

工科高等学校教学用书

概率论与数理统计

高雷阜 柴岩 主编



NEUPRESS
东北大学出版社

工科高等学校教学用书

概率论与数理统计

主 编 高雷阜 柴 岩

东北大学出版社

· 沈 阳 ·

4502

© 高雷阜 柴 岩 2005

图书在版编目 (CIP) 数据

概率论与数理统计 / 高雷阜, 柴岩主编. — 沈阳: 东北大学出版社, 2005.8 (2006.7 重印)

ISBN 7-81102-191-9

I. 概… II. ①高… ②柴… III. 概率论; 数理统计
IV. O21

中国版本图书馆 CIP 数据核字 (2005) 第 098880 号

出 版 者: 东北大学出版社

地址: 沈阳市和平区文化路 3 号巷 11 号

邮编: 110004

电话: 024—83687331 (市场部) 83680267 (社务室)

传真: 024—83680180 (市场部) 83680265 (社务室)

E-mail: neuph @ neupress.com

http: // www. neupress. com

印 刷 者: 沈阳市光华印刷厂

发 行 者: 东北大学出版社

幅面尺寸: 140mm × 203mm

印 张: 10.125

字 数: 272 千字

出版时间: 2005 年 8 月第 1 版

印刷时间: 2006 年 7 月第 2 次印刷

责任编辑: 刘乃义

责任校对: 文 浩

封面设计: 唐敏智

责任出版: 杨华宁

定 价: 15.00 元

前 言

20 世纪下半叶以来，人们对事物认识演化的表现之一是从单纯的确定性思维模式进入确定-随机性思维模式。相应地，概率论与数理统计由于自己别开生面的课题、独特的概念和方法成为数学一个极具特色和活跃的分支之一。

本书是作者在多年教学实践和教材修改的基础上，根据国家教育部高等学校工科本科生《概率论与数理统计》课程的教学基本要求，并参考工科硕士研究生入学考试大纲编写而成的。在编写中，充分考虑了概率论的公理化结构和随机模式认识的训练以及数学应用软件在统计计算中的强大辅助作用。可作为高等学校工科、理科（非数学专业）概率论与数理统计的教材，也可供工程技术人员参考。

全书分为三大部分：概率论部分（第 1 章至第 5 章）作为基础知识，为读者提供了必要理论基础；数理统计部分（第 6 章至第 9 章）讲述了数理统计的基本理论和方法；统计计算部分（第 10 章）介绍了基于 MATLAB 的统计数据处理方法。

本书第 1, 2, 3, 4 章由高雷阜编写，第 6, 7, 8 章

由柴岩编写，第5、9章由齐俊玲编写，第10章由胡行华编写，最后由高雷卓统稿。

在编写过程中，得到了辽宁工程技术大学教务处和基础科学部的大力支持和帮助。本校石素英教授阅读了全部原稿并提出了宝贵的意见和建议，王磊、李卓同志承担了本书的校对工作。谨此一并致谢。

限于编者的水平，疏漏之处在所难免，恳请广大读者批评指正。

编 者

2005年7月

目 录

第 1 章 随机事件与概率	1
1.1 样本空间与随机事件	2
1.1.1 样本空间	2
1.1.2 随机事件	3
1.1.3 事件的关系和运算	4
1.2 古典概率	7
1.2.1 古典概率定义	7
1.2.2 排列与组合	9
1.2.3 古典概率计算举例	11
1.2.4 概率的性质	15
1.3 几何概率	18
1.4 统计概率	22
1.5 概率的公理化定义	24
1.6 条件概率	27
1.6.1 条件概率	27
1.6.2 乘法定理	29
1.6.3 全概率公式和贝叶斯公式	31
1.7 独立性	34
练习题	38

第 2 章 随机变量及其分布	43
2.1 随机变量	43
2.2 离散型随机变量的概率分布	44
2.2.1 退化分布	45
2.2.2 (0—1)分布	45
2.2.3 伯努利试验与二项分布	45
2.2.4 泊松分布	47
2.3 随机变量的分布函数	48
2.4 连续型随机变量的概率密度	53
2.4.1 均匀分布	55
2.4.2 正态分布	57
2.5 随机变量的函数的分布	62
练习题	67
第 3 章 多维随机变量及其分布	72
3.1 二维随机变量	72
3.2 边缘分布	77
3.3 条件分布	81
3.4 相互独立的随机变量	86
3.5 两个随机变量的函数的分布	91
3.5.1 $Z = X + Y$ 的分布	91
3.5.2 $Z = \frac{X}{Y}$ 的分布	95
3.5.3 $M = \max\{X, Y\}$ 及 $N = \min\{X, Y\}$ 的分布	97
练习题	100
第 4 章 随机变量的数字特征	105
4.1 数学期望	105

4.2 方 差	114
4.3 几种重要随机变量的数学期望及方差	119
4.4 协方差及相关系数	122
4.5 矩与协方差矩阵	126
练习题	130
第 5 章 极限定理	135
5.1 切比雪夫不等式	135
5.2 大数定理	137
5.3 中心极限定理	141
练习题	145
第 6 章 数理统计的基本概念	147
6.1 总体与样本	147
6.1.1 总体与个体	147
6.1.2 抽样与样本	149
6.1.3 样本的分布	150
6.2 统计量和样本的数字特征	151
6.2.1 统计量	151
6.2.2 样本的数字特征	152
6.3 抽样分布	154
6.3.1 χ^2 分布	154
6.3.2 t 分布	156
6.3.3 F 分布	158
6.3.4 正态总体的样本均值与样本方差的分布	160
练习题	165
第 7 章 参数估计	168
7.1 点估计	168

7.1.1	点估计的概念	168
7.1.2	矩估计法	169
7.1.3	极大似然估计法	174
7.2	估计量的评价标准	180
7.2.1	无偏性	181
7.2.2	有效性	183
7.2.3	一致性 (相合性)	184
7.3	区间估计	184
7.3.1	置信区间的概念	185
7.3.2	单个正态总体的期望与方差的区间估计	186
7.3.3	两个正态总体的均值差和方差比的置信区间	191
7.3.4	单侧置信区间	195
	练习题	198
第 8 章	假设检验	204
8.1	假设检验	204
8.1.1	什么是假设检验	204
8.1.2	单边检验	208
8.2	单个正态总体均值的检验	212
8.3	单个正态总体方差的检验	213
8.4	两个正态总体均值的检验	216
8.5	两个正态总体方差的检验	220
8.6	总体分布的检验	224
8.6.1	皮尔逊 χ^2 检验法	224
8.6.2	秩和检验法	229
	练习题	231
第 9 章	方差分析和回归分析	236
9.1	单因子方差分析	236

9.2 二因子方差分析	243
9.3 一元线性回归	253
9.3.1 回归直线方程的求法	253
9.3.2 回归方程的显著性检验	257
9.3.3 利用回归方程进行预测	260
9.4 某些一元非线性回归的线性化处理	261
练习题	266
第 10 章 基于 MATLAB 的统计数据处理	272
10.1 统计的基本概念	272
10.1.1 总体和样本	272
10.1.2 频数表和直方图	273
10.1.3 统计量	275
10.1.4 统计中几个重要的概率分布	277
10.1.5 正态总体统计量的分布	283
10.2 参数估计	284
10.2.1 点估计	284
10.2.2 均值的区间估计	284
10.2.3 方差的区间估计	286
10.2.4 参数估计的 MATLAB 实现	287
10.3 假设检验	288
10.3.1 均值的假设检验	288
10.3.2 方差 (或标准差) 的假设检验	291
10.3.3 两总体的假设检验	292
10.3.4 (0—1) 分布总体均值的假设检验	294
10.3.5 总体分布正态性检验	295
练习题	296
附表 A 标准正态分布函数表	297
附表 B $N(0, 1)$ 常用临界值表	298

附表 C 泊松分布累计概率表	298
附表 D t 分布临界值表	301
附表 E χ^2 分布临界值表	302
附表 F F 分布临界值表	304
附表 G 秩和检验临界值表	313
附表 H 相关系数检验临界值($r_{\frac{\alpha}{2}}$)表	314

第 1 章 随机事件与概率

概率论与数理统计是研究客观现象统计规律性的数学学科。在大量同类客观现象中，就个别现象而言，其结果是不肯定的，但从整个集体现象来说，却遵从一定的规律，这种规律性叫做统计规律性。概率论与数理统计的任务，就在于透过大量表面的偶然性发现内部隐藏着的规律，通过随机性去认识决定性，通过偶然性去认识必然性。随机性和决定性、偶然性和必然性的矛盾，是概率论与数理统计研究的主要矛盾。

赌博现象有一种独特的性质曾引起了概率论学者的研究：它的不肯定性使得人们在一次特定的赌博中不能预测结果，但若多次进行下去，情形就不同了，人们可以预测平均赢利，可以谈论两种赌注中哪一种更为有利。其他许多现象也具有同样性质，例如某地区种植某种庄稼的收成，大规模生产中废品的件数，某种零件的寿命等，这些现象在单独一次观测中其结果是不能预测的，但经多次重复观测或试验后，就会呈现某种规律性，概率论的研究目的就在于构造这类随机现象的数学模型。为了构造这种模型，必须列出一切可能的结果来精确地描述一个试验，例如掷一枚硬币，人们关心的是它落下时出现“正面”还是“反面”，当然用不着去考虑那些偶发事件（如硬币侧立，或者掉入洞里等）。又如掷一颗骰子，有六种可能的结果，可以用朝上那一面的点数来表示，试验的结果称为“事件”。在掷一颗骰子时，人们还可以按出现奇数点或偶数点来打赌，“出现偶数点”这一事件可以以三种不同的方式发生（即出现 2 点、4 点或 6 点），但“出现 2 点”这一事件却只能以一种方式发生。通常把前者称为复合事件，后者称为简单事件（或基本事件）。假设掷一颗骰子 n 次，而出现 6 点的情况有 m 次，那么，就称 m/n 为 n 次试验中“出现 6 点”这一事件

的相对频率. 再假定骰子是均匀的, 将会看到当 n 很大时, m/n 接近 $1/6$. 所谓的统计规律性就表现在这种相对频率的稳定性中, 所构造的数学模型应当反映事件的这种性质. 因此, 对每个事件指定的一个数, 作为相对频率的理想化数字(稳定值), 称为这个事件的**概率**.

1.1 样本空间与随机事件

1.1.1 样本空间

研究随机现象, 总是要进行观察、测量或做各种科学试验. 为了叙述方便起见, 将这一切活动统称为**试验**.

一个试验, 若满足下述条件:

① 试验可以在相同条件下重复地进行;

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.

就称其为**随机试验**, 简称**试验**, 以字母 E 表示.

如掷硬币的试验, 可以在相同条件下重复进行, 试验的可能结果有两个, 即“正面”和“反面”, 每次试验结果必出现其中之一, 但投掷之前不可能预言出现正面还是反面.

本书中以后提到的试验都是指随机试验.

一般通过研究随机试验来研究随机现象. 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. 通常将随机试验 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**, 记为 w .

【例 1.1】 随机试验 E_1 : 将一枚硬币抛掷三次, 观察正面 H 、反面 T 出现的情况.

E_1 的样本空间 $S_1: \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.

【例 1.2】 随机试验 E_2 : 将一枚硬币抛掷三次, 观察出现正面的次数.

E_2 的样本空间 $S_2: \{0, 1, 2, 3\}$.

【例 1.3】 随机试验 E_3 : 记录电话交换台一分钟内接到的呼唤次数.

E_3 的样本空间 $S_3: \{0, 1, 2, 3, \dots\}$.

【例 1.4】 随机试验 E_4 : 在一批灯泡中任意抽取一次, 测试它的寿命.

E_4 的样本空间 $S_4: \{t \mid t \geq 0\}$.

【例 1.5】 随机试验 E_5 : 记录某地一昼夜的最高温度和最低温度.

E_5 的样本空间 $S_5: \{(x, y) \mid T_0 \leq x \leq y \leq T_1\}$, 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低于 T_0 , 也不会高于 T_1 .

【例 1.6】 随机试验 E_6 : 一尺之棰, 折成三段, 观察其长度.

E_6 的样本空间 $S_6: \{(x, y) \mid 0 < x < 1, 0 < y < 1, 0 < x + y < 1\}$.

用 x, y 分别表示折成的第一段和第二段的长度, 则第三段的长度为 $1 - x - y$, 且各段长度都必须大于 0 小于 1.

应该特别注意的是: 样本空间的元素是由试验目的所确定的, 例如, 在 E_1 和 E_2 中同是将一枚硬币连抛三次, 由于试验目的不同, 其样本空间也不一样.

1.1.2 随机事件

实际进行随机试验时, 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(h)小于 500 为

次品,则在 E_4 中人们关心灯泡的寿命是否有 $t \geq 500$, 满足这一条件的样本点组成 S_4 的一个子集: $A = \{t \mid t \geq 500\}$. 称 A 为试验 E_4 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 $t \geq 500$.

一般地, 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.

特别地, 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 8 个基本事件 $\{HHH\}, \{HHT\}, \dots, \{TTT\}$; 试验 E_2 有 4 个基本事件 $\{0\}, \{1\}, \{2\}, \{3\}$.

样本空间 S 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**; 空集不含任何样本点, 它也作为样本空间的子集, 在每次试验中都不发生, 称为**不可能事件**.

【例 1.7】 在 E_1 中, 事件 A_1 : “三枚硬币出现同一面”, 即

$$A_1 = \{HHH, TTT\};$$

事件 A_2 : “恰有两枚硬币出现同一面”, 即

$$A_2 = \{HHT, HTH, THH, HTT, THT, TTH\};$$

事件 A_3 : “至少两枚硬币出现正面”, 即

$$A_3 = \{HHH, HHT, HTH, THH\};$$

事件 A_4 : “恰有两枚硬币出现正面”, 即

$$A_4 = \{HHT, HTH, THH\}.$$

在 E_4 中, 事件 A_5 : “寿命小于 1000h”, 即

$$A_5 = \{t \mid 0 \leq t < 1000\}.$$

在 E_5 中, 事件 A_6 : “最高温度与最低温度相差 10°C ”, 即

$$A_6 = \{(x, y) \mid y - x = 10, T_0 \leq x \leq y \leq T_1\}.$$

1.1.3 事件的关系和运算

事件既然是一个集合, 因而事件间的关系与事件的运算自然

可以按照集合论中集合之间的关系和集合运算相应处理,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 $A, B, A_k (k = 1, 2, \dots)$ 是 S 的子集.

(1) 事件的包含与相等

若 $A \subset B$, 则称事件 B 包含事件 A , 此时, 事件 A 发生必然导致事件 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) 事件的和

事件 $A \cup B = \{x \mid x \in A \text{ 或 } x \in B\}$ 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 $\bigcup_{k=1}^n A_k$ 为 n 个事件 A_1, A_2, \dots, A_n 的和事件, 称 $\bigcup_{k=1}^{\infty} A_k$ 为可列个事件 A_1, A_2, \dots 的和事件.

(3) 事件的积

事件 $A \cap B = \{x \mid x \in A \text{ 且 } x \in B\}$ 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 $\bigcap_{k=1}^n A_k$ 为 n 个事件 A_1, A_2, \dots, A_n 的积事件, 称 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件 A_1, A_2, \dots 的积事件.

(4) 事件的差

事件 $A - B = \{x \mid x \in A \text{ 且 } x \notin B\}$ 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) 互不相容事件

若 $A \cap B = \emptyset$, 则称事件 A 与 B 是互不相容的, 或称互斥的. 即事件 A 与事件 B 不能同时发生, 基本事件是两两互不相容的.

(6) 对立事件

若 $A \cup B = S$ 且 $A \cap B = \emptyset$, 则称事件 A 与事件 B 互为对立事件, 又称事件 A 与事件 B 互为逆事件. 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 \bar{A} , 且 $\bar{A} = S - A$.

用图 1.1 ~ 1.6 可直观地表示以上事件之间的关系与运算, 如图 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 . 又在如图 1.2 中正方形表示样本空间 S , 圆 A 与 B 分别表示事件 A 与事件 B , 而阴影部分表示和事件 $A \cup B$.

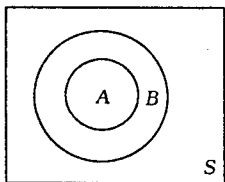


图 1.1

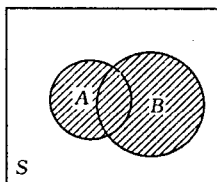


图 1.2

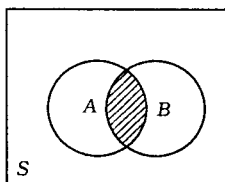


图 1.3

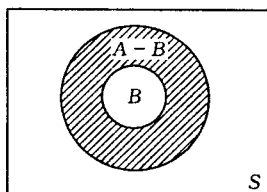
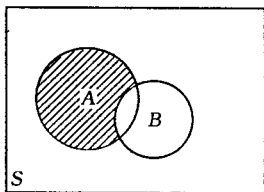


图 1.4

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

交换律: $A \cup B = B \cup A$;

$A \cap B = B \cap A$.

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