概率论与数理统计三十三讲

魏振军 编

中国统计出版社
China Statistics Press

概率论与数理统计

三十三讲

魏振军 编

中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

概率论与数理统计三十三讲/魏振军编.
—北京:中国统计出版社,2000.7
ISBN 7-5037-3290-3

Ⅰ. 概…
Ⅱ. 魏…
Ⅲ. ①概率论 ②数理统计
Ⅳ. O21

中国版本图书馆 CIP 数据核字(2000)第 63007 号

责任编辑/张美华
封面设计/张建民
出版发行/中国统计出版社
办公地址/北京市丰台区西三环南路甲 6 号
电　　话/(010)63459084、63266600—22500(发行部)
印　　刷/科伦克三莱印务(北京)有限公司
经　　销/新华书店
开　　本/787×1092mm　1/18
字　　数/400 千字
印　　张/16.5
印　　数/1—3000 册
版　　别/2000 年 8 月第 1 版
版　　次/2000 年 8 月北京第 1 次印刷
书　　号/ISBN 7-5037-3290-3/0·35
定　　价/28.50 元

中国统计版图书,版权所有,侵权必究。
中国统计版图书,如有印装错误,本社发行部负责调换。

前　言

《概率论与数理统计》是研究随机现象规律性的一门学科。它有着深刻的实际背景，在自然科学、社会科学、工程技术、军事和工农业生产等领域中有广泛的应用。

作者从事概率统计教学二十年，在这些年的教学过程中，作者深感这门课的教学难度较大。为了提高教学效果，激发学生的学习热情，作者一直在教学实践中进行改革探索。本书就是这种探索的初步结果。

本书将概率论与数理统计课程的基本内容划分为六章、33讲。前五章是概率论部分，共26讲；第六章为数理统计部分，共7讲。每章开始，给出教学基本要求、内容提要、重点和难点。每讲围绕一个中心展开，后面配有思考与练习。在讲述基本概念和方法时，力求给出它们的实际背景、基本思想以及结果的解释。在讲述时插入一些具体事例，以帮助学生理解。作者特意选择了一些联系实际的习题，目的是引起学生思考讨论，进而让学生了解这门课程的广泛应用。书末附有习题参考答案。

为了借助现代计算机技术辅助概率统计课程教学，作者在十年前就开始研制这门课程的系统教学软件。现在这套教学软件已正式出版发行，可与本书配套使用。

本书可作为50学时左右讲授或自学概率统计使用的教材。限于篇幅，概率论中的几何概率、部分重要分布；数理统计中的描述统计、抽样设计、方差分析等内容放在教学软件中介绍，供读者自学。书中“*”号部分可选学。

值此本书出版发行之际，我发自内心地感谢所有对我的教学和研究工作给予过无私帮助的人们。原国家教委高教司、国家统计局教育中心和中国统计出版社的同志们对我的工作给予了大力

支持;概率统计学界的前辈们对我的工作给予了悉心指导;我校有关领导对我的工作给予了热情关怀;本教研室的张新建老师一遍又一遍帮助审阅书稿。在这里,我一并向他们表示诚挚的谢意。

受本人水平所限,摆在读者面前的这一本书,肯定会有错误和疏漏之处,作者恳请各位老师和同学们给予批评指正。

魏振军

2000 年 5 月于郑州

此为试读,需要完整PDF请访问: www.ertongbook.com

目　录

第一章
随机事件与概率

一、教学基本要求

1. 较好地理解随机事件、概率、条件概率及独立性四个基本概念.

2. 较好地掌握事件的关系和运算，着重理解事件的“和”、“积”、“对立事件”及相应的概率性质.

3. 能熟练运用概率的加法公式和乘法公式.

4. 了解加法公式和乘法公式的综合运用：全概率公式和贝叶斯公式.

5. 会解简单的古典概型问题.

二、内容提要

1. 讲述随机事件及其概率的概念，介绍古典概型.

2. 通过几个简单的随机试验，揭示随机事件的一个极其重要的特征 — 频率稳定性.

3. 介绍事件的关系和运算，概率的公理、概率的有关性质；给出计算概率的一个重要公式：加法公式，并说明其应用.

4. 介绍条件概率、独立性的概念，给出计算概率的又一个重要公式：乘法公式，并说明其应用.

5. 介绍全概率公式和贝叶斯公式.

三、重点与难点

本章重点是：

1. 四个基本概念：随机事件、概率、条件概率、独立性.
2. 两个公式：加法公式和乘法公式.

本章难点是：

1. 古典概型中的概率计算.
2. 条件概率.
3. 全概率公式和贝叶斯公式的应用.

第一讲 有趣的偶然世界(概率论序言)

我们生活于其中的大自然和社会，是一个千姿百态、五光十色、充满生机的世界.浩瀚宇宙、茫茫人海、大千世界蕴涵了多少美妙的故事与哲理.冬去春来，周而复始，大自然又赐予我们多少奥秘，等待有志者去探索、去研究.

让我们把眼光投向大自然，投向社会，去探求、去思索，去揭示其中的奥秘！

同学们！不知你们可曾留心观察过自然界和社会的各种现象？你们知道这些现象又有什么特点和规律吗？现在，就让我们一起来看看吧！

春天到了，万物复苏，百花盛开，大自然呈现出一派勃勃生机.秋风吹来，枝叶凋零；上抛的物体一定会下落；在标准大气压下，水加热到 100℃ 就会沸腾；无论是什么形状的三角形，它的两边之和总要大于第三边.

总结上面这些列举的现象，我们可以发现，它们都有着共同的特点.如果用比较科学的语言来表达，那就是：服从特定的因果规律，从一定的条件出发，一定可以推出某一结果.这一类现象我们把它称作“确定性现象”，也叫作必然现象.在自然界和社会中还大量存在着另一类现象，我们称之为随机现象或者说是偶然现象.

比如，在十字交叉路口，每天都要通过许多人和车辆，但是，我们无法事先预测每天确切的人数及车辆数.

你考上了大学,即将到新的班级生活,在报到前你能准确地预知新同学中的最大身高是多少吗?

当你和同学划着小船,荡漾在美丽的湖上,你们能让小船每一次都精确地沿着上次所走的路线行驶吗?

美国“挑战者”号航天飞机升空不久便发生爆炸,震惊了整个世界,可事先谁也无法料到……

山崩、地震、海啸、森林失火、火山爆发,使地球上的人们感到神秘和恐怖.飞机失事,火车颠覆、名人遇刺、意外爆炸,象这样偶然的事件几乎每天都在发生着……

好吧!让我们来做一个试验.

这里有十个签,其中只有一张是彩票,抽到彩票的人可以获得许多的好处.

小张:“当然啦,这类活动总是少不了我的,而且到目前为止,我一直运气不赖,不信您瞧!”

(没抽到彩票)“真倒霉!运气太坏了!”

那么,随机现象又有什么特点呢?

这就是:当人们在一定的条件下对它加以观察或进行试验时,观察或试验的结果是许多可能结果中的某一个.而且在每次试验或观察前都无法确知其结果,即呈现出偶然性.或者说,出现哪个结果“凭机会而定”.后面我们将会看到,这个或那个结果出现的机会是这个现象本身所固有的性质.我们把这种现象叫做随机现象.

研究随机现象,首先要对研究对象进行观察试验.如果每次试验的可能结果不止一个,且事先不能肯定会出现哪一个结果,这样的试验称为随机试验.

在随机试验中,可能发生也可能不发生的事件,叫做随机事件,通常用大写英文字母 A、B、C 等表示随机事件.例如,掷一颗骰子,“掷出 6 点”是一个随机事件;再如,100 个灯泡中,有 3 个次品,现任取一个,则“取出的是正品”也是一个随机事件.

又比如上面抽签的例子里,在十张签中抽到一张彩票就是一个随机事件.随机事件可以简称“事件”.当然,刚才小张手气不太好!“不过,这也是很偶然的嘛!”

其实,很多现象看起来偶然,实际上有一定的必然性.

让我们再来看一个例子.

在掷一枚质地均匀的硬币时,既可能出现正面,也可能出现反面,但事先我们无法知道到底会出现哪一面.但是在大量的重复试验后,人们发现,出现正面与出现反面的次数竟然几乎各占一半.

法国科学家蒲丰,做过试验,他掷硬币 4 040 次,结果出现正面的频率为 50.69%

英国科学家皮尔逊,通过两组上万次重复试验,进一步得出,出现正面的频率分别为 50.16% 和 50.05%.

事实告诉我们,抛掷次数越多,一般来说,出现正面的频率就越接近于 1/2.

读者一定听说过福尔摩斯探案的故事.

一次,福尔摩斯受理破译一份密码,密码写在一张纸条上(见图 1-1-1).这份密码与以往的方式完全不同,上面画的全是些会跳舞的小人.

图 1-1-1 第一张纸条

怎样破译这份密码呢?请看福尔摩斯的回答:

"只要看出了这些符号是代表字母的,再应用字母的规律来分析,就不难找到答案."

字母的规律是什么呢?

人们通过大量的试验证明,在英文文章中,每个字母出现的频率是相当稳定的.比如,据一份英文频率的统计表给出(不同的统计表可能有些差异,但大致相同):字母 E 出现的频率最高,是 12.7%,T 为 9.78%,A 为 7.88%,O 为 7.76%,I 为 7.07%,N 为 7.06%,字母 Z 出现的频率最小,只有 0.6%.

福尔摩斯恰恰就是运用了这种规律,破出了那份神秘的密码,抓住了凶手.

此为试读,需要完整PDF请访问: www.ertongbook.com

让我们从《福尔摩斯探案全集》中找出有关的内容：

“第一张纸条上的话很短，我只能稍有把握假定符号  代表 E，E 是英文中使用最多的字母，它出现的次数即使在一个短的句子中也是最常见的．第一张纸条上的 15 个符号，其中有 4 个完全一样，因此把它估计为 E．这些图形中，有的还带一面小旗，有的没有小旗．从小旗的分布来看，带旗的图形可能是用来把这个句子分成一个一个的单词．我把这看作一个可以接受的假设，同时记下 E 是用  来代表的．”

“可是，现在最难的问题来了，因为，除了 E 之外，英文字母出现次数的顺序并不很清楚，这种顺序，在平常一页印出的文字里和一个短句子里，可能正相反．一般来说，字母按出现次数排列的顺序是

$$T,A,O,I,N,S,H,R,D,L,\cdots\cdots$$

但是 T,A,O,I 出现的次数几乎不相上下．要是把每一种组合都试一遍，直到得出一个意思来，那会是一项无止境的工作．所以我们只好等来了新材料再说．”

之后，又有几张纸条来到福尔摩斯手中，它们是这样的(见图 1-1-2)：

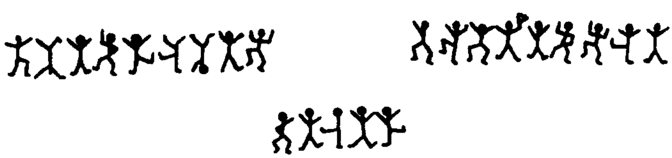
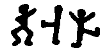


图 1-1-2

对上面五个小人的这幅图，福尔摩斯说：“在这由五个符号组合的单词中，我找出了第二个和第四个都是 E．这个单词可能是 sever(切断)，也可能是 lever(杠杆)，或者 never(决不)．”根据对情况的分析，认为这张纸条是回答一项请求的，而使用 never 这个词

来回答一项请求的可能性极大.于是福尔摩斯判断出：
三个符号分别代表字母 N，V 和 R.

甚至在这个时候，破译密码困难仍很大.经过福尔摩斯对案情的分析，加上他对字母规律及对各种密码的研究经验，对上述各幅图进行了分析，最后把第一张纸条上的一句话译为：

AM HERE ABE SLANE.（我已到达.阿贝．斯兰尼）

把第二张纸条上的一句话译为：

AT ELRIG ES.（住在埃尔里奇）

在得到最后画的(见图 1-1-3) 小人后，福尔摩斯用已经知道的字母再加上分析，将其译成这样一句话：

ELSIE PREPARE TO MEET THY GOD.（埃尔茜，准备见上帝）

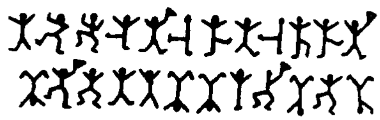


图 1-1-3

字母使用频率的研究，对于打字机键盘的设计，印刷铅字的铸造，信息的编码，密码的破译都有着十分重要的意义.

事实告诉我们：随机现象有其偶然性的一面，也有其必然性的一面，这种必然性表现在大量重复试验或观察呈现出的固有规律性，称为随机现象的统计规律性.

研究和揭示随机现象的统计规律的学科，就是我们下面要和大家一起来学习的“概率论与数理统计”.

它将把我们引向一个千百年来被人们视为“风云莫测”的随机世界！

思考与练习(1)

1. 从资料查找或自己给出两个随机现象的例子，说明大量随机现象蕴涵的规律．

2. 汉字在文章中的使用频率有规律吗?人们如何利用这些规律?

3. 圆周率 $\pi = 3.141\ 592\ 6\cdots$ 是一个无限不循环小数,我国数学家祖冲之第一次把它计算到小数点后七位,这个记录保持了1 000 多年. 以后有人不断把它算得更精确 .1873 年,英国学者沈克士公布了一个 π 的数值,它的数目在小数点后一共有 707 位之多!

但是,经过几十年后,曼彻斯特的费林生对它产生了怀疑. 原因是他统计了 π 的 608 位小数,得到下面的表:

数字	0	1	2	3	4	5	6	7	8	9
出现次数	60	62	67	68	64	56	62	44	58	67

你能猜出他怀疑的理由吗?

第二讲 随机事件及其概率

上一讲中,我们已经了解到,随机现象有其偶然性的一面,也有其必然性的一面,这种必然性表现在大量重复试验或观察中呈现出的固有规律性,称为随机现象的统计规律性. 而概率论正是研究随机现象统计规律性的一门学科. 现在,就请你和我们一起,步入这充满随机性的世界,开始第一步的探索和研究.

以下,我们从观察试验中引出几个基本概念.

一、随机试验与事件

在本章第一讲中,我们已经接触到了好几个随机试验. 现在,让我们再来看几个随机试验.

掷一颗骰子,观察掷出的点数. 所看到的是 6 种可能结果中的某一个,而事先无法肯定掷出的是几点.

将一枚硬币抛掷两次,观察出正出反的情况,可能结果为:{正,反},{正,正},{反,正},{反,反},但抛掷之前亦不能预言出现哪一种结果.

从一批灯泡中任意抽取 1 只,测试它的寿命. 可以知道寿命

$t \geqslant 0$,但在测试之前不能确定它的寿命究竟有多长.

从以上的随机试验中,我们看到,试验是在一定条件下进行的,试验有一个需要观察的目的,根据这个目的,每次试验只能观察到多个可能结果中的某一个.试验的全部可能结果,是在试验前就明确的;或者虽不能确切知道试验的全部可能结果,但可知道它不超过某个范围.而且,每次试验的结果事先不可预言.也就是说,试验的结果具有随机性.由于我们这里只研究可以大量重复的随机现象,因此只考虑可以在相同条件下重复进行的随机试验.以下,随机试验也简称为试验.

在随机试验中,可能发生也可能不发生的试验结果称为随机事件.例如,在掷硬币试验中,"掷出正面"是一随机事件.又如,在掷骰子试验中,"掷出 6 点"、"掷出偶数点"也都是随机事件."随机"的意思无非是说,事件是否在某次试验中发生,随机会而定.随机事件也简称为事件,通常用大写英文字母 A、B、C 等表示.

事件可分为基本事件和复合事件.我们把相对于观察目的不可再分解的事件,称为基本事件.例如,在掷骰子试验中,我们的目的是要观察掷出的点数,则"掷出点数为 1","掷出点数为 2",…,"掷出点数为 6"都是基本事件.两个或一些基本事件并在一起,就构成一个复合事件.例如,在掷骰子试验中,事件"掷出偶数点"就是由"掷出点数为 2","掷出点数为 4","掷出点数为 6"三个基本事件构成的.

有两个特殊的事件必须说明一下,一个是必然事件,即在试验中必定发生的事件,常用 S 或 Ω 表示;另一个是不可能事件,即在试验中不可能发生的事件,常用 ϕ 表示.例如,在掷骰子试验中,"掷出点数小于 7"是必然事件;而"掷出点数 8"则是不可能事件.

必然事件与不可能事件都是确定性的,但为了今后讨论问题方便,不妨将它们视为随机事件的特例.

二、事件的概率

对于一个随机事件来说,它在一次试验中可能发生,也可能不发生,既然有可能性,就有可能性大小问题.我们常常希望知道某些事件在一次试验中发生的可能性究竟有多大,或者说,希望找到

一个合适的数来表示事件发生的可能性大小.

对于一个随机事件来说,它发生可能性大小的度量是由它自身决定的,并且是客观存在的.就好比一根木棒有长度,一块土地有面积一样,概率是随机事件发生可能性大小的度量.也就是说,事件 A 发生的可能性大小就是事件 A 的概率.我们用 $P(A)$ 表示事件 A 的概率.

例如,在掷硬币试验中,用 A 表示"出现正面",用 B 表示"出现反面".如果硬币质地均匀,形状对称,那么人人都会说事件 A 和事件 B 出现的可能性一样大.也就是说

$$P(A)=1/2, P(B)=1/2$$

这就把事件发生可能性的大小数量化了.

由于必然事件在每次试验中必定发生,或者说,它发生的可能性是百分之百,所以它的概率是 1.不可能事件发生的可能性是零,所以它的概率是 0,即有

$$P(S)=1, P(\phi)=0 \qquad (1.2.1)$$

而任一事件 A 发生的可能性不会小于 0,也不会大于百分之百,所以 A 的概率介于 0 与 1 之间,即有

$$0 \leqslant P(A) \leqslant 1 \qquad (1.2.2)$$

现在,让我们看一个从死亡线上生还的故事.

相传古代有个王国,由于崇尚迷信,世代沿袭着一条奇特的法规:凡是死囚,在临刑前都要抽一次"生死签",即在两张小纸片上分别写着"生"和"死"的字样,由执法官监督,让犯人当众抽签.如果抽到写有"死"字的签,则立即处刑;如果抽到写有"生"字的签,则被认为这是神的旨意,应予当场赦免.

有一次,国王决定处死一名敢于"犯上"的大臣,为了不让这名囚臣得到半点获赦的机会,他与几名心腹密谋暗议,想出了一条狠毒的计策:暗中嘱咐执法官,把"生死签"的两张签都写成"死"字.这样,不管犯人抽得的是哪张签,终难免一死.

在国王一伙看来,这个"离经叛道"的臣子的"死"是必然事件,因为他们让这位臣子抽的生死签实际是"两死抽一".然而,聪明的囚臣正是巧妙地利用了这一点,使自己死里逃生.

当执法官宣布抽签的办法后,只见囚臣以极快的速度抽出一

张签，并迅速塞进嘴里．等到执法官反应过来，嚼烂的纸早已吞下．执法官赶忙追问："你抽到'死'字签还是'生'字签?"囚臣故作叹息说："我听从天意的安排，如果上天认为我有罪，那么这个咎由自取的苦果我业已吞下，只要查看剩下的签是什么字就清楚了．"

剩下的签当然写着"死"字，这意味着囚臣已抽到"生"字签．国王和执法官有苦难言，由于怕触犯众怒，只好当众赦免了囚臣．

本来，这位囚臣抽到"生"还是"死"是一个随机事件．抽到每一种的可能性各占一半，也就是各有 1/2 的概率．但由于国王一伙"机关算尽"，通过偷换试验条件，想把这种概率只有 1/2 的"抽到死签"的随机事件，变为概率为 1 的必然事件，终于搬起石头砸了自己的脚，反使囚臣得以死里逃生．

三、样本空间与事件

上面这个故事中的从"生""死"两签抽一签的试验与掷一枚均匀硬币观察出正反情况的试验一样，都是只有两个基本事件的试验，而且每个基本事件出现概率都是 1/2．现代集合论为表述随机试验提供了一个方便的工具．把随机试验的每个基本结果称为样本点，记作 e 或 ω．全体样本点的集合称为样本空间．样本空间用 S 或 Ω 表示．这样，抽生死签试验的样本空间为 $S=\{生,死\}$；而掷硬币试验的样本空间为 $S=\{正,反\}$．去掉具体含义，它们都是包含两个样本点的样本空间．

如果试验是抛掷一颗骰子，则样本空间是

$$S=\{i:i=1,2,3,4,5,6\}$$

这里结果 i 表示掷出 i 点，$i=1,2,3,4,5,6$．

如果骰子是均匀的，则每个结果出现的概率都是 1/6．

样本空间在如下意义上提供了一个理想试验的数学模型：在每次试验中必有一个样本点出现(样本空间包括了试验的所有可能结果)且仅有一个样本点出现(不可能同时发生两个结果)．

如果试验是将一枚硬币抛掷两次，则样本空间由如下四个样本点组成：

$$S=\{(H,H),(H,T),(T,H),(T,T)\}$$

结果 (H,H) 表示两次都掷出正面，(H,T) 表示第一次掷出正面

此为试读，需要完整PDF请访问：www.ertongbook.com