概率与统计

赖景耀
王春林　编著

兰州大学出版社

前 言

随着社会的发展，对随机现象规律性的研究逐步广泛地渗透到自然科学、社会科学与人们的日常生活中。作为以随机现象为研究对象的概率论与数理统计，对于现代社会的每个成员来说，学习它、了解它、掌握它已成为时代发展的要求。基于此，我们撰写了这部书。

为使《概率与统计》有自己的特色，在写作中，我们注意了以下几个方面：在理论上，力求严谨，并要求反映相关领域中的最新成果；在阐述上，追求深入浅出，既考虑到理论的完备，又兼顾实际背景的介绍，增强全书的趣味性、可读性；在实用上，尽量照顾不同职业、不同专业人士的要求，大量选用来自各行各业的应用实例，体现概率与统计应用的广泛性、灵活性。

作者长期在高等学校从事概率论与数理统计教学、研究与应用，这本书对多年来的经验积累作了一定的总结和反映，愿为社会所接受，希望对同仁、对读者有所启迪。

在撰写、出版过程中，得到西北师范大学数学与信息科学学院霍念祖、郝黎仁、马国顺、姚兵等几位老师的帮助，特此致谢！

全书付梓之际，谨向多年来支持、关心我们的亲人、老师、同学深表谢意！

书中存在的问题，诚望读者批评指正。

赖景耀　王春林

2001 年 7 月 1 日

目　录

第一章　随机事件与概率

§ 1.1　随机事件

在概率论中，我们把实现一组条件称为一次“试验”，记为 E.

若某试验 E 满足：1）试验可以重复进行；2）试验的所有结果不止一个，且在试验之前能预计到可能出现的每个结果；3）每次试验仅能有一个结果出现，但试验前无法预知，则称 E 为一个**随机试验**.

例如 E_1：掷两枚钱币，用 H 表示“出正面”，T 表示“出反面”．试验前知道可能出现的所有结果：$(H、H)$、$(H、T)$、$(T、H)$、$(T、T)$，但只有试验之后方可知晓四个结果中哪一个会出现.

E_2：记录电话交换站单位时间接到的呼叫次数，试验前能预计到次数可能为 0, 1, 2, 3, …，但只有试验结束方可知晓本次试验的具体“次数”.

E_3：观察炮弹落点位置，炮弹落点在某一平面之内是可能预先估计到的：

$$\{(x, y)|\, a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$$

但只有炮弹落下之后才能知晓落点确切位置.

以上 E_1, E_2, E_3 均满足上述三条，故是随机试验．以下，我们所指的“试验”，均指“随机试验”.

应当指出的是，我们所讲的试验，是一个含义广泛的术语，某一具体的科学实验是一次试验，对某一事物的某一特征的观察也是一次试验.

随机试验的一个结果称为事件，事件可分为：

必然事件：在一定条件下一定会发生的事件叫必然事件，记为Ω；

不可能事件：在一定条件下一定不会发生的事件叫不可能事件，记为 ∅；

随机事件：在一定条件下可能发生，也可能不发生的事件叫随机事件上，记为 A、B、C、…. 随机事件也简称事件.

基本事件：在一定条件下，不可能再分的结果叫基本事件，记为 e_i, $i=1, 2, \cdots$.

在一个试验 E 中，基本事件是最小单元，事件均由基本事件构成. 例如：掷一枚均匀的骰子，其基本事件有 $e_1=1, e_2=2, \cdots, e_6=6$（$e_i=i$ 表示出 i 点）. 而"掷一枚均匀的骰子出点为偶数"."掷一枚均匀的骰子出点为不大于 7"，"掷一枚均匀的骰子出点为−3"都是事件，若用集合论的表述方法，它们分别写为：

$$A=\{e_2, e_4, e_6\};$$

$$\Omega=\{e_1, e_2, e_3, e_4, e_5, e_6\};$$

$$\varnothing=\{\quad\}.$$

一个事件是必然事件，不可能事件或随机事件，是相对的，随条件的变化而变化.

例如　一个工人每天生产零件 3000 只左右（条件&），则 1）"每天生产零件至少生产 1 只"事件是必然事件；2）"每天生产零件超过 5000 只"事件是不可能事件；3）"每天生产零件 3001 只左右"事件是随机事件.

若条件&变为&′：经技术革新，工效提高，工人每天生产零件 6000 只，则以上三个事件分别是：1）必然事件；2）必然事件；3）不可能事件.

为了便于研究，我们将随机试验的所有可能结果组成的集合称为 E 的**样本空间**，记为Ω，Ω中的元素亦即试验 E 的基本事件，基本事件在样本空间中也可称为**样本点**. 即有：

$$\Omega=\{e_1, e_2, \cdots\}$$

在前边列举的例中，E_1, E_2, E_3 的样本空间分别写为

$$\Omega_1=\{(H、H),(H、T),(T、H),(T、T)\};$$

$$\Omega_1=\{0, 1, 2, \cdots\};$$

$$\Omega_3=\{(x, y) \mid a \leqslant x \leqslant b, c \leqslant y \leqslant d\}.$$

样本空间中的元素是基本事件，而基本事件的确定不是唯一的，它是由试验目的确定的. 例如试验 E_4 是将一枚钱币掷三次，观察出现正、反面的情况，则共有 $2^3=8$ 个基本事件，

$$\Omega_4=\{(H、H、H),(H、H、T),(H、T、H),(T、H、H),(H、T、T),(T、H、T),(T、T、H),(T、T、T)\}.$$

又若试验 E_5 是将一枚钱币掷三次，观察其反面出现的次数. 设 e_i 表示出三次中反面出现 i 次，则仅有 4 个基本事件，且

$$\Omega_5=\{e_0, e_1, e_2, e_3\}.$$

§ 1.2 事件之间的关系及运算

事件之间的关系及运算可按照集合论中集合之间的关系及运算来描述.

一、事件之间的关系

定义 1 若事件 A 发生必然导致事件 B 发生，则称 B 包含事件 A，或称事件 A 含于事件 B，记为 $A \subset B$ 或 $B \supset A$.

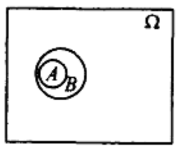

图 1-1

对定义 1 可作这样的理解：

$\forall e_i, \in A, i=1, 2, \cdots$, 有 $e_i \in B$, 则 $A \subset B$.

也可用“**文氏图**”来帮助理解（图 1-1）.

定义 2 对事件 A 与 B，若 $A \subset B$，且 $A \supset B$，则称事件 A 与 B 相等，记为 $A=B$.

定义 3　事件 A、B 中至少有一个发生而组成的新事件 C，称为 A 与 B 的**和事件**，或**并事件**，记为 $C=A+B$ 或 $C=A\cup B$.

对定义 3 可作这样的理解（图 1-2）：

$\forall e_i\in C,\ i=1, 2, \cdots$，则 $e_i\in A$ 或 $e_i\in B$，反之亦然.

定义 3 可推广至 n 个事件的情形. 对事件 $A_i,\ i=1, 2, \cdots, n$，其有限和（并）为 $A=\sum\limits_{i=1}^{n}A_i$（或 $\bigcup\limits_{i=1}^{n}A_i$），表示 $A_1, A_2, \cdots, A_n$ 中至少有一个事件发生. 还可推广到可列个事件的情形. 其可列和（并）为 $A=\sum\limits_{i=1}^{\infty}A_i$（或 $\bigcup\limits_{i=1}^{\infty}A_i$）.

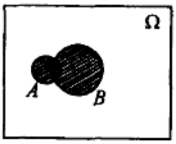

图 1-2

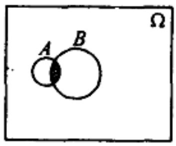

图 1-3

定义 4　事件 A、B 同时发生而组成的新事件 C，称为 A 与 B 的**积事件**，或**交事件**，记为 $C=A\cdot B\ (AB)$ 或 $C=A\cap B$.

定义 4 可作这样的理解（图 1-3）：$\forall e_i\in C,\ i=1, 2, \cdots$，则 $e_i\in A$ 且 $e_i\in B$，反之亦然.

定义 4 可推广到 n 个事件及可列个事件的情形. 对事件 $A_i,\ i=1, 2, \cdots, n$ 及 $A_i,\ i=1, 2, \cdots$，其有限积（交）及可列积（交）为：

$$A=\prod_{i=1}^{n}A_i \quad (\text{或 } A=\bigcap_{i=1}^{n}A_i);$$

$$A=\prod_{i=1}^{\infty}A_i\ (\text{或 } A=\bigcap_{i=1}^{\infty}A_i).$$

定义 5　事件 A 与 B 不能同时发生，则称 A 与 B **互不相容**，或**互斥**，记为 $A\cdot B=\varnothing$.

定义 5 可作这样的理解（图 1-4）：$\forall e_i$、e_j，$e_i\in A$, $e_j\in B$, $i, j=1, 2, \cdots$，则 $e_i\notin B$ 且 $e_j\notin A$.

定义 5 可推广到 n 个事件的情形. 对事件 $A_1, A_2, \cdots, A_n$，当且仅当 $A_1A_2\cdots A_n=\varnothing$，则称这 n 个事件**彼此互不相容（彼此互斥）**；又，当且仅当 $\forall i, j$，有 $A_i\cdot A_j=\varnothing$（$i\neq j$, $i, j=1, 2, \cdots, n$,），则称这 n 个事件**两两互不相容（两两相斥）**.

易见，彼此互不相容与两两不互相容有以下关系：若 n 个事件两两互不相容，则必有彼此互不相容，反之未必.

定义 6 如果事件 A 发生，则事件 B 一定不会发生，且在一次试验中，两者必发生其一，则称 B 为 A 的**逆事件**，记为 $\overline{A}=B$.

B 为 A 的逆事件，亦即 A 为 B 的逆事件，常称 A 与 B 互为逆事件，逆事件也叫**对立事件**.

定义 6 可作这样理解（图 1-5）：

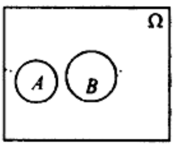

图 1-4

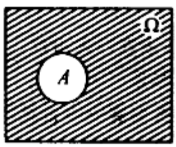

图 1-5

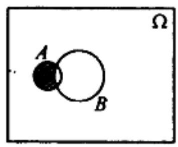

图 1-6

对事件 A、B，若 $A+B=\Omega$，且 $A\cdot B=\varnothing$，则 $\overline{A}=B$ 且 $\overline{B}=A$.

对于两个事件 A、B，定义 5、6 中的两个概念有以下关系：A 与 B 互为逆事件，则 A 与 B 互不相容，反之未必.

定义 7 事件 A 发生但事件 B 不发生而组成的新事件 C，称为 A 与 B 的**差事件**，记为 $A-B$.

定义 7 可作这样理解（图 1-6）：$\forall e_i\in C$, $i=1, 2, \cdots$，则 $e_i\in A$ 且 $e_i\notin B$.

综合使用以上事件之间的各种关系，可以更好地表述事件，例如，$\overline{A}=\Omega-A$, $A-B=A\overline{B}$. 证明这类等式并不难，常使用一种递推法，现举第二式证明.

求证 $A-B=A\overline{B}$.

证 $\forall e_i\in A-B$, $i=1, 2, \cdots$, 则 $e_i\in A$ 且 $e_i\notin B$，亦即 $e_i\in A$ 且 $e_i\in\overline{B}$，故 $e_i\in A\overline{B}$，有 $A-B\subset A\overline{B}$；又$\forall e_i\in A\overline{B}$，$i=1, 2, \cdots$, 则 $e_i\in A$，且 $e_i\in\overline{B}$，亦即 $e_i\in A$ 且 $e_i\notin B$，故 $e_i\in A-B$, 有 $A-B\supset A\overline{B}$.

综合 $A-B\subset A\overline{B}$ 与 $A-B\supset A\overline{B}$，有 $A-B=A\overline{B}$.

事件与集合之间的对应关系可见下表：

表 1-1

记　号	概率论	集合论
Ω	必然事件，基本事件空间	抽象点集
$\varnothing$	不可能事件	空集
e	基本事件	点（元素）
A	事件	Ω的子集
$e\in A$	事件 A 出现	e 是 A 中的点
$A\subset B$	事件 A 出现事件 B 一定出现	A 是 B 的子集
$A=B$	二事件 A、B 相等	二集合 A、B 相等
$A\cup B$	二事件 A B 至少有一个出现	二集合 A、B 的并集
$A\cap B$	二事件 A、B 同时出现	二集合 A、B 的交集
$A-B$	事件 A 出现而 B 不出现	二集合 A、B 的差集
$\overline{A}$	A 的对立事件	A 对Ω的补集
$A\cap B=\varnothing$	二事件 A、B 互不相容	二集合 A、B 不相交

二、事件的运算

事件的运算，常用到以下运算律：

交换律：$A+B=B+A$；

$AB=BA$.

结合律：$A+(B+C)=(A+B)+C$;

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

分配律：$A+(B\cdot C)=(A+B)\cdot(A+C)$;

$$A\cdot(B+C)=A\cdot B+A\cdot C.$$

摩根（De Morgen）律：$\overline{A+B}=\overline{A}\cdot\overline{B}$；

$$\overline{A\cdot B}=\overline{A}+\overline{B}.$$

以上运算律可推广到 n 个事件的情形，如对摩根律，有

$$\overline{A_1+A_2+\cdots+A_n}=\overline{A_1}\ \overline{A_2}\cdots\overline{A_n};$$

$$\overline{A_1A_2\cdots A_n}=\overline{A_1}+\overline{A_2}+\cdots+\overline{A_n}.$$

除以上运算律外，还有一些显然的运算规则：

$\overline{\overline{A}}=A,\ \overline{\overline{\overline{A}}}=\overline{A},\cdots$；

$A+A=A,\quad AA=A$；

$\Omega+A=\Omega,\quad \Omega A=A$；

$A\supset B$，则 $A+B=A, AB=B$.

以上各式的证明，或用恒等变换，或用前述递推方法，现举两例加以证明.

求证 $A+(BC)=(A+B)(A+C)$

证 $(A+B)(A+C)=(A+B)A+(A+B)C=AA+BA+AC+BC$，注意到 $AA=A, BA+AC\subset A$. 有

$$(A+B)(A+C)=A+(BA+AC)+BC=A+BC$$

求证 $\overline{A+B}=\overline{A}\overline{B}$

证 $\forall e\in\overline{A+B}$，则 $e\notin A+B$，

即 $e\notin A$ 且 $e\notin B$，亦即 $e\in\overline{A}$ 且 $e\in\overline{B}$，有 $e\in\overline{A}\ \overline{B}$.

故 $\overline{A+B}\subset\overline{A}\overline{B}$；

又，$\forall e\in\overline{A}\ \overline{B}$，则 $e\in\overline{A}$ 且 $e\in\overline{B}$，

即 $e\notin A$ 且 $e\notin B$，亦即 $e\notin A+B$，有 $e\in\overline{A+B}$，

故 $\overline{A}\overline{B}\subset\overline{A+B}$.

综合 $\overline{A+B}\subset\overline{A}\overline{B}$ 和 $\overline{A+B}\supset\overline{A}\overline{B}$，得 $\overline{A+B}=\overline{A}\overline{B}$.

利用事件之间的关系及运算，可以根据提出的问题写出事件，并能将事件简化.

例 2.1 设 A、B、C 为三个事件，则

事件“A、B 发生但 C 不发生”可以表示为 $AB\overline{C}$ 或 $AB-C$;

事件“A、B、C 中至少有两个发生”;可以表示为 $A\overline{B}C+AB\overline{C}+\overline{A}BC+ABC$. 经过化简后可写为 $AB+AC+BC$;

事件“A、B、C 中恰有一个发生”可以表示为 $A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C$;

事件“A、B、C 中发生不多于一个”可以表示为 $A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C+\overline{A}\overline{B}\overline{C}$. 或简化为 $\overline{AB+AC+BC}$.

例 2.2 在图书馆书库中任取一本书，A 表示“取到数学书，”B 表示“取到中文版书”，C 表示“取到 1990 年以前出版的书”. 以下事件及关系的实际意义为：

ABC 表示取到的书是 1990 年以前出版的中文版数学书；

$A\overline{B}\overline{C}$ 表示取到的书是 1990 年以后出版的非中文版数学书；

$ABC=C$ 表示取到 1990 年以前出版的书都是中文版数学书；

$\overline{C}=B$ 表示图书馆的中文版书都不是 1990 年以前出版的书，反之亦然.

§ 1.3 概率的统计定义

概率，作为概率论的最基本概念，需要从不同角度去描述它. 概率的统计定义，就是从寻求事件发生可能性大小统计规律的意义下，以频率描述概率.

频数 在 n 次重复试验中，事件 A 发生的次数叫频数，记为 n_A.

频率 称 $\dfrac{n_A}{n}$ 为事件 A 发生的频率，记为 $f_n(A)$.

频率具有以下基本性质：

性质 1° $f_n(A)\geqslant 0$ （非负性）；

性质 2° $f_n(\Omega)=1$ （规范性）；

性质 3° 若事件 A 与 B 互不相容，则

$$f_n(A+B)=f_n(A)+f_n(B) \quad \text{（可加性）}$$

性质的证明是显然的，略. 对性质 3°，可推广到有限个事件的情形. 对事件 $A_1, A_2, \cdots, A_n$ 两两互不相容，则

$$f_n\left(\sum_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} f_n(A_i) \quad \text{（有限可加性）}$$

频率还具有一种特殊的性质，**叫频率的稳定性**. 当试验次数不多时，事件出现的频率往往有很大的随机性，而当试验次数增多时，频率逐渐失去了随机性，呈现出稳定于某一数值的趋势，这一性质叫频率的稳定性.

例 3.1 掷一枚钱币，若该钱币是"均匀"的（自然对称性），则"出正面"的频率随着试验次数的增多，与 0.5 的差别很小. 表 1-2 是抛掷次数分别为 5 次、50 次、500 次的试验记录.

由表 1-2 可见，当 $n=5$ 时，随机性太大，而当 $n=500$ 时，统计规律性就显示出来了. 表 1-3 是历史上几个有名的试验记录.

表 1-2

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

表 1-3

实验者	掷的次数 n	正面出现的频数 n_A	正面出现的频率 $f_n(A)$
蒲　丰	4040	2048	0.5069
摩　根	4092	2048	0.5005
杰万斯	20480	10379	0.5068
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005
费　勒	10000	4979	0.4979
罗曼诺夫斯基	80640	39699	0.4923

例 3.2　记录一位母亲的一次生育生男生女情况，研究男、女出生率．我国早在公元前 2238 年，就根据人口普查得出生男生女的频率各占 0.5 左右．下表反映瑞典某年的统计资料．

表 1-4

月　份	总　数	男　孩	女　孩	女孩频率
1	7280	3743	3537	0.486
2	6957	3550	3407	0.489
3	7883	4017	3866	0.490
4	7884	4173	3711	0.471
5	7892	4117	3775	0.478
6	7609	3944	3665	0.482
7	7585	3964	3621	0.462
8	7393	3797	3596	0.484
9	7203	3712	3491	0.485
10	6903	3512	3391	0.491
11	6552	3392	3160	0.482
12	7132	3761	3371	0.473
全年	88273	45682	42591	0.4825

生男生女各占一半这一规律，几乎对于所有国家和地区都是稳定的．瑞典统计 1871-1900 年三十年间，得出 f_n(女)=0.4860；

数学家拉普拉斯统计了欧洲许多国家，几乎都有 $f_n(女)=0.4884$. 我国第三次全国人口普查得到的结果为 $f_n(女)=0.4848$.

例 3.3 英文字母出现的频率对于不同性质的书籍，只要使用的字母数多，也几乎是稳定的（表 1-5）.

表 1-5

字 母	频 率	字 母	频 率	字 母	频 率
空格(标点)	0.2	H	0.047	W	0.012
E	0.105	D	0.035	G	0.011
T	0.072	L	0.029	B	0.0105
O	0.0654	C	0.023	V	0.008
A	0.063	F	0.0225	K	0.003
N	0.059	U	0.0225	X	0.002
I	0.055	M	0.021	J	0.001
R	0.054	P	0.0175	Q	0.001
S	0.052	Y	0.012	Z	0.001

由随机试验中事件发生频率的稳定性可见事件概率的存在性，因为频率与概率都是刻划随机事件发生可能性大小的一种度量. 由此，我们给出概率的统计定义.

定义 1 在一定条件下，当试验次数增多，事件 A 出现的频率 $f_n(A)$稳定，则称 $f_n(A)$为事件 A 发生的概率，记为 $P(A)$[①]，即有

$$P(A)\approx f_n(A).$$

概率的统计定义为我们提供了通过大量试验求概率的方法，这就是使用频率的稳定值. 但应当指出，这种方法求得的数值是近似值，如掷一枚均匀钱币，实际频率为 0.5005，（皮尔逊的试验结果），而后边我们会求得，理论概率值为 0.5.

历史上，有人认为（见 R.Von Mises 的《概率论学说与物理统计学原理》，载《物理学成就》，卷 IX，第二期），概率的统计定义是在极限场合下有

[①]英文“概率”为 Probability，记号用第一字母“p”.

$$P(A)=\lim_{n\to\infty} f_n(A)=\lim_{n\to\infty}\frac{n_A}{n}.$$

这是不对的，读者可思考其理由，关于这一点，我们在第七章中还要从理论上进行讨论.

这样定义的概率，具有以下性质：

性质 1° $P(A)\geqslant 0$ （非负性）；

性质 2° $P(\Omega)=1$ （规范性）；

性质 3° 若事件 A 与 B 互不相容，则

$$P(A+B)=P(A)+P(B) \quad (\text{可加性}).$$

性质 3°可推广到多个两两互不相容的事件，即有有限可加性

$$P\left(\sum_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i);$$

可列可加性

$$P\left(\sum_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i).$$

性质的证明是显然的，可直接应用频率的性质推得，略.

§1.4 概率的古典定义

在引进古典定义之前，我们先了解一下**古典概型**. 所谓古典概型，是一种具有有限性和等可能性的概率模型.

一试验若满足：1）仅有有限个结果；2）每个基本事件出现的可能性都相同，则称此概型为古典概型. 亦即，设$\Omega=\{e_1, e_2, \cdots, e_n\}$，有 $P(e_1)=P(e_2)=\cdots=P(e_n)$.

前例中，掷钱币、掷骰子，若钱币、骰子都是“均匀”的（即满足自然对称性），它们都属于古典概型. 若不“均匀”，则不是