应用数理统计

钱尚玮 陈兆顺 编著

YINGYONG SHULITONGJI

中国商业出版社

应用数理统计

钱尚玮　陈光顺　编著

中国商业出版社

应 用 数 理 统 计

钱尚玮 陈兆顺 编著

*

中国商业出版社出版发行

杭州商学院印刷厂印刷

*

787×1092毫米 32开 18,375印张 430千字

1987年12月第1版 1988年9月第1次印刷

印数 1—3000册 定价: 5.80元

ISBN 7-5044-0096-3/F·51

前　言

《应用数理统计》是为高等商业、财经类院校统计专业以及经济信息、企业管理、会计、物价等专业编写的教学用书，也是各类成人高等教育及统计工作者、经济管理人员学习、进修的参考书。

本书共分七章，前三章介绍了概率论的基本知识，后四章讲述数理统计的基本原理和方法，各章均有习题，且附有习题答案。编写此书力求做到理论联系实际，基本概念从实际例子引进，定理、公式尽量阐明实际意义，叙述深入浅出，便于自学。书中例子较多，有利于培养读者运用数理统计理论与方法分析、解决实际问题的能力。

全书七章，其中第一、二、四、五、六章由钱尚玮同志编写，第三章、第七章及各章习题由陈兆顺同志编写，最后由钱尚玮同志总纂定稿。

王思立、黄树颜两位教授曾审阅过书稿，提出了很好的意见，特此表示衷心感谢。

由于编者水平所限，书中难免有不妥之处，在使用过程中，敬请读者提出宝贵意见。

编　者

一九八七年十二月

目　录

第一章　事件及其概率

§1　事　件

一、确定性现象与随机现象

在自然界及人类社会中，人们所观察到的现象大体可归结为两类。一类是事前可以断定它们的结果的。例如，在标准大气压下，水加热到100℃，就会沸腾。这类现象称为确定性现象或必然现象；另一类是事前无法断定其结果的现象。例如，在一分钟内，一个电话交换台接到多少次呼唤，这是事先无法断定的。这种在一定条件下，具有多种可能结果，但究竟发生哪一种结果事先并不能肯定的现象，称为随机现象或偶然现象。

随机现象是广泛存在的，研究它的规律性具有重要意义。就个别的随机现象而言，并无规律性可言，但大量的随机现象却显现出某种规律性，并称之为随机现象的统计规律性。

二、随机事件

在一定条件下，可能发生也可能不发生的事件称为随机事件，简称为事件。

例 1　“某电话总机在一分钟内接到15次呼唤”是一个随机事件。因为在一分钟内，可能接到不足15次呼唤，也可能正好接到15次呼唤或多于15次呼唤。

例 2　“考察人们去商店购买甲种牌号的电视机”是一个随机事件。因为我们事先无法确切知道顾客到底挑选那一种牌号的电视机。

例 3 “在桌面上旋一枚硬分币正面朝上”是一个随机事件，因为我们无法事先断言该分币究竟是正面朝上还是反面朝上。

例 4 “检查15件商品的质量时发现有1件次品”是一个随机事件，因为检查的结果可能正好发现为1件次品，也可能没有发现次品或发现次品不止1件。

这类例子可以举很多。

随机事件是随机现象的某一种结果或表现。例如“明天天气如何”是随机现象，而“明天下雨”则是随机事件。“一种商品的销售量”是随机现象，而“销售量为一万件”则是一个随机事件。

三、必然事件与不可能事件

在一定条件下必然发生的事件称为必然事件，而在一定条件下必然不发生的事件称为不可能事件。

例 5 “同性电荷必然互相排斥”。

例 6 “没有水份，种籽不能发芽”。

例 7 “就目前世界上的人的身高来说，身高小于4米”。

例 8 “在现代医学条件下，人只能活到200岁以下”。

上述是四个必然事件的例子，在这四个例子中的同样条件下，那个事件的反面就是不可能事件。

在一定条件下，必然事件，不可能事件是确定性的。把它们看作随机事件的两种特例，这对讨论问题是方便的。

还需指出，每一事件的发生总离不开一定的条件，离开一定条件来谈一个事件，是没有什么意义的。

此外，我们常常通过随机试验来观察事件。例如事件“旋一枚硬分币正面朝上”是随机试验“旋一枚硬分币”的一个可

能结果。事件“出现1件次品”是随机试验“检查15件商品”的一个结果。一般地说，如果一个试验的可能结果不止一个，且在相同条件下，可以重复地进行这个试验，而每次试验结果事前不可预言，我们称此试验为随机试验。

于是我们也可以说，随机事件是随机试验的一个结果。

以后我们用U表示必然事件，用V表示不可能事件，用A，B，C，…，或A_1，A_2，A_3，…表示随机事件。

四、基本事件和复合事件

事件有两类：一类是基本事件，它是不能再分解的随机事件。另一类是复合事件，它可以分解成同一随机现象中的一些基本事件。

例如事件A：“电话总机在一分钟内接到的呼唤不多于10次”，在这个事件中，“没有呼唤”，“有1次呼唤”等均为电话总机在一分钟内接到的呼唤次数这个随机试验的每一个可能出现的结果，由于它们不能再分解，故为基本事件。而事件A由“没有呼唤”，“1次呼唤”……“10次呼唤”等11个基本事件组成，要事件A出现，必须而且只须这11个基本事件之一出现。所以A是复合事件。

又如事件B：掷两颗骰子点子总和为6”可分解为（1，5），（2，4），（3，3），（4，2），（5，1）等5个基本事件，所以事件B也是复合事件。

应注意，我们所谈及的事件是“基本”的或“复合”的，是对一定的研究范围而言的。

例如“三粒油菜籽的发芽粒数”这个随机现象，包含有下列几种结果：

“全部不发芽”，记为事件A_0；

“有1粒发芽”，记为事件A_1；

“有 2 粒发芽”，记为事件 A_2；

“有 3 粒发芽”，即“全部发芽”记为事件 A_3。

在不考虑次序的范围内考察这四个事件，由于都不能再分解，故均为基本事件。

但是，如果现在把这三粒油菜籽给定一个次序（不妨假定是放在“1”，“2”，“3”这三只培养皿中作发芽试验），并要考虑“是第几粒发芽”，这样，研究范围就从原来不考虑次序，变成要考虑次序了。在这个新的研究范围内，“有 1 粒发芽”，就不是基本事件，它还可以区分为“只有第一粒发芽”，“只有第 2 粒发芽”和“只有第3粒发芽”这三种情况，这时，基本事件有表 1～1 中的 8 个。

表 1～1

事件名称	第一粒	第二粒	第三粒
B_1	不发芽	不发芽	不发芽
B_2	发芽	不发芽	不发芽
B_3	不发芽	发芽	不发芽
B_4	不发芽	不发芽	发芽
B_5	发芽	发芽	不发芽
B_6	发芽	不发芽	发芽
B_7	不发芽	发芽	发芽
B_8	发芽	发芽	发芽

从（表 1～1）中看到，“有 1 粒发芽”这个事件是由基本事件 B_2，B_3，B_4复合而成的，或者说事件“有 1 粒发芽”可以分解为 B_2，B_3，B_4三个基本事件。

分析（表 1～1）中的 8 个基本事件，我们看到，基本事

件有一个很重要的性质：在一次试验中只能发生基本事件中的一个。换句话说，任何两个或两个以上的基本事件，不可能在一次试验中同时发生。

五、样本空间

为了研究随机试验E，我们把E的所有可能的试验结果（或所有基本事件）的全体所构成的集合叫做E的样本空间，（或称为基本空间）记为U，而把每一个可能出现的结果（或每一个基本事件）称为一个样本点。

由于随机事件或是基本事件，或是由若干个基本事件组合而成的复合事件，因而引入样本空间U以后，随机事件便是样本空间U中的子集。事件的发生，就是当且仅当子集中的一个样本点发生。特别地，必然事件就是样本空间U，不可能事件就是空集V。

例9　试验E_1——一枚硬币抛掷一次。
样本空间为：$U=\{$出现正面，出现反面$\}$

此例中，样本点有两个：出现正面，出现反面。

例10　试验E_2——计算某电话总机在$[0,t]$内的呼唤次数。
样本空间为：$U=\{0,1,2,\cdots\cdots\}$

此例中，样本点有可列无穷多个。

例11　E_3——测试某种灯泡的寿命t。
样本空间为：$U=\{t\mid t\geqslant 0\}$，此例中，样本点有不可列无穷多个。

§2　概　率

研究随机现象，不仅要知道它可能出现哪些事件，更重要的是要研究各种事件出现可能性的大小。我们要求有一个衡量事件出现可能性大小的数量指标，而且事件发生可能性大的，

这个值就大；事件发生可能性小的，这个值就小，必然事件的值最大且等于1，不可能事件的值最小且等于零。

我们把衡量事件发生可能性大小的数值称为事件的概率，事件A的概率用$P(A)$表示，并且规定$0\leqslant P(A)\leqslant 1$。

然而，对于事件A，应该用哪个数值作为A的概率呢？下面我们分别介绍概率的统计定义、古典概型，在§3中介绍概率的公理化定义。

一、概率的统计定义

下面先引入事件的频数与频率的概念。

定义1.1 设随机事件A在相同条件下的n次试验中发生了k次，称k为事件A发生的频数，比值$\frac{k}{n}$称为事件A发生的频率，记作$W(A)$，用公式可表述如下：

$$W(A)=\frac{k}{n}$$

显然，任何随机事件的频率是介于0与1之间的一个数。

即 $$0\leqslant W(A)\leqslant 1$$

如果A是必然事件，则$W(A)=1$；如果A是不可能事件，则 $W(A)=0$。

例12 某工厂生产某种产品，一般是合格品，但也可能出现少量次品，故"生产的产品是合格品"是一随机事件。为观察该随机事件出现的规律，可将生产情况记录如下，这里每生产1件产品就相当于进行1次试验。

表1～2

产品件数(n)	5	10	60	150	600	900	1200	1800
合格品数(k)	5	7	53	131	548	820	1091	1631
频　　率$\left(\frac{k}{n}\right)$	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

例13　历史上，有些人作过成千上万次投掷钱币的试验，下表列出了蒲丰（Buffon）和皮尔逊（Pearson）试验记录：

实 验 者	投掷次数 n	出现正面次数(k)	频　率 $\left(\frac{k}{n}\right)$
Buffon	4,040	2,048	0.5069
Pearson	12,000	6,019	0.5016
Pearson	24,000	12,012	0.5005

从上面两个例子的数据可以看出，在每次重复试验中，同一事件的频率有波动，带有偶然性，当试验重复多次时，随机事件的频率却稳定在一个固定的数值附近。如例12稳定在数0.9附近，例13稳定在数0.5附近，而且随着试验次数增加，频率越来越明显地呈现出稳定性。

频率的稳定性揭示出一个随机事件出现的可能性有一定的大小，频率稳定在较大数值时，表明相应事件出现的可能性大，频率稳定在较小数值时，表明相应事件出现的可能性小。因此，很自然会想到用频率所稳定的那个数值作为衡量事件出现可能性大小的一个量的标志。这个数值就称为相应事件的概率，这就是所谓概率的统计定义。

定义1．2　在相同条件下，重复进行 n 次试验，其中事件 A 出现了 k 次．如果当试验次数 n 足够大时，事件 A 出现的频率

$\frac{k}{n}$稳定在某定数p附近，则称数p为事件A发生的概率，记作$P(A)=p$。

在不少实际问题中，当概率不易求出时往往用频率$\frac{k}{n}$作为概率的近似值。

由于频率$\frac{k}{n}$总介于0与1之间，由概率统计定义可得任何随机事件A的概率$P(A)$有以下性质：

1） $0\leqslant P(A)\leqslant 1$

2） 必然事件U的概率等于1，即$P(U)=1$；
不可能事件V的概率等于0，即$P(V)=0$。

应当注意，我们不能把事件A的概率看作是频率$\frac{k}{n}$当$n\to\infty$时的极限。例如，例10中投掷硬币试验，出现正面的概率是$\frac{1}{2}$，但是如果我们把概率$\frac{1}{2}$看作是频率$\frac{k}{n}$当$n\to\infty$时的极限，那么应当证明，对任意的$\varepsilon>0$，可以找到一个N，使得当$n\geqslant N$时，不等式

$$\left|\frac{k}{n}-\frac{1}{2}\right|<\varepsilon$$

永远成立。然而，这无论从理论上考察，或者就具体的实验而言，都不可能找到这样一个N。因为n充分大时，频率虽然具有稳定性，但仍然是摆动的。所以把概率看作频率当$n\to\infty$时所趋于的极限是不妥的。

二、概率的古典定义

上面所谈的概率统计定义为我们了解概率概念的实际含义提供了依据，同时又给出了近似计算的一般方法。但在考虑一类

常见而又简单的特殊模型时，并不需要进行反复试验，根据对试验的条件作具体分析就可以直接计算事件的概率。下面介绍古典概率的定义，这种定义和“等可能性”的概念相联系，我们先通过具体例子来说明。

例14 设有一批电子产品，总数为1500个，其中有50个是废品，现从中任取1个产品，问取到的产品是废品的可能性多大？

解：我们把每抽取1个产品看成1个基本事件，因为1500个产品，每个产品都有可能被抽到，所以有1500个不同的基本事件。而且每个产品被抽到的可能性都是一样的，因此，这1500个基本事件是等可能的，我们把事件“抽取的产品为废品”记为A，因为共有50个废品，所以A中包含的基本事件个数只有50种可能性，很自然，事件A出现的可能性为$\frac{50}{1500}\approx 0.033$。

从上例可以看到，有一类简单的随机现象，它有下列两个特征：

1、在试验中，试验的可能结果只有有限个，记为E_1，E_2，…，E_n，称它们为基本事件，每次试验中，必出现其中的一个，而且只出现其中的一个。

2、基本事件E_1，E_2，…，E_n，出现的可能性相等，即它们发生的概率都相等。

具有上述两个特征的随机现象，属于一类特殊的模型，在数学上称为古典概率模型，简称古典概型。

概率的古典定义如下：

定义1.3 在古典概型下，若所有基本事件个数为n，而事件A包含了其中的m个，那么事件A的概率

$$P(A)=\frac{m}{n} \qquad (1.1)$$

m 有时也称为有利于事件 A 发生的基本事件的个数。

上述概率的古典定义给出了计算事件的概率的一种方法，即事件 A 的概率 $P(A)$ 可由事件 A 中包含的基本事件的个数 m 与全体基本事件个数 n 之比来计算。

例15 设有一批产品，总数为 N 个，其中 R 个是废品，现从中任取 n 个（$n \leqslant R$），问 n 个中有 r 个废品的概率是多少？

解：从 N 个产品中每取 n 个产品作为 1 个基本事件。从 N 个中取 n 个的取法共有 $C_N^n = \dfrac{N!}{n!\,(N-n)!}$ 种，因此基本事件的总数为 C_N^n 个，在 N 个产品中共有废品 R 个，合格品 $N-R$ 个。现在要求在抽取的 n 个产品中有废品 r 个，合格品有 $n-r$ 个。废品 r 个只能从总废品数 R 中选取，取法 C_R^r 种，合格品 $n-r$ 个只能从总合格品 $N-R$ 中选取，取法有 C_{N-R}^{n-r} 种。因此，从 N 个中抽取 n 个，其中包含有 r 个废品的方法有 $C_{N-R}^{n-r}C_R^r$ 种，所以若以 A 表示事件“n 个中有 r 个废品”，则根据概率的古典定义有

$$P(A) = \frac{C_{N-R}^{n-r}\ C_R^r}{C_N^n}$$

例16 一个小组有 8 个学生，问这 8 个学生的生日都不相同的概率是多少？（设一年为365天）

解：用 B 表示事件“这 8 个学生生日都不相同”。基本事件总数为 365^8，而 B 包含的基本事件总数为 A_{365}^8，故所求概率

$$P(B) = \frac{A_{365}^8}{365^8}$$