YIYAOYINGYONGSHULITONGJI

医药应用数理统计

主编 杨继泰 副主编 周全喜 徐 锐

河南大学出版社

前 言

随着我国医药卫生事业的发展，数理统计已成为药理学、毒理学、药物动力学以及临床医学和预防医学的一门基础理论课．同时，它在疾病的分类、防治、预报预测、计量诊断、处方选择，以及制药的工艺改进、质量控制、药物评价、生物检定及试验设计等方面都已得到广泛、深入的应用，并取得显著的成效．

为了深化教育改革，贯彻理论联系实际的原则，更好地提高教学质量，促进医药事业的发展，根据药学专业数理统计方法教学大纲的要求及学校实行“三一制”教学计划和立足于培养实用型人才的精神，我们特组织编写了《医药应用数理统计》一书．该书教学时数约60学时，使用时可根据本校具体情况适当增减．附有星号的内容，超出教学大纲要求，可以不讲．每章都有小结和习题，书末附有答案．

编写时我们突出了以下几点：(1)突出体现医药应用的特点，尽量根据现代医药研究的需要选编典型的例题和习题．(2)突出基本概念和方法．对于基本理论，一般作了必要的推导和论证．对于不便证明的，也尽量给予直观解释，尽力说明原理．(3)突出简明扼要、深入浅出、通俗易懂、容易掌握、方法适用的特点．因此，本书可作为医药院校师生以及医药卫生工作者和科研人员学习参考用书．

本书内容共分八章，包括随机事件的概率，随机变量及其分布，随机变量的数字特征，随机抽样与抽样分布，假设检验，方差分析，相关与回归分析，正交试验设计与分析．各章的执笔人分别为徐锐(第一章)、杨保华(第二章)、唐宗贤(第三章)、洪仁杰(第四章一、二节)、杨继泰(第五、六、七章)、周全喜(第四章三、四、五节，第八章)．在编写出版过程中曾得到张大录教授、姚智敏副教授的大力支持，在此一并表示衷心的感谢．

由于编者水平有限，经验不足，书中难免有不当之处，恳请读者指正．

编　者

1994年5月于开封

目　录

第一章　随机事件与概率

第一节　随机事件的概率

一、随机现象

在生产实践、科学实验和日常生活中，人们观察到的现象有各种不同的类型：一类是在相同条件下重复进行试验，每次结果不尽相同，对一次试验来说，究竟出现哪一种结果，事前不能确定．这种现象称作随机现象．如：在同一条件下生产的一批针剂中，有的是合格品，有的是不合格品；某种疾病的患者，服剂量相同的一种药物，有的痊愈，有的无效．另一类是在相同条件下重复进行试验，每次结果完全相同，对一次试验来说，其结果可以准确预料，这就是必然现象．如：在标准大气压下，将水加热到100℃，水必然沸腾等．

二、随机试验与随机事件

对随机现象进行实验或观察称为随机试验，简称为试验．随机试验的各种可能结果称为事件．

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．如：{明天下雨}、{这批将制成的药丸的合格率不低于99%}等都是随机事件．随机事件通常用字母A，B，C，…表示．

在一定条件下，试验结果中必然出现的事件，称为必然事件．如：{导体通电后发热}就是必然事件．必然事件通常用字母U表示．

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．如：{从黄连里提出青霉素}即是不可能事件．不可能事件通常用字母V表示．

三、频率与概率

1.统计概率

随机事件在一定条件下，可能发生，也可能不发生．这是随机事件所具有的偶然性的一面，其实它也具有规律性的一面．随机事件的规律性可通过大量试验而观察到．

如果在一组相同的条件下，重复进行n次试验，事件A发生了m次，则称$\frac{m}{n}$为事件A的频率，记为

$$f_n(A)=\frac{m}{n}. \tag{1.1.1}$$

例1　投掷一枚硬币，把其中一面定为正面，结果如下：

试验者	掷币次数	正面出现次数	频　率
蒲　丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

从表上看，在一次掷币中，出现正、反哪一面，事先是不能预料的．但是，随着试验次数的大量增加就可以看出，出现正面的次数接近掷币次数的50％．

例2 考虑某种子的发芽率．从一大批种子中任取10批种子做发芽试验，其结果如下：

种子粒数	2	5	10	70	130	310	700	1500	2000	3000
发芽粒数	2	4	9	60	116	282	639	1339	1806	2715
发芽率	1	0.8	0.9	0.857	0.892	0.910	0.913	0.893	0.903	0.905

从上表看出，发芽率在0.9附近摆动．

从上面几个例子中我们可以看到，当试验的次数n增加时，事件A的频率总是表现出一定的稳定性——其频率值总会在某一个确定的数值附近摆动，并且，试验次数越多，事件A的频率就越接近于这个数值．这个稳定值正好表达了随机事件A出现的可能性的大小．

定义1 若随着试验次数n的充分增加，事件A的频率$f_n(A)$稳定地在某一个常数p附近摆动，则称p为事件A的统计概率，记为

$$P(A)=p. \tag{1.1.2}$$

由此定义显然有，任何事件A满足

$$0\leqslant P(A)\leqslant 1. \tag{1.1.3}$$

必然事件U的概率为1，不可能事件V的概率为0，即

$$P(U)=1,\quad P(V)=0.$$

2.古典概型与古典概率计算公式

从事件的统计概率定义中可以看出，通过大量的重复试验是可以找到事件的概率的．但是，在实际工作中做大量的试验是不容易的，有时甚至是不可能的或者是不允许的．在某些特殊情况下，不需要通过大量的重复试验，只须利用问题本身的某些性质，即能得到事件的概率．

例3 有编号为1，2，…，10的10件同种产品．现从中任取1件，很容易想到{取得偶数号码的产品}的可能性是5/10，{取得号码大于2的产品}的可能性是8/10，而{取得号码为3的产品}的可能性是1/10等．

例4 在一个口袋中装有大小相同的12个球，其中7个红球，5个白球．从中任取一个，{取得的为白球}的可能性有多大？人们同样很容易得到结论：这个可能性为5/12．

以上两个例子具有两个共同点：

(1) 试验中所有的最基本的可能结果为有限个．如：例3共有10件同类的产品，每次只取1件，所以可划分出10个最基本的可能结果，每次只能出现其中的一个．同样的道理，例4有12个大小相同的球，每次取1个，就可划分出12个最基本的可能结果．

(2) 每一个基本的可能结果，在一次试验中发生的可能性均等．如：例3中哪一件产品都不比其它产品特殊，被取到的可能性也就和其它产品相等．例4同理．

具有上述两个特点的随机试验称为古典概型．

通常我们把表示试验中一个最基本可能结果的事件称为基本事件．

在古典概型中，如果基本事件的总数为 n，事件 A 由 m 个基本事件组成（即 A 包含 m 个基本事件），则事件 A 的概率由下式计算：

$$P(A)=\frac{A\text{包含的基本事件数}}{\text{基本事件的总数}}=\frac{m}{n}. \tag{1.1.4}$$

例 5 瓶中装有 50 片同类的药片，其中有 3 片次品．今从瓶中任取 5 片，求所取 5 片中有 2 片是次品的概率．

解： 由于是从 50 片中任取 5 片，所以每 5 片为一组就是一个基本事件．基本事件的总数为 C_{50}^5，而且每个基本事件发生的可能性均等，每次只有一个基本事件发生．其中 5 片中有 2 片次品的情况占 $C_3^2C_{47}^3$ 个，故所求概率为：

$$p=\frac{C_3^2C_{47}^3}{C_{50}^5}=\frac{9}{392}=0.023.$$

例 6 一部为四分册的文集，按任意的次序放到书架上去．各分册自右向左或自左向右恰成1，2，3，4 的顺序的概率为多少？

解： 基本事件就是四分册自左到右的一种排列方式，有多少种不同的排列方式就有多少个基本事件，所以基本事件的总数为 $P_4^4=4!$．而自右向左或自左向右恰为 1，2，3，4 顺序的占其中的两种排法，故所求概率为

$$p=\frac{2}{P_4^4}=\frac{2}{4!}=\frac{1}{12}.$$

练　习　1.1

1．某医院眼科某年做“金针拨内障”手术，资料如下：

月　份	1	2	3	4	5	6	7	8	9	10	11	12
手术眼数（只）	2	4	7	9	9	7	7	6	6	17	17	19
成功眼数（只）	2	3	6	7	7	4	5	5	4	16	17	18

求一次手术（做一只眼）成功的概率（近似值）．

2．电话号码由 5 个数字组成，每个数字可以是 0，1，2，…，9 中任意一个数．求电话号码由完全不同数字组成的概率．

3．从一付扑克牌（52 张）中任取 4 张，求 4 张牌的花色各不相同的概率．

4．全班共有 36 人，其中男生、女生各一半．现将其随机分成两组，每组各 18人．求两组中男、女生各一半的概率．

第二节　事件之间的关系

一、包含关系

设 A 与 B 为两个事件．若事件 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 $B\supseteq A$．

若事件 A 包含 B，同时事件 B 又包含 A，即 $A\supseteq B$ 且 $B\supseteq A$，则称事件 A 与事件 B 相等，记作 $A=B$．

例如，$B=$｛取出的球为白色或为红色｝，$A=$｛取出的球为红色｝，则$B\supseteq A$.

二、事件的和（并）

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（并），记作$A+B$或$A\cup B$.

一般地，若事件$A_1,A_2,\cdots,A_n$中至少有一个发生，则这一事件被称为事件$A_1,A_2,\cdots,A_n$的和，记作$A_1+A_2+\cdots+A_n$或$\sum\limits_{i=1}^{n}A$.

例如，$A=$｛取出的球为白色或为红色｝，$B=$｛取出的为红球｝，$C=$｛取出的为白球｝，则$A=B+C$.

三、事件的积（交）

事件A与事件B同时发生，这一事件称为事件A与事件B的积（交），记作AB或$A\cdot B$或$A\cap B$.

一般地，若事件$A_1,A_2,\cdots,A_n$同时发生，则这一事件被称为该n个事件的积，记作$A_1A_2\cdots A_n$或$\prod\limits_{i=1}^{n}A_i$.

四、事件之间的互不相容关系

若事件A与事件B不能同时发生，即$AB=V$，则称这两个事件互不相容.

例如，$B=$｛任取一个为红球｝，$C=$｛任取一个为白球｝，B与C就是互不相容的.

五、事件之间的互逆关系

若事件A与事件B满足条件：$A+B=U$，$AB=V$，则称A与B为互逆的，A和B互称为对方的逆事件，A的逆事件记作$\bar{A}$.

六、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B的差，记作$A-B$.

如：$B=$｛取出的球为红球或为白球｝，

$C=$｛取出的球为红球或为黄球｝，

$A=$｛取出的球为白球｝，

则
$$A=B-C.$$

事件之间的关系，完全类似于集合之间的关系，如下表所示：

符　号	事　件	集　合
U	必然事件	全集
V	不可能事件	空集
e	基本事件	元素
A	事件	子集
$\bar{A}$	A的逆事件	A的余集
$A\subseteq B$	A发生必然导致B发生	A是B的子集
$A=B$	A与B相等	A与B相等
$A+B(A\cup B)$	事件｛A与B至少有一个发生｝	A与B的并集
$AB(A\cap B)$	事件｛A与B同时发生｝	A与B的交集
$A-B$	事件｛A发生而B不发生｝	A与B的差集
$AB=V$	事件A与事件B互不相容	A与B没有公共元素

事件之间的关系也可用韦恩图作直观理解，如图 1 - 1．

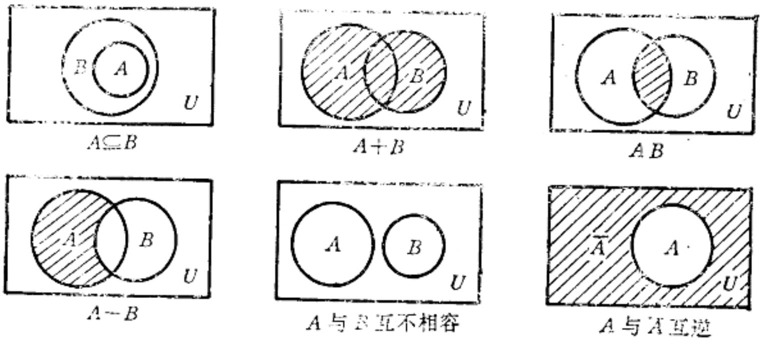

图 1 - 1

既然事件是一种集合，那么事件之间的运算律便服从集合的运算律，主要性质如下：

(1) $A+B=B+A$；　　(2) $A+(B+C)=(A+B)+C$；

(3) $A+A=A$；　　(4) $A+\bar{A}=U$；

(5) $A+U=U$；　　(6) $A+V=A$；

(7) $AB=BA$；　　(8) $(AB)C=A(BC)$；

(9) $A(B+C)=AB+AC$；　　(10) $A+BC=(A+B)(A+C)$；

(11) $\overline{A+B}=\bar{A}\bar{B}$；　　(12) $\overline{AB}=\bar{A}+\bar{B}$．

例 1　设 A，B，C 为三个事件，试将下列事件用 A，B，C 表示：

(1) A 出现，B 和 C 不出现；　　(2) 三个事件都出现；

(3) 三个事件都不出现；　　(4) 三个事件中至少出现一个；

(5) 不多于两个事件出现．

解：(1) $A\bar{B}\bar{C}=A-(B+C)$；　　(2) ABC；

(3) $\bar{A}\bar{B}\bar{C}=\overline{A+B+C}$；　　(4) $A+B+C$；

(5) $\overline{ABC}=\bar{A}+\bar{B}+\bar{C}$．

练　习　1.2

1. 何谓两个事件互不相容？何谓两个事件互逆？两个事件具有互不相容关系，这两个事件互逆吗？

2. 若事件 A，B，C 有关系 $C=A-B$，那么 C 与 A 有何关系？C 与 B 有何关系？

3. 对三人作舌诊，设 $A=\{$三人正常$\}$，$B=\{$至少一人不正常$\}$，$C=\{$只有一人正常$\}$，$D=\{$只有一人不正常$\}$．指出这四个事件中的互不相容事件、互逆事件．又，$A+D$，BD 各表示什么意思？

第三节　概率的加法法则

定理 1　若事件 A 和 B 是两个互不相容的事件，则

$$P(A+B)=P(A)+P(B). \tag{1.3.1}$$

证：设一个随机试验的基本事件共有 n 个，而事件 A 包含 m_1 个基本事件，事件 B 包

含 m_2 个基本事件．由于 A 与 B 互不相容，所以 A 与 B 包含的基本事件各不相同，和事件 $A+B$ 就包含 m_1+m_2 个基本事件．按古典概率的概率公式，有

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

推论 1 若事件 $A_1,A_2,\cdots,A_n$ 两两互不相容，则

$$P\left(\sum_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}P(A_k). \tag{1.3.2}$$

推论 2 若事件 A 与事件 $\overline{A}$ 是互逆事件，则

$$P(A)=1-P(\overline{A}). \tag{1.3.3}$$

例 1 在 20 片外观一样的药片中，有黄连素 15 片，穿心莲 5 片．今从中任取 3 片，求至少有 1 片穿心莲的概率．

解： 设 $B=\{$ 3 片中至少有 1 片穿心莲 $\}$，$A_k=\{$ 3 片中恰有 k 片穿心莲 $\}$，$k=0,1,2,3$，则

$$P(A_0)=\frac{C_{15}^3}{C_{20}^3}=\frac{91}{228},\qquad P(A_1)=\frac{C_{15}^2C_5^1}{C_{20}^3}=\frac{105}{228},$$

$$P(A_2)=\frac{C_{15}^1C_5^2}{C_{20}^3}=\frac{30}{228},\qquad P(A_3)=\frac{C_5^3}{C_{20}^3}=\frac{2}{228}.$$

法一：因为 A_1,A_2,A_3 互不相容且 $B=A_1+A_2+A_3$，所以

$$\begin{aligned}P(B)&=P(A_1+A_2+A_3)=P(A_1)+P(A_2)+P(A_3)\\&=\frac{105}{228}+\frac{30}{228}+\frac{2}{228}=\frac{137}{228}.\end{aligned}$$

法二：$\overline{B}=\{$3 片中无穿心莲$\}=A_0$，从而

$$P(B)=1-P(\overline{B})=1-\frac{91}{228}=\frac{137}{228}.$$

定理 2 对于任意两个事件 A 与 B，有

$$P(A+B)=P(A)+P(B)-P(AB). \tag{1.3.4}$$

证： 因为 $A+B=AB+A\overline{B}+\overline{A}B$，而 $AB,A\overline{B},\overline{A}B$ 三事件两两互不相容，所以

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

而

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

所以

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

同理

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

于是

$$\begin{aligned}P(A+B)&=P(AB)+P(A)-P(AB)+P(B)-P(AB)\\&=P(A)+P(B)-P(AB).\end{aligned}$$

定理 1 是定理 2 的特例，因为当 A 和 B 互不相容时，$P(AB)=0$．

推论 对于任意的 n 个事件 $A_1,A_2,\cdots,A_n$，有

$$\begin{aligned}P\left(\sum_{k=1}^{n}A_k\right)=&\sum_{k=1}^{n}P(A_k)-\sum_{1\leq i<j\leq n}P(A_iA_j)+\sum_{1\leq i<j<k\leq n}P(A_iA_jA_k)\\&+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n).\end{aligned}$$

证明从略．

例 2 一个口袋内有 2 个红球，3 个白球，4 个黑球．每次摸取 1 个，有放回地取两次，求取得的球中无红球或无黑球的概率．

解：设 $A=\{$两次取球无红球$\}$，$B=\{$两次取球无黑球$\}$，则

$$P(A+B)=P(A)+P(B)-P(AB)=\frac{7^2}{9^2}+\frac{5^2}{9^2}-\frac{3^2}{9^2}=\frac{65}{81}.$$

第四节 条件概率与乘法法则

一、条件概率

在实际问题中，经常要同时考虑几个事件发生的概率问题，其中一类重要的问题是：在同一随机试验下，有 A 和 B 两个事件，当需要计算事件 A 发生的概率时，已经得到了事件 B 发生的信息，有时事件 B 的发生对于事件 A 的发生有一定的影响，这就是所谓的条件概率．

定义 1 在事件 B 发生的条件下，事件 A 发生的概率叫做条件概率，记作 $P(A|B)$．

例 1 有 100 件圆柱形零件，其中 95 件直径合格，97 件长度合格，两个指标都合格的 94 件．今从这 100 件中任取 1 件．求：

(1) 取出的这件为两项指标都合格的概率；

(2) 已知取出的为长度合格，它的直径也合格的概率．

解：设 $A=\{$直径合格$\}$，$B=\{$长度合格$\}$．

(1) $P(AB)=\dfrac{94}{100}=0.94.$

(2) 因为已知取出的长度合格，即在 B 已经发生的条件下，再考虑直径是否合格，这时只有 97 种可能情形．而这 97 种可能情形中，直径也合格者只有 94 件，所以

$$P(A|B)=\frac{94}{97}.$$

经过如下简单换算，可得到

$$P(A|B)=\frac{94}{97}=\frac{94/100}{97/100}=\frac{P(AB)}{P(B)}.$$

一般地，条件概率 $P(A|B)$ 与事件 B 和事件 AB 的概率有如下关系：

$$P(A|B)=\frac{P(AB)}{P(B)}.\qquad (P(B)>0) \tag{1.4.1}$$

二、乘法法则

由公式(1.4.1)可得

$$P(AB)=P(A)P(B|A) \quad 或 \quad P(AB)=P(B)P(A|B).$$

定理1 若 $A_1,A_2,\cdots,A_n$ 为 $n(n\geqslant 2)$ 个事件且 $P(A_1A_2\cdots A_n)>0$，则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

证：因为 $P(A_1A_2\cdots A_n)>0$，所以有

$$P(A_1)\geqslant P(A_1A_2)\geqslant\cdots\geqslant P(A_1A_2\cdots A_n)>0,$$

$$P(A_1A_2\cdots A_n)=P(A_1)\cdot\frac{P(A_1A_2)}{P(A_1)}\cdot\frac{P(A_1A_2A_3)}{P(A_1A_2)}\cdots\frac{P(A_1A_2\cdots A_n)}{P(A_1A_2\cdots A_{n-1})}.$$

在上式各项中用(1.4.1)式，可得

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

例2 某药检所从送检的 10 件药品中先后抽检 2 件. 如果 10 件中有 3 件次品，求：

(1) 第一次检得次品后，第二次检得次品的概率；

(2) 两次都检得次品的概率.

解： 设 $A_1=\{$第一次检得次品$\}$，$A_2=\{$第二次检得次品$\}$，则

(1) $P(A_2|A_1)=\dfrac{P(A_1A_2)}{P(A_1)}=\dfrac{2}{9}$；

(2) $P(A_1A_2)=P(A_1)P(A_2|A_1)=\dfrac{3}{10}\cdot\dfrac{2}{9}=\dfrac{1}{15}$.

三、事件的独立性

从条件概率的意义可知，条件概率 $P(A|B)$ 和 $P(A)$ 一般情况下不相等. 也就是说，事件 B 的发生通常对事件 A 的发生可能性有影响. 但是，在有些情形下，事件 B 的发生对事件 A 的发生并不影响（或者影响甚微，可忽略不计）. 因此，事件 A 与事件 B 之间有无影响可以通过条件概率来研究.

定义 2 若事件 B 的发生对事件 A 的发生没有影响，即

$$P(A|B)=P(A), \tag{1.4.2}$$

则称事件 B 对事件 A 独立.

两个事件的独立总是相互的. 因为若 $P(A|B)=P(A)$，则有

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

定理 2 事件 A 与 B 独立的充要条件是它们积的概率等于各自的概率之积，即

$$P(AB)=P(A)P(B). \tag{1.4.3}$$

证： 必要性 因为 A 与 B 独立，所以 $P(A|B)=P(A)$，从而

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

充分性 因为有

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

所以 $P(A)=P(A|B)$，即 A 与 B 独立.

推论 1 若事件 A 与 B 相互独立，则事件 A 与 $\overline{B}$，$\overline{A}$ 与 B，$\overline{A}$ 与 $\overline{B}$ 也相互独立.

推论 2 若 $A_1,A_2,\cdots,A_n$ 是 n 个相互独立的事件，则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n). \tag{1.4.4}$$

例 3 为了研究某种药品对某种病的疗效，随机选取 400 名患者进行观察. 经过一段时间后，情况记录如下：

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

此药品对这种病的疗效如何？

解： 如果事件 A(好转)与事件 B(服药)独立，就说明此药没有什么作用.

$$P(A|B)=\frac{127}{160}=0.794,\quad P(A)=\frac{317}{400}=0.793.$$

所以 $P(A)\approx P(A|B)$. 两者几乎相等，事件 A 与事件 B 几乎独立. 这说明，此药对这种病没有什么疗效.

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

解： 设 $A_k=\{$第 k 个人血清中有肝炎病毒$\}$，$\bar{A}_k=\{$第 k 个人血清中无肝炎病毒$\}$，则

$$P(A_k)=0.004,\quad P(\bar{A}_k)=1-P(A_k)=1-0.004=0.996.$$

显然 $\bar{A}_1,\bar{A}_2,\cdots,\bar{A}_{100}$ 是相互独立的，所以混和血清无肝炎病毒的概率为

$$P(\bar{A}_1\bar{A}_2\cdots\bar{A}_{100})=P(\bar{A}_1)P(\bar{A}_2)\cdots P(\bar{A}_{100})=0.996^{100}\approx0.67.$$

从上例可知，当混和的份数减少时，混和血清无病毒的可能性就会加大.

若要求混合血清无肝炎病毒的概率为 0.95 以上，那么混合份数不能超过多少份呢？

设最多允许 n 份. 因为 $0.996^n=0.95$，所以

$$\ln0.996^n=n\ln0.996=\ln0.95,$$

从而

$$n=\frac{\ln0.95}{\ln0.996}\approx12.8.$$

故当不超过 12 份时，混和血清无肝炎病毒的可靠程度就可达 95%，这一计算体现了可靠性的思想.

练 习 1.4

1. 某药房有包装相同的六味地黄丸 100 盒，其中 5 盒为去年产品，95 盒为今年产品. 现随机发出 4 盒，求：

(1) 有 1 盒或 2 盒是陈药的概率；

(2) 有陈药的概率.

2. 在某医院住院病人中，有哮喘病者占 20%，有脾胃病者占 65 %. 从他们中随机抽 1 人，求下列事件的概率：

(1) 这人至少有这两种病中的一种

(2) 这人既无哮喘病又无脾胃病.

3. 一个盒子中有 4 只坏晶体管和 6 只好晶体管. 在其中取二次，每次取一只作不放回抽样，发现第一只是好的，求第二只也是好的概率.

4. 一个盒里有 5 个乒乓球，其中 3 个旧球，2 个新球. 每次取一个，有放回地取两次，求下列事件的概率：

(1) 两次都取到新球；

(2) 第一次取到新球，第二次取到旧球.

(3) 至少有一次取到新球.

5. 甲、乙两个反应罐在一小时内需要工人照顾的概率分别为 0.1 和 0.2 . 求在一小时内：

(1) 甲、乙两罐都需要照顾的概率；

(2) 甲、乙两罐都不需要照顾的概率；

(3) 有一个罐需要照顾而另一个罐不需要照顾的概率.

第五节　全概率公式与贝叶斯公式

一、全概率公式

如果直接计算一个复杂事件的概率遇到困难，可以把该事件进行分解，化成若干互不相容事件的和，再联合使用加法定理、乘法定理来解决.

例1　一个口袋内有5个球，其中3个红球，2个白球. 每次任取1个，无放回地取两次，求第二次取到红球的概率.

解：设 $A=\{$第一次取到红球$\}$，$B=\{$第二次取到红球$\}$. 第二次取球与第一次取球的结果有一定的关系. 对于第二次取到红球而言，有两种情形：(1) 第一次取到的是白球，第二次取到红球；(2) 第一次取到红球，第二次取到红球. 所以 $B=AB+\bar{A}B$，而事件 AB 与 $\bar{A}B$ 互不相容，从而

$$\begin{aligned}P(B)&=P(AB+\bar{A}B)=P(AB)+P(\bar{A}B)\\&=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})\\&=\frac{3}{5}\cdot\frac{2}{4}+\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{5}.\end{aligned}$$

从上例可看到，将 B 分解为 AB 和 $\bar{A}B$，也就是把 B 这个较复杂的事件的概率问题分解转化为 $AB,\bar{A}B$ 的概率问题. 一般地来讲有下述定理：

定理1（全概率公式）　如果事件 B 能且只能与 n 个互不相容的事件组 $A_1, A_2, \cdots, A_n$ 之一同时发生，即

$$B=B(A_1+A_2+\cdots+A_n)=BA_1+BA_2+\cdots+BA_n,$$

则 B 发生的概率为

$$P(B)=\sum_{k=1}^{n}P(A_k)P(B|A_k). \tag{1.5.1}$$

证：由于在 $A_1,A_2,\cdots,A_n$ 中任意两个事件 A_i 与 A_j 互不相容，所以 BA_i 与 BA_j 也互不相容. 由加法定理得

$$P(B)=P\left(\sum_{k=1}^{n}A_kB\right)=\sum_{k=1}^{n}P(A_kB).$$

由乘法定理可得 $P(A_kB)=P(A_k)P(B|A_k)$，所以

$$P(B)=\sum_{k=1}^{n}P(A_k)P(B|A_k).$$

例2　某药房的某种药品由甲、乙、丙三药厂提供，其中甲厂生产的占1/2，乙、丙两厂各占1/4. 已知甲、乙两厂的次品率都为2%，丙厂的次品率为4%. 现从药房中任取一份这种药品，求取到的为次品的概率.

解：设 $B=\{$取得的是次品$\}$，$A_1=\{$是由甲厂生产的$\}$，$A_2=\{$是由乙厂生产的$\}$，$A_3=\{$是由丙厂生产的$\}$，则事件 A_1,A_2,A_3 是互不相容的，

$$P(A_1)=\frac{1}{2},\quad P(A_2)=\frac{1}{4},\quad P(A_3)=\frac{1}{4},$$

$$P(B|A_1)=0.02,\quad P(B|A_2)=0.02,\quad P(B|A_3)=0.04.$$

由题给条件可知，B 能且只能与 A_1,A_2,A_3 之一同时发生．由全概率公式，得

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$

$$=0.02\times\frac{1}{2}+0.02\times\frac{1}{4}+0.04\times\frac{1}{4}=0.0225.$$

二、贝叶斯公式（逆概率公式）

条件与例 2 相同，现在把问题改成：已知拿到的药品是次品，问该次品为乙厂生产的概率．这就是说，在拿到的产品为次品的条件下，求这次品是由乙厂生产的条件概率，即

$$P(A_2|B)=\frac{P(A_2B)}{P(B)}=\frac{P(A_2)P(B|A_2)}{\sum_{k=1}^{3}P(A_k)P(B|A_k)}$$

$$=\frac{0.005}{0.025}=0.2.$$

一般地，有以下定理：

定理 2 如果事件 B 能且只能与 n 个互不相容的事件组 $A_1,A_2,\cdots,A_n$ 之一同时发生，则在事件 B 发生的条件下，事件 A_k 发生的条件概率为

$$P(A_k|B)=\frac{P(A_k)P(B|A_k)}{\sum_{k=1}^{n}P(A_k)P(B|A_k)}. \tag{1.5.2}$$

证明从略．

公式(1.5.2)解决的问题与公式(1.5.1)相反，所以称其为逆概率公式，也称为贝叶斯公式．

例 3 用某种检验方法检查癌症．根据临床记录，癌症患者施行此项检查结果为阳性的概率为 95%，非癌症患者此项检查的结果为阴性的概率为 90%．根据以往的统计，某地区癌症的发病率为 0.0005．若用此法在该地区检查癌症，效果如何？

解： 检查效果的好坏决定于判断的准确程度，即需要求出检查结果为阳性者患癌症的概率．设 $A=\{$检验结果为阳性$\}$，$B=\{$是癌症患者$\}$，则

$$P(A|B)=0.95,\quad P(A|\bar{B})=0.10,$$

$$P(B)=0.0005,\quad P(\bar{B})=0.9995.$$

由贝叶斯公式得

$$P(B|A)=\frac{P(B)P(A|B)}{P(B)P(A|B)+P(\bar{B})P(A|\bar{B})}$$

$$=\frac{0.0005\times0.95}{0.0005\times0.95+0.10\times0.9995}=0.0047.$$

即此检验法的正确性很小．

***三、贝叶斯公式在医疗诊断中的应用**

例 4 某病人开始上腹部疼痛后转至下腹部疼痛，伴以恶心、呕吐、腹泻，入院时体温

39℃，脉搏 120 次/分，全腹肌紧张，有压痛，白血球 19350．据此疑为阑尾炎，为采取合适的治疗，需鉴别以下三种情况：(1)慢性阑尾炎，(2)急性阑尾炎，(3)阑尾穿孔．

解：设 $A_1=\{$慢性阑尾炎$\}$，$A_2=\{$急性阑尾炎$\}$，$A_3=\{$阑尾穿孔$\}$，B 表示患者的症候．如果能求出具有症候 B 患病 $A_k(k=1,2,3)$ 的概率

$$P(A_k|B)=\frac{P(A_k)P(B|A_k)}{P(B)},\quad (k=1,2,3)$$

那么根据其中最大的概率可判定患者患哪一种病．

(1) 通过分析历史资料求出 $m\times n$ 个条件概率 $P(B_j|A_i)$，把条件概率和对应的判别指数列成矩阵表待用(也可贮存在计算机内备用)．

(2) 按已出现的指标 B_{ij} 在表中查找对应的条件概率或判别指数，得到 n 个后验概率．

(3) 诸后验概率值中最大的那一个所对应的类型 A_k，就以最大可能性成为导致该指标出现的原因．

对 1002 例病历以其病理诊断为标准，按慢性、急性和穿孔三种情形进行症候统计，结果如下表：

征候 B_{ij} (体征、症状、化验指标)		慢性 A_1(392例) 例数	慢性 A_1(392例) 频率	急性 A_2(494例) 例数	急性 A_2(494例) 频率	穿孔 A_3(116例) 例数	穿孔 A_3(116例) 频率
1.腹痛开始部位	(1) 右下腹	260	66%	84	17%	13	11%
	(2) 下腹	7	2%	20	4%	7	6%
	(3) 上腹	59	15%	143	29%	49	42%
	(4) 脐周	47	12%	188	38%	30	26%
	(5) 全腹	19	5%	59	12%	17	15%
2.恶心呕吐	(1) 恶心(－)呕吐(－)	129	33%	104	21%	14	12%
	(2) 恶心(＋)呕吐(＋)	204	52%	193	39%	32	28%
	(3) 呕吐(＋)	59	15%	197	40%	70	60%
3.大便	(1) 正常	337	86%	366	74%	61	53%
	(2) 不正常	43	11%	64	13%	29	25%
	(3) 腹泻	12	3%	64	13%	26	22%
4.压痛	(1) 右下腹	384	98%	450	91%	71	61%
	(2) 大于右下腹	8	2%	44	9%	45	39%
5.肌紧张和反跳痛	(1) 肌紧张(＋)反跳痛(＋)	39	10%	282	57%	106	92%
	(2) 肌紧张(－)反跳痛(＋)	145	37%	158	32%	5	4%
	(3) 肌紧张(－)反跳痛(－)	208	53%	54	11%	5	4%
6.体温	(1) ≤37℃	274	70%	143	29%	10	9%
	(2) ＞37℃且＜38℃	106	27%	267	54%	37	32%
	(3) ≥38℃	12	3%	84	17%	69	59%
7.白血球计数	(1) ≤10000	274	70%	44	9%	19	16%
	(2) ＞10000且＜15000	79	20%	203	41%	32	28%
	(3) ≥15000	39	10%	247	50%	65	56%

本例患者具有症候群 $B=B_{13}B_{12}B_{42}B_{51}B_{63}B_{73}$，若可认为各症候是相互独立的，由上表，在患慢性阑尾炎的条件下出现症候 B 的概率为：

$$P(B|A_1)=P(B_{13}|A_1)P(B_{22}|A_1)P(B_{33}|A_1)P(B_{42}|A_1)P(B_{51}|A_1)$$
$$\cdot P(B_{63}|A_1)P(B_{73}|A_1)$$
$$=15\%\times25\%\times3\%\times8\%\times39\%\times3\%\times10\%=10^{-7}.$$

同理,可求出在患急性阑尾炎、阑尾穿孔的条件下,出现症候 B 的概率:

$$P(B|A_2)=3.13439\times10^{-4},\quad P(B|A_3)=3.5389\times10^{-3}.$$

由贝叶斯公式有

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{\frac{392}{1002}\times2.19024\times10^{-7}}{P(B)}$$
$$=\frac{8.56860\times10^{-8}}{P(B)},$$
$$P(A_2|B)=\frac{\frac{494}{1002}\times3.13439\times10^{-4}}{P(B)}=\frac{1.54530\times10^{-4}}{P(B)},$$
$$P(A_3|B)=\frac{\frac{116}{1002}\times3.53892\times10^{-3}}{P(B)}=\frac{4.09695\times10^{-4}}{P(B)}.$$

这里,$P(A_1)$,$P(A_2)$,$P(A_3)$是用相应的频率估计的.由上面结果可以看出 A_3(穿孔)的可能性最大.

在上面的方法中,应指出的是,公式要求满足两个条件:(1) 各种类型 A_k 之间应互不相容,即,一名患者只能出现一种类型而不能同时兼有另一种类型.(2)症候 B_{ij} 独立,即,各种指标之间互不影响.在医学上对各类型的定义都是按病理组织分类,区分严格,界限清楚,可以认为基本互不相容.而对症候B的独立性,在实际上很难满足,经验表明尽量注意选择相对独立的症候,采用上述方法仍能取得较满意的诊断效果.如果用以产生后验概率统计表的病例(训练样本)不是一段时间内的全部病例,也不是从全部病例中随机抽出的,那么 $P(A_1)$,$P(A_2)$,$P(A_3)$便不能用相应的频率估计,需寻求别的来源.

练　习　1.5

1.把甲、乙两种外观一样、数量相等的药片混在一起.若甲种药片的次品率为 0.05,乙种药片的次品率为 0.0025,求

(1) 从混和药片中任取一片为次品的概率;

(2) 从混和药片中取出一片发现是次品,该片药品来自甲、乙种的概率分别为多少.

2.有甲、乙、丙三人加工同一种产品,他们出废品的概率分别为:0.03,0.04,0.06.他们加工的产品放在一起,每人加工的数量分别占总数的 1/4,1/4,1/2,求:

(1) 从中任取一件为废品的概率;

(2) 已知取出一件为废品,这件废品是谁加工的可能性最大?

第六节　贝努里概型

为了认识随机现象的规律性,往往需要进行相同条件下的多次试验.在实际中有一种