

高等学校试用教材

工程数学

概率论与数理统计

浙江大学数学系高等数学教研组编

人民教育出版社

高等学校试用教材

工 程 数 学

概率论与数理统计

浙江大学数学系高等数学教研组编

人 民 教 育 出 版 社

本书是按照 1977 年高等学校工科数学教材编写会议上确定的编写大纲编写的，包括概率论、随机过程、数理统计三部分，每章均附有相应的习题，可作高等学校工科试用教材，也可供工程技术人员参考。

高等学校试用教材

工程数学

概率论与数理统计

浙江大学数学系高等数学教研组编

*

人民教育出版社出版

新华书店北京发行所发行

北京印刷一厂印装

*

开本 787×1092 1/32 印张 11 4/16 字数 271,000

1979 年 3 月第 1 版 1981 年 9 月第 4 次印刷

印数 353,001—413,000

书号 13012·0258 定价 0.82 元

前 言

本书是按照1977年高等学校工科数学教材编写会议上确定的编写大纲编写的，可作为高等学校工科工程数学——概率论与数理统计课程的试用教材，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，是全书的重点，为读者提供必要的理论基础；随机过程部分(第六章至第八章)在讲清基本知识的基础上，主要讨论了平稳随机过程；数理统计部分(第九章至第十二章)介绍了常用的统计推断方法：参数估计、假设检验、方差分析和回归分析。随机过程和数理统计这两部分内容是相互独立的，可根据各专业的需要选用。打*号的内容可删去。

我们希望本书能为读者在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。

本书由浙江大学数学系高等数学教研组盛骤，黄纪青，谢式千，潘承毅编写。

本书由主审单位上海师范大学魏宗舒教授主审，参加审稿的还有华中工学院林少官教授，吉林工业大学沈恒范副教授，上海师范大学茆诗松，上海交通大学颜钰芬等同志，他们提出了很多宝贵意见，主审单位的同志对我们的帮助尤大，对此，我们表示衷心的感谢。

由于我们水平有限，书中一定存在着不少缺点和错误，希望读者批评和指正。

编 者

1979.1

目 录

前言	iv
第一章 概率论的基本概念	1
§ 1 随机试验	2
§ 2 随机事件、样本空间	3
§ 3 频率与概率	9
§ 4 等可能概型(古典概型)	14
§ 5 条件概率	21
§ 6 独立性	28
习题	32
第二章 随机变量及其分布	38
§ 1 随机变量	38
§ 2 离散型随机变量的概率分布	40
§ 3 随机变量的分布函数	49
§ 4 连续型随机变量的概率密度	54
§ 5 随机变量的函数的分布	62
习题	65
第三章 多维随机变量及其分布	71
§ 1 二维随机变量	71
§ 2 边缘分布	77
§ 3 条件分布	80
§ 4 相互独立的随机变量	86
§ 5 两个随机变量的函数的分布	89
习题	96
第四章 随机变量的数字特征	101
§ 1 数学期望	101
§ 2 方差	111

§3	几种重要随机变量的数学期望和方差	115
§4	协方差和相关系数	120
§5	矩、协方差矩阵	123
	习题	126
第五章	大数定律和中心极限定理	131
§1	大数定律	131
§2	中心极限定理	133
	习题	137
第六章	随机过程的基本知识	139
§1	随机过程的概念	139
§2	随机过程的分布函数以及随机过程的分类	144
§3	随机过程的数字特征	153
§4	两个或两个以上随机过程的联合分布和数字特征	157
	习题	160
第七章	平稳随机过程	161
§1	平稳随机过程的数字特征. 宽平稳随机过程	161
§2	各态历经性	166
§3	相关函数的性质	176
§4	平稳过程的功率谱密度	180
	习题	192
*第八章	线性系统对随机输入响应	195
§1	线性系统的基本知识	195
§2	线性系统输出的均值和自相关函数	199
§3	系统输出的谱密度	202
§4	输入输出之间的互相关函数与互谱密度	205
	习题	206
第九章	样本及其分布	208
§1	随机样本和统计量	208
§2	抽样分布	215
	习题	226
第十章	参数估计	228

§1 点估计	228
§2 极大似然估计法	231
§3 估计量的评选标准	237
§4 区间估计	242
§5 正态总体均值与方差的区间估计	244
§6 (0-1)分布参数的区间估计	250
§7 单侧置信限	252
习题	253
第十一章 假设检验	257
§1 假设检验	257
§2 t 检验	263
§3 F 检验	267
§4 χ^2 检验	269
§5 曲线拟合	270
习题	279
第十二章 方差分析和回归分析	284
§1 单因素试验	284
§2 双因素试验	295
§3 线性回归	306
习题	320
习题答案	323
附表1 标准正态分布表	337
附表2 泊松分布表	339
附表3 t 分布表	341
附表4 χ^2 分布表	342
附表5 F 分布表	344

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(或必然不发生)，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。过去我们学过的微积分学和线性代数等就是研究这类现象的数学工具。然而在自然界和社会中也还存在着另一类现象，例如，在相同条件下抛同一枚硬币，其结果可能是花这面朝上，也可能是数字一面向上，并且不论怎样控制抛掷条件，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，并且不论怎样控制射击条件，在一次射击以前无法预测弹着点的确切位置。这类现象归纳起来可以看作在相同条件下一系列的试验或观察，而每次试验或观察的可能结果不止一个，在每次试验或观察之前无法预知确切的结果，即呈现出不确定性。

人们经过长期实践并深入研究之后，发现这类现象虽然就每次试验或观察结果来说，它具有不确定性，但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花这面朝上大致有半数，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。

一类现象，在个别试验中呈现出不确定性；在大量重复试验中，又具有统计规律性，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的统计规律性的一门数学学科。

概率统计的理论与方法在应用上是很广泛的，目前它已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之中。

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器件或装置的使用可靠程度及平均寿命的估计；在自动控制中可用以给出数学模型以便通过电子计算机来控制工业生产；在通讯工程中可以提高信号的抗干扰性和分辨率等。

§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。下面举一些例子来说明。

E_1 : 抛一枚硬币，观察正面 H (有花的一面)、反面 T 出现的情况。

E_2 : 将一枚硬币抛二次，观察正、反面出现的情况。

E_3 : 掷一颗骰子，观察出现的点数。

E_4 : 记录某电话交换台一分钟内接到的呼唤次数。

E_5 : 一口袋中装有红白二种颜色的乒乓球。从袋中任取一只球，观察其颜色。

E_6 : 一射手进行射击，直到击中目标为止，观察其射击情况。

E_7 : 在一批灯泡中任意抽取一只，测试它的寿命。

E_8 : 记录某地一昼夜的最高温度($^{\circ}\text{C}$)和最低温度。

上面举了 8 个试验的例子，它们有着共同的特点。例如试验 E_1 ，它有二种可能结果：出现 H 或者出现 T ，但在投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_3 ，它有 6 种可能结果，即出现点数为 1, 2, ..., 6 中之一，但在投掷之前不能确定会出现几点，这个试验可以在相同的条件下重复地进行。再如试验 E_7 ，我们知道灯泡的寿命(以小时

计) $t \geq 0$, 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, 这些试验都具有以下特性:

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性的试验称为随机试验, 简称试验。

我们就是通过研究随机试验研究随机现象的。

§ 2 随机事件、样本空间

(一) 随机事件 一个随机试验的所有可能的结果是可以知道的, 但在每次试验之前却不能确定哪个会出现。例如在试验 E_1 中“出现 H ”这件事情可能发生也可能不发生, 但如果重复抛许多次, 就能看出它的发生是具有某种规律性的(这一点将在下一节说明)。在试验 E_3 中“出现 2 点”这件事情也是这样。

在随机试验中, 对一次试验可能出现也可能不出现, 而在大量重复试验中却具有某种规律性的事情, 称为此随机试验的随机事件, 简称事件。

在大量重复试验中随机事件的发生具有某种规律性, 揭示和研究这种规律性就是概率论所要研究的问题。

在一随机试验中, 它的每一个可能出现的结果都是一个随机事件, 它们是这个试验的最简单的随机事件, 我们称这些简单的随机事件为基本事件。

例如, 在试验 E_1 中, “出现 H ”、“出现 T ”就是试验的基本事件; 在试验 E_3 中, “出现 1 点”、“出现 2 点”、…、“出现 6 点”就是

基本事件.

一试验中,除基本事件以外还有其它的随机事件.例如,在 E_3 中“出现偶数点”也是一个随机事件,它是由“出现2点”、“出现4点”、“出现6点”这三个基本事件所组成的,当且仅当这三个基本事件中有一个发生,“出现偶数点”这一事件发生.又如“点数小于3”也是随机事件,它由“出现1点”、“出现2点”这两个基本事件所组成,当且仅当这两个基本事件中有一个发生,“点数小于3”这一事件就发生.

在试验 E 中必然会发生的事情叫做必然事件;不可能发生的事情叫做不可能事件.例如,在 E_3 中“点数不大于6”是必然事件;“点数大于6”是不可能事件.必然事件和不可能事件本来没有不确定性,也就是说它们不是随机事件,但为了今后讨论方便起见,我们把它们当作一种特殊的随机事件.

(二)样本空间 为了便于研究随机试验 E ,我们将随机试验 E 的所有基本事件所组成的集合叫做 E 的样本空间,记为 S . S 中的元素就是试验 E 的基本事件.基本事件也称样本点.

下面写出了§1中试验 $E_k(k=1, 2, \dots, 8)$ 的样本空间 S_k .

$$S_1: \{H, T\}$$

$$S_2: \{(H, H), (H, T), (T, H), (T, T)\}^{\textcircled{1}}$$

$$S_3: \{1, 2, 3, 4, 5, 6\}$$

$$S_4: \{0, 1, 2, \dots\}$$

$$S_5: \{\text{白色}, \text{红色}\}$$

$S_6: \{+, -+, --+, \dots\}$, 这里“+”表示击中,“-”表示没有击中

$$S_7: \{t | t \geq 0\}$$

$S_8: \{(x, y) | T_0 < x < y < T_1\}$, 这里 x 表示最低温度, y 表示

^① 这里所用的记号,例如 (H, T) 表示第一次出现 H ,第二次出现 T .

最高温度,并设这一地区的温度不会小于 T_0 , 不会大于 T_1 .

要注意的是: 样本空间中的元素是由试验的内容所确定的. 例如在试验 E_5 中, 如果将球自 1 到 n 编号(设袋中装有 n 只球), 若试验为在袋中任取一球观察其号码, 那么样本空间不再是 {白色, 红色} 而是 $\{1, 2, \dots, n\}$.

由于随机事件是基本事件, 或是由基本事件所组成的, 引入了样本空间 S 之后, 我们看到试验 E 的事件是样本空间 S 中的子集, 而且事件发生, 当且仅当子集中的一个样本点发生. 例如, 上面所说的事件 A : “出现偶数点”是由基本事件“2”“4”“6”所组成的, A 是 S_3 的子集, 即 $A = \{2, 4, 6\}$. 又如事件 B : “点数小于 3”是子集 $\{1, 2\}$, 即 $B = \{1, 2\}$. 特别, 必然事件就是样本空间 S ; 不可能事件就是空集 ϕ .

下面再举几个事件的例子.

例 1 在 E_2 中, 事件 A_1 : “第一次出现 H ”, 即

$$A_1 = \{(H, H), (H, T)\}$$

事件 A_2 : “两次出现同一面”, 即

$$A_2 = \{(H, H), (T, T)\}$$

事件 A_3 : “只有一次出现 H ”, 即

$$A_3 = \{(H, T), (T, H)\}$$

在 E_7 中, 事件 A_4 : “寿命小于 5 小时”, 即

$$A_4 = \{t | 0 \leq t < 5\}$$

在 E_8 中, 事件 A_5 : “最高温度与最低温度相差 10 度”, 即

$$A_5 = \{(x, y) | y - x = 10, T_0 < x < y < T_1\}$$

(三) 事件之间的关系与事件的运算 为了研究事件的需要, 下面介绍事件之间的关系与事件的运算.

设试验 E 的样本空间为 S , $A, B, A_k (k=1, 2, \dots)$ 是 E 的事件.

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

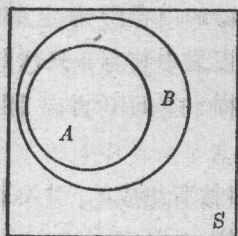


图 1-1

这可用图 1-1 来直观地说明, 图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 .

若事件 B 包含事件 A , 事件 A 也包含事件 B , 即 $B \supset A, A \supset B$, 则称事件 A 与事件 B 相等, 记为 $A = B$.

2° 事件 A 与事件 B 至少有一个发生, 这一事件称为事件 A 与事件 B 的和, 记为 $A \cup B$.

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 , 图中阴影部分即表示事件 A 与事件 B 的和.

类似地, 事件 A_1, A_2, \dots, A_n 中至少有一个发生, 这一事件称为 A_1, A_2, \dots, A_n 的和, 记为 $A_1 \cup A_2 \cup \dots \cup A_n$, 简记为 $\bigcup_{k=1}^n A_k$; 事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生, 这一事件称为 $A_1, A_2, \dots, A_n, \dots$ 的和, 记为 $A_1 \cup A_2 \cup \dots \cup A_n \dots$, 简记为 $\bigcup_{k=1}^{\infty} A_k$.

3° 事件 A 与事件 B 同时发生, 这一事件称为事件 A 与事件 B 的积, 记为 $A \cap B$ 或 AB .

图 1-3 中阴影部分表示 $A \cap B$.

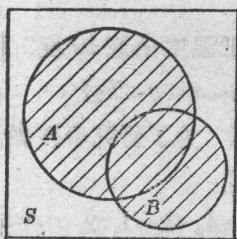


图 1-2

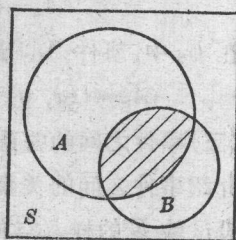


图 1-3

类似地,可以定义 $A_k (k=1, 2, \dots, n)$ 的积:

$$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{k=1}^n A_k;$$

以及 $A_k (k=1, 2, \dots)$ 的积:

$$A_1 \cap A_2 \cap \dots \cap A_n \dots = \bigcap_{k=1}^{\infty} A_k.$$

4° 事件 A 发生而事件 B 不发生, 这一事件称为事件 A 与事件 B 的差, 记为 $A-B$.

图 1-4 中的阴影部分表示 $A-B$.

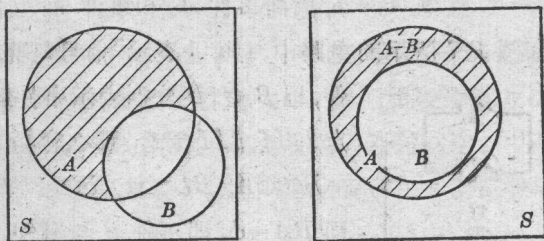


图 1-4

5° 若事件 A 与事件 B 不能同时发生, 亦即 $AB = \phi$, 则称事件 A 与事件 B 是互不相容的. 基本事件是互不相容的.

图 1-5 直观地表示了事件 A 与事件 B 是互不相容的.

若在试验中, 事件 A 与事件 B 中必然有一个发生, 且仅有一个发生, 亦即, 事件 A 和事件 B 满足条件

$$A \cup B = S, \quad AB = \phi.$$

则称事件 A 与事件 B 互逆, 又称 A 是 B 的对立事件 (或 B 是 A 的对立事件) 记为 $A = \bar{B}$ (或 $B = \bar{A}$).

图 1-6 直观地表示事件 A 与事件 B 互逆.

例 2 在例 1 中

$$A_1 \cup A_2 = \{(H, H), (H, T), (T, T)\}$$

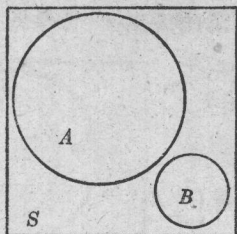


图 1-5

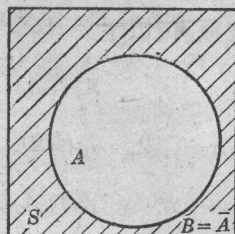


图 1-6

$$A_1 A_2 = \{(H, H)\}$$

$$A_1 - A_2 = \{(H, T)\}$$

因为 $A_2 \cup A_3 = S$, 且 $A_2 A_3 = \phi$, 因而事件 A_2 和事件 A_3 互逆.

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

$$BC \subset A, \quad BD \subset A, \quad BC \cup BD = A,$$

而 $\bar{B}A = \phi$, 即事件 \bar{B} 与事件 A 互不相容.

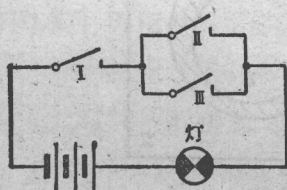


图 1-7

从上面的讨论可以看到, 概率论

中事件之间的关系与运算和集合论中集合之间的关系与运算是一致的. 为了便于对照, 我们列出下面的表格.

记号	概 率 论	集 合 论
S	样本空间, 必然事件	全集
ϕ	不可能事件	空集
e	基本事件	元素
A	事件	子集
\bar{A}	A 的对立事件	A 的余集
$A \subset B$	事件 A 发生导致事件 B 发生	A 是 B 的子集
$A = B$	事件 A 与事件 B 相等	A 与 B 相等
$A \cup B$	事件 A 与事件 B 至少有一个发生	A 与 B 的和集
AB	事件 A 与事件 B 同时发生	A 与 B 的交集
$A - B$	事件 A 发生而事件 B 不发生	A 与 B 的差集
$AB = \phi$	事件 A 和事件 B 互不相容	A 与 B 没有相同的元素

这样,我们常常可以把对事件的分析转化为对集合的分析,利用集合间的运算来分析事件间的关系。

§3 频率与概率

(一)频率 一个随机试验有许多可能结果,我们常常希望知道某些结果出现的可能性有多大。例如,要在某河流上建筑一座防洪水坝,为了确定水坝的高度,就要知道该河流在造水坝地段每年最大洪水达到某高度的可能性的。最大洪水达到某一高度是随机事件,我们希望能将一个随机事件发生的可能性的。大小用一个数来表达。

先来分析一个简单的例子。考虑“抛硬币”这个随机试验。为了要知道 H 出现的可能性的。大小,我们将硬币抛 n 次,观察在 n 次试验中 H 出现的次数。现在将硬币连抛 5 次、50 次、500 次各做了十遍,得数据如表 1 (表中, n_H 表示在 n 次试验中 H 出现的次数,比值 $\frac{n_H}{n} = f_n(H)$ 叫做在这 n 次试验中出现 H 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

从表 1 可以看出, 抛硬币次数较少时, H 出现的频率差异较大, 但是随着抛硬币的次数的增多, H 出现的频率呈现出稳定性.

这种试验从前也有人做过, 得出如表 2 所示的数据.

表 2

实验者	n	n_H	$f_n(H)$
蒲丰	4040	2048	0.5070
K. 皮尔逊	12000	6019	0.5016
K. 皮尔逊	24000	12012	0.5005

从表 2 更能看出, 不管什么人去抛, 当试验的次数逐渐增多时, $f_n(H)$ 总是在 $0.5 = \frac{1}{2}$ 附近摆动而逐渐稳定于 0.5. 这个数能反映出现 H 的可能性的.

一般, 设随机事件 A 在 n 次试验中出现 n_A 次, 比值

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

叫做事件 A 在这 n 次试验中出现的频率.

下面我们再来分析随机试验 E : 在一口袋中装有 6 只乒乓球, 其中 4 只是白球 2 只是红球, 从袋中任意取一只球, 现在来考察取到白球的可能性有多大. 设 W 为取出白球这一事件. 现在来作许多试验, 观察取到白球的次数, 并算出它出现的频率. 做了 600 次试验 (从袋中任意取一球, 观察其颜色后放回袋中, 搅匀后再任意抽取一球观察其颜色, 这样继续下去), 得数据如表 3 所示.

表 3

n	100	200	300	400	500	600
n_W	69	139	198	261	337	401
$f_n(W)$	0.690	0.695	0.660	0.653	0.674	0.668