统 计 分 析 方 法

朱仁海　杨琪瑜　沈文瑛

中国林业出版社

编 者 的 话

本书是为农业、林业科技工作者而编写的。重点介绍农林业生产和科研中观测数据常用的各种统计分析方法及其所依据的统计学基本原理。内容包括概率论基础，数理统计及多元统计方法三篇，各篇相互衔接，而各自又具有较完整的体系。

全书在编写中，力求文字通俗易懂，内容由浅入深，引进概念从具体到抽象，对公式和定理不强调理论上的严格推证，但尽可能将各种统计方法的基本思想和计算步骤介绍清楚，并配有较多的农林业上的应用实例。每章都有适量的练习题，以配合读者掌握本书内容，提高解决实际问题的能力。书末还附有统计分析方法各类数表，以备查用。为了适应高低不同层次的读者需要，书中加“*”号的部分，可以略去不读，不会影响后面内容的学习。阅读本书除了注有*号内容外，只需具有微积分初步、线性代数的基本知识即可。学完本书前二篇，则可达到高等农林业院校本科对本门课程的要求。

本书可作为农林业工程师和科技人员的培训教材、高等林业院校函授教材，亦可作为农林业科技人员自学用书。

本书的编写工作由南京林业大学干训部组织，在干训部主任瞿斌同志的领导下进行，其中第一篇概率论基础部分由朱仁海编写，第二篇数理统计方法部分由杨琪瑜编写，第三篇多元统计方法部分由沈文瑛编写，本书由南京大学张守智副教授主审，并对本书提出了许多宝贵意见；束仁荣副教授阅读过部分原稿，提出了宝贵意见；杨振华、王守根和蔡观华等同志在工作上给予了大力支持，谨此一并致谢。

由于编者水平有限，书中难免有不少缺点和错误，恳请读者批评、指教。

编　　者

1989年12月于南京

目　录

第三篇　多元统计分析方法

附　表

第一章 随机事件及其概率

我们在生产实践和日常生活中，经常可以观察到两种不同类型的现象，一种是确定性现象，即在一定条件下必然发生的现象。例如在标准大气压下，水加热到100℃必然沸腾；到秋冬季节，梧桐树必然落叶，等等。另一种是非确定性现象，或称为随机现象，即在一定条件下，它具有多种可能的结果，而事先不能确定会出现哪一种结果的现象。例如投掷一枚硬币，其结果可能是“徽花向上”，或是“币值向上”，事先不能确定；从含有二个次品、八个正品的十个同类产品中，任意抽取三个，三个中所含的次品个数可能是0，1，或2，抽取前不能确定，等等。

对确定性现象，我们可以用初等数学、微积分等数学分支来研究其数量规律性。对随机现象，又如何寻找其规律性呢？人们经过长期实践和深入研究后，发现这类现象，虽然就每次观察来讲，具有不确定性，但在多次重复下，其结果就呈现出某种规律性了，我们称此为统计规律性。例如多次重复投掷一枚均匀硬币，得到“徽花向上”的次数，大致为总投掷数的一半；对含有部分次品的一大批产品，进行多次重复抽查，就可确定它所含次品的次品率大致为多少，等等。概率论与数理统计，就是研究随机现象统计规律性的数学分支，它在生产实践和社会科学中，都有广泛的应用。本章主要介绍随机事件及其概率的一些基本知识。

§1 预备知识——排列与组合

排列与组合是一种数学方法。它是研究古典概率的重要工具，故在此先作一些简要的介绍。

一、乘法原理

如果完成一个事件的整个过程，可以分成两个阶段进行，第一阶段有m种不同方法，第二阶段有n种不同方法，并且第一阶段的任何一种方法，都可以与第二阶段的任何一种方法相搭配，成为整个过程的一种方法，那么整个过程就应该有$m\times n$种不同的方法。

对于整个过程可以分成多于两个阶段（有限个）的情况，亦可按上述方法作相应的推广。

例1.1 设有A、B、C三类书籍，其中A类8本，B类6本，C类5本，现从三类书籍中各取一本，问共有多少种不同的取法？

解 根据题意，按乘法原理，共有不同取法的种数为

$$S=8\times 6\times 5=240\text{（种）}$$

二、排 列

从n个不同的元素中，任意取出r个（$r\leqslant n$）不同的元素，按照一定的顺序排成一列，叫

做从n个不同元素中取r个不同元素组成的一种排列。所有这种排列的种数，通常记作 P_n^r，并有计算公式：

$$P_n^r = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!} \qquad (1.1)$$

这是因为从n个不同的元素中，任意取出r个（$r \leqslant n$）不同的元素，相当于从n个不同的元素中，每次取一个元素且不放回，共取r次。而取第一个元素时有n种取法，取第二个元素时有$n-1$种取法，…，取第r个元素时有$n-r+1$种取法，故按照乘法原理，共有排列的种数为

$$P_n^r = n(n-1)\cdots(n-r+1)$$

如果用阶乘来表示，并规定$0! = 1$，则为

$$P_n^r = \frac{n!}{(n-r)!}$$

当$r = n$时，上述排列即为将n个不同的元素任意排成一列，这种排列叫做全排列。记作P_n，并有

$$P_n = n! \qquad (1.2)$$

例1.2　从1，2，…，7七个数字中任取三个不同的数字组成三位数，问这样的三位数共有多少个？又其中属于偶数的占多少？

解　三个数字都不同的三位数的个数，即为从所给的七个不同数字中，任取三个的排列种数。故共有

$$P_7^3 = 7 \times 6 \times 5 = 210(\text{个})$$

上述三位数中属于偶数的，则要求个位上应是2，4，6中的一个，因此在个位上的数只有三种不同的取法，当个位上的数取定后，十位、百位上的数共有6×5种不同的取法，从而，所求的三位数为偶数的个数为

$$S = 3 \times 6 \times 5 = 90\ (\text{个})$$

例1.3　用a、b、c、d、e五个字母，可以组成多少种无重复字母的排列？若规定a与b必须排在一起，则有多少种排法？

解　五个不同字母组成无重复的排列，即为五个不同元素的全排列，其种数为

$$P_5 = 5! = 120\ (\text{种})$$

如果a与b必须排在一起，则可以先把a与b合在一起看作一个元素，与其余三个字母进行全排列，其种数为P_4，而a与b排在一起时，还可以互换位置，即有P_2种排法，因此，所求的排列种数为

$$P_4 \times P_2 = 4! \times 2! = 48\ (\text{种})$$

二、可重复排列

从n个不同的元素中，任取r个元素进行排列，并且每个元素都可以重复选取，叫做从n个不同元素中取r个元素的可重复排列。

在可重复排列中，任取r个元素，相当于从n个不同的元素中，每次任取一个元素且取后仍放回，共取r次。故所取每个元素都有n种取法，因此，所有排列的种数为

$$N=\underbrace{n\times n\times\cdots\times n}_{r\text{个}}=n^r\ (\text{种}) \tag{1.3}$$

例1.4 某城市的电话号码由五位数字组成，问最多可以安装多少台不同号码的电话？

解 因为组成电话号码的五个数字中，每个数字都可以是0，1，2，…，9这十个数字中的任一个，所以这是在十个不同的元素中，任取五个的可重复排列，其种数为

$$N=10^5=100000\ (\text{个})$$

即最多可以安装十万台不同号码的电话。

四、组 合

从n个不同的元素中，任取r个（$r\leqslant n$）元素组成一组，而不考虑元素的顺序，叫做从n个不同的元素中，取r个元素的一个组合，所有不同组合的种数，叫做组合数，通常记作C_n^r或$\binom{n}{r}$。

组合与排列的不同之处，在于排列要考虑所取元素的前后顺序，而组合不考虑这种顺序。下面我们来推导组合种数C_n^r的计算公式。

从n个不同元素中任取r个元素，得到一种组合，再对这r个元素进行全排列，排列数为$r!$，因为这$r!$个排列是由一种组合变化来的，所以从n个不同的元素中任取r个元素的排列种数P_n^r是组合种数C_n^r的$r!$倍，即有

$$P_n^r=C_n^r\times r!$$

从而得到组合种数为

$$C_n^r=\frac{P_n^r}{r!}=\frac{n(n-1)\cdots(n-r+1)}{r!}=\frac{n!}{r!(n-r)!} \tag{1.4}$$

组合有一个重要性质，即

$$C_n^r=C_n^{n-r} \tag{1.5}$$

这是因为

$$C_n^{n-r}=\frac{n!}{(n-r)![n-(n-r)]!}=\frac{n!}{(n-r)!\,r!}=C_n^r$$

故当r接近n时，可利用这个性质，把C_n^r化成C_n^{n-r}来计算。

例1.5 已知12株林木中有2株为染病林木，现从中任意抽取4株，问

（1）共有多少种不同的取法？

（2）恰好取得1株为染病林木的取法有多少种？

（3）至少取得1株为染病林木的取法有多少种？

解　因为任取4株林木，不讲顺序，故它们都是属于组合问题。

（1）共有不同取法的种数为

$$C_{12}^{4}=\frac{12\times 11\times 10\times 9}{4!}=495\ （种）$$

（2）恰好取得1株为染病林木，这可看作是从2株染病林木中任取一株，同时从10株无病林木中任取3株，前者有C_2^1种取法，后者有C_{10}^3种取法，合起来即得所求取法的种数为

$$C_2^1\times C_{10}^3=2\times 120=240\ （种）$$

（3）因为12株林木中只有2株染病林木，所以至少取得1株为染病林木，就是恰好取得1株为染病林木和恰好取得2株为染病林木这两种情况，因此所求取法的种数为

$$C_2^1\times C_{10}^3+C_2^2\times C_{10}^2=240+45=285\ （种）$$

本题也可按以下方法计算：从所有取法的种数中，减去未取得染病林木的取法种数，即

$$C_{12}^4-C_{10}^4=495-210=285\ （种）$$

习　　题

1．由数字1，2，3，4，5，6能组成多少个没有重复数字的五位数？

2．由数字0，1，2，3，4，5能组成多少个没有重复数字的五位数？

3．某种产品加工中，需要经过四道工序？问

（1）若各道工序可不分先后，加工顺序共有多少种不同的排法？

（2）若其中有一道工序必须最先开始，加工顺序有多少种不同排法？

（3）若其中有一道工序不能安排在最后，加工顺序有多少种不同的排法？

4．七个不同的字母排成一排，如果某一个字母必须放在中间，共有多少种排法？

5．电话号码由六位数字组成，问最多可以有多少个号码？其中数字都不相同的号码有多少？

6、从100株苗木中任取4株进行检查，有多少种不同的取法？其中某一株恰好被取到的取法有多少种？

7、有三本政治书、五本科技书、四本文艺书，现要从中取二本政治书、三本科技书、三本文艺书，问有多少种不同的取法？

8、某试验小组共有13人，其中男同志8人，女同志5人，现要选出3人准备作介绍，并且3人中至少要有一位女同志，问共有多少种选法？

§2　随机事件、事件的关系及运算

一、随机事件的概念

首先，为了叙述的方便，我们把对随机现象进行观察、测试、实验等统称为随机试验，简称试验。例如投掷一枚硬币，观察其结果是“徽花向上”，还是“币值向上”是随机试验；测量某一片林地树木的胸径，也是随机试验，因为设被测的某株树木的胸径为x，则x一定落

在该林地上树木的最小胸径a和最大胸径b之间，即$a \leqslant x \leqslant b$，但每次测量前不能确定x会取何值。

当然，作为随机试验，应具有下列三个特性：

（1）试验可在相同条件下重复进行；

（2）每次试验的结果不止一个，但事先能够知道会出现哪些可能的结果；

（3）每次试验前，不能确定出现哪个结果。

其次，我们把在随机试验中可能出现，也可能不出现的事情，称为随机事件，简称事件，通常用大写字母A、B、C、…等来表示，例如，在投掷一枚硬币的试验中，出现“徽花向上”可记作A，出现“币值向上”可记作B，它们都是随机事件，又如，从含有二个次品、八个正品的十个同类产品中，任意抽取三个，我们可记A＝{三个都是正品}，B＝{二个正品、一个次品}，C＝{一个正品、二个次品}，A、B、C都是随机事件，又D＝{至少有一个次品}也是随机事件。

在随机试验中，每一个可能出现的结果，都是一个随机事件，它们是这个试验的最简单的随机事件，称为基本事件。而有些随机事件由若干基本事件组成，称为复合事件。例如，在投掷一枚硬币的试验中，“徽花向上”与“币值向上”都是基本事件。若将投掷一枚硬币改为投掷一颗骰子，则“出现1点”，“出现2点”，…，“出现6点”都是基本事件，而“出现偶数点”这一随机事件就是复合事件，它由“出现2点”、“出现4点”、“出现6点”三个基本事件组成。同样，在上述从十个产品中抽取三个产品的试验中，A、B、C都是基本事件，而D是复合事件，它由B和C两个基本事件组成。

在随机试验中必然出现的事情，称为必然事件，记作U；必然不出现的事情，称为不可能事件，记作V，例如，在投掷一颗骰子的试验中，“点数不大于6”是必然事件，而“点数大于6”是不可能事件。同样，在上述从十个产品中抽取三个产品的试验中，“至少一个是正品”是必然事件，而“三个都是次品”是不可能事件。必然事件和不可能事件，本来没有不确定性，即不是随机事件，但为了今后讨论方便，我们把它们作为一种特殊的随机事件来看待。

为了研究随机试验，我们还把所有基本事件的全体组成的集合，称为样本空间（或基本空间），每一个基本事件可以看作是这个集合中的一个元素，称为一个样本点。由于任何一次试验，其结果必然出现于全部基本事件中，故样本空间作为一个事件是必然事件，亦记作U，而不可能事件V是一个空集。例如：

投掷一枚硬币，其样本空间为

U＝{徽花向上，币值向上}；

投掷一颗骰子，其样本空间为

$$U=\{1,\ 2,\ 3,\ 4,\ 5,\ 6\};$$

测量某一片林地树木的胸径，设该林地上树木的最小胸径为a，最大胸径为b，则其基本事件为区间$[a,\ b]$中的任一数值，故样本空间为

$$U=\{x \mid a \leqslant x \leqslant b\}。$$

二、事件的关系及运算

在一个随机试验中，往往有多个随机事件，其中有些事件之间存在着一定的联系。例如

某单位要求一批长度和直径都达到一定标准的木材，若我们把这两个指标都符合的，称为合格品，则在检测（试验）中，我们有："合格品"，"不合格品"，"长度符合"，"长度不符合"，"直径符合"，"直径不符合"，"长度符合但直径不符合"等各个事件。显然，这些事件相互之间是有联系的。下面就来介绍事件之间的几种主要关系及运算，并结合本例予以说明。

（一）事件的包含与相等

如果事件A的发生，必然导致事件B的发生，则称事件B包含事件A，记作$A \subset B$或$B \supset A$。

例如设A表示"直径不符合"，B表示"不合格品"，则有$A \subset B$。

如果$A \subset B$，且$B \subset A$，则称事件A与B相等，记作$A = B$。

（二）事件的和与积

事件A与事件B中至少有一个发生，这样的事件称为事件A与事件B的和事件，记作$A \cup B$。

例如，设A表示"直径不符合"，C表示"长度不符合"，B表示"不合格品"，则有$A \cup C = B$。

事件A与事件B同时发生，这样的事件称为事件A与事件B的积事件，记作$A \cap B$或AB。

例如，设D表示"直径符合"，E表示"长度符合"，F表示"合格品"，则有$D \cap E = F$或$DE = F$。

事件的和与积都可推广到多于两个事件的情形：

事件$A_1, A_2, \cdots, A_n$中至少有一个发生，这样的事件称为事件A_1，A_2，…，A_n的和事件，记作$A_1 \cup A_2 \cup \cdots \cup A_n$，简记为$\bigcup\limits_{i=1}^{n} A_i$。

事件$A_1, A_2, \cdots, A_n$同时发生，这样的事件称为事件A_1，A_2，…，A_n的积事件，记作$A_1 \cap A_2 \cap \cdots \cap A_n$或$A_1 A_2 \cdots A_n$，简记为$\bigcap\limits_{i=1}^{n} A_i$。

（三）事件的互斥与互逆

如果事件A与事件B不可能同时发生，即$AB = V$，则称事件A与事件B是互斥的（或互不相容的）。

例如"直径符合"与"直径不符合"是互斥事件，"直径不符合"与"合格品"也是互斥事件。

对于互斥事件A、B，它们的和事件$A \cup B$，可以记作$A + B$。

如果一组事件A_1，A_2，…，A_n中的任意两个都互斥，则称这组事件为两两互斥。

对于一个两两互斥的事件组A_1，A_2，…，A_n，它们的和事件$A_1 \cup A_2 \cup \cdots \cup A_n$可以记作$A_1 + A_2 + \cdots + A_n$。

如果事件A与事件B同时满足$A \cup B = U$及$AB = V$（即A、B中必有一发生，但不能同时发生），则称事件A与事件B是互逆的（或对立的）。并称A是B的逆事件（或对立事件），同样，B也是A的逆事件。事件A的逆事件记作$\overline{A}$。

例如"直径符合"与"直径不符合"是互逆事件。但"直径符合"与"不合格品"不是

互逆事件。

必须注意，事件A与事件B互斥，并不一定互逆，因为它们只满足了条件$AB=V$，而不一定满足条件$A\cup B=U$。又如在投掷一颗骰子的试验中，“掷得1点”与“掷得偶数点”是互斥的，但它们不是互逆的，而“掷得奇数点”与“掷得偶数点”才是互逆的。

对于任一事件A，显然有

$$\overline{\overline{A}}=A,\qquad A\overline{A}=V,\qquad A+\overline{A}=U$$

（四）事件的差

事件A发生而事件B不发生，这样的事件称为事件A与事件B的差事件，记作$A-B$。例如，设D表示“直径符合”，C表示“长度不符合”，G表示“直径符合但长度不符合”，则有$D-C=G$。

由对立事件可得

$$A-B=A\overline{B},\qquad \overline{A}=U-A$$

为使读者对上述关系及运算有一个直观的理解，我们还可用几何图形表示，如图1—1所示。图中平面上的一个矩形，表示样本空间U，矩形内的每一点表示样本点，两个小圆表示事件A与事件B，阴影部分表示A与B的关系及运算。

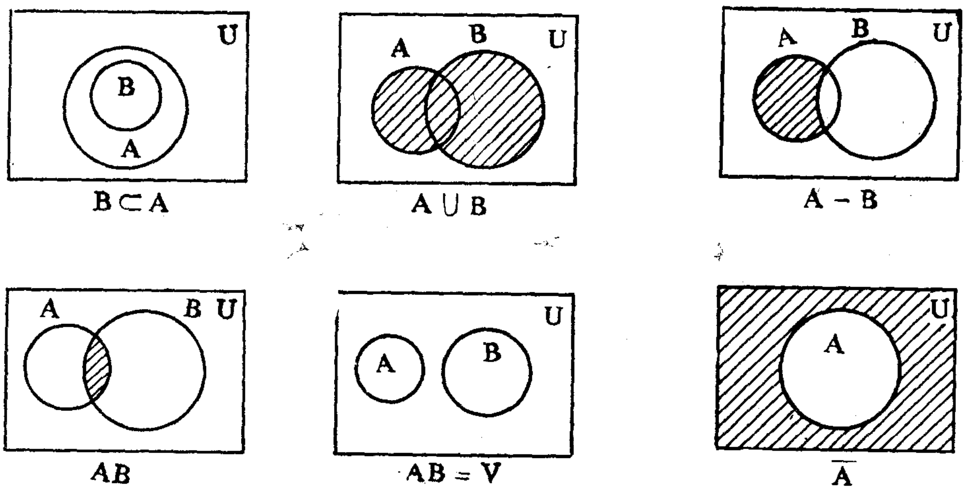

图1—1

从上面的讨论可知，概率论中事件之间的关系及运算，与集合论中集合之间的关系及运算是完全类似的。概率论中的基本事件和样本空间，即相当于集合论中的元素和全集。故一个随机事件，也可看作是全集中的某个元素或某些元素组成的一个小集合。

例1.6　设A，B，C为三个事件，试用事件的运算关系表示下列各事件：

（1）A发生，而B、C都不发生；

（2）A、B、C都发生；

（3）A、B、C都不发生；

（4）A、B、C中恰好有一个发生；

（5）A、B、C中最多有一个发生；

（6）A、B、C中至少有一个发生。

解 （1）{A发生，而B、C都不发生}$=A\bar{B}\bar{C}$；

（2）{A、B、C都发生}$=ABC$；

（3）{A、B、C 都不发生}$=\bar{A}\bar{B}\bar{C}$；

（4）{A、B、C中恰好有一个发生}$=A\bar{B}\bar{C}+\bar{A}B\bar{C}+\bar{A}\bar{B}C$；

（5）{A、B、C中最多有一个发生}$=\bar{A}\bar{B}\bar{C}+A\bar{B}\bar{C}+\bar{A}B\bar{C}+\bar{A}\bar{B}C$；

（6）{A、B、C中至少有一个发生}$=A\cup B\cup C$；或$=A\bar{B}\bar{C}+\bar{A}B\bar{C}+\bar{A}\bar{B}C+AB\bar{C}+A\bar{B}C+\bar{A}BC+ABC$；或$=U-\bar{A}\bar{B}\bar{C}$。

有的事件可以有几种不同的表示方法，如上例中的（6），可根据问题的具体情况选择使用。上例中的（3）与（6）为对立事件，故（6）亦可用（3）来表示，且较简单。应注意（2）与（3）不是对立事件。

§3 随机事件的概率

随机事件在一次试验中，可能发生，也可能不发生，但在大量重复试验中，人们发现它还是具有内在规律的，也就是说，它发生的可能性的大小，还是可以度量的。概率就是表示这种大小的数量化指标。下面我们对概率的基本概念和不同的定义，作一些简单的介绍。

一、概率的统计定义

设事件A在n次重复试验中发生了m次，则$\frac{m}{n}$称为事件A发生的频率。显然，任何随机事件的频率，都是介于0与1之间的一个数。

经验表明，多次重复进行同一试验，当试验次数n很大时，随机事件A发生的频率，总具有一定的稳定性——其数值在某个确定的数值附近摆动，并且一般说来，试验次数越多，摆动的幅度就越小。例如，历史上有人作过成千上万次投掷一枚均匀钱币的试验，表1.1为他们试验的结果。

表1.1

试 验 者	投掷次数n	“正面向上”的发生次数m（即频数）	频率$=\frac{m}{n}$
德莫根	2048	1061	0.5181
蒲 丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

容易看出，随着投掷次数的增多，“正面向上”这一事件发生的频率，就更稳定地在数值0.5附近摆动。

在现实生活中，频率稳定性的例子很多，例如，有人从世界各国人口统计中发现，在人口出生中，男婴出生的频率稳定于$\frac{22}{43}$；又如，某林场在一定条件下，进行落叶松种子的发芽试验，随着试验中种子数量的增加，发现其发芽率逐渐稳定于0.6。一般，我们给出如下定义。

定义1.1 在一个随机试验中，如果事件A出现的频率为$\frac{m}{n}$，且随着试验次数n的增大，

$\frac{m}{n}$逐渐稳定地在某个常数 p 附近摆动，则称常数 p 为事件A的概率，记作

$$P(A) = p$$

上述定义，称为概率的统计定义。其中，数值p就是事件A发生的可能性大小的数量描述。

由概率的统计定义，可推得下列性质：

（1）对任一事件A，有

$$0 \leqslant P(A) \leqslant 1 \tag{1.6}$$

（2）对必然事件U及不可能事件V，有

$$P(U) = 1, \quad P(V) = 0 \tag{1.7}$$

概率的统计定义，直观易懂，且为我们提供了一种近似计算概率的一般方法。实际上，人们常常在许多随机试验中，把大量重复试验所得的频率，来作为概率的近似值。如人口统计和产品检验中的许多指标，就是这样得来的。但这种方法，有时做起来毕竟费时费力，而在某些特殊情况下，可以由问题本身的条件和性质，来直接计算其概率，这就是古典概型的问题。

二、古典概型、概率的古典定义

所谓古典概型，是指试验结果具有下列两个特点的一类随机试验：

（1）基本事件的个数是有限的（有限性）；

（2）每个基本事件发生的可能性相同（等可能性）。

对于古典概型，我们给出如下的概率定义。

定义1.2 设一随机试验是古典概型，基本事件的总数为n，如果事件A包含m个基本事件，则事件A的概率为

$$P(A) = \frac{\text{事件A包含的基本事件个数}}{\text{基本事件的总数}} = \frac{m}{n} \tag{1.8}$$

这叫做概率的古典定义，也是古典概型的概率计算公式。

例1.7 从0，1，2，…，9十个数字中，任取一个数字，求取得的数不超过 2 的概率。

解 从十个数字中任取一个数字，共有十种取法，故该试验基本事件的总数n＝10，以A表示“取得的数不超过 2”这一事件，则A所包含的基本事件个数m＝3，故所求概率为

$$P(A) = \frac{3}{10} = 0.3$$

例1.8 在五位数字组成的电话号码中，求五个数字都不相同的概率。

解 从十个数字中，可重复地任取五个，组成一个五位数的电话号码，共有10^5个，故基本事件总数$n = 10^5$，而五个数字都不相同的五位数电话号码，共有P_{10}^5个，即$m = P_{10}^5$，仍以A表示这一事件，则所求概率为

$$P(A) = \frac{P_{10}^5}{10^5} = \frac{10 \times 9 \times 8 \times 7 \times 6}{10^5} = 0.3024$$

例1.9 已知12株林木中，有 2 株为染病林木，现从中任意抽取 4 株进行观察，求恰好取得 1 株为染病林木和至少取得 1 株为染病林木的概率。

解 从12株林木中任意抽取 4 株，共有 $C_{12}^{4}=495$种取法，即基本事件总数$n=495$。设A表示“恰好取得 1 株为染病林木”这一事件，又设B表示“至少取得 1 株为染病林木”这一事件。则由§1例1.5中的计算结果可知，事件A所包含的基本事件数$m_1=C_2^1\times C_{10}^3=240$，事件B所包含的基本事件数$m_2=C_2^1\times C_{10}^3+C_2^2\times C_{10}^2=285$，故所求的概率分别为

$$P(A)=\frac{240}{495}=0.485;$$

$$P(B)=\frac{285}{495}=0.576$$

例1.10 设有一批产品，总数为N个，其中有M个次品，现从中任取n个（$n\leqslant N-M$），求其中恰好有m个（$m\leqslant M$）次品的概率。

解 该试验的基本事件总数为C_N^n，设A表示“取出的n个产品中恰好有m个次品”这一事件，则A所包含的基本事件数为$C_M^m\times C_{N-M}^{n-m}$，故所求概率为

$$P(A)=\frac{C_M^m\times C_{N-M}^{n-m}}{C_N^n} \tag{1.9}$$

这个概率称为超几何分布概率，它在林业统计中应用较多。

从以上的几个例子可以看到，求解古典概型的问题，关键在于正确算出基本事件总数及所求事件所包含的基本事件数。在求解中常用到一些排列、组合的有关知识，故了解这方面的知识是必要的，我们在§1中已作了介绍。

*三、几何概型、概率的几何定义

概率的古典定义，是在假设试验的基本事件的个数是有限的，且各个基本事件发生的可能性是相同的情况下给出的。对于试验结果有无穷多个的情形，概率的古典定义显然就不适用了。为了克服这个局限性，我们仍以等可能性为基础，把古典定义作必要的推广，使得推广后的定义，能够适用于基本事件的个数为无穷多个的情形。

设在平面上有某一区域G，在G内又有一小区域g。（图1—2），现在区域G内随机地投掷一点M，假设M点落在G内的任何位置都是等可能的，这也就是说，落在G内任何一部分的概率，只与这部分的面积大小有关，而与其位置和形状无关。于是M点落在小区域g内的概率，就很自然地可以定义为g的面积与G的面积之比。即

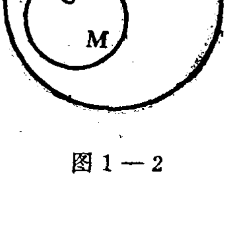

图1—2