

概率论与 数理统计教程

交通部六院校数学教研室 合编



西北工业大学出版社

概率论与数理统计教程

交通部六院校数学教研室合编
西安公路学院数学教研室主编
大连海运学院数学教研室主审

西北工业大学出版社

2000年1月 西安

(陕)新登字 009 号

编辑委员会

主任编委：杨慰祖

编 委：曲修轩 王存林 李之后 李晓风 程期远
 张世杰 敬福田 孙华荣 袁天俊 张宏哲
 钟忠奎

概率论与数理统计教程

交通部六院校数学教研合编

责任编辑 郑永安

责任校对 樊 阳

*

西北工业大学出版社出版发行

(西安市友谊西路 127 号)

全国新华书店经销

西安市高陵县印刷厂印装

*

开本：850 毫米 × 1 168 毫米 1/32

1991 年 8 月第 1 版

印数：13 301 ~ 18 301 册

印张：10.5 字数：266 千字

2000 年元月第 3 次印刷

定价：10.00 元

购买本社出版的图书，如有缺页、错页的，本社发行部负责调换。

前 言

本教程是在长期教学实践的基础上,由交通部部属六所高等院校的数学教师合作编写的。内容的深度和广度,符合国家教委工科数学课程指导委员会颁布的“概率论与数理统计”的基本要求。

本书的各部分内容具有相对的独立性,使用时可按照需要和可能随意撷取,适用于各种不同学时数的教学要求。

该书注重应用。对一些重要的概念、定理及公式的概率意义和实际背景,不吝笔墨,论述清晰。对各类工程技术人员及经济管理工作,也不无参考价值。

本教程第一章由南京航务工程专科学校陈节禄编写,第二章由西安公路学院敬福田编写,第三章由大连海运学院张彦超编写,第四章由长沙交通学院周世琼、杨润生编写,第五章和第九章由重庆交通学院陈庆轩编写,第六章由大连海运学院李彩荣编写,第七、八章由武汉水运工程学院戴梁编写,第十章由西安公路学院李晓风编写。

该教程由西安公路学院数学教研室主编,由大连海运学院数学教研室主审。

在本书的编写过程中,西安公路学院和重庆交通学院的有关领导,交通部各兄弟院校数学教研室的同志们,给予了热情的关怀、积极的支持和具体的帮助。为使本教程尽快和读者见面,西北工业大学出版社的领导和同志们,倾注了不少心血,付出了艰辛的劳动,克服了许多困难,在此谨表示衷心感谢。

因编者水平有限,难免会有不妥之处。希望使用本教程的教师及广大读者多加指正。

编 者

1991年5月

目 录

第一章 随机事件及其概率	(1)
第一节 随机事件和样本空间	(1)
第二节 事件间的关系及运算	(4)
第三节 随机事件的概率	(9)
第四节 古典概型	(16)
第五节 条件概率、乘法公式及事件的独立性	(21)
第六节 全概率公式和贝叶斯公式	(30)
习题一	(35)
第二章 随机变量及其概率分布	(42)
第一节 随机变量	(42)
第二节 离散型随机变量	(45)
第三节 连续型随机变量	(54)
第四节 分布函数	(62)
第五节 随机变量的函数的分布	(67)
习题二	(72)
第三章 随机变量的数字特征	(77)
第一节 随机变量的均值	(77)
第二节 随机变量的方差	(85)
习题三	(90)

第四章 随机向量	(94)
第一节 随机向量的分布与边缘分布	(94)
第二节 两个随机变量的函数的分布	(109)
第三节 随机向量的数字特征	(112)
习题四	(120)
第五章 大数定律和中心极限定理	(124)
第一节 大数定律	(124)
第二节 中心极限定理	(129)
习题五	(134)
第六章 子样的抽样分布	(136)
第一节 母体和子样	(137)
第二节 一些常用的抽样分布	(147)
习题六	(162)
第七章 参数估计	(166)
第一节 点估计和估计量的求法	(167)
第二节 估计量的好坏标准	(180)
第三节 区间估计	(183)
习题七	(192)
第八章 假设检验	(196)
第一节 假设检验的基本概念和基本思想	(196)
第二节 单个母体的参数假设检验	(201)
第三节 两个正态母体的参数假设检验	(212)
第四节 分布假设检验	(218)
习题八	(224)

第九章 方差分析和正交试验设计	(230)
第一节 单因素方差分析.....	(230)
第二节 双因素方差分析.....	(238)
第三节 正交试验设计.....	(246)
习题九.....	(256)
第十章 回归分析	(262)
第一节 一元线性回归.....	(264)
第二节 非线性回归.....	(283)
习题十.....	(290)
习题答案	(294)
附表 1 标准正态分布函数表	(312)
附表 2 t 分布上侧分位数表	(314)
附表 3 χ^2 分布上侧分位数表	(315)
附表 4 F 分布上侧分位数表	(317)
附表 5 泊松分布表	(326)

第一章 随机事件及其概率

第一节 随机事件和样本空间

客观世界中发生的现象,可以分为两大类:一类为**必然现象**,另一类为**随机现象**。

所谓必然现象,是指在一定条件下必然发生的现象。例如,物体在重力作用下必然向下落;在标准大气压下,纯水加热到 100°C ,必然会沸腾,而到 0°C 时必然会结冰等等。经典的数学理论如代数、几何、微积分等都是用来研究必然现象中的数量关系的。

所谓随机现象,即通常所称的偶然现象,是指在一定条件下可能发生这样的结果,也可能发生那样的结果,且预先不能断言发生哪种结果的现象。例如,一枚硬币,我们把有花的一面叫“正面”,有字的一面叫“反面”。抛一枚硬币,可能出现正面,也可能出现反面;新生的婴儿,可能是男孩,也可能是女孩;一门炮对一目标射击,尽管经过瞄准,炮弹却可能落在目标附近的不同位置。概率论与数理统计就是研究随机现象数量规律的一门数学学科。

对随机现象进行研究,就要进行试验或者观察。如果试验(或观察)满足下述三个条件,就称为**随机试验**:

- (1) 试验可以在相同条件下重复进行;
 - (2) 试验的所有可能结果不止一个,但能事先明确试验的所有可能结果;
 - (3) 每次试验之前不能确定哪一个结果将会出现。
- 今后,把随机试验简称为试验,记为 E 。

随机试验的每一个可能出现的简单直接结果,称为基本事件。基本事件的全体所构成的集合,称为样本空间,通常用字母 Ω 表示。 Ω 中的基本事件,有时也称作样本点,常用 e 或 ω 表示。

在概率论中,讨论一个随机试验时,首先要求明确它的样本空间。对于一个具体的随机试验来说,样本空间可根据试验的内容来决定。

例 1. 一个盒子中有 10 个大小、形状及质地相同的球,其中 5 个白球,5 个黑球。从中任意摸取一球,观察颜色,可能出现的直接结果有两个:“取得白球”、“取得黑球”,即有两个基本事件。令

$$e_1 = \text{“取得白球”}, \quad e_2 = \text{“取得黑球”}$$

则这个试验的样本空间 $\Omega = \{e_1, e_2\}$ 。

例 2. 一个盒子中有 10 个完全相同的球,分别标以号码 1, 2, 3, $\dots$ 10。从中任取一球,观察号码,其可能出现的直接结果有 10 个,亦即有 10 个基本事件。令

$$i = \text{“取得的球标号为 } i\text{”} (i = 1, 2, \dots, 10)$$

则

$$\Omega = \{1, 2, \dots, 10\}$$

例 3. 将一枚硬币抛两次,观察正反面出现的情况,记正面朝上为 H ,反面朝上为 T ,则

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

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 。注意, (H, T) 表示一个基本事件。

例 4. 记录某电话交换台在某段时间内接到的呼唤次数。令

$$i = \text{“接到的呼唤次数为 } i\text{”}$$

则

$$\Omega = \{0, 1, 2, \dots\}$$

例 5. 对某一目标发射一发炮弹,测量弹着点到目标的距离。
令

$$\omega = \text{“弹着点到目标的距离为 } \omega \text{ 单位”}$$

$$\Omega = \{\omega | 0 \leq \omega < +\infty\}.$$

从上面的这些例子可以看到, 样本空间 Ω 可以只含有限个基本事件, 如例 1, 例 2; 也可以含有无穷多个基本事件, 如例 4, 例 5.

在随机试验中, 通常更为关心的是具有某些特征的基本事件是否发生. 如在例 2 中, 我们可以研究

$A = \text{“取得球的标号为 6”}$

$B = \text{“取得球的标号是偶数”}$

$C = \text{“取得球的标号} \leq 5 \text{”}$

这些结果是否发生? 其中 A 是一个基本事件, 而 B 和 C 是由多个基本事件所组成, 相对于基本事件, 称它们为**复合事件**. 基本事件和复合事件在试验中是否发生, 都带有随机性, 所以都叫做**随机事件**. 简称为**事件**, 并以大写英文字母 $A, B, C, \dots$ 来表示. 在试验中, 如果事件 A 所包含的某一个基本事件 e 出现了, 就说 A 发生或出现了.

样本空间 Ω 包含了全体基本事件, 而随机事件是由具有某些特征的基本事件组成. 从集合论的观点来看, 随机事件是样本空间 Ω 的子集. 随机事件 A 发生, 当且仅当 A 所包含的某一个基本事件发生.

如在例 2 中, $\Omega = \{1, 2, \dots, 10\}$. 显然, 前述事件 A, B, C 都是 Ω 的子集, 它们可分别表示为 $A = \{6\}, B = \{2, 4, 6, 8, 10\}, C = \{1, 2, 3, 4, 5\}$.

我们把样本空间 Ω 也作为一个事件, 由于 Ω 是由全体基本事件所组成, 所以在任一次试验中必然会发生, 因此称 Ω 为**必然事件**. 类似地, 把空集 $\emptyset$ 也作为一个事件, 因它在每次试验中都不会发生, 故称它为**不可能事件**. 如在例 2 中, “取得球的标号不大于 10”就是一个必然事件, 这是因为在试验中, 不论哪一个基本事件发生, 均导致“取得球的标号不大于 10”这一事件的发生. 而“取得球的标号大于 10”这一事件就是一个不可能事件. 须注意的是, 必

然事件与不可能事件本来都不是随机事件,但为研究方便,我们还是把它们作为随机事件的两个极端情形来统一处理。

第二节 事件间的关系及运算

在概率论中,常常要求我们把一个复杂的事件用某些简单的事件来表示,于是研究事件之间的关系和运算就显得非常重要了。

为了方便起见,在以下的叙述中,总认为试验 E 的样本空间 Ω 已经给定了,并且 $A, B, A_i, B_i (i=1, 2, \dots)$ 等都表示 E 的事件。

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

例如,在第一节提到过的 A = “取得球的标号为 6” 这一事件发生就导致事件 B = “取得球的标号是偶数” 的发生,因为摸到标号为 6 的球意味着标号为偶数的球出现了,所以 $B \supset A$ 。

我们用图 1-1 直观地表示事件的包含关系,矩形表示样本空间 Ω , A 与 B 是 Ω 的两个子集。“ A 发生必然导致 B 发生” 意味着 “属于 A 的样本点必然属于 B ”, 即 A 中的点全在 B 中。由此可知, $A \subset B$ 的含义与集合论中包含的意义是一致的。

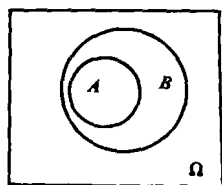


图 1-1

因为不可能事件 $\emptyset$ 不含有任何样本点,所以对任一事件 A , 约定 $\emptyset \subset A$ 。

2. 若 $A \subset B, B \subset A$ 同时成立,则称事件 A 与事件 B 相等,记为 $A = B$ 。相等的两个事件,总是同时发生或同时不发生。

3. “事件 A 与事件 B 至少有一个发生” 是一个事件,称它为 A 与 B 的和,记为 $A \cup B$ 。

例如,某种产品的合格与否是由该产品的长度与直径是否合格所决定,因此事件“产品不合格”是事件“长度不合格”与事件“直径不合格”的和。图 1—2 中的阴影部分表示事件 $A \cup B$ 。

类似地,“ $A_1, A_2, \dots, A_n$ 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 $A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{k=1}^n A_k$; “ $A_1, A_2, \dots, A_n, \dots$ 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 $A_1 \cup A_2 \cup \dots \cup A_n \cup \dots = \bigcup_{k=1}^{\infty} A_k$ 。

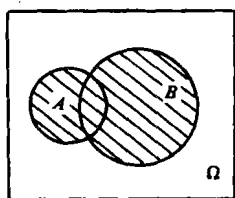


图 1—2

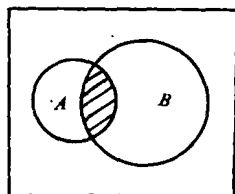


图 1—3

4. “事件 A 与事件 B 同时发生”这一事件称为 A 与 B 的积,记为 $A \cap B$ 或 AB 。图 1—3 中的阴影部分表示事件 AB 。

如上例,“产品合格”是“长度合格”与“直径合格”的积。

类似地,可定义 n 个事件 $A_k (k=1, 2, \dots, n)$ 的积:

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

以及无穷多个事件 $A_k (k=1, 2, \dots)$ 的积为

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

5. “事件 A 发生而事件 B 不发生”这一事件称为 A 与 B 的差,记为 $A - B$,它由图 1—4 中的阴影部分表示。

如在第一节例 2 中,若令 A = “取得球的标号是 2, 4, 6, 8, 10,” B = “取得球的标号 ≤ 3 ”,则

$$A - B = \text{“取得球的标号是 4, 6, 8, 10”}$$

又如,“直径合格但长度不合格”是“直径合格”与“长度合格”

的差。

6. 若事件 A 与事件 B 不能同时发生, 也就是说 AB 是一个不可能事件, 即 $AB = \emptyset$, 则称 A 与 B 互不相容(或互斥)。图 1-5 表示 A 与 B 互不相容。

如在第一节例 2 中, 若令 A = “取得球的标号为偶数”, B = “取得球的标号为 3,” 则 $AB = \emptyset$, 从而 A 与 B 是互不相容的。

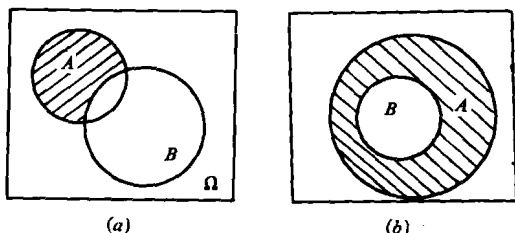


图 1-4

对于互不相容的事件 A, B , 可以把和事件 $A \cup B$ 记作 $A+B$ 。如果一组事件 $A_1, A_2, A_3, \dots, A_n, \dots$ 中任意两个都互不相容, 则称这组事件两两互不相容, 简称这组事件互不相容。

对于一组互不相容的事件 $A_1, A_2, \dots, A_n, \dots$ 的和可以记作 $A_1 + A_2 + \dots = \sum_{k=1}^{\infty} A_k$ 。

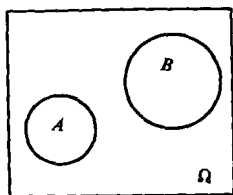


图 1-5

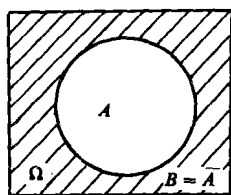


图 1-6

7. 若在试验中, 事件 A 与事件 B 必然有一个发生, 且仅有一个发生, 即 A, B 满足条件

$$A \cup B = \Omega, \quad AB = \emptyset$$

则称 A 与 B 互逆, 而称 A 是 B 的对立事件或逆事件 (或 B 是 A 的对立事件), 记为 $A = \bar{B}$ (或 $B = \bar{A}$)。图 1—6 表示了 A 与 B 互逆。

由定义可知, 若 B 是 A 的对立事件, 即 $B = \bar{A}$, 则 A 也是 B 的对立事件, $\bar{B} = A$, 从而 $\bar{\bar{A}} = A$ 。另外, 大家不难看出: 相互对立的事件一定是不相容的, 但反过来则不一定成立。

如在第一节例 2 中, 若令 A = “取得球的标号为偶数”, B = “取得球的标号为 3”, 则因 $AB = \emptyset$, 故 A, B 互不相容。但 $A \cup B \neq \Omega$, 可见 A 与 B 不是相互对立的事件。显然, A 的逆事件应是 “取得球的标号为奇数”, 而 B 的逆事件应是 “取得球的标号为 ‘1, 2, 4, 5, 6, 7, 8, 9, 10’”。

利用对立事件的定义, 可将差事件 $A - B$ 表示为事件 A 与事件 $\bar{B}$ 的积, 即 $A - B = A\bar{B}$ 。

8. 若在任一次试验中, 事件 $A_1, A_2, \dots, A_n$ 至少有一个发生, 即 $\bigcup_{i=1}^n A_i = \Omega$, 则称这 n 个事件构成一完备事件组 (或完备组)。

以后对我们特别重要的是互不相容事件构成的完备组。设 n 个事件 $A_1, A_2, \dots, A_n$ 满足下面的关系式:

$$\bigcup_{i=1}^n A_i = \Omega, A_i A_j = \emptyset (i \neq j)$$

则称这 n 个事件构成互不相容事件的完备组 (或称互不相容的完备事件组)。

例如, 投掷一颗骰子, 设事件 A_i 表示 “出现 i 点” ($i = 1, 2, \dots, 6$), 则这六个事件满足

$$A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6 = \Omega$$

$$A_i A_j = \emptyset (1 \leq i < j \leq 6)$$

可见这六个事件构成互不相容事件的完备组。

从以上的讨论可以看到, 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, 列表如下

记 号	概率论	集合论
Ω	样本空间, 必然事件	全集
$\emptyset$	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 的和	A 与 B 的并集
AB	事件 A 与事件 B 的积	A 与 B 的交集
$A - B$	事件 A 与事件 B 的差	A 与 B 的差集
$AB = \emptyset$	事件 A 与事件 B 互不相容	A 与 B 不相交

例 6. 若 A, B, C , 是三个事件, 则

(1) 事件“ A 发生, 而 B, C 都不发生”, 可以表示为: $A\bar{B}\bar{C}$ 或 $A - (B \cup C)$ 。

(2) 事件“ A, B 都发生, 而 C 不发生”可以表示为: $AB\bar{C}$ 或 $AB - C$ 。

(3) 事件“ A, B, C 恰有一个发生”可以表示为: $A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C$ 。

(4) “事件 A, B, C 至少有一个发生”可以表示为: $A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C + A\bar{B}C + \bar{A}BC + \bar{A}\bar{B}C + ABC$, 或 $A \cup B \cup C$ 。

(5) 事件“ A, B, C 均不发生”可以表示为: $\bar{A}\bar{B}\bar{C}$, 或 $\overline{A \cup B \cup C}$ 。

这里事件 $A \cup B \cup C$ 表示“事件 A, B, C 至少有一个发生”, 它的对立事件 $\overline{A \cup B \cup C}$ 当然就是 A, B, C 都不发生了。

(6) 事件“ A, B, C 至少有一个不发生”可以表示为: $\overline{ABC} + \overline{ABC} + \overline{ABC} + \overline{ABC} + \overline{ABC} + \overline{ABC} + \overline{ABC}$, 或 $\overline{A} \cup \overline{B} \cup \overline{C}$, 或 $\overline{ABC}$ 。

这里事件 ABC 表示“事件 A, B, C 都发生”, 它的对立事件 $\overline{ABC}$ 表示“ A, B, C 不都发生”, 亦即“ A, B, C 至少有一个不发生”。

从(5), (6)可以看到

$$\overline{ABC} \neq \overline{A} \cup \overline{B} \cup \overline{C}, \overline{A \cup B \cup C} \neq \overline{A} \cup \overline{B} \cup \overline{C}$$

关于事件的运算有下列基本关系式:

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 $(A \cup B)C = AC \cup BC, AB \cup C = (A \cup C)(B \cup C)$
4. 德·莫根定理(对偶公式)

$$\overline{A \cup B} = \overline{A} \overline{B}, \quad \overline{AB} = \overline{A} \cup \overline{B}$$

对于 n 个事件, 甚至可列个事件, 德·莫根定理也成立。

5. 幂等律 $A \cup A = A, AA = A$
6. 零壹律 $A \cup \emptyset = A, A \cap \emptyset = \emptyset, A \cup \Omega = \Omega, A\Omega = A$

第三节 随机事件的概率

一、概率的统计定义

在随机试验中, 有些事件发生的可能性大些, 有些事件发生的可能性小些。随机事件发生可能性的大小是能够量化的, 即可以用一个数值来刻画一个事件发生的可能性的大小。为此, 我们先介绍随机事件的频率的概念。

设随机事件 A 在 n 次试验中出现 n_A 次, 则 n_A 称作 A 在这 n 次试验中的频数, 比值

$$f_n(A) = \frac{n_A}{n} \quad (1-1)$$

称作 A 在这 n 次试验中的频率。

显然,在 n 次试验中随机事件 A 的频数具有随机性,从而它的频率也具有随机性。但随着试验次数 n 的增大,频率却呈现出明显的规律性,即所谓频率的稳定性。下面通过例题加以说明。

例 1. 将一枚均匀的硬币连抛 n 次,观察在 n 次试验中正面(事件 A)出现的频率。现将硬币连抛 5 次、50 次、500 次各做了 10 轮,所得数据如表 1-1 所示。

表 1-1

试验轮次	$n=5$		$n=50$		$n=500$	
	n_A	$f_n(A)$	n_A	$f_n(A)$	n_A	$f_n(A)$
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	264	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 1-1 可以看出,当抛硬币的次数较少时(如 $n=5$),事件 A (正面)在各轮中的频率差异较大,即频率是不稳定的。但是,当抛硬币的次数增多时(如 $n=500$),事件 A 在各轮中的频率虽然并不完全相同,却明显地在一个固定的数值 0.5 附近摆动,这种性质称作频率的稳定性,这个固定的数值称作频率的稳定值。历史上有许多人进行过抛硬币试验,从他们的试验结果中(表 1-2),可进一步看出事件 A (正面)的频率所具有的稳定性。