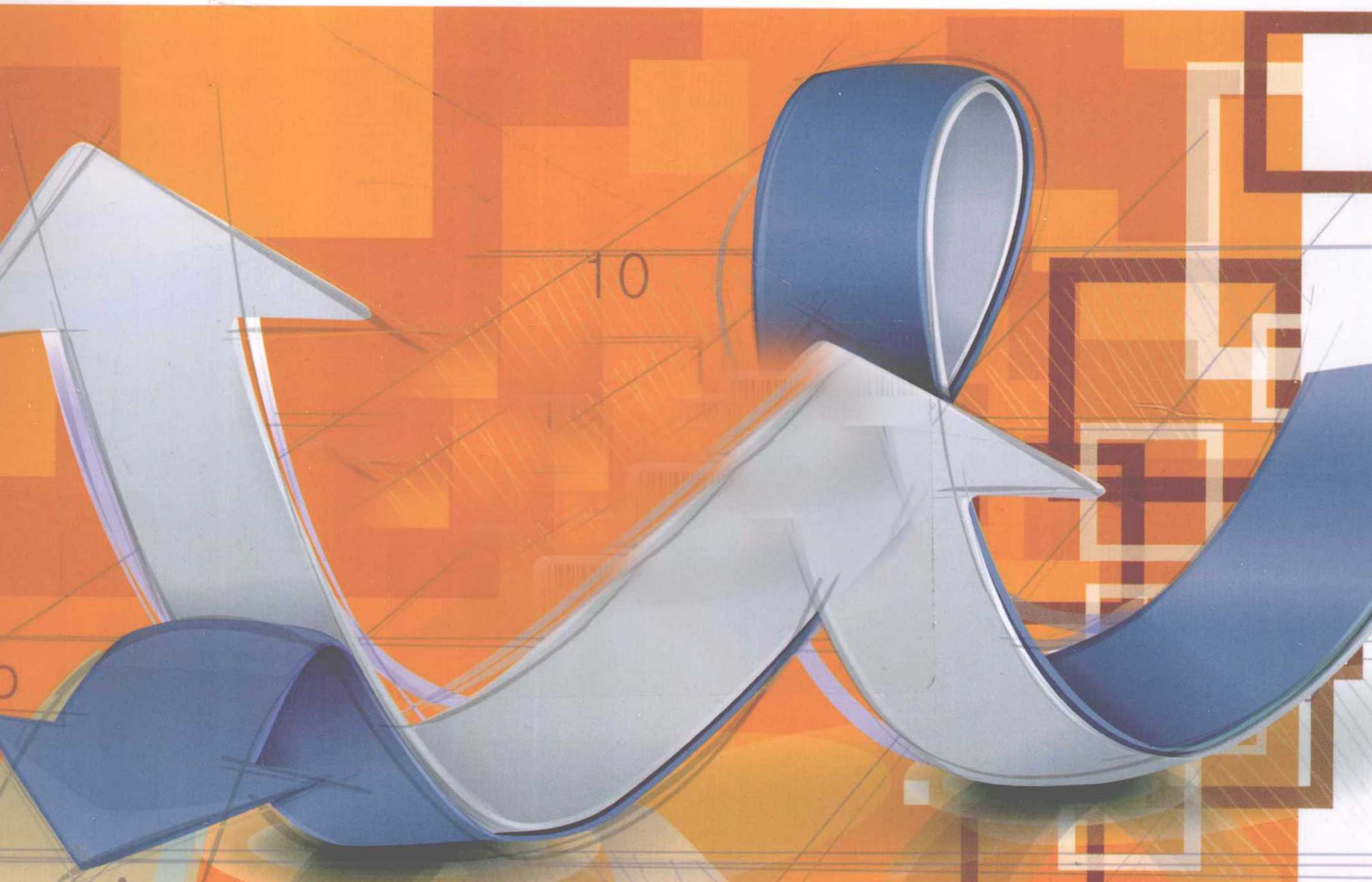
普通高等教育“十二五”规划教材
全国高等医药院校规划教材

医药数理统计

第4版

马志庆　周介南　主编



科学出版社

普通高等教育“十二五”规划教材
全国高等医药院校规划教材

医药数理统计

第4版

马志庆　周介南　主编
周永治　主审

科学出版社
北京

《医药数理统计》(第4版)编写人员

主　编　马志庆　周介南

副主编　杨松涛　尹立群　钱微微　王世钦
郑洁钢　胡灵芝　赵文峰　陈丽君

主　审　周永治

编　委　(按姓氏笔画排序)

马志庆	山东中医药大学	郑洁钢	湖南中医药大学
王世钦	甘肃中医学院	赵　莹	上海中医药大学
韦　杰	贵阳中医学院	赵文峰	河南中医学院
尹立群	天津中医药大学	赵聪俐	天津中医药大学
冯天义	宁夏医科大学	郝　涛	山东中医药大学
吕佳萍	南京中医药大学	胡灵芝	陕西中医学院
许华萍	浙江中医药大学	钟小芳	安徽中医学院
李上达	湖北中医药大学	莫修明	北京城市学院
杨松涛	安徽中医学院	钱微微	浙江中医药大学
沈宗山	云南中医学院	黄　翔	安徽中医学院
沈晓婧	南京中医药大学	黄爱武	湖南中医药大学
陈丽君	湖北中医药大学	覃　洁	广西中医药大学
周　菲	甘肃中医学院	傅　爽	山东中医药大学
周介南	南京中医药大学	魏国强	福建中医药大学

第 4 版编写说明

《医药高等数学》、《医药数理统计》、《医药数学实验》是全国 19 所中医院校联合编写的科学出版社 2001 年 4 月出版的数学系列教材. 相继,2004 年 9 月《医药高等数学》、《医药数理统计》第 2 版出版,同时《医药高等数学学习辅导》、《医药数理统计学习辅导》(辅导教材比配套的理论教材迟一版)第 1 版也相继出版;2009 年 5 月又再版了第 3 版的理论教材与第 2 版的辅导教材。该套教材自 2001 年出版以来,发行面广，发行量大，在中医院校受到广大师生的欢迎. 为了进一步提高该套教材的质量,努力打造为"十二五"精品教材，编写组对前几版教材进行分析、总结,根据科学出版社的"五性"、"三基"精品教材的要求认真地进行了修改、补充，编写了第 4 版的《医药高等数学》、《医药数理统计》与第 3 版的《医药高等数学学习辅导》、《医药数理统计学习辅导》. 该套教材将更适应医药院校的医药类、管理类、信息类、人文类等专业的需要，定于 2012 年 6 月由科学出版社正式出版.

《医药数理统计》是应用数理统计方法研究医药、生物等领域中的随机现象的一门学科. 全书共 9 章,包括概率论基本知识、统计方法的原理与步骤、正交试验设计等内容,考虑到中医院校数理统计课时数普遍较少的具体情况,为此对上版的内容进行了紧缩,删去第十章的内容,有的章节则打了" * "号。本教材每章后配有习题，书后有解答. 另外还有配套辅导教材《医药数理统计学习辅导》。本教材授课需 50～60 学时，不同专业可根据需要对加" * "号的内容有所选择，课时偏少的专业还可重点讲授前六章的内容.

参加本版教材编写的有以下院校:天津中医药大学、山东中医药大学、甘肃中医学院、北京城市学院、河南中医学院、陕西中医学院、宁夏医科大学、安徽中医学院、南京中医药大学、上海中医药大学、浙江中医药大学、福建中医药大学、湖北中医药大学、湖南中医药大学、广西中医药大学、贵阳中医学院、云南中医学院.

本教材编写过程中得到许多同行专家的关心与支持,在此一并表示感谢.

本教材尚有不少不足之处,恳请读者与同行批评指正.

编　者

2012 年 4 月

目　录

第一章

事件与概率

数理统计方法是以概率论为理论基础，通过一定的设计来收集数据和进行整理分析，以部分资料推断总体的一种方法，用它去研究大量随机现象的规律性. 由于概率和随机事件是联系在一起的，故事件和概率都是数理统计中最基本的概念.

本章将介绍随机事件、事件的概率及其运算.

§1-1　随机事件及其运算

1-1.1　随机事件

当我们多次观察自然现象和社会现象后，会发现许多事情在一定条件下必然会发生或者必然不会发生. 例如，纯净的水在一个大气压下，温度是0℃时必然结冰，在20℃时必然不会结冰，在100℃时必然沸腾，在80℃时必然不会沸腾；又如，把锌放入稀硫酸一定会逸出氢气，而永动机存在是不可能的. 这种完全可以预言其结果的现象是一种确定性现象，叫**必然现象**.

另一类现象，在一定条件下，不可能事前完全准确地预言其结果，也就是它有多种可能发生的结果，是一种不确定性现象，这类现象称为**偶然现象**. 例如，抛起一枚硬币落地时究竟哪一面朝上？从一批针剂中抽取一支来检验，其结果可能是正品，也可能是次品，在抽取之前是无法肯定的. 偶然现象也称为**随机现象**.

对各种现象的“观察”称为试验，对随机现象的“观察”就称为**随机试验**. 随机试验具有下列特征：

(1) 在相同条件下，可以重复进行；

(2) 各次试验结果不一定相同，而且每次试验之前不能预先判断哪一个结果发生；

(3) 所有可能的试验结果是预先可以明确的，并且在每一次试验中必有其中一个结果出现.

对某种现象的“观察”而得到的结果就称为**事件**. 在一定条件下，试验结果中必然出现的事件称为**必然事件**，记为Ω. 例如，{纯净的水在一个大气压下，加热到100℃沸腾}$=\Omega$，{物体会热胀冷缩}$=\Omega$. 反之，那种在一定条件下试验结果中必然不出现的事件称为**不可能事件**，记为$\varnothing$. 例如，{$x^2+1=0$有实数解}$=\varnothing$，{人的寿命可达200岁}$=\varnothing$等.

随机试验观察的是随机现象，在一定条件下，试验结果中可能出现，也可能不出现的事件称为**随机事件**，简称**事件**. 随机事件一般用大写字母A,B,C等表示. 例如，投掷一个硬币，这个随机试验中有两个事件$A=${出正面}和$B=${出反面}. 必然事件与不可能事件可以说不是随机事件，但为了研究方便起见，把必然事件与不可能事件作为随机事件的两个极端来统一处理.

1-1.2 事件之间的关系及运算

在各种现象中，往往要求同时考察几个随机事件及它们之间的联系，下面就来讨论事件的关系及运算.

一、包含

设有事件 A 及 B，如果事件 A 发生必然导致事件 B 发生，则称事件 A 包含于事件 B 或事件 B 包含事件 A，并记为 $A\subset B$ 或 $B\supset A$. 例如，$A=\{$乙肝患者$\}$，$B=\{$乙肝病毒携带者$\}$，则有 $A\subset B$.

二、等价

若事件 A 包含事件 B，事件 B 也包含事件 A，即 $A\supset B$ 且 $B\supset A$，就称事件 A 与 B 等价(或称相等)，记作 $A=B$.

三、并事件

若事件 $C=\{A$ 或 B 中至少有一个发生$\}$，则称 C 为 A，B 两事件的并事件，记为 $C=A+B$.

例如，

$$A_1=\{\text{甲份血清含乙肝病毒}\},\quad A_2=\{\text{乙份血清含乙肝病毒}\}$$
$$A=\{\text{甲、乙两份混合血清含乙肝病毒}\}$$

则有 $A=A_1+A_2$.

n 个事件的并事件记为 $A=\sum\limits_{i=1}^{n}A_i$.

四、交事件

若事件 $C=\{A$ 与 B 同时发生$\}$，则称 C 为 A，B 两事件的交事件，记为 $C=AB$.

例如，

$$A_1=\{\text{甲份血清不含乙肝病毒}\},\quad A_2=\{\text{乙份血清不含乙肝病毒}\}$$
$$A=\{\text{甲、乙两份混合血清不含乙肝病毒}\}$$

则有 $A=A_1A_2$.

n 个事件的交事件记为 $A=\prod\limits_{n=1}^{n}A_i$.

五、互不相容事件

若事件 A 与事件 B 不能同时发生，则称 A 与 B 为**互不相容事件**，记作 $AB=\varnothing$. 互不相容事件也称为**互斥事件**. n 个事件互斥，是指它们两两互斥.

例如，三人做体检，$A=\{$三人正常$\}$，$B=\{$只一人不正常$\}$，A 与 B 是互斥事件.

若 n 个互斥事件的并事件是必然事件，即 $A_iA_j=\varnothing(1\leqslant i<j\leqslant n)$ 且 $\sum\limits_{i=1}^{n}A_i=\Omega$，则称这 n 个事件构成**互斥完备群**.

例如，治疗某种疾病，其疗效标准分为 4 个等级：痊愈、显效、微效和无效. 那么，就一次试验(治疗一个患者的结果)而言，事件$\{$痊愈$\}$、$\{$显效$\}$、$\{$微效$\}$、$\{$无效$\}$是互斥事件，而且这 4 个事件构成互斥完备群.

六、对立事件

若在任一次试验中，事件 A 与事件 B 二者必有一个发生且仅有一个发生，亦即 A，B 同时满足 $A+B=\Omega$ 及 $AB=\varnothing$ 两个条件，也就是互斥完备群仅由两事件 A 与 B 构成，则称**事件 A 与事件 B 对立**，如果治疗某种疾病，只考虑有效和无效两个等级，那么事件{有效}与{无效}就是对立事件，**或事件 B 是事件 A 的对立事件**，当然**事件 A 也是事件 B 的对立事件**. A 的对立事件记作 $\overline{A}$，那么就有 $B=\overline{A}$ 或 $\overline{A}=B$.

不难理解，对立事件必为互斥事件，而互斥事件不一定是对立事件.

例如，投掷一枚骰子，事件{出 1 点}与{出 2 点}互斥，但不对立，而事件{出偶数点}与{出奇数点}对立且互斥.

事件之间的这些关系，读者可以通过熟知的韦恩图作直观理解，图 1-1 给出几种常见情况.

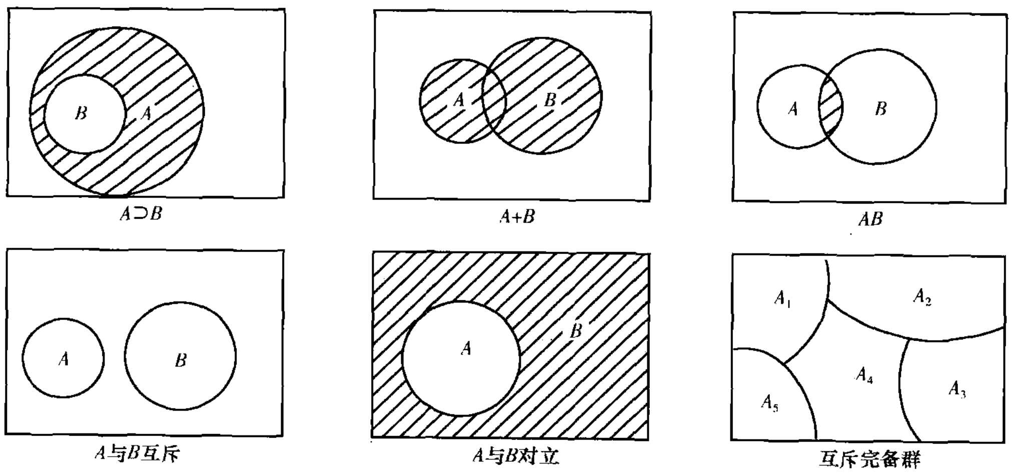


图 1-1　韦恩图

由事件的定义可知，事件之间的关系与运算同集合的关系与运算是一致的，因此在进行事件运算时，经常遇到下述定律：

设 A，B，C 三事件，则有

交换律：$A+B=B+A$；$AB=BA$.

结合律：$(A+B)+C=A+(B+C)$；$(AB)C=A(BC)$.

等幂律：$A+A=A$；$AA=A$.

分配律：$A(B+C)=AB+AC$；$(A+B)(A+C)=A+BC$.

补余律：$A+\overline{A}=\Omega$；$A\overline{A}=\varnothing$.

同一律：$A+\varnothing=A$；$A+\Omega=\Omega$.

零律：$A\Omega=A$；$A\varnothing=\varnothing$.

德摩根律：$\overline{A+B}=\overline{A}\,\overline{B}$；$\overline{AB}=\overline{A}+\overline{B}$.

一个比较复杂的事件常常包含若干个简单事件，把一个复杂事件划成几个简单事件的并、交或混合形式以及找出构成互斥完备群的全部事件是必要的，因为这是讨论事件间关系进而施行运算的重要途径.

例 1　依次检查黄芩、黄连、人参三种中药材质量作为一次试验. 令 $A=${黄芩合格}，$B=${黄连合格}，$C=${人参合格}. 试用 A，B，C 三个事件表示下列在一次试验中出现的事件：

(1) 只有黄芩质量合格；

(2) 只有一种中药质量合格；

(3) 三种中药质量都不合格；

(4) 至少有一种中药质量合格；

(5) 构成互斥完备群的全部事件.

解 令 $\overline{A}$={黄芩质量不合格}，$\overline{B}$={黄连质量不合格}，$\overline{C}$={人参质量不合格}.

(1) {只有黄芩质量合格}={黄芩合格且黄连、人参不合格}$=A\overline{B}\overline{C}$.

(2) 因为

$$\{只有黄芩质量合格\}=A\overline{B}\overline{C}$$
$$\{只有黄连质量合格\}=\overline{A}B\overline{C}$$
$$\{只有人参质量合格\}=\overline{A}\overline{B}C$$

所以

$$\{只有一种中药质量合格\}=A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C$$

(3) {三种中药质量都不合格}$=\overline{A}\overline{B}\overline{C}$.

(4)

$$\{至少有一种中药质量合格\}=A+B+C$$

或者

$$\{至少有一种中药质量合格\}=A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C+\overline{A}BC+A\overline{B}C+AB\overline{C}+ABC$$

(5) 构成互斥完备群的全部事件有 8 个，即

$$\Omega=\overline{A}\overline{B}\overline{C}+A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C+\overline{A}BC+A\overline{B}C+AB\overline{C}+ABC$$

例 2 事件 A_K 表示某射手第 K 次($K=1,2,3$)击中目标，试叙述下列事件的具体含义：A_1+A_2；$A_1+A_2+A_3$；$A_1A_2A_3$；$\overline{A}_2$；$\overline{A_1+A_2}$；$\overline{A}_1\ \overline{A}_2$；$\overline{A}_2+\overline{A}_3$；$\overline{A_2A_3}$；$A_1A_2A_3+A_1A_2\ \overline{A}_3+A_1\ \overline{A}_2\ A_3+\overline{A}_1A_2A_3$；$A_1\overline{A}_2\overline{A}_3+\overline{A}_1A_2\overline{A}_3+\overline{A}_1\overline{A}_2A_3$.

解 A_1+A_2：前两次中至少有一次击中目标；

$A_1+A_2+A_3$：三次射击中至少有一次击中目标；

$A_1A_2A_3$：三次射击都击中了目标；

$\overline{A}_2$：第二次射击未击中目标；

$\overline{A_1+A_2}=\overline{A}_1\overline{A}_2$：前两次射击均未击中目标；

$\overline{A}_2+\overline{A}_3=\overline{A_2A_3}$：后两次射击中至少有一次未击中目标；

$A_1A_2A_3+A_1A_2\overline{A}_3+A_1\overline{A}_2A_3+\overline{A}_1A_2A_3$：三次射击中至少有两次击中目标；

$A_1\overline{A}_2\overline{A}_3+\overline{A}_1A_2\overline{A}_3+\overline{A}_1\overline{A}_2A_3$：三次射击中仅有一次击中目标.

注意，把一个事件化成若干个事件的并事件时，必须明白是有交并（不互斥），还是无交并（互斥），这对后面的计算尤为重要. 例如，例 1 的问题(2)表示成 3 个事件的无交并，问题(4)可表示成 3 个事件的有交并，或者 7 个事件的无交并，问题(5)构成互斥完备群的 8 个事件当然是无交并的.

§1-2 事件的概率

通俗地说，所谓概率是某一随机事件在试验中发生的可能性大小的数值表示，通常用$P(A)$来表示事件 A 的概率. $P(A)$越大，说明事件 A 发生的可能性越大. 下面给出概率论中关于概率的两个定义，从中可以了解概率的特性和计算方法.

1-2.1 概率的统计定义

随机事件是一种可能发生，也可能不发生的事件，看起来似乎没什么规律可循，当我们在同

一条件下进行大量重复试验时，就会显现某种规律性. 若进行条件相同的 n 次试验，事件 A 出现 m 次，则称 m 为事件 A 的**频数**，称比值 m/n 为事件 A 的频率，记为

$$f(A)=\frac{m}{n} \tag{1-1}$$

显然，事件 A 的频率是通过特定的试验获得的，每做 n 次试验，所得到的频率可以各不相同，但经验证明，在同一条件下进行多次重复试验时，事件出现的频率会在某一常数附近左右摆动，这种性质叫做**频率的稳定性**.

在历史上，这种频率的稳定性是在人口统计方面最先被注意到的. 例如，世界上一些国家通过多年观察，发现男婴的出生率稳定在 22/43 附近，而女婴的出生率稳定在 21/43 附近.

再如，著名的投币试验. 表 1-1 列出试验记录. 容易看出，投掷次数逐渐增多时，{出现正面}这个事件的频率 m/n 总是在 0.5 这个数附近摆动而逐渐稳定于 0.5.

表 1-1

试验者	投掷次数 n	正面次数 m	频率 m/n
德摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

由此可见，频率的稳定性充分说明随机事件发生的可能性大小是事件本身固有的一种客观属性，并为我们衡量一个随机试验中随机事件发生的可能性提供了客观基础.

概率的统计定义　在条件相同的 n 次试验中，事件 A 发生 m 次，如果加大 n 时，A 的频率 m/n 逐渐稳定在一个常数 p 附近，就把这个常数 p 称为事件 A 的概率，记为 $P(A)=p$. 在此定义下有

$$0\leqslant P(A)\leqslant 1,\quad P(\Omega)=1,\quad P(\varnothing)=0 \tag{1-2}$$

概率的统计定义实际上给出了一个近似地计算随机事件概率的方法，即当试验次数 n 足够大时，一个事件的频率与概率应充分接近，所以用事件的频率作为概率的近似值. 在医药学中，这种估计经常用到. 需要注意的是：不要把频率和概率相混淆. 频率是已经进行的试验结果，其数值随着试验次数的不同而变化，具有偶然性；而概率是一种客观存在，是个确定的数值，具有必然性.

1-2.2　概率的古典定义

有些事件的概率不用进行大量重复试验也能确定，如表 1-1 中的事件，因为硬币是比较均匀的，可以认为投掷一次时，只会出现正面和反面，具有等可能性，谁也没有优先出现的理由. 而每次只能出现其中的一个结果，所以出现正面的可能性大小是 1/2，即事件 $A=$ {出现正面}的概率，$P(A)=1/2$. 这就是说，可以用划分等可能事件的个数方法求得事件的概率.

为此，先给出**等概率基本事件组**的定义.

定义 1　如果一组事件 $A_1,A_2,\cdots,A_n$ 满足以下条件，则称该事件组为等概率基本事件组：

(1) N 个事件，每一个事件出现的概率是相等的(等可能性)；

(2) 任一次试验中，N 个事件中只能出现 N 个事件中的一个(互不相容性)；

(3) 任一次试验中，N 个事件中必然会出现一个(完备性).

等概率基本事件组记为 $\Omega=\{A_1,A_2,\cdots,A_n\}$.

定义 2　如果一组等概率基本事件 $A_1,A_2,\cdots,A_n$ 中，事件 A 包含 $m(m\leqslant n)$ 个等概率基本

事件，则事件 A 的概率

$$P(A)=\frac{A\text{ 所包含的基本事件个数}}{\text{等概率基本事件的总个数}}=\frac{m}{n} \tag{1-3}$$

且有

$$0\leqslant P(A)\leqslant 1,\quad P(\Omega)=1,\quad P(\varnothing)=0$$

这种用等概率基本事件的个数来计算概率的方法称为古典概率定义，它是概率论发展初期的主要研究对象. 古典概率的大部分问题都能形象化地归结为抽球问题.

例 1 在盒子中有 6 个相同的球，分别标号码为 1，2，…，6，从中任取一球，求此球的号码为偶数的概率.

解 $\Omega=\{1,2,3,4,5,6\}$，基本事件总数 $n=6$. 令 $A=\{$所取球的号码为偶数$\}$，显然，$A=\{2\}+\{4\}+\{6\}$，所以 A 中含有 $m=3$ 个基本事件，从而

$$P(A)=\frac{m}{n}=\frac{3}{6}=\frac{1}{2}$$

例 2 某厂生产 50 件产品，其中，有 3 件次品，求

(1) 一次取一件，取得次品的概率；

(2) 一次取 5 件，5 件中有 2 件是次品的概率.

解 (1) 50 件产品中取一件，其可能结果有 50 个基本事件(每件产品被取到的可能性相等)，即 $n=50$.

设 $A=\{$取到次品$\}$，则 A 包含 3 个基本事件，即 $m=3$. 由古典定义得

$$P(A)=\frac{m}{n}=\frac{3}{50}=0.06$$

(2) 50 件产品中任取 5 件，其可能结果有 C_{50}^5 个基本事件(C_{50}^5 种机会均等的取法)，即 $n=C_{50}^5$.

设 $B=\{5$ 件中有 2 件次品$\}$，则事件 B 包含的基本事件数 $m=C_3^2C_{47}^3$，故所求概率

$$P(B)=\frac{C_3^2C_{47}^3}{C_{50}^5}=\frac{9}{392}=0.023$$

例 3 袋中有 2 个白球和 8 个黑球，现在无放回地一个个抽出来，求第 k 次抽到的是白球的概率($1\leqslant k\leqslant 10$).

解法一 把 10 个球当成是有区别的，即设想把它们按 1，2，…，10 进行编号，若将抽出的球依次排成一排，则全部可能的结果相当于把 10 个元素进行全排列，即全部基本事件数为 10!.

第 k 次抽到白球，即排在第 k 号位置上的那一个白球，只能在 2 个白球中取得，故有 2 种抽法. 而另外 9 次抽的球可在余下的 9 个中任取，故有 9! 种抽法. 以事件{第 k 次抽到白球}包含的基本事件数为 $2\times 9!$，故第 k 次抽到白球的概率 $p=\frac{2\times 9!}{10!}=\frac{2}{10}$.

解法二 把 2 个白球看成一样，8 个黑球看成一样，把抽出的球仍依次放在 10 个位置上，由于白球看成一样，黑球看成一样，所以当白球位置选好，其他位置必放黑球，故总的排法即总的基本事件数为 C_{10}^2，而事件{第 k 次抽到白球}所包含的基本事件数为 $C_{10-1}^{2-1}=C_9^1$(因为 2 个位置中已有 1 个位置，即第 k 号位置固定放了白球)，所以

$$p=\frac{C_9^1}{C_{10}^2}=\frac{2}{10}$$

两种解法结果一样，抽到白球的概率 $p=2/10$，与次数无关，这正好说明广泛应用于生产和生活中的抽签方法是公平合理的，先抽后抽都一样，机会均等.

需要说明的是，无论是概率的统计定义，还是古典定义，都在概率计算中起一定的作用，但又有着各自的局限性. 古典概率是以试验的所有可能结果只有有限个且具有等可能性为基础，

实际上这种条件很难满足. 至于统计概率,则要求试验次数 n 充分大,并以事件频率的稳定值近似地作为该事件的概率,这里的 n 大到什么程度,稳定值是什么都是不确切的. 因此,人们需要对概率有个严格的定义,使之能适用于一般的随机试验. 经过人们不断地探索和总结,终于在 1933 年,由前苏联数学家科尔莫戈罗夫提出了概率公理化结构,明确定义了基本概念,使概率论成为严谨的数学分支. 至于公理化体系的内容,有兴趣的读者可参阅概率论专著,此处不再赘述.

§ 1-3 概率的运算

把复杂事件的概率分解成简单事件的概率来计算,可以借助概率的运算法则.

1-3.1 加法定理

一、互斥事件加法定理

若事件 A 与 B 互斥,则

$$P(A+B)=P(A)+P(B) \tag{1-4}$$

证 设试验的全部结果包含 n 个基本事件,而事件 A 包含其中 m_1 个基本事件. 事件 B 包含其中的 m_2 个基本事件. 由于 A 与 B 互斥,因而它们各包含的基本事件应该完全不同,所以事件 $A+B$ 所包含的基本事件数为 m_1+m_2,按古典定义有

$$P(A+B)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}=P(A)+P(B)$$

这个定理给出了计算两个互斥事件的并事件的概率的方法. 它不难推广到 n 个互斥事件的情形

$$P(A_1+A_2+\cdots+A_n)=P(A_1)+P(A_2)+\cdots+P(A_n) \tag{1-5}$$

简记为

$$P\left(\sum_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$$

如果 n 个事件构成互斥完备群,就有

$$P\left(\sum_{i=1}^{n}A_i\right)=P(\Omega)=1.$$

特别地,若 A 与 B 对立,则

$$P(A+B)=P(A)+P(B)=1$$

从而

$$P(A)=1-P(B)$$

即

$$P(A)=1-P(\overline{A}) \tag{1-6}$$

例 1 8 个乒乓球队中,有两个强队,将 8 个球队任意分为两组(每组 4 个队)进行比赛,求这两个强队被分在一个组内的概率是多少?

解法一 把 8 个球队任意分成两组(每组 4 个队)的分法有 C_8^4 种. 设 $A=\{$两个强队分在同一组$\}$,$A_1=\{$两个强队在第一组$\}$,$A_2=\{$两个强队在第二组$\}$,则 $A=A_1+A_2$ 且 A_1 与 A_2 互斥. 又

$$P(A_1)=\frac{C_2^2\cdot C_6^2}{C_8^4}=\frac{C_6^2}{C_8^4},\quad P(A_2)=\frac{C_2^2\cdot C_6^2}{C_8^4}=\frac{C_6^2}{C_8^4}$$

于是

$$P(A)=P(A_1)+P(A_2)=\frac{2C_6^2}{C_8^4}=\frac{3}{7}$$

解法二 设 $A=\{$两个强队分在同一组$\}$，则 $\overline{A}=\{$两个组中各有一个强队$\}$，

$$P(\overline{A})=\frac{C_2^1\cdot C_6^3}{C_8^4}=\frac{4}{7}$$

$$P(A)=1-P(\overline{A})=1-\frac{4}{7}=\frac{3}{7}$$

如果不是互斥两事件，而是任意两事件，那么它的并事件的概率则应由下面的定理计算.

二、一般加法定理

对于任意两事件 A 与 B 有

$$P(A+B)=P(A)+P(B)-P(AB) \tag{1-7}$$

证 事件 $A+B$ 可以表示成三个互斥事件 $A\overline{B}$，$\overline{A}B$，AB 的并事件，即

$$A+B=A\overline{B}+\overline{A}B+AB$$

按互斥事件加法定理得

$$P(A+B)=P(A\overline{B})+P(\overline{A}B)+P(AB) \tag{1-8}$$

因为 $A=AB+A\overline{B}$，而 AB 与 $A\overline{B}$ 互斥，所以

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

由此得

$$P(A\overline{B})=P(A)-P(AB)$$

同理可得

$$P(\overline{A}B)=P(B)-P(AB)$$

把最后两式代入(1-8)式得

$$P(A+B)=P(A)+P(B)-P(AB)$$

不难把这一定理推广到有限个事件的情形.

例如，A_1，A_2，A_3 为任意三个随机事件，则有

$$P(A_1+A_2+A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3)$$

一般地，对于任意 n 个事件 A_1，A_2，$\cdots$，A_n，可以用归纳法证得

$$\begin{aligned}P(A_1+A_2+\cdots+A_n)&=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_j)\\&\quad+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)\end{aligned} \tag{1-9}$$

例 2 某药厂自动生产线上有两个料仓，在一天内甲料仓装满需清理的概率为 0.15，乙料仓装满需清理的概率为 0.25，两料仓同时装满需清理的概率为 0.08，问一天至少有一个料仓装满需清理的概率是多少？

解 令 $A=\{$甲料仓装满需清理$\}$，$B=\{$乙料仓装满需清理$\}$，则有

$$AB=\{\text{两料仓同时装满需清理}\}$$

$A+B=\{$至少有一个料仓装满需清理$\}=P(A)+P(B)-P(AB)=0.15+0.25-0.08=0.32$

即一天至少有一个料仓装满需清理的概率是 0.32.

1-3.2 条件概率、概率的乘法定理

一、条件概率

定义 1 在事件 B 已发生的条件下，事件 A 发生的概率称为 A 的条件概率，记为 $P(A|B)$，

读作在条件 B 下事件 A 的概率.

条件概率当然具有普通概率的性质,即

$$0 \leqslant P(A|B) \leqslant 1, \quad P(\Omega|B)=1, \quad P(\varnothing|B)=0$$

相对地,$P(A)$可以称为无条件概率. 在一般情况下,无条件概率 $P(A)$ 与条件概率 $P(A|B)$ 是不相等的.

例如,在 1-2.2 例 3 中我们算得第 k 次抽到白球的概率是 2/10,与先后次序无关. 现在考虑一下,如果已知第一个人抽得了白球,那么第二个人抽得白球的概率是多少?

设 $A=\{$第一个人抽得白球$\}$,$B=\{$第二个人抽得白球$\}$,于是上面的问题就可写成 $P(B|A)$. 由于 A 已发生且 $P(A)=2/10$,故剩下的 9 个球中只有一个白球,所以 $P(A)=1/9$.

显然,就事件{抽到白球}来说,无条件概率和条件概率不相等,亦即 $P(A)\neq P(A|B)$.

二、独立事件

在某些情况下,若无条件概率和条件概率相等,即 $P(A)=P(A|B)$. 这说明事件 A 的概率与事件 B 出现与否无关,也就是说,A 与 B 是相互独立的.

定义 2 若 $P(A)=P(A|B)$,就称事件 A 与 B 相互独立. 由对称性,此时必有 $P(B)=P(B|A)$.

如果 A 与 B 独立,易知 A 与$\overline{B}$,$\overline{A}$与 B,$\overline{A}$与$\overline{B}$也独立,四对事件中有一对独立,则其余三对也独立.

例 3 为研究某种方剂对风热外感证的疗效,随机选取 400 名患者,有的服药,有的不服药,经过一段时间后,有的有疗效,有的无效,结果见表 1-2. 试判断用此方剂治疗风热外感证是否有效.

表 1-2

	B(服药)	$\overline{B}$(未服药)	合计
A(有效)	127	190	317
$\overline{A}$ (无效)	33	50	83
合计	160	240	400

解 如果事件 A(有效)与事件 B(服药)独立,就说明有效与服药无关,方剂未起作用.

$P(A)=317/400=0.793$,$P(A|B)=127/160=0.794$,可见 $P(A)\approx P(A|B)$,两者几乎相等. 由定义 2,认为事件 A 与 B 相互独立,即该方剂对风热外感证没有确实疗效.

需要注意的是,如果单看条件概率,该方剂对风热外感证的有效率高达 0.794,效果似乎不错. 但一经比较,发现无条件概率已高达 0.793,当然不能认为方剂确实有效. 这说明判断一种医学方案的客观效果,往往不能只凭单方面的数据下结论,而应当进行必要的对照.

三、乘法定理

一般乘法定理 对任意两事件 A 与 B 有

$$P(AB)=P(A)P(B|A)=P(B)P(A|B) \tag{1-10}$$

证 设实验的全部结果包含有 n 个基本事件,而事件 A,B,AB 分别包含其中的 m_1 个,m_2 个,m 个基本事件,显然这 m 个基本事件就是 A 所包含的 m_1 个和 B 所包含的 m_2 个基本事件中共有的基本事件. 按古典定义有 $P(A)=m_1/n,P(B)=m_2/n,P(AB)=m/n$.

在事件 A 已经发生的前提下,事件 B 所包含的基本事件就是事件 AB 所包含的那些基本事件,有且仅有 m 个,所以

$$P(B|A)=\frac{m}{m_1}=\frac{m/n}{m_1/n}=\frac{P(AB)}{P(A)}$$

由此得

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

同理可得

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

对于前一个式子要求 $P(A)\neq 0$，对于后一个式子要求 $P(B)\neq 0$.

设 A,B,C 是三个事件，如果它们满足等式

$$P(AB)=P(A)P(B)$$
$$P(AC)=P(A)P(C)$$
$$P(BC)=P(B)P(C)$$
$$P(ABC)=P(A)P(B)P(C)$$

中的前三个，则称事件 A,B,C 两两独立，如果满足上述四个等式，则称事件 A,B,C 是互相独立的.

一般地，设 $A_1,A_2,\cdots,A_n$ 是任意 n 个事件，如果对任意的 $k(1\leqslant k\leqslant n)$ 与任意的 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$，满足等式 $P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$，则称事件 $A_1,A_2,\cdots,A_n$ 为相互独立.

独立事件乘法定理 若事件 A 与 B 独立，则

$$P(AB)=P(A)P(B) \tag{1-11}$$

因为此时

$$P(B|A)=P(B),\quad P(A|B)=P(A)$$

将其代入(1-10)式，即得

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

这个定理其逆亦真.

对于 n 个独立事件，容易推出

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n) \tag{1-12}$$

还应指出，实际应用中，事件的独立性常常不是根据定义而是根据实际意义来作出判断的.

例 4 若每人血清中有肝炎病毒的概率为 0.4%，今混合 100 人的血清，求混合血清无肝炎病毒的概率.

解 设 $A_i=\{$第 i 个人血清中有病毒$\}$，则 $\overline{A}_i=\{$第 i 个人血清中无病毒$\}$.

$$P(A_i)=0.004,\quad P(\overline{A}_i)=1-P(A_i)=0.996$$

因为 100 个事件 $\overline{A}_1,\overline{A}_2,\cdots,\overline{A}_{100}$ 独立，所以混合血清无病毒的概率为

$$P(\text{混合血清无病毒})=P(\overline{A}_1\overline{A}_2\cdots\overline{A}_{100})=P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_{100})=0.996^{100}=0.67$$

应用概率的加法和乘法定理时，必须注意到事件的互斥性和独立性，并且要注意到如下命题成立：具有非零概率的两事件，互斥就不独立，独立就不互斥.

例 5 由例 4 知，当混合的份数减少时，混合血清无病毒的概率就会加大. 如果要求混合血清无病毒的概率在 95%以上，那么混合的份数 n 应当不超过多少？

解 因为 $0.996^n=0.95$，所以

$$n=\frac{\lg 0.95}{\lg 0.996}\approx 12.8$$

故应不超过 12 份. 也就是说，只要用不超过 12 份的血清混合，所得血清无病毒的可靠程度就高达 95%以上.

例 6 某药厂针剂车间灌装一批注射液需用 4 道工序. 已知由割锯(安瓿割口)时掉入玻璃屑而造成废品的概率为 0.5%，由于安瓿洗涤不洁而造成废品的概率为 0.2%，由于灌装药时污染而造成废品的概率为 0.1%，由于封口不严而造成废品的概率为 0.8%，试求产品