一组事件的独立性的直观背景与两个事件的独立性是一样的. 独立性意味着下面一个事实：设把一组事件任意地分成两个小组, 一个小组中的任意个数的事件的出现与不出现, 都不会带来另一个小组中的事件的任何信息. 例如, 事件 $A_1A_2A_3A_4$ 是独立的事件组, 则下面一类等式都是成立的

$$\mathrm{P}(A_1 \cup A_2 | A_3 \cap A_4) = \mathrm{P}(A_1 \cup A_2),$$
$$\mathrm{P}(A_1 \cup A_2^c | A_3^c \cap A_4) = \mathrm{P}(A_1 \cup A_2^c).$$

证明可见本章末的习题.

1.5.3 可靠性

在由多个元件组合成的一个复杂系统中, 通常假定各个元件的表现是相互独立的. 下面的例子说明做了这样的假定以后, 计算和分析将变得十分简单.

例 1.24(网络连接) 在计算机网络中, A 和 B 两个结点通过中间结点 C、 D、 E、F 相互连接(见图1.15a). 图上直接连接的两个点 i 和 j 表示 i 和 j 之间有一个元件运行着, 当这个元件失效时两个点之间就失去连接. 我们假定 i 和 j 之间具有给定的连接概率 p_{ij}.[7]假定各点之间的连接与否独立于其他各点之间连接与否. 问 A 和 B 之间相互连接的概率有多大?

> [7]图1.15a中两个结点之间的箭头旁边的数字就是结点之间的连接概率. ——译者注

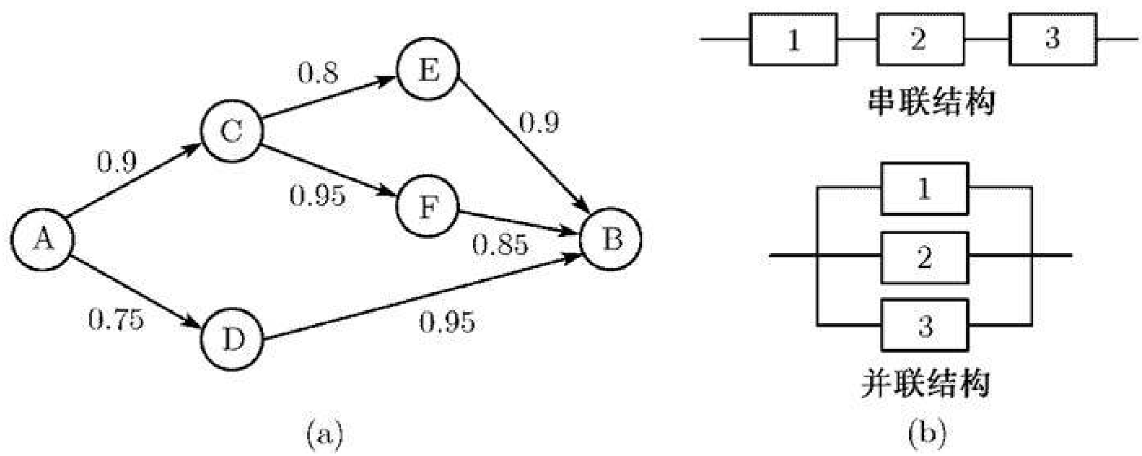

图 1.15 (a)例1.24的网络. 箭头旁边的数字表示相应的结点之间的元件有效的概率. (b) 在可靠性问题中由三个元件组成的串联和并联系统的图示

这是一个典型的系统可靠性的估计问题. 系统由元件组合而成, 而各元件的失效与否是相互独立的. 这些系统通常能够分解成若干子系统, 而每个子系统又由若干元件组成, 这些元件可以以**串联**方式或**并联**方式相互连接(见图1.15b).

设系统由元件 $1, 2, \cdots, m$ 组成, 令 p_i 为元件 i 有效(运行)的概率. 串联系统只有在所有元件均有效的情况才是有效的. 即

$$\mathrm{P}(\text{串联系统有效}) = p_1 p_2 \cdots p_m.$$

在并联系统中只需诸元件中有一个元件有效,系统就有效,即

$$\begin{aligned}\mathrm{P}(\text{并联系统有效}) &= 1 - \mathrm{P}(\text{并联系统失效}) \\ &= 1 - (1-p_1)(1-p_2)\cdots(1-p_m).\end{aligned}$$

现在回到图1.15a的网络连通的概率(A 和 B 之间连通的概率)的计算.我们用 $X \to Y$ 表示“由 X 到 Y 是连通的”这一随机事件.我们有

$$\begin{aligned}\mathrm{P}(C \to B) &= 1 - \big(1 - \mathrm{P}(C \to E\text{和}E \to B)\big)\big(1 - \mathrm{P}(C \to F\text{和}F \to B)\big) \\ &= 1 - (1 - p_{CE}p_{EB})(1 - p_{CF}p_{FB}) \\ &= 1 - (1 - 0.8 \cdot 0.9)(1 - 0.95 \cdot 0.85) \\ &= 0.946,\end{aligned}$$

$$\mathrm{P}(A \to C \text{ 和 } C \to B) = \mathrm{P}(A \to C)\mathrm{P}(C \to B) = 0.9 \cdot 0.946 = 0.851,$$

$$\mathrm{P}(A \to D \text{ 和 } D \to B) = \mathrm{P}(A \to D)\mathrm{P}(D \to B) = 0.75 \cdot 0.95 = 0.712.$$

最后,我们得到所需的概率

$$\begin{aligned}\mathrm{P}(A \to B) &= 1 - \big(1 - \mathrm{P}(A \to C \text{ 和 } C \to B)\big)\big(1 - \mathrm{P}(A \to D\text{和}D \to B)\big) \\ &= 1 - (1 - 0.851)(1 - 0.712) \\ &= 0.957.\end{aligned}$$

1.5.4　独立试验和二项概率

现在设试验由一系列独立并且相同的小试验组成,称这种试验为**独立试验序列**.当每个阶段的小试验只有两种可能结果的时候,就称为独立的**伯努利试验序列**,此处的两种可能结果可以是任何结果,例如“下雨”和“不下雨”.但是,在学术讨论中,我们通常用抛掷硬币的两个结果“正面”(H)和“反面”(T)作为代表.

现在考虑连续n次独立地抛掷硬币的试验,每次抛掷的结果为正面的概率为 p ,其中 p 是在0和1之间的数.此处“独立”意味着事件 $A_1, A_2, \cdots, A_n$ 是独立的,事件第 A_i={第 i 次抛掷的结果为“正面”}.

我们可以用序贯树形图来直观上刻画独立伯努利试验序列.图1.16中显示的是 n=3的情况.由于独立性,不管前面的抛掷结果是什么,每次抛掷得到正面的条件概率都是 p.这样,每个试验结果(长度为3的正面和反面的序列)的概率只与序列中的正面出现次数有关.设试验结果中有 k 个正面,3-k 个反面,则这个试验结果的概率为 $p^k(1-p)^{3-k}$.这个公式可以推广到任何 n 次抛掷硬币的试验结果的计算.在长度为 n 的独立伯努利试

验序列中,任何具有 k 个正面和 $n-k$ 个反面的试验结果的概率为 $p^k(1-p)^{n-k}$, 其中 k 的取值可以从0变到 n.

现在我们要计算概率

$$p(k) = \mathrm{P}(n\text{ 次抛掷中有 }k\text{ 次出现正面}),$$

这个概率在概率论中处于十分重要的地位. 由于任何包含 k 次正面向上的结果的概率都是 $p^k(1-p)^{n-k}$, 我们得到

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

此处记号

$$\binom{n}{k} = n\text{次抛掷硬币的试验中出现}k\text{次正面的试验结果数.}$$

数 $\binom{n}{k}$ 就是有名的**二项式系数**, 称为 n 选 k 的组合数, 概率 $p(k)$ 就是有名的**二项概率**. 在1.6节将介绍计数法, 利用计数法可以得到

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad k = 0, 1, \cdots, n,$$

此处记号 $i!$ 表示正整数 i 的阶乘,

$$i! = 1 \cdot 2 \cdots (i-1) \cdot i,$$

按传统, 记0!=1. 在本章末的习题中给出了这个公式的另一证明. 由于二项式概率 $p(k)$ 的总和必须为1, 这样我们得到**二项式公式**

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

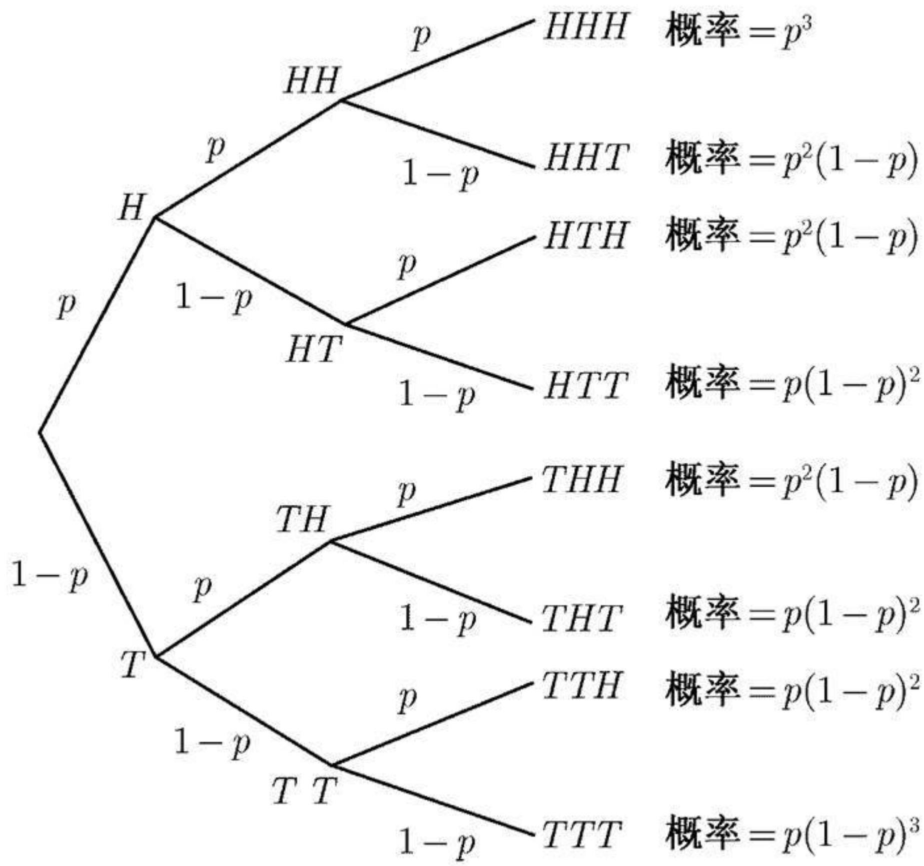

图 1.16　连续三次抛掷硬币试验的序贯树形图表示.在树枝上已经标明相应的条件概率.作为顺序三次抛掷硬币的结果的概率是在树形图的相应路径上的条件概率的乘积

例 1.25(服务等级)　设一个互联网服务器备有 c 个调制解调器以满足 n 个用户的需要.设在给定时刻,每一个用户相互独立地以概率 p 需要与服务器连接,当连接的时候,服务器需要有一个调制解调器以供使用.现在的问题是调制解调器不够用的概率有多大?

当同一时刻需要调制解调器的用户数多于 c 的时候,服务器就不能够满足用户的需要.它的概率为

$$\sum_{k=c+1}^{n} p(k),$$

其中

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

是二项概率.例如 n=200、p=0.1和 c=15,相应的概率为0.039 9.

这是一个典型的满足用户需求的设备规模问题.这批用户是一群具有相同需求并且独立行动的用户.现在的问题是要选择服务设备的规模,使得满足用户需求(指所有需要使用设备的用户都能得到服务)的概率超过给定的门限值(有时候,给概率值设立若干门限,称为**服务等级**).

1.6 计数法

在计算概率的时候,通常需要数清楚有关事件中的试验结果数(或基本事件数).我们已经遇到两种情况,需要这样的计数法.

(a) 当样本空间 Ω 只有有限个等可能的试验结果,因此这是一个等概率模型.事件 A 的概率可由下式给出

$$\mathrm{P}(A)=\frac{A \text{ 中元素的数目}}{\Omega \text{ 中元素的数目}},$$

公式中涉及 A 和 Ω 中元素的计数问题.

(b) 当我们需要计算事件 A 的概率,且 A 中的每一个试验结果具有相同的概率 p(p 已知)时,那么

$$\mathrm{P}(A)=p\cdot(A\text{中元素的数目}).$$

此时,也涉及事件 A 中的元素的计数问题.前面提到的 n 次抛掷硬币的试验中出现 k 次正面的事件的概率(二项概率)的计算就是这样一类的计算问题.这个概率的计算过程显示,每个试验结果的概率的计算是比较容易的,但是要数清楚具有 k 次正面向上的试验结果的个数,却有一些复杂.

计数问题原则上很简单,但是真正计算起来却不简单.计数的艺术属于**组合数学**的一部分.本节将介绍一些计数的基本准则,并将之应用到概率模型中经常遇到的计算问题.

1.6.1 计数准则

这是计数的最基本的方法.计数准则基于分阶段计数的原则,因此可以借助序贯树形图进行计数.例如,考虑一个由两个相继阶段组成的试验.第1阶段试验的可能结果为 $a_1,a_2,\cdots,a_m$,而第2阶段的结果为 $b_1,b_2,\cdots,b_n$.这样两阶段的试验结果为所有的有序对 $(a_i,b_j),i=1,\cdots,m,j=1,\cdots,n$.这些有序对的个数总和为 mn.这种计数方法可以进行推广 r 个阶段试验的情况(也可见图1.17的说明).

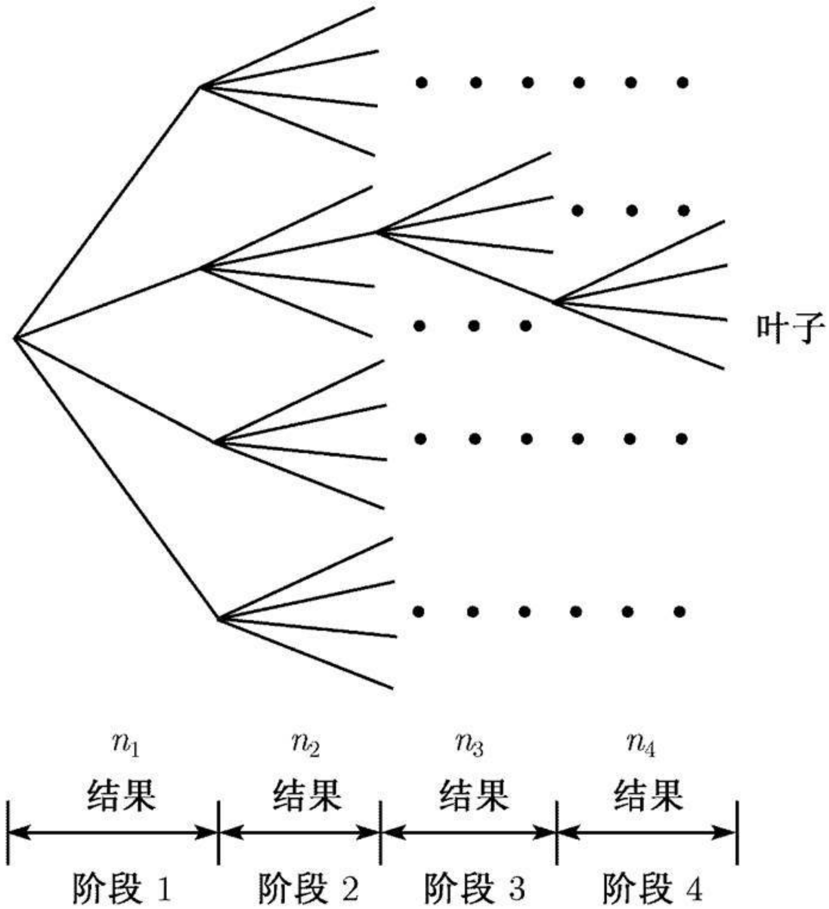

图 1.17　基本的计数准则的序贯树形图说明.通过 r 个阶段进行计数(图中r =4).第一个阶段有 n_1 个可能的结果.前r-1个阶段的每一个可能的结果,在第 r 阶段都对应着 n_r 个可能结果.总共的叶子数目为 $n_1n_2\cdots n_r$

计数准则[8]

考虑由 r 个阶段组成的一个试验}.假设:

(a) 在第1阶段有 n_1 个可能的结果;

(b) 对于第1阶段的任何一个结果,在第2阶段有 n_2 个可能的结果;

(c) 一般地,在前 r -1个阶段的任何一个结果,在接下来的第 r 阶段有 n_r 个结果,则在 r 个阶段的试验中一共有

$$n_1n_2\cdots n_r$$

个试验结果.

[8]国内称为“计数的乘法准则”或“乘法准则”,这个名称更通俗易懂.——译者注

例 1.26(电话号码数) 电话号码由7位数字组成,但第一位不能是0或1.一共有多少个不同的号码呢?我们可以将之看成序贯地选择数字的过程,但每次只选一位.总共有7个阶段,第1个阶段一共有8种选择,从第2阶段开始,每次都从10个数字中任选一个.因此电话号码的个数为

$$8 \cdot \underbrace{10 \cdot 10 \cdots 10}_{6\ 次} = 8 \cdot 10^6.$$

例 1.27(n 元素集合的子集的个数) 考虑一个 n 元素集合 $\{s_1, s_2, \cdots, s_n\}$.这个集合有多少个子集(包括这个集合本身和空集)呢?我们可以用序贯的方法选择一个子集.我们可以对每一个元素做一个选择,并判断它是否属于这个子集.这样一共分成 n 个阶段,每一个阶段有两种选择.这样子集的总数为

$$\underbrace{2 \cdot 2 \cdots 2}_{n\ 次} = 2^n.$$

可以对这个计数准则做一些小修改.对于不同的第一阶段的结果后面可以接着不同的第二阶段的试验,只要各个第二阶段的可能结果的数目相同.

下面我们将讨论从 n 个对象中选取 k 个对象的计数问题.若选取的对象与次序有关,则选出来的一组对象称为**排列**,若选出来的一组对象是形成一个集合,与选取的对象的次序无关,则这一组对象称为**组合**.以后我们还会讨论更一般的**分割**的计数问题.所谓分割就是将 n 个对象分成多个子集.

1.6.2 n 选 k 排列

首先假定 n 个不同的对象组成一个集合.令 k 是一个正整数,$k \le n$.现在我们希望找出从 n 个对象中顺序地选出 k 个对象的方法数,或 k 个不同对象的序列数.作为第一阶段,我们可以从 n 个对象中任意选一个.当第一个对象选定以后,在第二阶段,我们只可能从剩下的 $n-1$个对象中选择一个.当前两个对象选定以后,在第三阶段,只可能从剩下的 $n-2$个对象中选择一个,等等.最后,当我们选择第 k 个对象的时候,只能从剩下的 $n-(k-1)$个对象中选择了.利用计数准则,所有可能的序列数为

$$\begin{aligned} n(n-1)\cdots(n-k+1) &= \frac{n(n-1)\cdots(n-k+1)(n-k)\cdots 2 \cdot 1}{(n-k)\cdots 2 \cdot 1} \\ &= \frac{n!}{(n-k)!}. \end{aligned}$$

这些序列称为 **n 取 k 排列**.特别当 $k=n$ 的时候,简称为**排列**[9],此时所有可能的序列数为

> [9]此处的排列、组合和分割在中英文中均有双重意义,一个排列是指 n 个元素的一个顺序,同时又可以指排列数 $n!$,具体指哪种内容要看行文.——译者注

$$n(n-1)\cdots 2 \cdot 1 = n!.$$

(在 n 取 k 排列的序列数公式中令 k=n,并回忆我们已经约定0!=1.)

例 1.28 现在计算由4个不同字母组成的字的个数.这是26选4的排列数.按排列公式为

$$\frac{n!}{(n-k)!}=\frac{26!}{22!}=26\cdot 25\cdot 24\cdot 23=358\ 800.$$

排列计数法可以与计数的乘法准则联合起来解决更复杂的排列问题.

例 1.29 你有 n_1 张古典音乐CD盘,n_2 张摇滚音乐CD盘,n_3 张乡村音乐CD盘.有多少种排列方法将这些CD盘排在CD架上,使得相同种类的CD盘是排在一起的?

我们将问题分成两步解决.首先选择CD盘类型的次序,然后选择每种CD盘内部的次序.一共有3!种类型次序(例如古典/摇滚/乡村,乡村/古典/摇滚等),一共有 $n_1!$ (或 $n_2!$,或 $n_3!$)种古典(或摇滚,或乡村)CD的排列.这样对每一种CD类型的排列,有 $n_1!n_2!n_3!$ 种CD盘的排列方式.从而总的排列方法数为 $3!n_1!n_2!n_3!$.

现在假定,计划将每一类CD盘中选出 k_i 张(你原有 n_i 张 i 类CD)送给你的朋友.当你送出盘以后,你的CD架上有多少种排列法?这个问题与没有送出时的计算方法是一样的,只是将 $n_i!$ 换成 n_i 选 n_i-k_i 的排列数即可.所以可能的排列数为[10]

> [10]在计算排列方法数的时候,要顾及各种不同的送CD盘的方法.——译者注

$$3!\cdot\frac{n_1!}{k_1!}\cdot\frac{n_2!}{k_2!}\cdot\frac{n_3!}{k_3!}.$$

1.6.3 组合

一共有 n 个人,希望组织一个 k 个人的委员会.问有多少种不同的委员会?用抽象的语言说,给定的 n 个元素的集合中有多少种不同的 k 个元素的子集?注意,形成 k 子集不同于形成 n 选 k 排列,因为在选择子集的过程中,选出来的 k 个元素之间是没有次序的.例如4个字母 A、B、C 和 D 中选2个的排列有12种:

$$AB,\ AC,\ AD,\ BA,\ BC,\ BD,\ CA,\ CB,\ CD,\ DA,\ DB,\ DC,$$

而这4个字母的两个字母的组合有下列6种:

$$AB,\ AC,\ AD,\ BC,\ BD,\ CD.$$

(因为在组合中元素是没有次序的,AB 和 BA 是无法区别的.)

在上面的例子中,组合实际上是由排列归并而成的.例如,从组合的观点看来,AB 和 BA 是不可区分的,它们都对应于组合 AB.这种推导方法可以推广到一般的情况:在 n 个对象取 k 个对象的组合中,每一个组合对应了 $k!$ 个不同的排列.这样 n 个对象取 k 个对象的排列数 $n!/(n-k)!$ 等于组合数乘以 $k!$.因此,从 n 个元素的集合中选 k 个元素的组合数为

$$\frac{n!}{k!(n-k)!}.$$

现在回到二项式系数 $\binom{n}{k}$ 的表达式. 二项式系数定义为 n 次抛掷硬币时, 正面向上次数为 k 的可能的试验结果数. 我们注意到, 确定一个 k 次向上的试验结果等价于在所有 n 次抛掷结果(正面向上或反面向上)选出 k 次(正面向上)来. 因此二项式系数刚好等于从 n 个元素选择 k 个元素的组合数. 这样

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

例 1.30　A、 B、 C 和 D 四个字母中选出两个字母的组合数为

$$\binom{4}{2} = \frac{4!}{2!2!} = 6.$$

这个结果与前面列举的组合数相同.

值得指出的是, 有时候利用计数法能够导出一些在代数上很难证明的公式. 一个例子是1.5节讨论的二项式公式

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

作为特殊情况, 当 p=1/2时, 公式变成

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n,$$

上式还可以得到新的解释. 由于 $\binom{n}{k}$ 是 n 元素集合的所有 k 元素子集的个数, 将 $\binom{n}{k}$ 对所有的 k 求和得到这个集合的所有子集的个数, 而这个数刚好等于 2^n.

例 1.31　设有一群人, 一共有 n 个. 现在要组织一个个人爱好俱乐部, 俱乐部由一个主任和若干成员组成(成员人数可为0). 问有多少种方式组成一个俱乐部?我们用两种不同的计数法计算, 从而得到一个代数恒等式.

首先挑选一个俱乐部主任, 一共有 n 种不同的选法. 然后从剩下的 n-1个人员中挑选一般成员. 实际上, 这 n-1人中任意一个子集, 配上主任, 就成为一个俱乐部. 而不同的子集个数共有 2^{n-1} 个. 这样一共有 $n2^{n-1}$ 种不同的方式组成一个俱乐部.

另外, 我们可以这样考虑此问题. 首先选择 k 个人组成一个 k 人集体, 然后从中选择一个主任, 组成一个 k 人俱乐部. 这样一共有 $k\binom{n}{k}$ 种方式组成一个 k 人俱乐部. 对所有的 k $(k=1,\cdots,n)$, 将组成 k 人俱乐部的方式数相加, 就得到组成俱乐部的方式数. 由此得到代数恒等式

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}.$$

1.6.4 分割

注意到组合是从 n 元素集合中选出的一个元素个数为 k 的子集,因此可将一个组合看成将集合分成两个子集合的一个分划,其中一个子集的元素个数为 k ,另一个子集为补集,其元素的个数为 $n-k$.现在我们考虑将一个集合分成多于两个集合的分割.

给定一个元素个数为 n 的集合,并设 $n_1, n_2, \cdots, n_r$ 为非负整数,其总和为 n .现在考虑将具有 n 个元素的集合分解成 r 个不相交的子集,使得第 i 个子集元素个数刚好是 n_i.问一共有多少种分解的方法.

现在分阶段每次确定一个子集.一共有 $\binom{n}{n_1}$ 种方法确定第一个子集.当第一个子集确定以后,只剩下 $n-n_1$ 个元素可以用来确定第二个子集.这样在确定第二个子集的时候,一共有 $\binom{n-n_1}{n_2}$ 种方法,以此类推.对 r 个阶段的选择过程利用计数准则,得到总共的选择方法数目为

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-\cdots-n_{r-1}}{n_r},$$

上式等于

$$\frac{n!}{n_1!(n-n_1)!}\cdot\frac{(n-n_1)!}{n_2!(n-n_1-n_2)!}\cdots\frac{(n-n_1-\cdots-n_{r-1})!}{n_r!(n-n_1-\cdots-n_{r-1}-n_r)!}.$$

经过消去化简,上式等于

$$\frac{n!}{n_1!n_2!\cdots n_r!}.$$

这个数称为**多项式系数**,并且用下列记号表示:

$$\binom{n}{n_1, n_2, \cdots, n_r.}$$

例 1.32(相同字母异序词) 将 TATTOO 这个英文单词的字母颠倒排列可得到多少个不同的单词?这里有6个位置供这些字母去填充.每一种重新排列方式可以看成一个6个位置的分割,分割的一个小组的大小为3,用于放置字母 T,另一个小组的大小为2,用于放置字母 O,第三个小组的大小为1,用于放置字母 A.这样一共有

$$\frac{6!}{1!2!3!}=\frac{1\cdot2\cdot3\cdot4\cdot5\cdot6}{1\cdot1\cdot2\cdot1\cdot2\cdot3}=60$$

个单词.

也可以用另一种方法导出这个结果(这种方法也可以用于导出多项式系数的公式,见本章后习题).我们将 TATTOO 写成 $\mathrm{T_1AT_2T_3O_1O_2}$ 的形式,假装这6个字母是不相同的.这样一共有6!种不同的排列.然而有3!种 $\mathrm{T_1T_2T_3}$ 的排列和2!种 $\mathrm{O_1O_2}$ 的排列形成同一个单词,这样当下标去掉以后,一共有6!/(3!2!)个不同的单词.

例 1.33 一个班由4个研究生和12个本科生组成. 将这个班随机地分成4个小组, 每组4人. 问每个组刚好包含一个研究生的概率有多大?这个问题就是1.3节例1.11的问题. 但是现在我们要利用计数方法解答这个问题.

首先应该确定样本空间. 我们将分小组的问题设想成将16个学生随机地放入4个房间, 每个房间4个人, 这是一个分割问题. 由于16个人是随机地分派到各个房间里去的, 故每个分割的概率是相等的. [11]

> [11]这样, 样本空间由全体分割组成, 并且概率律是等概率的. ——译者注

按照分割的定义, 分割数为

$$\binom{16}{4,4,4,4} = \frac{16!}{4!4!4!4!}.$$

现在考虑每一个房间只分配一个研究生的分割数. 我们可以分两个阶段完成学生的分派问题.

(a) 第一阶段, 将4个研究生分派到4个房间中去, 每个房间1人. 这是一个只有4个人的分割问题, 分割数为4!.

(b) 第二阶段, 将12个本科生分派到4个房间中去, 每个房间分派3人. 这也是一个分割问题, 分割数为

$$\binom{12}{3,3,3,3} = \frac{12!}{3!3!3!3!}.$$

利用乘法准则, 每个房间分派1个研究生和3个本科生的方法一共有

$$\frac{4!12!}{3!3!3!3!}$$

种. 这样, 按古典概型的定义, 每个小组分派到一个研究生的概率为

$$\frac{\dfrac{4!12!}{3!3!3!3!}}{\dfrac{16!}{4!4!4!4!}}.$$

经过化简, 这个数为

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

这个结果与例1.11的结果相符合.

下面是计数法的汇总.

计数法汇总

- n 个对象的**排列数**：$n!$.
- n 个对象中取 k 个对象的**排列数**：$n!/(n-k)!$.
- n 个对象中取 k 个对象的**组合数**：$\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
- 将 n 个对象分成 r 个组的**分割数**, 其中第 i 个组具有 n_i 个对象

  $$\binom{n}{n_1, n_2, \cdots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}.$$

1.7 小结和讨论

解决一个概率问题通常分成下列几个步骤:

(a) 描述样本空间, 样本空间是一个试验的所有可能的试验结果的集合;

(b) (可能不直接地)列出概率律(每个事件的概率);

(c) 计算各种事件的概率和条件概率.

概率律必须满足非负性、可加性和归一性公理. 对于试验结果的总数有限的重要特例, 我们只需列出每一个可能试验结果的概率, 而任何事件的概率的计算, 只需将组成这个事件的所有可能的试验结果的概率相加, 就得到这个事件的概率.

给定一个概率律, 我们经常需要计算条件概率, 这是因为条件概率涉及得到部分信息以后的概率计算问题. 我们也可以将条件概率看成特殊的概率律, 在这个概率律之下, 只有包含于由条件所确定的事件内的事件才有正的条件概率. 条件概率可以通过公式 $\mathrm{P}(A|B)=\mathrm{P}(A\cap B)/\mathrm{P}(B)$ 进行计算. 然而在应用中, 更常见的是利用条件概率来计算无条件概率.

我们已经用例子说明了计算概率的如下三种方法.

(a) **计数法**. 这种方法适用于古典概型, 即试验只有有限个可能的试验结果, 而试验结果是等可能的. 为计算一个事件的概率, 只需数清楚这个事件中的基本事件个数, 再除以基本事件总数, 就得到这个事件的概率.

(b) **序贯树形图方法**. 在试验具有序贯特征的情况下, 可以利用序贯树形图方法. 这种方法的关键是必须计算相应树枝事件的条件概率. 这些条件概率或者是已知的或者是利用各种方法(包括计数法)计算得到的. 利用乘法规则将相应路径上的事件的条件概率相乘, 就可以得到相应事件的概率.

(c) **全概率公式**. 利用全概率公式可以计算事件B的概率 $\mathrm{P}(B)$, 关键是要找到样本空间的一个分割 $A_i, i=1,\cdots,n$, 使得相应的概率 $\mathrm{P}(A_i)$ 和条件概率 $\mathrm{P}(B|A_i)$ 为已知或比较容易计算, 然后利用全概率公式计算 $\mathrm{P}(B)$.

最后, 我们还讨论了若干问题, 这些问题或者扩大了概率论的应用范围, 或者提高了利用主要定理进行计算的能力. 我们引入了贝叶斯准则, 这是概率论的一个重要应用领域. 同时, 为了加强计算能力, 我们讨论了计数方法的一些基本规则, 包括组合、排列等.

习题

1.1 节　集合

1. 考虑掷一个具有6个面的骰子. 令事件 A 为掷出偶数. 令 B 表示点数大于3的事件. 验证下面的德摩根公式：

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c.$$

2. 设 A 和 B 是两个集合.

(a) 证明

$$A^c = (A^c \cap B) \cup (A^c \cap B^c), \quad B^c = (A \cap B^c) \cup (A^c \cap B^c).$$

(b) 证明

$$(A \cap B)^c = (A^c \cap B) \cup (A^c \cap B^c) \cup (A \cap B^c).$$

(c) 考虑掷一个均匀的、具有6个面的骰子. 令事件 A 为掷出奇数. 令 B 表示点数小于4的事件. 求出(b)中公式两边的集合并验证集合等式.

3.* 证明恒等式

$$A \cup (\cap_{n=1}^{\infty} B_n) = \cap_{n=1}^{\infty} (A \cup B_n).$$

解　若 x 为左边的集合的元素, 则有两种可能性. (i) $x \in A$, 此时对一切 $n \geq 1$, $x \in A \cup B_n$, 从而 x 属于等式右边的集合; (ii) 对一切 $n \geq 1$, $x \in B_n$, 此时对一切 $n \geq 1$, $x \in A \cup B_n$, 这样, x 也属于等式右边的集合.

反过来, 若 x 是等式右边的集合的元素, 说明对一切 $n \geq 1$, $x \in A \cup B_n$. 若 $x \in A$, 显然 x 是等式左边的集合的元素. 若 $x \notin A$, 此时, 对一切 $n \geq 1$, x 必须是 B_n 的元素, 这再一次证明 x 是等式左边的集合的元素.

4.* 康托尔的三角论证方法　指出单位区间[0,1]是不可数集合, 即[0,1]中的数不可能排成一个数列.

解　每一个[0,1]区间中的数, 都有十进制表达式, 例如 $1/3 = 0.3333\cdots$. 注意, 绝大部分数具有唯一的表达式, 但也有例外, 例如1/2可以表为 $0.5000\cdots$ 或 $0.4999\cdots$. 可以证明这些数是仅有的例外, 即只有结尾是无限个0的数或结尾是无限个9的数才有两种表达式.

现在用反证法. 假设所有的[0,1]区间中的数, 可以排成一列, $x_1, x_2, x_3, \cdots$, 即[0,1]区间中的每一个数都在这个序列中. 考虑 x_n 的十进制表达式

$$x_n = 0.a_n^1 a_n^2 a_n^3 \cdots,$$

其中 a_n^i 为集合 $\{0,1,\cdots,9\}$ 中的一个数. 现在构造一个数 y, 它的第 n 位小数取成1或2, 但是它不等于 x_n 的第 n 位数 $a_n^n, n=1,2,\cdots$. 由于 y 的第 n 位与 x_n 的第 n 位数不同, y 与 x_n 是不同的. 这样 y 不可能在 $x_1,x_2,x_3,\cdots$ 中, 与假设矛盾. 从而[0,1]区间中的数是不可数的.

1.2 节　概率模型

5. 在一个班上, 有60%的学生是天才, 70%的学生喜欢巧克力, 40%的学生既是天才又喜欢巧克力. 现在从班上随机地选择一位同学, 请问他既不是天才学生又不爱好巧克力的概率有多大?

6. 一个有6个面的骰子是这样设计的：在抛掷骰子的时候, 所有偶数面出现的概率比奇数面出现的概率大一倍, 不同的偶数面出现的概率是相同的, 不同的奇数面出现的概率也是相同的. 现在将骰子抛掷一次, 为这个试验建立概率律, 并求出点数小于4的概率.

7. 将一个有4个面的骰子持续地抛掷若干次, 直到第一次出现偶数面为止. 这个试验的样本空间是什么?

8. 你参加一个象棋比赛, 必须与三个对手下象棋. 按规定, 只有赢两场比赛, 才算你得胜. 假定, 与每个对手比赛的时候, 你赢棋的概率是已知的. 另外, 你成为得胜者的概率与比赛的次序有关. 证明将三位比赛对手中的最弱者排在第二位的时候, 你成为得胜者的概率最大, 而与其他两位对手的比赛次序无关.

9. 样本空间 Ω 的分割是一组互不相容的事件组 $\{S_1,\cdots,S_n\}$, 满足条件 $\Omega=\cup_{i=1}^n S_i$.

(a) 证明对任何事件 A , 下式成立

$$\mathrm{P}(A)=\sum_{i=1}^n \mathrm{P}(A\cap S_i).$$

(b) 利用(a)的结论, 证明对任何事件 A、B 和 C , 下式成立

$$\mathrm{P}(A)=\mathrm{P}(A\cap B)+\mathrm{P}(A\cap C)+\mathrm{P}(A\cap B^c\cap C^c)-\mathrm{P}(A\cap B\cap C).$$

10. 证明公式

$$\mathrm{P}\left((A\cap B^c)\cup(A^c\cap B)\right)=\mathrm{P}(A)+\mathrm{P}(B)-2\mathrm{P}(A\cap B),$$

这个公式给出 A 和 B 中间恰有一个事件发生的概率. (与公式 $\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$ 相比较, 后者给出 A 和 B 中间至少有一个事件发生的概率.)

11.* 邦费罗尼不等式.

(a) 对于任何两个事件 A 和 B , 证明

$$\mathrm{P}(A\cap B)\ge \mathrm{P}(A)+\mathrm{P}(B)-1.$$

(b) 将上式推广到 n 个事件 $A_1, A_2, \cdots, A_n$ 的情况,证明

$$\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \geq \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots + \mathrm{P}(A_n) - (n-1).$$

解 由等式 $\mathrm{P}(A \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$ 和不等式 $\mathrm{P}(A \cup B) \leq 1$ 立即可得 (a). 至于 (b),利用德摩根公式可得到下面的结果

$$\begin{aligned}
1 - \mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n) &= \mathrm{P}\left((A_1 \cap A_2 \cap \cdots \cap A_n)^c\right) \\
&= \mathrm{P}(A_1^c \cup A_2^c \cup \cdots \cup A_n^c) \\
&\leq \mathrm{P}(A_1^c) + \mathrm{P}(A_2^c) + \cdots + \mathrm{P}(A_n^c) \\
&= (1 - \mathrm{P}(A_1)) + (1 - \mathrm{P}(A_2)) + \cdots + (1 - \mathrm{P}(A_n)) \\
&= n - \mathrm{P}(A_1) - \mathrm{P}(A_2) - \cdots - \mathrm{P}(A_n),
\end{aligned}$$

由这个公式可得到(b).

12.* **容斥恒等式**.将下面的公式推广

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

(a) 设 A、B、C 为三个事件,则下列恒等式成立

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C) - \mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) + \mathrm{P}(A \cap B \cap C).$$

(b) 设 $A_1, A_2, \cdots, A_n$ 为 n 个事件. 记 $S_1 = \{i | 1 \leq i \leq n\}, S_2 = \{(i_1, i_2) | 1 \leq i_1 < i_2 \leq n\}$, 一般地,令 S_m 为满足条件 $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ 的 m 维指标 $(i_1, \cdots, i_m)$ 的集合,则下列恒等式成立

$$\begin{aligned}
\mathrm{P}(\cup_{k=1}^n A_k) = &\sum_{i \in S_1} \mathrm{P}(A_i) - \sum_{(i_1, i_2) \in S_2} \mathrm{P}(A_{i_1} \cap A_{i_2}) \\
&+ \sum_{(i_1, i_2, i_3) \in S_3} \mathrm{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots + (-1)^{n-1} \mathrm{P}(\cap_{k=1}^n A_k).
\end{aligned}$$

解 (a) 利用公式 $\mathrm{P}(X \cap Y) = \mathrm{P}(X) + \mathrm{P}(Y) - \mathrm{P}(X \cap Y)$ 和集合等式 $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ 得到

$$\begin{aligned}
\mathrm{P}(A \cup B \cup C) &= \mathrm{P}(A \cup B) + \mathrm{P}(C) - \mathrm{P}((A \cup B) \cap C) \\
&= \mathrm{P}(A \cup B) + \mathrm{P}(C) - \mathrm{P}((A \cap C) \cup (B \cap C)) \\
&= \mathrm{P}(A \cup B) + \mathrm{P}(C) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) + \mathrm{P}(A \cap B \cap C) \\
&= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) + \mathrm{P}(C) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) \\
&\quad + \mathrm{P}(A \cap B \cap C) \\
&= \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C) - \mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) \\
&\quad + \mathrm{P}(A \cap B \cap C).
\end{aligned}$$

(b) 利用归纳法.其主要推断部分可以模仿(a)中的推导步骤.另一种证明方法可以参考第2章末的习题.

13.∗ 概率的连续性.

(a) 设 $A_1, A_2, \cdots$ 是一个单调递增的事件序列,即对每一个 n,$A_n \subset A_{n+1}$. 令 $A = \cup_{n=1}^{\infty} A_n$. 证明 $\mathrm{P}(A) = \lim_{n\to\infty} \mathrm{P}(A_n)$. **提示**:将 A 表示成可数无限个不相交的事件之和.

(b) 设 $A_1, A_2, \cdots$ 是一个单调递减的事件序列,即对每一个 n,$A_n \supset A_{n+1}$. 令 $A = \cap_{n=1}^{\infty} A_n$. 证明 $\mathrm{P}(A) = \lim_{n\to\infty} \mathrm{P}(A_n)$. **提示**:将(a)的结果应用于事件的补集.

(c) 考虑一个概率模型,其样本空间是实数集合. 证明

$$\mathrm{P}([0,\infty)) = \lim_{n\to\infty} \mathrm{P}([0,n]) \quad 和 \quad \lim_{n\to\infty} \mathrm{P}([n,\infty)) = 0.$$

解 (a) 令 $B_1 = A_1$, 对 $n \geq 2$, 令 $B_n = A_n \cap A_{n-1}^c$. 这样定义的事件序列 B_n 是互不相容的事件序列,并且 $\cup_{k=1}^{n} B_k = A_n$, $\cup_{k=1}^{\infty} B_k = A$. 利用可加性公理得到

$$\mathrm{P}(A) = \sum_{k=1}^{\infty} \mathrm{P}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathrm{P}(B_k) = \lim_{n\to\infty} \mathrm{P}(\cup_{k=1}^{n} B_k) = \lim_{n\to\infty} \mathrm{P}(A_n).$$

(b) 令 $C_n = A_n^c$和$C = A^c$. 由于 $A_{n+1} \subset A_n$, 可知 $C_n \subset C_{n+1}$, 即事件序列 C_n 是上升的序列. 进一步 $C = A^c = (\cap_{n=1}^{\infty} A_n)^c = \cup_{n=1}^{\infty} A_n^c = \cup_{n=1}^{\infty} C_n$. 将(a)用于事件序列 C_n, 得到

$$1 - \mathrm{P}(A) = \mathrm{P}(A^c) = \mathrm{P}(C) = \lim_{n\to\infty} \mathrm{P}(C_n) = \lim_{n\to\infty} (1 - \mathrm{P}(A_n)),$$

由此可得结论:$\mathrm{P}(A) = \lim_{n\to\infty} \mathrm{P}(A_n)$.

(c) 令 $A_n = [0,n]$ 和 $A_n = [n,\infty]$ 和 $A = \cap_{n=1}^{\infty} A_n = \varnothing$, 再利用结论(b),就可以得到第二式.

1.3 节　条件概率

14. 将一个均匀的具有6个面的骰子连续抛掷两次. 36个可能的结果是等概率的.

(a) 找出抛掷出“一对”的概率;

(b) 已知抛掷得到的点数总和小于或等于4, 求抛掷出“一对”的概率;

(c) 求出至少一个骰子得6点的概率;

(d) 已知抛掷得到两个骰子的点数不同的条件下, 求出至少一个骰子得6点的概率.

15. 将一枚硬币抛掷两次. 爱丽丝声称在已知头一次得到正面朝上的条件下, 抛掷得到两次正面的可能性比已知两次中至少有一次正面朝上的条件下的可能性大. 这个结论对吗?当硬币为对称和不对称的条件下结论会不会不同?能不能将爱丽丝的推论方法推广呢?