高等学校教材

概率论与数理统计

（第二版）

浙江大学　盛骤　谢式千　潘承毅　编

高等教育出版社

高等学校教材

概率论与数理统计

第二版

浙江大学　盛　骤　谢式千　潘承毅　编

高等教育出版社

（京）112号

内 容 提 要

本书是在1979年3月出版的《概率论与数理统计（工程数学）》第一版的基础上修订而成的，内容包括概率论、数理统计、随机过程三部分，每章附有习题，可以作为高等工业学校本科的教材使用，也可供工程技术人员参考.

图书在版编目（CIP）数据

概率论与数理统计/盛骤等编．—2版．—北京：高等教育出版社，1989.8（1997重印）
ISBN 7-04-001968-X

Ⅰ.概… Ⅱ.盛… Ⅲ.①概率论-高等教育-教材②数理统计-高等教育-教材 Ⅳ.021

中国版本图书馆CIP数据核字（96）第00056号

*

高等教育出版社出版
新华书店北京发行所发行
河北新华印刷一厂印装

*

开本850×1168 1/32 印张12.875 字数310 000
1979年3月第1版 1989年8月第2版 1997年8月第11次印刷
印数653 222—661 731
定价：11.70元

前　言

本书是在1979年出版的第一版的基础上修订的，可作为高等学校工科、理科（非数学专业）概率论与数理统计课程的教材，也可供工程技术人员参考.

本书分三部分. 概率论部分（第一章至第五章）作为基础知识，为读者提供了必要的理论基础. 数理统计部分（第六章至第九章）主要讲述了参数估计和假设检验，并介绍了方差分析和回归分析. 随机过程部分（第十章至第十二章）在讲清基本知识的基础上主要讨论了平稳随机过程，还介绍了马尔可夫过程. 数理统计和随机过程这两部分内容是相互独立的，可根据专业的需要选用.

在本书第一版出版后，我们经过进一步的教学实践，积累了不少的经验，并吸收了广大读者的意见，修订稿是在这一基础上写出的. 我们修改了第一版中存在的不当之处，并致力于教材质量的提高. 我们在选材和叙述上尽量做到联系工科专业的实际，注重应用，力图将概念写得清晰易懂，做到便于教学. 我们在例题和习题的选择上作了努力，这些题目既具有启发性，又有广泛的应用性，从题目的广泛性也可看到本门课程涉及生活和技术应用领域的广泛性. 读者将会发现，这些例题和习题是饶有趣味的. 为适应经济建设的需要，我们加强了数理统计的内容，例如编写了“矩估计法”、“样本容量的选取”和“正态分布的偏度、峰度检验”等，并有意识地加强学习者统计计算能力的培养.

书中的一部分内容能直接应用于解决实际课题，另一部分内

容为读者今后进一步学习有关课程或在实际应用方面提供一定的基础.

黄纪青同志曾参加过本书第一版编写大纲的讨论，撰写过第一版第一章的初稿.

本书的全部插图是由张礼明同志描绘的.

本书第二版承魏宗舒教授、林少宫教授、沈恒范教授、范大茵副教授、樊孝述副教授和汪振鹏副教授审阅，他们提出了很多宝贵意见，对此我们表示衷心的感谢.

书中不足之处，诚恳地希望读者批评指正.

盛　骤　谢式千　潘承毅

1988年1月

目　录

第一章　概率论的基本概念

自然界和社会上发生的现象是多种多样的．有一类现象，在一定条件下必然发生，例如，向上抛一石子必然下落，同性电荷必不相互吸引，等等．这类现象称为**确定性现象**．在自然界和社会上也还存在着另一类现象，例如，在相同条件下抛同一枚硬币，其结果可能是正面朝上，也可能是反面朝上，并且在每次抛掷之前无法肯定抛掷的结果是什么；用同一门炮向同一目标射击，各次弹着点不尽相同，在一次射击之前无法预测弹着点的确切位置．这类现象，在一定的条件下，可能出现这样的结果，也可能出现那样的结果，而在试验或观察之前不能预知确切的结果．但人们经过长期实践并深入研究之后，发现这类现象在大量重复试验或观察下，它的结果却呈现出某种规律性．例如，多次重复抛一枚硬币得到正面朝上大致有一半，同一门炮射击同一目标的弹着点按照一定规律分布，等等．这种在大量重复试验或观察中所呈现出的固有规律性，就是我们以后所说的统计规律性．

这种在个别试验中其结果呈现出不确定性；在大量重复试验中其结果又具有统计规律性的现象，我们称之为**随机现象**．概率论与数理统计是研究和揭示随机现象统计规律性的一门数学学科．

概率统计的理论与方法的应用是很广泛的，几乎遍及所有科学技术领域、工农业生产和国民经济的各个部门中．例如，使用概率统计方法可以进行气象预报、水文预报及地震预报，产品的抽样验收；在研制新产品时，为寻求最佳生产方案可用以进行试

验设计和数据处理；在可靠性工程中，使用概率统计方法可以给出元件或系统的使用可靠性及平均寿命的估计；在自动控制中可用以给出数学模型以便通过计算机控制工业生产；在通讯工程中可用以提高信号的抗干扰性和分辨率等等.

§1 随机试验

我们遇到过各种试验. 在这里，我们把试验作为一个含义广泛的术语. 它包括各种各样的科学实验，甚至对某一事物的某一特征的观察也认为是一种试验. 下面举一些试验的例子.

E_1：抛一枚硬币，观察正面 H、反面 T 出现的情况.

E_2：将一枚硬币抛掷三次，观察正面 H、反面 T 出现的情况.

E_3：将一枚硬币抛掷三次，观察出现正面的次数.

E_4：抛一颗骰子，观察出现的点数.

E_5：记录电话交换台一分钟内接到的呼唤次数.

E_6：在一批灯泡中任意抽取一次，测试它的寿命.

E_7：记录某地一昼夜的最高温度和最低温度.

上面举出了七个试验的例子，它们有着共同的特点. 例如，试验 E_1 有两种可能结果，出现 H 或者出现 T，但在抛掷之前不能确定出现 H 还是出现 T，这个试验可以在相同的条件下重复地进行. 又如试验 E_6，我们知道灯泡的寿命（以小时计）$t \geqslant 0$，但在测试之前不能确定它的寿命有多长. 这一试验也可以在相同的条件下重复地进行. 概括起来，这些试验具有以下的特点：

1° 可以在相同的条件下重复地进行；

2° 每次试验的可能结果不止一个，并且能事先明确试验的所有可能结果；

3° 进行一次试验之前不能确定哪一个结果会出现.

在概率论中，我们将具有上述三个特点的试验称为**随机试验**.

本书中以后提到的试验都是指随机试验.

我们是通过研究随机试验来研究随机现象的.

§2 样本空间、随机事件

(一) 样本空间 对于随机试验，尽管在每次试验之前不能预知试验的结果，但试验的所有可能结果组成的集合是已知的. 我们将随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**，记为 S. 样本空间的元素，即 E 的每个结果，称为**样本点**.

下面写出§1中试验 E_k ($k=1, 2, \cdots, 7$) 的样本空间 S_k：

S_1：$\{H, T\}$

S_2：$\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

S_3：$\{0, 1, 2, 3\}$

S_4：$\{1, 2, 3, 4, 5, 6\}$

S_5：$\{0, 1, 2, 3, \cdots\}$

S_6：$\{t \mid t \geqslant 0\}$

S_7：$\{(x, y) \mid T_0 \leqslant x \leqslant y \leqslant T_1\}$，这里 x 表示最低温度，y 表示最高温度. 并设这一地区的温度不会小于 T_0 也不会大于 T_1.

要注意的是：样本空间的元素是由试验的目的所确定的. 例如，在 E_2 和 E_3 中同是将一枚硬币连抛三次，由于试验的目的不一样，其样本空间也不一样.

(二) 随机事件 在实际中，在进行随机试验时，人们常常关心满足某种条件的那些样本点所组成的集合. 例如，若规定某种灯泡的寿命（小时）小于500为次品，则在 E_6 中我们关心灯泡的寿命是否有 $t \geqslant 500$. 满足这一条件的样本点组成 S_6 的一个子集：$A=\{t \mid t \geqslant 500\}$. 我们称 A 为试验 E_6 的一个随机事件. 显然，当且仅当子集 A 中的一个样本点出现时，有 $t \geqslant 500$.

一般，我们称试验 E 的样本空间 S 的子集为 E 的**随机事件**，

简称**事件**. 在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一**事件发生**.

特别，由一个样本点组成的单点集，称为**基本事件**. 例如，试验 E_1 有两个基本事件$\{H\}$和$\{T\}$；试验 E_4 有 6 个基本事件$\{1\}$，$\{2\}$，…，$\{6\}$.

样本空间 S 包含所有的样本点，它是 S 自身的子集，在每次试验中它总是发生的，称为**必然事件**. 空集$\varnothing$不包含任何样本点，它也作为样本空间的子集，它在每次试验中都不发生，称为**不可能事件**.

下面举几个事件的例子.

例 1 在 E_2 中事件 A_1："第一次出现的是 H"，即

$$A_1=\{HHH，HHT，HTH，HTT\}.$$

事件 A_2："三次出现同一面"，即

$$A_2=\{HHH，TTT\}.$$

在 E_6 中，事件 A_3："寿命小于 1000 小时"，即

$$A_3=\{t\,|\,0\leqslant t<1000\}.$$

在 E_7 中，事件 A_4："最高温度与最低温度相差 10 度"，即

$$A_4=\{(x,y)\,|\,y-x=10,T_0\leqslant x\leqslant y\leqslant T_1\}.$$

（三）事件间的关系与事件的运算 事件是一个集合，因而事件间的关系与事件的运算自然按照集合论中集合之间的关系和集合运算来处理. 下面给出这些关系和运算在概率论中的提法. 并根据"事件发生"的含义，给出它们在概率论中的含义.

设试验 E 的样本空间为 S，而 A，B，A_k（$k=1$，2，…）是 S 的子集.

1° 若 $A\subset B$，则称事件 B 包含事件 A. 这指的是事件 A 发生必导致事件 B 发生.

若 $A\subset B$ 且 $B\subset A$，即 $A=B$，则称事件 A 与事件 B **相等**.

2° 事件 $A\cup B=\{x|x\in A$ 或 $x\in B\}$ 称为事件 A 与事件 B 的**和事件**. 当且仅当 A，B 中至少有一个发生时，事件 $A\cup B$ 发生.

类似地，称 $\bigcup\limits_{k=1}^{n}A_k$ 为 n 个事件 A_1，A_2，…，A_n 的和事件；称 $\bigcup\limits_{k=1}^{\infty}A_k$ 为可列个事件 A_1，A_2,…的和事件.

3° 事件 $A\cap B=\{x|x\in A$ 且 $x\in B\}$ 称为事件 A 与事件 B 的**积事件**. 当且仅当 A，B 同时发生时，事件 $A\cap B$ 发生. $A\cap B$ 也记作 AB.

类似地，称 $\bigcap\limits_{k=1}^{n}A_k$ 为 n 个事件 A_1，A_2，…，A_n 的积事件；称 $\bigcap\limits_{k=1}^{\infty}A_k$ 为可列个事件 A_1，A_2，…的积事件.

4° 事件 $A-B=\{x|x\in A$ 且 $x\notin B\}$ 称为事件 A 与事件 B 的**差事件**. 当且仅当 A 发生、B 不发生时事件 $A-B$ 发生.

5° 若 $A\cap B=\varnothing$，则称事件 A 与 B 是**互不相容的**，或**互斥的**. 这指的是事件 A 与事件 B 不能同时发生. 基本事件是两两互不相容的.

6° 若 $A\cup B=S$ 且 $A\cap B=\varnothing$，则称事件 A 与事件 B 互为**逆事件**. 又称事件 A 与事件 B 互为**对立事件**. 这指的是对每次试验而言，事件 A、B 中必有一个发生，且仅有一个发生. A 的对立事件记为 $\overline{A}$. $\overline{A}=S-A$.

用图 1-1～1-6 可直观地表示以上事件之间的关系与运算. 例如，在图 1-1 中正方形表示样本空间 S，圆 A 与圆 B 分别表示事件 A 与事件 B，事件 B 包含事件 A. 又如在图 1-2 中正方形表示样本空间 S，圆 A 与圆 B 分别表示事件 A 与事件 B，而阴影部分表示和事件 $A\cup B$.

在进行事件运算时，经常要用到下述定律. 设 A，B，C 为事件，则有

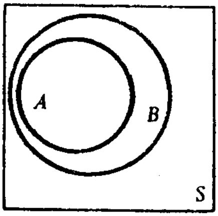

图 1-1

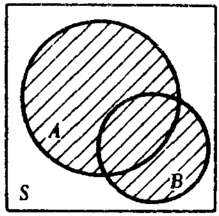

图 1-2

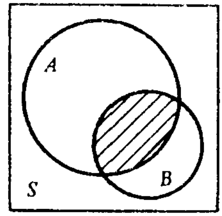

图 1-3

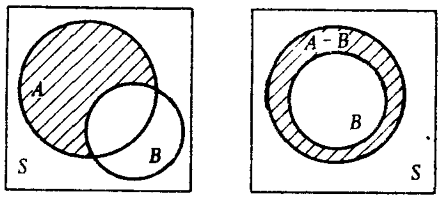

图 1-4

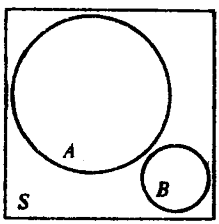

图 1-5

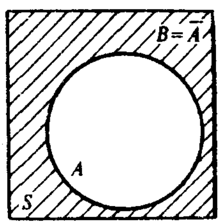

图 1-6

交换律：$A\cup B=B\cup A$；$A\cap B=B\cap A$.

结合律：$A\cup(B\cup C)=(A\cup B)\cup C$；

$A\cap(B\cap C)=(A\cap B)\cap C$.

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

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

德・摩根律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$；$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

例 2 在例 1 中有

$A_1\cup A_2=\{HHH,HHT,HTH,HTT,TTT\}$,

$$A_1 \cap A_2 = \{HHH\},$$
$$A_2 - A_1 = \{TTT\},$$
$$\overline{A_1 \cup A_2} = \{THT, TTH, THH\}.$$

例 3 如图 1-7 所示的电路中，以 A 表示“信号灯亮”这一事件，以 B,C,D 分别表示事件：继电器接点Ⅰ，Ⅱ，Ⅲ闭合，那么容易知道 $BC \subset A, BD \subset A, BC \cup BD = A$，而 $\overline{B}A = \varnothing$，即事件 $\overline{B}$ 与事件 A 互不相容.

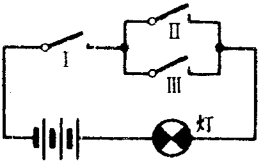

图 1-7

§3 频率与概率

对于一个事件（除必然事件和不可能事件外）来说，它在一次试验中可能发生，也可能不发生．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．例如，为了确定水坝的高度，就要知道河流在造水坝地段每年最大洪水达到某一高度这一事件发生的可能性大小．我们希望找到一个合适的数来表示事件在一次试验中发生的可能性大小．为此，首先引入频率，它描述了事件发生的频繁程度，进而引出表征事件在一次试验中发生的可能性大小的数——概率.

（一）**频率** **定义** 在相同的条件下，进行了 n 次试验，在这 n 次试验中，事件 A 发生的次数 n_A 称为事件 A 发生的**频数**．比值 n_A/n 称为事件 A 发生的**频率**，并记成 $f_n(A)$.

由定义，易见频率具有下述基本性质：

1° $0 \leqslant f_n(A) \leqslant 1$；

2° $f_n(S) = 1$；

3° 若 A_1，A_2，…，A_k 是两两不相容的事件，则

$$f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k).$$

由于事件 A 发生的频率是它发生的次数与试验次数之比，其大小表示 A 发生的频繁程度. 频率愈大，事件 A 发生愈频繁. 这意味着 A 在一次试验中发生的可能性愈大. 因而，直观的想法是用频率来表示 A 在一次试验中发生的可能性的大小. 但是否可行? 先看下面的例子.

例 1 考虑“抛硬币”这个试验，我们将一枚硬币抛掷 5 次、50 次、500 次，各做 10 遍. 得到 数据如表 1 所示（其中 n_H 表示 H 发生的频数，$f_n(H)$ 表示 H 发生的频率）.

表 1

实验序号	$n=5$		$n=50$		$n=500$	
	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

这种试验历史上有人做过，得到如表 2 所示的数据.

表 2

实 验 者	n	n_H	$f_n(H)$
德・摩根	2048	1061	0.5181
蒲 丰	4040	2048	0.5069
K・皮尔逊	12000	6019	0.5016
K・皮尔逊	24000	12012	0.5005

从上述数据可以看出：(1) 频率有随机波动性，即对于同样的 n，所得的 $f_n(H)$ 不尽相同；(2) 抛硬币次数 n 较小时，频率 $f_n(H)$ 随机波动的幅度较大，但随着 n 增大，频率 $f_n(H)$ 呈现出稳定性. 即当 n 逐渐增大时 $f_n(H)$ 总是在 0.5 附近摆动，而逐渐稳定于 0.5.

例 2 考察英语中特定字母出现的频率. 当观察字母的个数 n（试验的次数）较小时，频率有较大幅度的随机波动. 但当 n 增大时，频率呈现出稳定性. 下面就是一份英文字母频率的统计表[①]：

字母	频率	字母	频率	字母	频率
E	0.1268	L	0.0394	P	0.0186
T	0.0978	D	0.0389	B	0.0156
A	0.0788	U	0.0280	V	0.0102
O	0.0776	C	0.0268	K	0.0060
I	0.0707	F	0.0256	X	0.0016
N	0.0706	M	0.0244	J	0.0010
S	0.0634	W	0.0214	Q	0.0009
R	0.0594	Y	0.0202	Z	0.0006
H	0.0573	G	0.0187		

从上面两个例子中可以看出，当 n 较小时，频率 $f_n(A)$ 在 0 与 1 之间随机波动，其幅度较大. 因而，当 n 较小时用频率来表达事件发生的可能性的大小显然是不合适的. 而当 n 逐渐增大时，频率 $f_n(A)$ 逐渐稳定于某个常数. 对于每一个事件 A 都有这样一个客观存在的常数与之对应. 这种"频率稳定性"即通常所说的统计规律性，不断地为人类的实践所证实，它揭示了隐藏在随机现象中的规律性. 我们用这个频率稳定值来表示事件发生的可能性大小是合适的.

---

① 这是由 Dewey，G，统计了约 438023 个字母得到的. 引自 Relative Frequency of English Spellings (Teachers College Press, Columbia University, New York, 1970).