 的概率 p_B，计算

(a) A 没被击中的概率；(b) A，B 都被击中的概率；(c) n 局决斗后决斗停止的概率；

(d) A 没被击中的条件下，第 n 局决斗后停止的条件概率；

(e) 两人都被击中的条件下，第 n 局决斗后停止的条件概率.

3.64 在一个答题秀上，一对夫妇遇到了一道题，丈夫和妻子独立给出正确答案的概率都为 p. 对于这对夫妇，以下哪个策略更好?

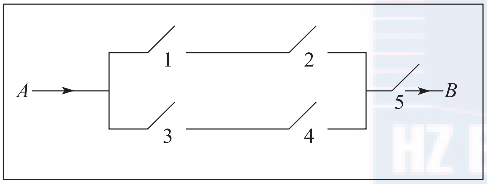
(a) 任选一个人并让其答题.

(b) 他们俩都给出问题的答案，如果答案一致，那么就采用这答案，如果答案不一致，那么掷硬币决定采取谁的答案.

3.65 假设在例 3h 中，64%的双胞胎是同性别的，已知一对新生的双胞胎是同性别的，请问这对双胞胎是同卵双生的条件概率是多少?

3.66 在图 3-4 所示的电路里，第 i 个继电器闭合的概率为 p_i，$i=1, 2, 3, 4, 5$. 如果所有继电器的功能相互独立，试对如下(a)和(b)两种情况，分别求出 A 和 B 之间是通路的概率.

提示：对于(b)，以继电器 3 是否闭合为条件.



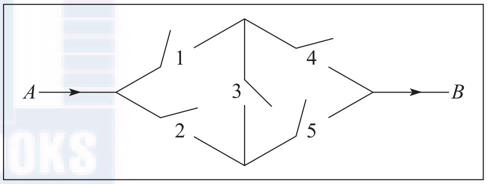


图 3-4　习题 3.66 中的电路示意图

3.67 一个由 n 个部件组成的系统称为"k-n"($k\leqslant n$)系统，如果此系统运行当且仅当它的 n 个部件中至少有 k 个部件运行. 假设它的各部件运行是相互独立的.

(a) 如果第 i 个部件运行的概率为 P_i，$i=1, 2, 3, 4$，求一个"2-4"系统运行的概率.

103 (b) 试对"3-5"系统求出(a)中的概率；

(c) 假设一个"k-n"系统所有的 P_i 都等于 p，即 $P_i=p$，$i=1, 2, \cdots, n$，对此系统试求出上述概率.

3.68 在习题 3.66a 中，已知 A 和 B 之间是通路的条件下，求继电器 1 和 2 均闭合的概率.

3.69 某种生物具有一对基因组，其中每个基因组由 5 个不同的基因组成(我们把这 5 个基因用 5 个英文字母表示). 每个基因有两种特性，分别用大小字母来区别，大写字母表示显性，小写字母表示隐性. 若某生物具有基因对(xX)，此时，该生物表现基因 X 的特征. 例如，X 代表棕色眼睛，x 代表蓝色眼睛，那么具有(X，X)或(x，X)的个体都是棕色眼睛，而具有(x，x)时为蓝色眼睛. 生物的表现特征(phenotype)与基因特征(genotype)是有区别的. 因此，一个生物具有基因特征 aA，bB，cc，dD，ee，另一个生物具有基因特征 AA，BB，cc，DD，ee，它们的基因特征不同，但表现特征是相同的. 对于每一种基因的基因对，交配双方都随机地贡献其中一个基因(5 种基因中每种各一个). 现在设交配双方的 5 种基因组分别为 aA，bB，cC，dD，eE 和 aa，bB，cc，Dd，ee. 分

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

别计算子女的(i)基因特征和(ii)表现特征与下列相同的概率：

(a) 第一个亲代；(b) 第二个亲代；(c) 与某一个亲代；(d) 都不相同.

3.70 女王有50%的可能携带有血友病的基因. 如果她是一个携带者，那么每个王子都有50%的可能患有血友病. 如果女王有3个王子，且都没有患血友病. 那么女王是携带者的概率有多大？如果有第4个王子，那么他患有血友病的概率有多大？

3.71 在1982年9月30日上午，美国全国棒球协会西部赛区积分榜排名前三名见表3-3. 每个队还剩三场比赛需要打，巨人队的所有三场比赛的对手都是躲闪者，而勇士队的所有三场比赛的对手都是圣迭戈教士队. 假设所有比赛结果是独立的，且每场比赛双方获胜的可能性是一样的. 那么各个队获得冠军的概率是多少？如果两个队并列第一，还需要进行附加赛，以决定名次，此时，假定各队获胜概率为1/2.

表 3-3

球队	赢	输
亚特兰大勇士	87	72
旧金山巨人	86	73
洛杉矶躲闪者	86	73

3.72 市政委员会由7人组成，其中包含一个由3人组成的核心委员会. 一个新的法律提案首先由核心会员会表决，只有核心委员会中有2人以上同意，才能拿到全体委员会上讨论. 一旦到了全体委员会，只要有4票以上通过，这个新的法律提案就生效. 现在有一新的法律提案. 设每个委员以概率 p 同意这个提案，并且相互独立地作出投票决定. 一个核心委员会成员起决定作用的概率是多少？所谓起决定作用是指若他的决定是相反的，其最后结果也是相反的. 一个不在核心委员会的委员起决定作用的概率是多少？

3.73 假设某对夫妇的每个孩子是男孩或女孩的可能性一样，且与这个家庭里的其他孩子的性别是独立的. 对于一个有5个孩子的家庭，计算如下事件的概率：

(a) 所有的小孩同一性别；(b) 最大的三个为男孩，其他为女孩；

(c) 正好三个男孩；(d) 最大的两个为女孩；(e) 至少有一个女孩.

3.74 A 和 B 轮流掷一对骰子，当 A 掷出"和为9"或 B 掷出"和为6"时停止. 假设 A 先掷，求最后一次掷是由 A 完成的概率.

3.75 某村子有一个传统，即家庭里的长子及其妻子负责照顾他们年迈的父母. 然而，近年来，这个村的女性不想承担这种责任，已经不愿意嫁给最大的男孩.

(a) 如果每个家庭都有两个孩子，问所有的儿子中长子占多大比例？

(b) 如果每个家庭都有三个孩子，问所有的儿子中长子占多大比例？

假设婴儿出生时是男孩或女孩的概率相同，并且家庭里各小孩的性别是独立的.

3.76 假设 E 和 F 为某次试验的互不相容的事件. 证明：如果独立重复进行这样的试验，那么 E 发生在 F 之前的概率为 $P(E)/[P(E)+P(F)]$.

3.77 考虑一个无穷的独立重复试验序列，每次试验都等可能地以1，2，3为结果. 假设3次试验后最后一次试验结果为3，求以下条件概率： 104

(a) 第一次试验结果为1；(b) 前两次试验结果都为1.

3.78 A和B进行一系列比赛，每局比赛A获胜的概率都为 p，B获胜的概率为 $1-p$，且每次比赛结果相互独立. 当其中一人比另一人多胜两局时，游戏停止，获胜局数多的选手赢得比赛.

(a) 求总共比赛了4局的概率；(b) 求A最后获得比赛胜利的概率.

3.79 连续地掷一对均匀的骰子，在出现6次"和为偶数"之前，出现2次"和为7"的概率是多少？

3.80 选手们水平相当，在每次比赛中任一方获胜的可能性都是1/2. 共有 2^n 个选手随机地一一配对比赛，随后的 2^{n-1} 个胜者再随机地一一配对比赛，等等，直到最后一个获胜者出现. 对指定的A，B

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

两人，定义事件 $A_i(i\leqslant n)$ 和 E 如下：

A_i：A 参与了 i 场比赛　　E：A，B 曾一起比赛

(a) 求 $P(A_i)$，$i=1$，…，n；(b) 求 $P(E)$.

(c) 令 $P_n=P(E)$，证明

$$P_n = \frac{1}{2^n-1}+\frac{2^n-2}{2^n-1}\left(\frac{1}{2}\right)^2 P_{n-1}$$

并利用此公式验证(b)中得到的答案.

提示：以事件 $A_i(i=1, \cdots, n)$ 发生为条件，计算 $P(E)$，再利用以下代数恒等式

$$\sum_{i=1}^{n-1} i x^{i-1} = \frac{1-nx^{n-1}+(n-1)x^n}{(1-x)^2}$$

来化简答案. 另外一个解决此问题的方法是，注意到一共进行了 2^n-1 场比赛.

(d) 解释为什么总共有 2^n-1 场比赛.

提示：给这些比赛进行编号，且令 B_i 表示 A 和 B 在第 i 场比赛中碰面，$i=1$，…，2^n-1.

(e) $P(B_i)$是多少？(f) 利用(e)计算 $P(E)$.

3.81 某股票投资者有一只股票，其现值为 25. 他决定当股价跌至 10 或涨至 40 时卖出股票. 如果股价的每次变化都是以 0.55 的概率上涨 1 点，以 0.45 的概率下跌 1 点，且连续的变动是相互独立的，那么投资者最后获利的概率是多少？

3.82 A 和 B 掷硬币，A 先开始并连续掷硬币，直到出现反面朝上. 此时，B 开始掷硬币直到出现反面朝上，然后又轮到 A 掷，等等. 令 P_1 和 P_2 分别表示 A 和 B 掷硬币时正面朝上的概率. 若谁先达到如下获胜条件谁就获胜，求各种情况下 A 获胜的概率.

(a) 在一轮中有两个正面朝上；(b) 双方总共 2 个正面朝上；

(c) 一轮中 3 个正面朝上；(d) 双方总共 3 个正面朝上.

3.83 骰子 A 有 4 面红 2 面白，而骰子 B 有 2 面红 4 面白. 先掷一枚均匀的硬币，如果正面朝上，那么用骰子 A 玩游戏，如果是反面朝上，那么用骰子 B.

(a) 证明每次掷出红色面的概率为 1/2；

(b) 若前两次投掷都出现红色面，那么第三次投掷也出现为红色面的概率是多大？

(c) 如果头两次投掷都出现红色面，那么使用的是骰子 A 的概率是多大？

3.84 坛子里有 12 个球，其中 4 个白球，三个选手 A，B，C 依次从坛中取球，A 最先，然后 B，然后 C，再是 A，依次进行下去. 第一个取出白球的人获胜. 求每个选手获胜的概率，如果

(a) 每个球取出后再放回；(b) 取出的球不放回.

3.85 当 3 个选手各自选择自己的坛子时，重做习题 3.84，即假设有 3 个不同的坛子，每个里面有 12 个球，其中 4 个白球.

3.86 令 $S=\{1, 2, \cdots, n\}$，假设 A 和 B 独立且等可能地为 2^n 个子集之一(包括空集和 S 本身).

(a)证明：

$$P\{A\subset B\} = \left(\frac{3}{4}\right)^n$$

提示：令 $N(B)$表示 B 中元素个数，利用

$$P\{A\subset B\} = \sum_{i=0}^{n} P\{A\subset B \mid N(B)=i\} P\{N(B)=i\}$$

(b) 证明 $P\{AB=\varnothing\}=\left(\frac{3}{4}\right)^n$.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

3.87 考虑例 2a，但是现在假定当钥匙在某一个口袋里时，有 10%的可能在搜那个口袋时没有找到钥匙．R 和 L 分别代表钥匙在右边和左边口袋的概率．同时，令 S_R 代表搜右边口袋会成功找到钥匙的概率，U_L 代表搜左边口袋不会成功找到钥匙的概率，试用如下(a)和(b)两种方法计算条件概率 $P(S_R \mid U_L)$： 105

(a) 利用等式

$$P(S_R \mid U_L) = \frac{P(S_R U_L)}{P(U_L)}$$

并以钥匙在或不在右口袋为条件计算 $P(S_R U_L)$，先用给定钥匙在或不在左口袋的条件概率计算 $P(U_L)$；

(b) 用下面的等式：

$$P(S_R \mid U_L) = P(S_R \mid RU_L)P(R \mid U_L) + P(S_R \mid R^c U_L)P(R^c \mid U_L)$$

3.88 在例 5e 里，已知前 n 次结果都是正面朝上的条件下，选中第 i 枚硬币的条件概率是多大？

3.89 在拉普拉斯继承准则(例 5e)中，各次投掷结果是否独立？试解释之．

3.90 由 3 名法官组成一个陪审团．某人被审讯后，若至少两名法官投“有罪”票，则判决此人有罪．假设对一名事实上有罪的被告，每个法官独立地投“有罪”票的概率为 0.7，而对事实上无罪的被告，这个概率下降到 0.2．如果 70%的被告事实上是有罪的，试对如下各条件求出 3 号法官投“有罪”票的条件概率．

(a) 1 号法官和 2 号法官都投“有罪”票；

(b) 在 1 号和 2 号法官所投的票中，一张“有罪”，另一张“无罪”；

(c) 1 号和 2 号法官全投“无罪”票．

以 $E_i(i=1, 2, 3)$表示 i 号法官投一张“有罪”票的事件，这些事件是否相互独立？是否条件独立？试说明理由．

3.91 假设进行 n 次独立重复试验，每次结果是 0，1 或 2 的概率分别是 p_0，p_1 和 p_2，$\sum_{i=0}^{2} p_i = 1$．求结果 1 和 2 都至少出现一次的概率是多大？

理论习题

3.1 若 $P(A)>0$，证明

$$P(AB \mid A) \geqslant P(AB \mid A \cup B)$$

3.2 令 $A \subset B$，尽可能简单地表达如下概率：

$$P(A \mid B) \qquad P(A \mid B^c) \qquad P(B \mid A) \qquad P(B \mid A^c)$$

3.3 考虑有 m 个家庭的社区，其中有 n_i 个家庭有 i 个孩子，$i=1, \cdots, k$，$\sum_{i=1}^{k} n_i = m$．考虑如下两种选择孩子方式：

(a) 随机选择一个家庭，然后再随机选择一个孩子；

(b) 从 $\sum_{i=1}^{k} i n_i$ 个孩子中随机挑选一个．

证明：第一种方法挑选出来的孩子是他家里第一个出生孩子的概率大于第二种方法挑选出来的．

提示：为了解此题，需要证明

$$\sum_{i=1}^{k} i n_i \sum_{j=1}^{k} \frac{n_j}{j} \geqslant \sum_{i=1}^{k} n_i \sum_{j=1}^{k} n_j$$

为了证明这点，将两边展开，证明对于任意一组(i, j)，式子左边的项 $n_i n_j$ 的系数比右边的大．

3.4 已知有一个球放在 n 个盒子中的一个内．已知球在第 i 个盒子的概率是 P_i．如果球在第 i 个盒子里，搜寻该盒子会以 α_i 的概率发现它．证明：已知搜索第 i 个盒子没有发现球的条件下，球在第 j 个盒子的条件概率是

HZ BOOKS 华章图书

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$\begin{cases} \dfrac{P_j}{1-\alpha_i P_i} & 若 j \neq i \\ \dfrac{(1-\alpha_i)P_i}{1-\alpha_i P_i} & 若 j = i \end{cases}$$

3.5 (a) 证明：若 E 和 F 互不相容，那么 $P(E\mid E\cup F)=\dfrac{P(E)}{P(E)+P(F)}$.

106 (b) 证明：如果 $E_i(i\geqslant 1)$互不相容，那么 $P\left(E_j \mid \bigcup\limits_{i=1}^{\infty} E_i\right) = \dfrac{P(E_j)}{\sum\limits_{i=1}^{\infty} P(E_i)}$.

3.6 证明：若 E_1，E_2，…，E_n 为相互独立事件序列，则

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) = 1 - \prod_{i=1}^{n}[1-P(E_i)]$$

3.7 (a) 坛子里有 n 个白球和 m 个黑球，每次随机从中取出一个，直到剩下的球为同一种颜色. 证明：剩下的球全为白球的概率为 $n/(n+m)$.

提示：设想试验一直进行直到所有球都被取走，并考虑最后取出的球的颜色.

(b) 池塘里有3种不同的鱼，我们分别称为红鱼、蓝鱼和绿鱼. 三种鱼分别有 r、b 和 g 条. 假设按一随机顺序将这些鱼全部移走(即每次选择对于剩下的鱼都是等可能的)，最先抓完红鱼的概率是多少？

提示：写 $P(\{R\})=P(\{RBG\})+P(\{RGB\})$，并以最后移走的是哪种颜色的鱼作为条件来计算右边的概率.

3.8 设一次掷两枚均匀的骰子，A，B 和 C 是与试验有关的三个事件.

(a) 若

$$P(A\mid C) > P(B\mid C) \text{ 且 } P(A\mid C^c) > P(B\mid C^c)$$

则证明 $P(A)>P(B)$，或者给出反例，定义事件 A，B，C 说明这个不等式不成立.

(b) 若

$$P(A\mid C) > P(A\mid C^c) \text{ 且 } P(B\mid C) > P(B\mid C^c)$$

则证明 $P(AB\mid C)>P(AB\mid C^c)$，或者给出反例，构造事件 A，B，C 说明这个不等式不成立.

提示：令 C 表示"掷一对骰子，点数之和为10"，令 A 表示"第一个骰子点数为6"，令 B 表示"第二个骰子点数为6".

3.9 考虑独立地掷两次均匀硬币. 令 A 表示第一次为正面朝上，B 表示第二次为正面朝上，C 表示两次朝向一样. 证明事件 A，B，C 为两两独立，即 A 和 B 独立，A 和 C 独立，B 和 C 独立，但 A，B，C 不独立.

3.10 在常规检查中2%的45岁女性患有乳腺癌. 90%患有乳腺癌的人会被检查出来，8%没患乳腺癌的女性也会检查出问题. 假定一个女性检查出有乳腺癌，请问她真正患乳腺癌的概率是多少？

3.11 在 n 次独立地掷硬币的试验中，每次正面向上的概率都是 p. 那么 n 等于多少时至少一次正面向上的概率是1/2?

3.12 若 $0\leqslant a_i\leqslant 1$，$i=1$，2，…，证明：

$$\sum_{i=1}^{\infty}\left[a_i \prod_{j=1}^{i-1}(1-a_j)\right] + \prod_{i=1}^{\infty}(1-a_i) = 1$$

提示：假设掷无限多个硬币，令 a_i 表示第 i 个硬币正面朝上的概率，考虑出现第一个正面朝上的时刻.

3.13 设有一枚硬币，在抛掷时正面朝上的概率为 p. 假设 A 开始连续掷硬币，直至出现了反面朝上，此时 B 接着连续掷硬币，直到掷出反面朝上为止，然后 A 再接着掷，如此进行下去. 令 $P_{n,m}$ 表示在 B 累计 m 次正面朝上之前 A 已累计 n 次正面朝上的概率. 证明：

$$P_{n,m} = pP_{n-1,m} + (1-p)(1-P_{m,n})$$

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

* **3.14** 假设你同一个无限富裕的人赌博，每一步你可能赢 1 个单位也有可能输 1 个单位，概率分别为 p 和 $1-p$. 证明：你最后输光的概率为

$$\begin{cases} 1 & \text{若 } p \leqslant 1/2 \\ (q/p)^i & \text{若 } p > 1/2 \end{cases}$$

其中 $q=1-p$，i 是你的最初资金.

3.15 独立重复地进行每次以概率 p 成功的试验，直到总共有 r 次成功为止. 证明：恰好需要进行 n 次试验的概率为

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

并利用此结果解决赌本分割问题(例 4j).

提示：为了在 n 次试验中得到 r 次成功，在前 $n-1$ 次试验中必须有多少次成功?

3.16 若某独立重复实验序列中，每次试验结果只有两种，成功或失败，则这种试验序列称为伯努利试验序列(Bernoulli trials). 现设一个 n 次独立重复的伯努利试验序列，每次试验成功的概率为 p(失败的概率为 $1-p$). 令 P_n 表示 n 试验中出现偶数次(0 认为是偶数)成功的概率，证明

$$P_n = p(1-P_{n-1}) + (1-p)P_{n-1}, \qquad n \geqslant 1$$ 107

并利用此公式(利用归纳法)证明

$$P_n = \frac{1+(1-2p)^n}{2}$$

3.17 考虑进行 n 次独立试验，第 i 次试验成功的概率为 $1/(2i+1)$. 令 P_n 表示总的成功次数为奇数的概率.

(a) 计算 P_n，$n=1, 2, 3, 4, 5$；

(b) 猜测 P_n 的一个一般公式；

(c) 导出用 P_{n-1} 表示的 P_n 的递推公式；

(d) 验证(b)中的猜测满足(c)中的递推公式. 因为递推公式只有一个解，这就证明了你的猜测是正确的.

3.18 令 Q_n 表示连续掷 n 次均匀硬币，没有出现 3 个连续的正面朝上的概率. 证明

$$Q_n = \frac{1}{2}Q_{n-1} + \frac{1}{4}Q_{n-2} + \frac{1}{8}Q_{n-3}$$

$$Q_0 = Q_1 = Q_2 = 1$$

并求 Q_8.

提示：以出现第一次反面朝上的时刻为条件.

3.19 考虑赌徒输光问题，不同的是 A 和 B 同意最多玩 n 局. 令 $P_{n,i}$ 表示 A 最后输光所有钱的概率，其中开始时 A 有 i 元而 B 有 $N-i$ 元. 推导用 $P_{n-1,i+1}$ 和 $P_{n-1,i-1}$ 表示的 $P_{n,i}$ 的公式，并计算 $P_{7,3}$，$N=5$.

3.20 假设有两个坛子，每个坛子中既有白球，又有黑球. 从两个坛子里取出白球的概率分别为 p 和 p'. 按照如下方式连续有放回地取球. 最初分别以概率 α 从第一个坛子里取球，以概率 $1-\alpha$ 从第二个坛子里取球. 接下来的取球按如下规则进行：一旦取出的是白球，则放回，然后再从同一个坛子里取球；如取出的是黑球，那么接下来从另一个坛子里取球. 令 α_n 表示第 n 次从第一个坛子里取球的概率，证明

$$\alpha_{n+1} = \alpha_n(p+p'-1) + 1 - p' \qquad n \geqslant 1$$

并用这个公式证明

$$\alpha_n = \frac{1-p'}{2-p-p'} + \left(\alpha - \frac{1-p'}{2-p-p'}\right)(p+p'-1)^{n-1}$$

令 P_n 表示第 n 次取出的球是白球的概率，求出 P_n. 并且计算 $\lim\limits_{n\to\infty}\alpha_n$ 和 $\lim\limits_{n\to\infty}P_n$.

3.21 **投票问题**. 在一次选举中，候选人 A 获得了 n 张选票，而 B 获得了 m 张选票，其中 $n>m$. 假定所有的 $(n+m)!/(n!\ m!)$ 种选票的顺序都是等可能的，令 $P_{n,m}$ 表示在统计选票时 A 一直领先的概率.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

(a) 计算 $P_{2,1}$，$P_{3,1}$，$P_{3,2}$，$P_{4,1}$，$P_{4,2}$，$P_{4,3}$.

(b) 求 $P_{n,1}$，$P_{n,2}$.

(c) 基于(a)和(b)的结果，猜测 $P_{n,m}$ 的值.

(d) 以谁获得了最后一张选票为条件，推导用 $P_{n-1,m}$ 和 $P_{n,m-1}$ 表示的 $P_{n,m}$ 的递推公式.

(e) 利用(d)，对 $n+m$ 用归纳法证明(c)中的猜测.

3.22 天气预报的一个简单模型，假设明天天气(湿润或者干燥)与今天相同的概率为 p. 如果1月1日天气为干燥，证明，n 天后的天气为干燥的概率 P_n，满足

$$P_n=(2p-1)P_{n-1}+(1-p)\qquad n\geqslant 1$$
$$P_0=1$$

并证明：

$$P_n=\frac{1}{2}+\frac{1}{2}(2p-1)^n\qquad n\geqslant 0$$

3.23 设一个包里有 a 个白球和 b 个黑球，根据以下方式从包里取球：

第1步. 随机取出一个球，然后扔掉；

第2步. 再随机取出一个球，若取出的球的颜色与前一次取出的不同，将球放回包内，转向第1步. 否则，将球扔掉，重复第2步.

换言之，随机取球并扔掉，直至颜色发生了变化，此时将取到的球放回包里重新开始. 令 $P_{a,b}$ 表示包里最后一个球为白球的概率，证明

$$P_{a,b}=\frac{1}{2}$$

提示：对 $k\equiv a+b$ 用归纳法.

* **3.24** n 个选手的单循环赛，$\binom{n}{2}$ 对选手只比赛一次，任何一场比赛的结果都分出输赢. 对于一个给定的

108 整数 $k(k<n)$，一个有趣的问题是：是否可能存在一种比赛结果，对于任意 k 个选手的集合，都有一位选手，他打败了该集合内的所有选手. 证明：若

$$\binom{n}{k}\left[1-\left(\frac{1}{2}\right)^k\right]^{n-k}<1$$

则这种结果是存在的.

提示：假设比赛结果是相互独立的，且每场比赛选手获胜的可能性都一样. 给 k 个选手的所有 $\binom{n}{k}$ 个组合标号，令 B_i 表示“在第 i 个组合里，没有人打败所有 k 个选手”，然后利用布尔不等式求出 $P\left(\bigcup_i B_i\right)$ 的界.

3.25 直接证明

$$P(E\mid F)=P(E\mid FG)P(G\mid F)+P(E\mid FG^c)P(G^c\mid F)$$

3.26 证明式(5.11)和式(5.12)的等价性.

3.27 将条件独立性的定义推广到2个以上的事件的情形.

3.28 证明或给出反例：如果 E_1 和 E_2 是独立的，那么给定 F 的条件下，它们也条件独立.

3.29 在拉普拉斯继承准则(例5e)中，证明：已知前 n 次掷的结果都是正面朝上的条件下，接下来 m 次掷的结果也是正面朝上的条件概率为 $(n+1)/(n+m+1)$.

3.30 在拉普拉斯继承准则(例5e)中，假设前 n 次投掷中，有 r 次正面朝上，$n-r$ 次反面朝上. 证明：第 $n+1$ 次投掷正面朝上的条件概率为 $(r+1)/(n+2)$. 为此，必须先证明恒等式

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

$$\int_0^1 y^n(1-y)^m \mathrm{d}y = \frac{n!m!}{(n+m+1)!}$$

提示：为了证明该恒等式，令 $C(n,m)=\int_0^1 y^n(1-y)^m \mathrm{d}y$．通过分部积分可得：

$$C(n,m)=\frac{m}{n+1}C(n+1,m-1)$$

从 $C(n,0)=1/(n+1)$开始，对 m 利用归纳法证明恒等式.

3.31 假定你的一个不懂数学但有哲学头脑的朋友声称拉普拉斯继承准则必定是错误的，理由是它可以导出荒谬的结论. 他说："例如，如果一个孩子 10 岁，按照这个准则，由于他已经活了 10 年，那么再活一年的概率为 11/12. 另一方面，如果孩子有位 80 岁的祖父，那么由拉普拉斯准则，祖父再活一年的概率为 81/82. 显然这是荒谬的，因为孩子比祖父更容易再活一年."你怎么回答这位朋友？

自检习题

3.1 在一局桥牌中，西家没有 A，以下事件的概率是多少：

(a) 他的搭档没有 A？

(b) 他的搭档有 2 个或更多的 A？

(c) 如果西家正好有一个 A，以上概率又各是多少？

3.2 一个新的汽车电池使用超过 10 000 英里的概率为 0.8，超过 20 000 英里的概率为 0.4，超过 30 000 英里的概率为 0.1. 如果一个新的电池在使用 10 000 英里以后仍能继续使用，求以下概率：(a) 其总寿命超过 20 000 英里；(b) 它还能继续使用 20 000 英里.

3.3 怎样将 10 个白球和 10 个黑球放入两个坛子内，使得随机选择一个坛子并随机从中取出一个球是白球的概率最大？

3.4 坛子 A 有 2 个白球 1 个黑球，而坛子 B 有 1 个白球 5 个黑球. 随机从坛子 A 中取一个球放入坛子 B，然后再随机从坛子 B 内取出一球. 已知取出的是白球，问从坛子 A 中取出的放入坛子 B 中的球是白球的概率是多大？

3.5 坛子里有 r 个红球和 w 个白球，每次随机取走一个. 令 R_i 表示"第 i 次取走的是红球"，计算

(a) $P(R_i)$；(b) $P(R_5 \mid R_3)$；(c) $P(R_3 \mid R_5)$.

3.6 坛子里有 b 个黑球和 r 个红球. 随机取出一个，再放回，同时还放入 c 个同样颜色的球. 现在，再取一个球，证明：已知第二个球是红球的条件下，第一个球是黑球的概率为 $b/(b+r+c)$.

3.7 你的一个朋友从一副 52 张牌里无放回地随机取出 2 张牌. 在下列情形下，计算两张都是 A 的条件概率： 109

(a) 你问是否其中一张是黑桃 A，你朋友的回答是肯定的；

(b) 你问是否第一张是 A，你朋友的回答是肯定的；

(c) 你问是否第二张是 A，你朋友的回答是肯定的；

(d) 你问是否其中有一张是 A，你朋友的回答是肯定的.

3.8 证明：

$$\frac{P(H \mid E)}{P(G \mid E)}=\frac{P(H)}{P(G)}\frac{P(E \mid H)}{P(E \mid G)}$$

假如在得到新的证据前，假设 H 成立的可能性是假设 G 成立的可能性的 3 倍. 如果 G 成立时新的证据出现的可能性是 H 成立时的两倍，那么当新证据出现时，哪个假设更有可能成立？

3.9 在你外出度假时，你请邻居给你的病树浇水. 如果没浇水的话，它死去的概率为 0.8. 如果浇水的话，它死去的概率为 0.15. 你有 90%的把握确定邻居记得浇水.

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

(a) 当你回来时，树还活着的概率是多少？

(b) 如果树死了，那么邻居忘记浇水的概率是多少？

3.10 坛子里有 8 个红球、10 个绿球和 12 个蓝球，现随意从中抽取 6 个.

(a) 至少有一个是红球的概率？

(b) 假设没有红球被抽到，刚好有 2 个绿球被抽到的条件概率是多少？

3.11 C 型号的电池会以 0.7 的概率正常工作．而 D 型号的电池会以 0.4 的概率正常工作．现从一个装有 8 个 C 型号和 6 个 D 型号的箱子中随机抽取一个电池.

(a) 请问电池正常工作的概率是多少？

(b) 假设电池没有正常工作，那么它是 C 型号的概率是多少？

3.12 玛利亚出差时会带两本书．假设她会带书 1 的概率是 0.6，带书 2 的概率是 0.5，她会带这两本书的概率是 0.4．求给定她不会带书 1 的条件下而她会带上书 2 的条件概率.

3.13 假设一个坛子里有 20 个红球和 10 个蓝球，现随机从坛子中一一取出.

(a) 所有红球在所有蓝球移出前先移出的概率是多少？

下面我们假定坛子里有 20 个红球，10 个蓝球，8 个绿球.

(b) 所有红球在所有蓝球移出前先移出的概率是多少？

(c) 球以蓝、红和绿的顺序完全移出的概率是多少？

(d) 蓝色球是这三组球中最先完全移出的概率是多少？

3.14 一枚硬币落地时正面向上的概率是 0.8．A 观察到了结果并急忙跑去告诉 B．然而，A 到 B 处时会有 0.4 的概率忘掉结果．如果 A 忘了，他告诉 B 正面或反面落地的概率相等．(如果他记得，那么他会正确地告诉 B.)

(a) A 告诉 B 正面向上的概率是多少？

(b) B 得知正确结果的概率是多少？

(c) 假设，B 得知硬币是正面向上，那么硬币真的是正面向上的概率是多少？

3.15 某种老鼠，黑色比棕色占优势(黑色对应显性基因，棕色对应隐性基因)．假设某只黑老鼠有个兄弟是棕色的，但其父母都是黑色的.

(a) 该老鼠是纯黑色的概率是多大(相对它的基因对是一个黑色，一个棕色的混合而言)？

(b) 假设当这只老鼠和一只棕色老鼠交配时，其所有 5 个后代都是黑色的，此时，这只老鼠为纯黑老鼠的概率是多少？

3.16 (a) 在习题 3.66b 中，以继电器 1 是否闭合为条件，计算 A 和 B 之间是通路的概率.

(b) 已知 A 和 B 之间是通路的条件下，继电器 3 是闭合的条件概率.

3.17 在习题 3.67 中描述的 k-n 系统，假定每个元件相互独立且正常工作的概率为 1/2，求已知系统正常运行的条件下，元件 1 工作正常的条件概率，当

(a) $k=1$，$n=2$；(b) $k=2$，$n=3$.

3.18 琼斯先生为了在赌场赢钱，设计了如下的赌博策略：当他押注时，只有当前面 10 次都出现黑色数字时才押注在红色数字上．他的理由是连续 11 次出现黑色的概率非常小．你认为他的策略如何？

3.19 A，B，C 三人同时各掷一枚硬币，掷出正面朝上的概率分别为 P_1，P_2，P_3，如果有一个人掷出的
110 结果与其他两人不一样，那么就称他为奇异人．如果没有出现奇异人，则继续掷硬币，直到出现奇异人，那么 A 被称为奇异人的概率是多大？

3.20 假设某次试验有 n 个可能的结果，结果 i 出现的概率为 p_i，$i=1$，…，n，$\sum_{i=1}^{n} p_i = 1$．如果观察两次试验，那么第二次的试验结果大于第一次结果的概率为多大？

互动出版网 China-Pub.com 专业图书网上第一专营店 www.china-pub.com 独家提供样章

3.21 如果 A 掷 $n+1$ 枚硬币，B 掷 n 枚均匀硬币，证明：A 得到的正面朝上数大于 B 得到的正面朝上数的概率为 1/2.

提示：以每人掷出 n 枚硬币后，谁具有更多的正面朝上数为条件(共有三种可能).

3.22 对于下列叙述，试证明或给出反例：

(a) 若 E 独立于 F，且 E 独立于 G，则 E 独立于 $F\cup G$；

(b) 若 E 独立于 F，E 独立于 G，且 $FG=\varnothing$，则 E 独立于 $F\cup G$；

(c) 若 E 独立于 F，F 独立于 G，且 E 独立于 FG，则 G 独立于 EF.

3.23 令 A 和 B 为具有正概率的事件，说明以下叙述是(i)必然对；(ii)必然错；(iii)可能对.

(a) 若 A 和 B 互不相容，则它们独立；(b) 若 A 和 B 独立，则它们互不相容；

(c) $P(A)=P(B)=0.6$，且 A 和 B 互不相容；(d) $P(A)=P(B)=0.6$，且 A 和 B 互相独立.

3.24 按照发生的概率的大小，将下列事件排序：

1. 掷一枚均匀硬币，正面朝上；
2. 3 次独立重复试验，每次成功的概率均为 0.8，三次都成功；
3. 7 次独立重复试验，每次成功的概率均为 0.9，7 次都成功.

3.25 有两个工厂生产收音机，工厂 A 生产的每台收音机是次品的概率为 0.05，而工厂 B 生产的每台收音机是次品的概率为 0.01. 假设你在同一家工厂购买了两台收音机，且这两台收音机来自 A 厂或 B 厂的概率是相等的. 如果第一台收音机经检测后是次品，另一台也是次品的条件概率是多少？

3.26 证明：若 $P(A\mid B)=1$，则 $P(B^c\mid A^c)=1$.

3.27 坛子里开始有 1 个红球和 1 个蓝球. 每步从其中随机地取出一个并同时放入两个同颜色的球.（比如，若开始取出了红球，那么在下次取球时，坛子里有 2 个红球和 1 个蓝球.）利用数学归纳法证明：n 步后，坛子里正好有 i 个红球的概率为 $1/(n+1)$，$1\leqslant i\leqslant n+1$.

3.28 共有 $2n$ 张牌，其中 2 张为 A. 将这些牌随机分给两位选手，每人 n 张. 然后两位选手按领牌次序声明自己是否有 A. 在第一人声称自己有 A 时，第二人没有 A 的条件概率是多大？其中

(a) $n=2$；(b) $n=10$；(c) $n=100$.

当 n 趋于无穷大时，该概率的极限值是多大？为什么？

3.29 市场上一共有 n 种不同的优惠券. 某人收集优惠券，每次收集一张，并且各次收集是相互独立的. 已知每次收集到第 i 种优惠券的概率为 p_i，$\sum_{i=1}^{n} p_i = 1$.

(a) 如果已经收集到 n 张优惠券，那么这 n 张优惠券都不相同的概率是多大？

(b) 现在假设 $p_1=p_2=\cdots=p_n=1/n$. 令 E_i 表示“在收集到的 n 张优惠券中没有第 i 种优惠券”，对 $P\left(\bigcup_i E_i\right)$ 利用容斥恒等式证明恒等式

$$n! = \sum_{k=0}^{n}(-1)^k\binom{n}{k}(n-k)^n$$

3.30 对任意事件 E 和 F，证明

$$P(E\mid E\cup F)\geqslant P(E\mid F)$$

提示：以 F 是否发生为条件计算 $P(E\mid E\cup F)$.

3.31 A 发生的概率是 60%，如果 A 没有发生，那么 B 发生的概率是 10%. 请问 A 或 B 至少有一个发生的概率是多少？ 111