高等学校教材

概率论与数理统计

内蒙古工业大学数学系 编

内蒙古教育出版社

前　言

概率论与数理统计是研究随机现象统计规律性的一门学科，它的思想与方法在农、工、商、医等社会、经济的许多领域以及自然科学的诸多学科中得到了广泛的应用。因而它已经成为高等理工科院校、经济类院校的一门重要的公共必修基础课程。诚然，有一本好的教材是学好这门课的首要前提。

本书根据教育部颁布的高等学校(工科)本科基础课教学基本要求(概率统计部分)，由内蒙古工业大学数学系几位教师在多年讲授《概率论与数理统计》课的基础上编写而成教材。前五章为概率论，约需 34 学时，后三章为数理统计，约需 20 学时。在编写中融入了内蒙古工业大学数学系教师多年积累的经验与资料，用更为直观、更易于学生接受的方式讲述概率论与数理统计的基本思想和方法，尽力做到叙述准确、简洁、通俗，选用例题与习题典型、规范、由浅入深、易于计算。书后附有总复习题，这些习题中的一些本身是正文的补充和一些有趣的命题。对于那些想考研究生及想深入学习概率论与数理统计的读者，这些习题有助于他们进一步掌握有关理论内容。

本书的编写得到了学校教务处、理学院的大力支持与鼓励。我系直接参加具体工作的有：谭福贵撰写了第二章、第八章；闫在在撰写了总复习题及附录；郑丽霞撰写了第一章、第四章、第五章；斯日古楞撰写了第三章、第六章、第七章，由闫在在统稿，校验了全书的内容与语言。

本书第一版面世，真诚希望得到读者的批评与指正。

内蒙古工业大学数学系

2004 年 6 月

内容简介

本书根据高等学校(工科)本科基础课程教学基本要求编写,内容包括随机事件与概率,随机变量及其分布,多维随机变量及其分布,随机变量的数字特征,大数定律和中心极限定理,抽样分布,参数估计,假设检验共八章。

本书可供高等理工科院校、经济类院校作为教材或教学参考书,也可供科技工作者参考使用。

目　　录

第一章　事件及其概率

§1.1　随机试验与随机事件

一、随机现象

客观世界中发生的各种各样的现象，大体上可分为两种：一为确定性现象，另一为不确定现象. 例如抛出一枚硬币，硬币必然要落地；水在标准大气压下达到100℃必然沸腾，变为气体，而低于0℃必然结冰，变为固体. 这些在一定条件下必然发生的现象称为**确定性现象**. 但是抛出的一枚硬币落地后，是正面(有币值的一面)向上，还是反面向上，在硬币落地之前是不确定的；一只灯泡能用多少小时，在使用之前是无法确定其寿命的；今年某地区十月份的平均气温是多少，在十月份结束之前是不能确定的. 这种在一定条件下，具有多种可能结果，但事先又不能确定究竟会发生哪一种结果的现象称为**随机现象**.

随机现象在一次观察或试验中，具有随机性，即偶然性. 但是人们在长期实践中发现，相同条件下，对随机现象进行大量的观察或试验时，随机现象的结果会呈现出某种确定的规律性. 例如抛出一枚硬币，是正面向上，还是反面向上，完全是偶然的、随机的，但在相同条件下，将一枚硬币抛掷100次，1000次，…，会发现出现正面的次数与出现反面的次数，几乎各占一半. 这就是说，看上去其结果具有不确定性的随机现象内部却蕴含着某种确定的规律性. 这种通过大量观察或试验总结出的规律性，称为**统计规律性**.

概率论与数理统计就是研究随机现象的统计规律性的一门数

学学科.

在自然界和人们的生产、生活实践中,存在着大量的随机现象.这种随机现象是受其内部蕴含着的规律所支配的.研究和发现这些规律,揭示偶然性与必然性之间的内在联系,是我们学习概率统计的方向和最终目的.这也就决定了概率统计的思想和方法在自然科学、社会科学,工农业生产及国民经济的各个部门中有着广泛的应用,并且与其它学科相互结合、渗透,推动和发展了许多边缘学科.因此,初步掌握处理随机现象的基本理论和方法,越来越成为一种基础知识和基本技能了.

二、随机试验

为了研究随机现象,就必须对其观察.这里,我们把对现象的观察、测量、记录、实验等统称为试验.而把具有下述三个特征的试验称为**随机试验**:

(1) 试验可以在相同的条件下重复进行;

(2) 每次试验的可能结果不止一个,且在试验前能预知试验的所有结果;

(3) 在每次试验之前无法断言哪个结果会出现,但若进行大量重复试验,其可能结果的出现具有一定的统计规律性.

随机试验,简称试验.今后我们说试验,就是指随机试验.下面给出随机试验的几个例子:

E_1 = {将一枚硬币投掷 3 次,观察其正、反面出现的情况};

E_2 = {将一枚硬币投掷 3 次,观察其正面出现的次数};

E_3 = {掷一颗骰子,观察其点数};

E_4 = {同时掷两颗骰子,观察其点数出现的情况};

E_5 = {炮击目标直到击中为止,记录发出的炮弹数};

E_6 = {一只灯泡,测试其使用寿命}.

三、样本空间

我们把根据观察要求所确定的随机试验最基本，不能再分解的结果叫做**基本事件**，其特点是：每次试验必出现一个而且只能出现一个基本事件，任何两个基本事件在一次试验中不能同时出现。把一切基本事件的集合 S 叫做**样本空间**。样本空间 S 的元素，即基本事件又称为**样本点**。

对应于上段给出的随机试验，样本空间分别为：

$S_1=\{HHH,HHT,HTH,HTT,THH,THT,TTH,TTT\}$

（其中 H 表示正面，T 表示反面）；

$S_2=\{0,1,2,3\}$；

$S_3=\{1,2,3,4,5,6\}$；

$S_4=\{(x,y)\mid 1\leqslant x\leqslant 6,1\leqslant y\leqslant 6,x\in N,y\in N\}$；

$S_5=\{1,2,3,4,\cdots\}$；

$S_6=\{x\mid 0\leqslant x<+\infty\}$.

四、随机事件

定义　设随机试验 E 的样本空间为 S，称 S 的任一子集合为随机试验 E 的一个**随机事件**，简称**事件**，记为 $A,B,C,\cdots$. 设 A 是 E 的一个事件，在一次试验中，当且仅当 A 中的某一个样本点出现时，则称事件 A 发生了.

只包含一个样本点的集合，就是基本事件. 显然，样本空间 S 包含多少个样本点，就包含多少个基本事件. 事件 A 包含几个样本点，也就相当于包含几个基本事件.

作为子集的特例，空集 $\Phi\subset S$，称 Φ 为**不可能事件**. 它在每次试验中，都不可能发生.

又因为 S 也是 S 自身的子集，$S\subset S$，称 S 为**必然事件**. 它在每次试验中，都必然发生.

例如前段 $E_4=\{$同时掷两颗骰子,观察其点数出现的情况$\}$,对应的 $S_4=\{(x,y)\mid 1\leqslant x\leqslant 6,1\leqslant y\leqslant 6,x\in N,y\in N\}$. 若令 $A=\{$两颗骰子点数和为 5$\}$,则 $A=\{(1,4),(2,3),(3,2)(4,1)\}\subset S_4$;$A$ 是一个随机事件. 若令 $B=\{$两颗骰子点数相同$\}$,则 $B=\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}\subset S_4$,$B$ 也是一个随机事件. 若令 $C=\{$两颗骰子点数和为 12$\}$,则 $C=\{(6,6)\}\subset S_4$,C 是一个基本事件. 若令 $D=\{$两颗骰子点数和为 13$\}$,则 $D=\Phi$,D 是不可能事件. 若令 $F=\{$两颗骰子点数之差不超过 5$\}$,则 $F=S_4$,F 是必然事件.

五、事件的关系与运算

设随机试验 E 的样本空间为 S,A,B,C 以及 $A_1,A_2,A_3,\cdots$ 均为 S 的子集,即均为随机事件. 用集合表示事件,从而集合间的某些关系与运算,也就表达了事件间的相应的关系与运算.

(1) $A\subset B$,称为事件 B 包含事件 A,或者说事件 A 是事件 B 的子事件. 在一次试验中事件 A 发生时,事件 B 必发生.

(2) $A=B$,称为事件 A 与事件 B **相等**. 其充分必要条件是 $A\subset B$ 且 $B\subset A$. 它表示,事件 A 与事件 B 是同一事件.

(3) $A\cup B$,称为事件 A 与事件 B 的**和(并)事件**. 在一次试验中,事件 $A\cup B$ 发生当且仅当事件 A 与事件 B 至少有一个发生.

(4) $A\cap B$ 也记为 AB,称为事件 A 与事件 B 的**积(交)事件**. 在一次试验中,事件 AB 发生当且仅当事件 A 与事件 B 同时发生.

和事件与积事件可以推广为:

$$A_1\cup A_2\cup\cdots\cup A_n=\bigcup_{k=1}^{n}A_k$$

$$A_1\cup A_2\cup\cdots\cup A_n\cup\cdots=\bigcup_{n=1}^{\infty}A_n$$

$$A_1A_2\cdots A_n=\bigcap_{k=1}^{n}A_k$$

$A_1A_2\cdots A_n\cdots=\bigcap\limits_{n=1}^{\infty}A_n$

其含义可以分别从 $A\cup B$,AB 的定义去理解.

(5) $A-B$,称为事件 A 与事件 B 的**差事件**.在一次试验中,$A-B$ 发生当且仅当事件 A 发生而事件 B 不发生.

(6) 若 $AB=\Phi$,则称事件 A 与事件 B 是**互斥事件**或**互不相容**的事件.它表达了事件 A 与事件 B 不会同时发生.同一试验中,基本事件彼此都是互斥的.

(7) 若 $A\cup B=S$ 且 $AB=\Phi$,则称事件 A 与事件 B 为**对立事件**或**互逆事件**.在一次试验中,事件 A 与事件 B 必有一个且仅有一个发生.此时常记 $B=\overline{A}$,即事件 A 与 $\overline{A}$ 互为对立事件.显然有 $\overline{A}=S-A$,以及 $A-B=A\overline{B}$.

可以借助于集合关系的维恩图来直观地表示上述事件间的关系与运算.见下面图 1-1 至图 1-6.其中矩形表示样本空间 S.

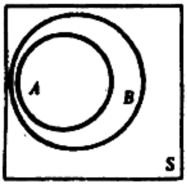

图 1-1

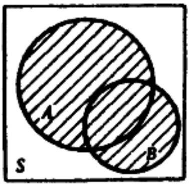

图 1-2

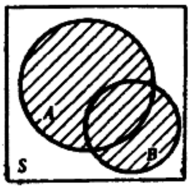

图 1-3

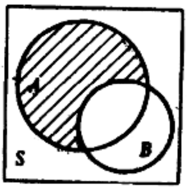

图 1-4

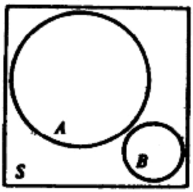

图 1-5

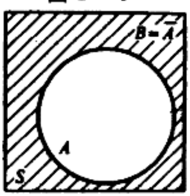

图 1-6

事件运算时，满足下列运算律：

(1) **交换律** $A \cup B = B \cup A, AB = BA$；

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

$(AB)C = A(BC)$；

(3) **分配律** $A(B \cup C) = AB \cup AC$，

$A \cup (BC) = (A \cup B)(A \cup C)$；

(4) **德·摩根律** $\overline{A \cup B} = \overline{A}\,\overline{B}$，　$\overline{AB} = \overline{A} \cup \overline{B}$.

以上运算律证明从略.

例 1　设 A,B,C 为三事件，则

(1) A 发生而 B 与 C 均不发生的事件应写为：

$$A\overline{B}\,\overline{C} \text{ 或 } A(\overline{B \cup C})；$$

(2) A,B,C 中不多于两个发生的事件应写为：

$$\overline{A} \cup \overline{B} \cup \overline{C} \text{ 或 } \overline{ABC}；$$

(3) A,B,C 中恰有一个发生的事件应写为：

$$A\overline{B}\,\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\,\overline{B}C；$$

(4) A,B,C 中至少有两个发生的事件为：

$$AB \cup AC \cup BC \text{ 或 } AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC \cup ABC.$$

例 2　图 1－7 所示电路. A 表示信号灯亮，B,C,D 表示对应开关闭合. 则事件 A,B,C,D 有关系 $BC \subset A, BD \subset A$，

$A = BC \cup BD$，

$\overline{B}A = \Phi$.

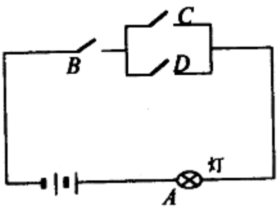

图 1－7

例 3　袋中有 10 个球，分别编号为 1 ～ 10，从中任取一球，设

A = {取出的球号码为偶数}；

B = {取出的球号码为奇数}；

$C=\{$取出的球号码小于 5$\}$.

则事件

(1) $A \cup B = S$ 为必然事件;

(2) $A \cap B = \Phi$ 为不可能事件;

(3) $A \cap C = \{$取出的球号码为 2 或 4$\}$;

(4) $\overline{AC} = \{$取出的球号码为 1,3,5,6,7,8,9,10,之一$\}$;

(5) $\overline{A}\ \overline{C} = \{$取出的球号码为 5,7,9 之一$\}$;

(6) $\overline{B \cup C} = \{$取出的球号码为 6,8,10 之一$\}$;

(7) $A - B = A$.

§1.2 频率与概率

一、频率

定义 设随机事件 A 在 n 次试验中发生了 n_A 次,则 n_A 称为事件 A 在 n 次试验中发生的**频数**,$\frac{n_A}{n}$ 称为事件 A 在 n 次试验中发生的**频率**,记作 $f_n(A)$,即事件 A 在 n 次试验中发生的频率为

$$f_n(A) = \frac{n_A}{n}.$$

频率 $f_n(A)$ 具有性质:

(1) $0 \leqslant f_n(A) \leqslant 1$;

(2) $f_n(S) = 1$;

(3) 若 $A_1, A_2, \cdots, A_k$ 两两互斥,则

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

性质(1)、(2) 是明显的. 对性质(3),不妨设 $k=2$,用 $n_{A\cup B}$ 表

示在 n 次试验中事件 $A\cup B$ 发生的频数，由于 A,B 互斥，所以 $n_{A\cup B}=n_A+n_B$，

$$f_n(A\cup B)=\frac{n_{A\cup B}}{n}=\frac{n_A+n_B}{n}=\frac{n_A}{n}+\frac{n_B}{n}$$
$$=f_n(A)+f_n(B).$$

频率除具有上述较直观的性质外，还具有一个内在的性质，即稳定性. 下面看一例子. 将一枚硬币抛掷 20 次，200 次，2000 次，各作 5 遍，得到如下表中的数据. 其中 n_H 表示 n 次抛掷中正面出现的频数，$f_n(H)$ 为 n 次抛掷中正面出现的频率.

试验序号	$n=20$		$n=200$		$n=2000$	
	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
1	11	0.55	91	0.455	990	0.4950
2	13	0.65	99	0.495	1012	0.5060
3	10	0.50	108	0.540	988	0.4940
4	6	0.30	98	0.490	1000	0.5000
5	10	0.50	101	0.505	998	0.4990

上述试验表明，当 n 较小时，出现正面的频率有较大的波动性. 但随着试验次数 n 的增大，出现正面的频率 $f_n(H)$ 逐渐稳定在常数 0.5 的附近. 这是随机现象固有的性质，即随机现象的统计规律性. 这个常数 0.5 是客观存在的，称为频率的稳定值，它反映了重复抛掷硬币时出现正面的频繁程度，也反映了在一次抛掷硬币时，出现正面的可能性的大小.

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，我们就用这个常数来衡量事件 A 发生的可能性大小，并称之为事件 A 的概率.

但是，在实际问题中，若对于每一个随机事件都要通过大量重

复试验来得到频率的稳定值，从而确定事件的概率是不现实的. 我们从频率的稳定性和频率的性质出发，给出度量事件发生可能性大小的概率的定义.

二、概率

定义 设随机试验 E 的样本空间为 S. 对于 E 中的每一个随机事件 A，赋予一个实数 $P(A)$ 与之对应. 如果集合函数 $P(\cdot)$ 具有如下性质：

(1) 对 E 中的每个事件 A，$P(A) \geqslant 0$；

(2) $P(S) = 1$；

(3) 若 $A_1, A_2, \cdots, A_n, \cdots$ 是两两互斥的事件序列，则有 $P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots$；

则称实数 $P(A)$ 为事件 A 的**概率**.

我们将会证明，当 $n \to \infty$ 时，事件 A 的频率在一定意义下收敛于事件 A 的概率. 因此，事件 A 的概率 $P(A)$ 度量了事件 A 在一次试验中发生的可能性的大小.

从上述概率的定义可以推得概率的一些基本性质：

(1) $P(\Phi) = 0$

证明 因 $\Phi = \Phi \cup \Phi \cup \cdots, \cup \Phi \cup \cdots$ 且 $\Phi \cap \Phi = \Phi$，有

$$P(\Phi) = P(\Phi \cup \Phi \cup \cdots \cup \Phi \cup \cdots)$$
$$= P(\Phi) + P(\Phi) + \cdots + P(\Phi) + \cdots$$

又因 $P(\Phi)$ 为实数，故 $P(\Phi)$ 必为 0.

(2) 若 $A_1, A_2, \cdots, A_n$ 两两互斥，则

$$P(A_1 \cup A_2 \cup \cdots \cup A_n)$$
$$= P(A_1) + P(A_2) + \cdots + P(A_n)$$

证明 因为 $A_1 \cup A_2 \cup \cdots \cup A_n$

$$= A_1 \cup A_2 \cup \cdots \cup A_n \cup \Phi \cup \Phi \cup \cdots$$

且右端随机事件序列两两互斥，所以有

$P(A_1 \cup A_2 \cup \cdots \cup A_n)$

$= P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \Phi \cup \Phi \cup \cdots)$

$= P(A_1) + P(A_2) + \cdots + P(A_n) + P(\Phi) + P(\Phi) + \cdots$

由性质(1)得证.

(3) 对任意事件 A,B,有 $P(A-B) = P(A) - P(AB)$

证明 因 $A = (A-B) \cup AB$,且右端 $A-B$ 与 AB 互斥,

故 $P(A) = P[(A-B) \cup AB]$

$= P(A-B) + P(AB)$

从而 $P(A-B) = P(A) - P(AB)$.

特别地 当 $B \subset A$ 时,$AB = B$,

故 当 $B \subset A$ 时,$P(A-B) = P(A) - P(B)$.

又因 $P(A-B) \geqslant 0$,

故 当 $B \subset A$ 时,$P(A) \geqslant P(B)$.

(4) 对任意事件 A,有 $P(A) \leqslant 1$

证明 因 $A \subseteq S$,有 $P(A) \leqslant P(S) = 1$.

(5) 对任意事件 A,有 $P(\overline{A}) = 1 - P(A)$

证明 因 $A \cup \overline{A} = S$,$A\overline{A} = \Phi$,故

$P(A \cup \overline{A}) = P(S) = 1$ 且 $P(A \cup \overline{A}) = P(A) + P(\overline{A})$

从而有 $P(A) + P(\overline{A}) = 1$,得证.

(6) 对任意事件 A,B 有

$P(A \cup B) = P(A) + P(B) - P(AB)$

证明 因 $A \cup B = A \cup (B-A)$,$A \cap (B-A) = \Phi$,

所以有 $P(A \cup B) = P[A \cup (B-A)]$

$= P(A) + P(B-A)$,

由性质(3) 有 $P(A \cup B) = P(A) + P(B) - P(AB)$.