简明概率统计教程

丁正育　薛　峰　编著

内 容 简 介

本书是概率论与数理统计的简明教材。概率部分以阐述基本概念和主要结论为重点，数理统计部分则强调统计思想和统计方法的应用。全书选材精当，叙述简明，深入浅出，通俗易懂，注重师范性。每章附有“基本要求”和“小结”，便于教学和自学。

本书可作为高等师范专科学校、教育学院、教师进修院校的全日制或函授用的概率统计教材或教学参考书。也是中学教师、工程技术人员、大学生和有兴趣的青年自学用的较好入门书。

简明概率统计教程
丁正育 薛 峰 编著
责任编辑 涂红

浙江大学出版社出版发行
浙江省良渚印刷厂印刷

开本：787×1092 1/32 印张：9 字数：202千字
1991年9月第1版 1991年9月第1次印刷
印数：0 001～2 060
ISBN 7-308-00630-1
O·084 定价：4.60 元

序 言

概率论与数理统计是近代数学的一个重要分支。它是一门定量地研究随机现象内部所蕴含的必然规律性的学科。理论严谨，应用广泛，近数十年来发展十分迅速。它的理论和方法已渗透到自然科学和社会科学的各个领域。在各级数学教学中也正在日益充实加强这一分支的课程。

本书是作者们在对初稿反复试讲的基础上几经修改后完成的。是一本师范专科学校数学专业（包括全日制、函授、教师进修）适用的教材或教学参考书，也可供有兴趣的青年或大学生自学参考。

该书选材适当，注意了该课程独特的概念和思想方法。概率部分以阐述基本概念和主要结论为重点；数理统计部分则强调统计思想和统计方法的应用。该书文字叙述简明扼要，注意了从具体特例说明一般的抽象概念。每章附有“基本要求”和“小结”，便于教学和自学。

本书的出版，对促进我国的师范专科学校的概率统计教学必将有所裨益，产生积极的影响。

林正炎

1989年11月22日于杭州大学

编者的话

长期以来，我们与其他师范专科《概率统计》教学工作者一样，都为无合适的教材而感到苦恼，本书就是从这种苦恼中冲出来的一本教材。本书编写的依据是国家颁发的师范专科全日制和成人教育《概率统计》教学大纲。编写的宗旨是“简明”，也就是取材精当，叙述简明。我们在编写时力求做到：既考虑到学生实际水平和工作需要，又注意到理论严谨性和科学系统性；充分体现该学科理论联系实际的特点；注意从具体到抽象、从特殊到一般的叙述方法；注重概率统计思想方法；论证简炼，深入浅出，通俗易懂。为便于读者学习，每章附有“基本要求”和“小结”。考虑师范专科不同类别教学的要求，对部分内容打上星号“*”，以供选用。

本书编写得到了杭州大学数学系主任林正炎教授的热忱鼓励和指导，他三次审阅了本书的初、复稿，提出许多真知灼识，并为本书写了《序言》。北京大学概率统计系主任陈家鼎教授和浙江湖州师专数学科主任何仲洛教授以及高教出版社高尚华副编审都看过手稿，给予较高的评价，并提过不少宝贵的意见。本书初稿于1986年11月浙江省高师系统《概率统计》学科协作会议上交流时，得到同行、专家们的支持。谨此一并表示衷心感谢。

本书是我们合作编写的。本书的初稿和修订稿讲义的一至四章是由薛峰同志执笔，五至八章是由丁正育同志执笔。丁正

育同志在总结试讲经验的基础上，对全稿作了统一的修改。薛峰同志对修改稿又作了仔细全面的修改。最后，由林正炎教授审定。

限于编者学识，书中难免存在疵误，诚望专家、同行和读者不啬赐正。

编　者

1989年11月

引　言

在自然界和人类社会中发生的各种现象虽然千变万化，但大体上可归结为两大类。一类是在一定条件下必然发生(或必然不发生)，这类现象称为**必然现象**（或**确定性现象**）。如水往低处流；上抛物体必下降；振兴经济必须依赖科学技术等。另一类是在一定条件下每次试验（或观察）的可能结果有多个，预先不能确定到底发生哪个结果，这类现象称为**随机现象**。如在掷硬币的试验中，预先不能确定每次试验出现“正面朝上”还是“反面朝上”。对随机现象作个别试验，好像无规律可寻，但正如恩格斯所指出：“在表面上是偶然性在起作用的地方，这种偶然性始终是受内部的隐蔽着的规律支配的，而问题只是在于发现这些规律。”大量实践表明，对随机现象作大量重复试验时，会呈现出某种规律性。这种规律性称为**统计规律性**。如预先不能确定一个孕妇是生男孩还是生女孩，但近代人口统计表明，男女孩出生的比例却稳定在22∶21。

概率论与数理统计(简称为概率统计)就是研究随机现象统计规律性的一门数学学科。它在长期发展进程中已形成理论严谨、概念独特、方法新颖、应用广泛的富具特色的学科。由于随机现象的普遍性，决定了它的应用的广泛性。目前它已广泛地应用于自然科学、社会科学以及国民经济各个部门的研究中。它的理论和应用的发展非常迅速，已成为近代数学的一个十分活跃的分支。本书就是这门学科的简明入门书。通过本书的学习，能使读者初步掌握随机现象的基础理论和基本方法，培养解决某些实际问题的能力。

目　　录

第一章 事件与概率

随机事件及其概率是概率论的最基本的概念。我们从随机试验出发，引入样本空间和随机事件的概念。接着讨论统计概率、古典概率和几何概率。然后介绍事件和概率的数学定义和概率的基本性质，并由此抽象出描述随机试验的数学模型——概率空间。进而讨论条件概率和独立性的概念。最后研究一类重要的独立重复试验概型。

§1.1 事件与样本空间

一、随机试验与随机事件

简单地说，我们把对现象（包括随机现象和必然现象）进行一次观察或科学试验统称为一次**试验**。如果试验满足下列条件：

(1) 可在相同条件下重复进行；

(2) 试验的所有可能结果不止一个，且事先能明确知道试验的所有可能结果；

(3) 每次试验结果事先不可预测，但是每次试验总是恰好出现上述这些可能结果中的一个。

则称该试验为**随机试验**，简称为**试验**，常用E表示。在概率统计中，总是通过随机试验来研究随机现象统计规律性的。应该指出，一个试验总是在一定条件下进行的，但通常不具体指出条件是什么。

随机试验的每一个可能结果，称为**基本事件**。由若干个基本事件组合而成的事件称为**复合事件**。基本事件和复合事件都是随机事件，常用大写字母A、B、C、……表示。我们称在一定条件下一定发生的事件为**必然事件**，常用Ω表示。称在一定条件下一定不发生的事件为**不可能事件**，常用$\varnothing$表示。必然事件和不可能事件，严格地说都不是随机事件，但为了研究方便，把它们看作是随机事件的两个极端情况。

例如从某厂的一批产品中任意抽取10件进行检验，每次抽样检验就是一次随机试验。如果试验的目的是要考察次品的情况，那末试验的所有可能结果有"无次品"，"1件次品"，……"10件次品"等11个基本事件。"至少有5个次品"的事件是一个复合事件；"次品不多于10件"的事件是必然事件；"次品超过10件"的事件是不可能事件。

二、样本空间与随机事件

为了用集合论的观点来研究随机试验，我们先引入样本空间的概念。随机试验结果是明确的。随机试验的每一个可能结果称为**样本点**（即基本事件），常用ω表示。由所有样本点（基本事件）构成的集合称为**样本空间**（或**基本事件空间**），用Ω表示。显然$\Omega=\{\omega\}$。样本空间是由试验目的和试验结果的含义来决定的。同一个试验可以构成不同的样本空间。在具体问题中，确定样本空间是描述随机试验的第一步。在今后的研究中，常把样本空间认作是预先给定的。

引入样本空间后，就可把**事件定义为样本空间的某个子集**。因为样本空间的子集是由样本点组成的，所以事件是样本点的集合。**一个事件发生（或出现）当且仅当它所包含的某一样本点出现**。我们把样本空间Ω看成是必然事件，把空集$\varnothing$看成是不可能事件。

例1.1.1 连续掷硬币二次，每次可能出现结果有"国徽"和"钱面"两种，分别用"反"、"正"表示。若考察全部可能出现的结果，则全部样本点有

$\omega_1=$（正，正），$\omega_2=$（正，反），$\omega_3=$（反，正），$\omega_4=$（反，反）。故$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$。

若考察二次结果是否相同，则全部样本点有

$$\omega_1'=(\text{相同}),\qquad \omega_2'=(\text{不同}),$$

故$\Omega'=\{\omega_1',\ \omega_2'\}$。显然$\omega_1'=\{\omega_1,\omega_4\}$，$\omega_2'=\{\omega_2,\omega_3\}$，而事件

$$A=\{\text{恰有一次出现正面}\}=\{\omega_2,\omega_3\},$$

$$B=\{\text{至少一次出现正面}\}=\{\omega_1,\omega_2,\omega_3\},$$

$$C=\{\text{至多一次出现正面}\}=\{\omega_2,\omega_3,\omega_4\},$$

$$D=\{\text{至多二次出现正面}\}=\Omega,$$

$$E=\{\text{三次以上出现正面}\}=\varnothing。$$

例1.1.2 考虑某电话交换台在单位时间内收到的呼唤次数，其样本点有

$\omega_i=\{\text{收到呼唤次数为}i\}$，$i=0,1,2,\cdots$。

则$\Omega=\{0,1,2,\cdots\}=\{\omega_0,\ \omega_1,\cdots\}$。

例1.1.3 距离为L千米的A、B两地间的电缆线断了，考察断点的位置。令$\omega_s=\{\text{断点离}A\text{地距离为}S\text{千米}\}(0\leqslant S\leqslant L)$，则$\Omega=[0,\ L]$。

最简单的是由有限多个样本点构成的样本空间称为**有限样本空间**。由可数个（有限个或无限可列个）样本点构成的样本空间称为**可数样本空间或离散样本空间**。由不可数个样本点构成的样本空间称为**不可数样本空间**。

三、事件的关系和运算

为了研究事件的规律，就必须研究事件间的关系和运算，

从而加深对事件本质的认识。我们知道事件是样本空间的某一子集，那么就可用集合的关系和运算来定义事件的关系和运算。表1.1列出两个事件间的关系和运算。学习概率论应该习惯于将集合论的概念和结论"翻译"成概率论的概念和结论。

表1.1 事件的关系和运算

事　件	记　号	说　明
A包含B（B含于A）	$A\supset B$ ($B\subset A$)	B发生必导致A发生
A与B等价（相等）	$A=B$	A发生当且仅当B发生
A与B之和（并）	$A\cup B$	A与B至少发生一个
A与B之积（交）	$A\cap B$（AB）	A与B同时发生
A与B互不相容(互斥)	$AB=\varnothing$	A与B不能同时发生
A与B之差	$A\setminus B=A\overline{B}$	A发生而B不发生
A之逆（对立事件）	$\overline{A}=\Omega\setminus A$	A不发生

对于事件的并、交运算可推广到有限个或可列个的情形：

$\bigcup\limits_{i=1}^{n}A_i=A_1\cup A_2\cup\cdots\cup A_n$ 表示这n个事件至少发生一个的事件；$\bigcap\limits_{i=1}^{n}A_i$表示这n个事件同时发生的事件；$\bigcup\limits_{i=1}^{\infty}A_i=A_1\cup A_2\cup\cdots$表示$A_1,A_2\cdots$中至少发生一个的事件。当事件$A_1,A_2,\cdots,A_n$两两互斥时，$\bigcup\limits_{i=1}^{n}A_i$可记为$\sum\limits_{i=1}^{n}A_i$；当$A\supset B$时，$A\setminus B$可记为$A-B$。

为了直观理解事件的关系和运算，下面给出文(Venn)图。

例1.1.4 在1,2,…,10十个数字中任选一个，设$A=\{$选的数字号为偶数$\}$，$B=\{$选的数字号为小于5的偶数$\}$，$C=\{$选

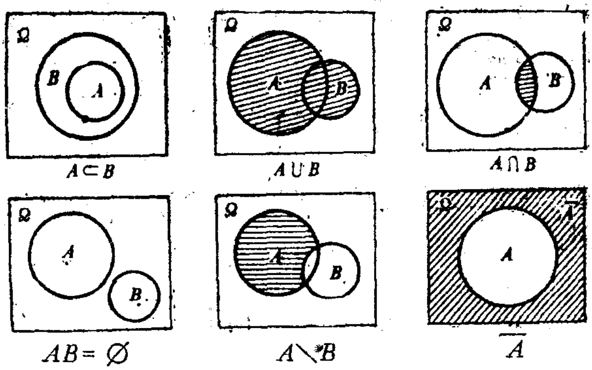

图 1.1.1

的数字号$\leqslant 5$}，则

$$A \cup C=\{1,2,3,4,5,6,8,10\}, \quad A \cap C=\{2,4\},$$

$$A \setminus B=\{6,8,10\}, \overline{A}=\{\text{选的数字号为奇数}\}。$$

例1.1.5 设A、B、C为事件，则事件

"A与B发生，但C不发生"可表示成$AB\overline{C}$；

"A、B、C中至少有两个发生"可表示成$AB \cup BC \cup CA$；

"A、B、C中恰有两个发生"可表示成$AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$。

与集合运算一样，事件运算有下面规律：

(1) **交换律** $A \cup B=B \cup A$；$A \cap B=B \cap A$。

(2) **结合律** $A \cup(B \cup C)=(A \cup B) \cup C$；

$A \cap(B \cap C)=(A \cap B) \cap C$。

(3) **分配律** $A \cap(B \cup C)=(A \cap B) \cup(A \cap C)$；

$A \cup(B \cap C)=(A \cup B) \cap(A \cup C)$。

(4) **对偶律** (De Morgen公式)

$$\overline{A \cup B}=\overline{A}\overline{B}; \qquad \overline{AB}=\overline{A} \cup \overline{B}。$$

事件的分配律和对偶律可推广到任意有限个甚至可列个的情形。例如对偶律

$$\overline{\bigcup_{i=1}^{n} A_i}=\bigcap_{i=1}^{n}\overline{A_i};\quad \overline{\bigcap_{i=1}^{n} A_i}=\bigcup_{i=1}^{n}\overline{A_i}。$$

上述这些事件运算性质的证明与集合运算相应性质的证明完全相同。

练　习　1.1

1. 袋中有5只同型号球，其中红、黄球各2只，白球1只，从中任取1只，写出它的样本空间，并写出下列事件的样本点集合：(1) 得红球；(2) 得黄球；(3) 得白球。

2. 从某班学生中任选一名代表，令$A=\{$被选学生为女生$\}$，$B==\{$被选学生是三好生$\}$，$C=\{$被选学生是班干部$\}$。试问：

(1) 事件$AB\overline{C}$的含义是什么？

(2) 在什么条件下$ABC=C$成立？

(3) 在什么条件下$\overline{A}=B$成立？

3. 设A、B、C为三个事件，试表示下列事件：

(1) A、B、C中至少有一个发生；

(2) A发生且B、C中仅有一个发生；

(3) A、B、C中恰有一个发生；

(4) A、B、C中至多有一个发生；

(5) A、B、C都不发生；

(6) A、B、C不都发生。

4. “A与B互斥”与“A与B互逆”有何关系？试举例说明。

5. 事件A、B、C两两互斥与$ABC=\varnothing$是否是一回事？为什么？

6. 试证：若$A\subset B$，则$\overline{A}\supset\overline{B}$。

7. 设A、B为事件，试证：

(1) $A=AB\cup A\overline{B}$；

(2) 若，$A \supset B$，则 $A = B \cup (A \setminus B)$，且 $B(A \setminus B) = \varnothing$。

8. 设 A、B 为事件，试问 $(A \cup B) \setminus B = A$，$(A \setminus B) \cup B = A$ 何时成立？

§1.2 统计概率

如何研究随机事件的规律性？在数学中只能从数量这个侧面来研究，也就是研究事件发生可能性的大小。直观地说，**事件A发生可能性大小的度量（数值）就是事件A发生的概率，**记作 $P(A)$。现在的问题是如何求事件的概率。最容易想到的一个方法就是通过大量重复的试验。

先介绍一个重要概念——事件的频率。设在 n 次重复试验中，事件A发生 μ_A 次，则称 μ_A 为事件 A 的**频数**，称 μ_A/n 为事件 A 的**频率**，记为 $f(A)$。例如，我们要求男婴出生的概率，以 A 表示男婴出生的事件，则可通过人口统计来求 A 的频率 $f(A)$。表1.2.1就是根据1964年浙江省台州地区人口普查资料统计的结果。从表中可以看出，当试验次数 n 充分大时，随着 n 的增大，事件 A 出现的频率围绕着某一个常数 p 附近摆动（当然

表1.2.1 1964年出生婴儿性别统计

县、区名	男 (μ_A)	女 (μ_B)	合计 (n)	男婴出生频率 $f(A) = \mu_A/n$
玉环	4606	4472	9078	0.5074
三门	4969	4693	9662	0.5143
天台	7266	6873	14139	0.5139
临海	16253	15423	31676	0.5131
黄岩	17179	16534	33713	0.5096
台州地区	72932	69224	142156	0.5130

也可能会出现少数次数较大的波动）。这种性质叫做频率的稳定性，而称p为稳定中心。那么p等于多少？前面我们已经指出，根据近代大量人口统计发现：男婴的出生率的稳定中心$p=22/43=0.5116$。从表中看出，各个频率与0.5116是很相近的。

历史上不少著名学者做过掷硬币试验。设A表示硬币出现正面向上的事件。从表1.2.2看出，A的频率的稳定中心$p=0.5$。

表1.2.2

试验者	掷币次数n	正面向上次数μ_A	频率μ_A/n
德莫根(De morgan)	4092	2048	0.5005
蒲丰(Buffon)	4040	2048	0.5069
费勒(W·Feller)	10000	4979	0.4979
皮尔逊(K·Pearson)	12000	6019	0.5016
皮尔逊(K·Pearson)	24000	12012	0.5005

人们通常总认为男、女婴出生与掷硬币出现正、反面一样，机会均等，其频率稳定中心是0.5。但事实表明，这种直观判断不是很精确的。事实上，事件频率的稳定性是不以人的意志为转移的事件固有的属性，在第四章还将给出严格的理论证明。因而，很自然地将频率的稳定中心作为事件发生可能性大小的一个客观的度量，也就是作为事件的概率。于是给出定义1.2.1。

定义1.2.1 如果事件A在n次重复试验中出现μ_A次，设频率$f(A)=\mu_A/n$的稳定中心为p，则称p为事件A发生的**概率**，即

$$P(A)=p$$