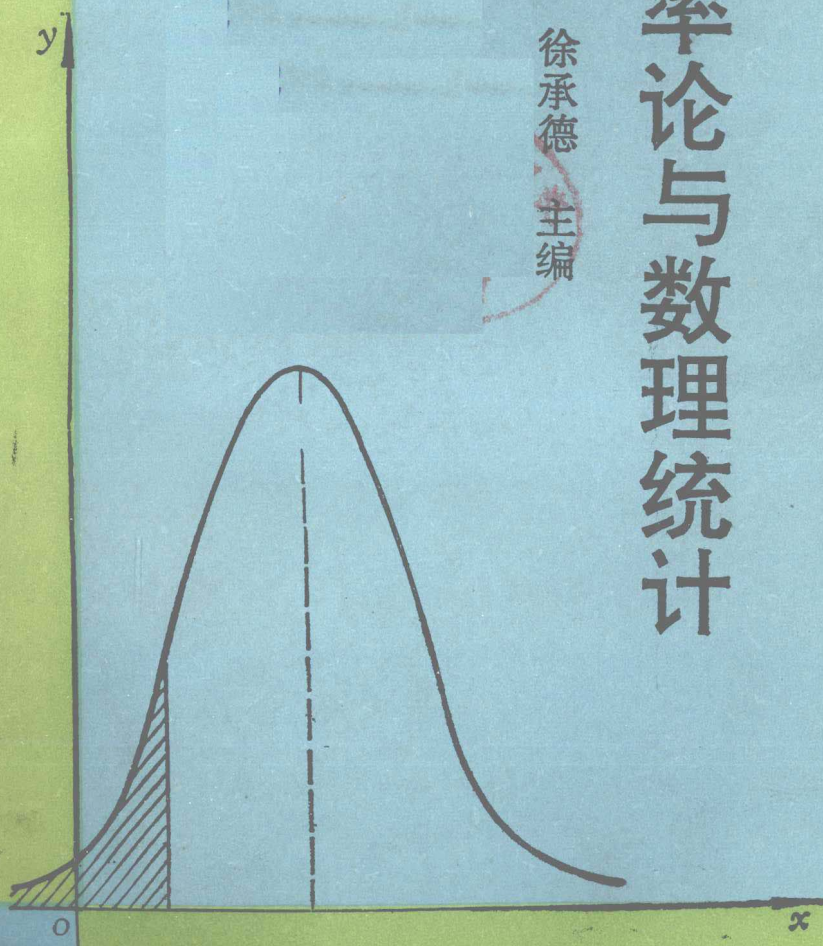


概率论与数理统计

曹彬

徐承德

主编



哈尔滨工业大学出版社

概率论与数理统计

曹 彬 许承德 主编

哈尔滨工业大学出版社

S00888
(黑) 新登字第4号

概率论与数理统计

曹 彬 许承德 主编

*

哈尔滨工业大学出版社出版
新华书店首都发行所发行
黑龙江省教委印刷厂印刷

*

开本 787×1092 1/32 印张 12.25 字数 275千字

1993年9月第1版 1993年9月第1次印刷

印数 1—7000

ISBN 7-5603-0530-X/O·42 定价: 8.50元

前 言

本书是为高等工科院校《概率论与数理统计》课程而编写的，它包括了1986年工科数学课程指导委员会制订的“概率论与数理统计教学基本要求”的全部内容。为了适应更多读者的需要，我们适当地增加了部分内容，并在这部分内容前打上了“*”号。本书除供工科各专业使用外，也可供某些非工科专业选用，还可作为工程技术人员自学参考书。

全书内容共分七个部分：第一部分为事件及其概率的概念与运算（第一、二章）；第二部分为随机变量及其分布（第三、四章）；第三部分为随机变量的数字特征（第五章）；第四部分为极限定理（第六章）；第五部分为数理统计的基本概念（第七章）；第六部分为估计和检验的基本方法（第八、九章）；第七部分为线性模型的统计分析初步（第十章）。每章后附有习题，书末附有习题答案。阅读本书只需具备高等工科院校微积分的数学基础。

本书原稿曾作为哈尔滨工业大学校内教材（曹彬、许承德编）使用多年，效果较好。在此基础上，陈桂林、王勇、关忠根据近年来的教学实践对本书又进行了认真修改、补充。由于编者水平有限，缺点错误在所难免，恳请读者批评指正。

武汉大学胡迪鹤教授和中国纺织大学吴让泉教授曾分别

编者

目 录

第一章 随机事件与概率	(1)
§ 1.1 随机事件	(1)
1.1.1 必然现象与随机现象	(1)
1.1.2 随机试验与事件、样本空间	(2)
§ 1.2 事件的关系与运算	(7)
§ 1.3 古典概率	(12)
1.3.1 古典概率定义	(13)
1.3.2 排列与组合	(14)
1.3.3 古典概率计算的例子	(17)
1.3.4 概率的性质	(20)
§ 1.4 几何概率	(25)
§ 1.5 统计概率	(27)
§ 1.6 概率的公理化定义	(30)
习题	(33)
第二章 条件概率与独立性	(38)
§ 2.1 条件概率、乘法定理	(38)
§ 2.2 全概率公式	(43)
§ 2.3 贝叶斯公式	(45)
§ 2.4 事件的独立性	(47)
2.4.1 两个事件的独立性	(47)
2.4.2 多个事件的独立性	(50)

§ 2.5	重复独立试验、二项概率公式	(54)
§ 2.6	泊松逼近	(57)
	习题	(63)
第三章	随机变量及其分布	(67)
§ 3.1	随机变量的概念	(67)
§ 3.2	离散型随机变量	(69)
3.2.1	概率分布列	(69)
3.2.2	0—1 分布 (贝努里分布、两点分布)	(70)
3.2.3	二项分布	(70)
3.2.4	泊松分布	(72)
§ 3.3	随机变量的分布函数	(76)
§ 3.4	连续型随机变量	(80)
3.4.1	连续型随机变量、概率密度	(80)
3.4.2	均匀分布	(84)
3.4.3	指数分布	(86)
§ 3.5	正态分布	(88)
§ 3.6	随机变量函数的分布	(94)
	习题	(103)
第四章	多维随机变量及其分布	(108)
§ 4.1	多维随机变量及其分布函数、边缘分布函数	(108)
§ 4.2	二维离散型随机变量	(111)
§ 4.3	二维连续型随机变量	(113)
4.3.1	概率密度及边缘概率密度	(113)
4.3.2	二维均匀分布	(116)
*4.3.3	二维正态分布	(118)

§ 4.4	条件分布	(120)
§ 4.5	随机变量的独立性	(125)
§ 4.6	二维随机变量函数的分布	(130)
4.6.1	和的分布	(130)
4.6.2	商的分布	(137)
4.6.3	瑞利分布	(139)
4.6.4	$\max(X, Y)$ 及 $\min(X, Y)$ 的分布	(140)
习题		(142)
第五章	随机变量的数字特征	(149)
§ 5.1	数学期望	(149)
5.1.1	离散型随机变量的数学期望	(149)
5.1.2	连续型随机变量的数学期望	(152)
5.1.3	随机变量函数的数学期望	(154)
5.1.4	数学期望的性质	(158)
§ 5.2	方差	(164)
5.2.1	方差的概念	(164)
5.2.2	方差的性质	(168)
§ 5.3	协方差和相关系数	(170)
§ 5.4	矩、协方差矩阵	(175)
5.4.1	矩	(175)
*5.4.2	协方差矩阵	(177)
习题		(179)
第六章	大数定律与中心极限定理	(185)
§ 6.1	大数定律	(185)
6.1.1	切比晓夫 (Tchebysheff) 不等式	(185)
6.1.2	大数定律	(187)

§ 6.2 中心极限定理	(190)
习题	(198)
第七章 数理统计的基本概念	(202)
§ 7.1 总体与样本	(202)
7.1.1 数理统计的基本问题	(202)
7.1.2 总体	(204)
7.1.3 样本	(205)
§ 7.2 直方图与经验分布函数	(208)
§ 7.3 χ^2 、 t 和 F 分布	(213)
7.3.1 χ^2 分布	(214)
7.3.2 t 分布	(217)
7.3.3 F 分布	(219)
§ 7.4 统计量及抽样分布	(221)
§ 7.5 $\bar{X}$ 和 S^2 的观察值的计算	(228)
习题	(231)
第八章 参数估计	(236)
§ 8.1 点估计	(236)
8.1.1 矩估计法	(237)
8.1.2 极大似然估计法	(240)
8.1.3 鉴定估计量的标准	(245)
§ 8.2 区间估计	(249)
8.2.1 单个正态总体参数的区间估计	(251)
8.2.2 两个正态总体参数的区间估计	(256)
*8.2.3 大样本区间估计	(258)
习题	(260)
第九章 假设检验	(266)
§ 9.1 假设检验的基本概念	(266)

9.1.1	问题的提出	(266)
9.1.2	假设检验的基本思想	(268)
9.1.3	假设检验中的两类错误	(269)
§ 9.2	单个正态总体参数的显著性检验	(270)
9.2.1	u 检验	(270)
9.2.2	t 检验	(275)
9.2.3	χ^2 检验	(277)
§ 9.3	两个正态总体参数的显著性检验	(280)
9.3.1	t 检验 (续)	(280)
9.3.2	F 检验	(281)
§ 9.4	非参数假设检验	(282)
9.4.1	正态概率纸检验	(283)
9.4.2	χ^2 拟合检验	(287)
9.4.3	秩和检验	(293)
习题		(296)

***第十章 单因素试验的方差分析及**

一元正态线性回归	(300)
§ 10.1 单因素试验的方差分析	(300)
§ 10.2 一元正态线性回归	(313)
10.2.1 一元正态线性回归的数学模型	(313)
10.2.2 未知参数的估计	(315)
10.2.3 α 和 δ 的数学期望与方差 以及 σ^2 的无偏估计	(318)
10.2.4 回归方程的显著性检验	(322)
10.2.5 利用回归方程进行预测和控制	(329)
10.2.6 一元非线性回归	(336)
习题	(339)

附录 1	定理 7.3 的证明	(343)
附录 2	定理 7.4 的证明	(344)
习题解答		(346)
参考书目		(366)
附表 1	泊松分布累计概率值表	(367)
附表 2	标准正态分布函数值表	(368)
附表 3	χ^2 分布表	(369)
附表 4	t 分布表	(371)
附表 5	F 分布表	(372)
附表 6	秩和检验表	(381)
附表 7	相关系数检验表	(382)

第一章 随机事件与概率

§1.1 随机事件

1.1.1 必然现象与随机现象

人们在实践活动中所遇到的现象，一般来说可分为两类：一类是**必然现象**，或称**确定性现象**；另一类是**随机现象**，或称**不确定性现象**。

必然现象是指在相同条件下重复试验，所得结果总是确定的现象；只要试验条件不变，试验结果在试验之前是可以预言的。例如：在标准大气压下，将水加热到 100°C ，水必然沸腾；用手向空中抛出的石子，必然下落；作等速直线运动的物体，如无外力作用，必然继续作等速直线运动等等，这些现象都是必然现象。

随机现象是指在相同条件下重复试验，所得结果不一定相同的现象，即试验结果是不确定的现象；对这种现象来说，在每次试验之前哪一个结果发生，是无法预言的。例如：新生婴儿，可能是男孩，也可能是女孩；向一目标进行射击，可能命中目标，也可能不命中目标；从一批产品中，随机抽检一件产品，结果可能是合格品，也可能是次品；测量某个物理量，由于许多偶然因素的影响，各次测量结果不一定相同等等，这些现象都是随机现象。

对随机现象，是否有规律性可寻呢？人们经过长期的反复实践，发现这类现象虽然就每次试验结果来说，具有不确定性，但大量重复试验，所得结果却呈现出某种规律性。例

如：

(a) 掷一枚质量均匀的硬币，当投掷次数很大时，就会发现正面和反面出现的次数几乎各占 $1/2$ 。历史上，蒲丰 (buffon) 掷过 4040 次，得到 2048 次正面；皮尔逊 (K. Pearson) 掷过 24000 次，得到 12012 次正面。

(b) 对一个目标进行射击，当射击次数不多时，对弹孔的分布看不出有什么规律性；但当射击次数非常多时，就可发现弹孔的分布呈现一定的规律性：弹孔关于目标的分布略呈对称性，且越靠近目标的弹孔越密，越远离目标的弹孔越稀。

(c) 从分子物理学观点来看，气体分子对器壁的压力是气体分子对器壁碰撞的结果。由于分子是时刻不停地、杂乱无章地运动着，速度和轨道都是随机的，因而对器壁的碰撞也是随机的。初看起来器壁所受的压力是不稳定的，可是实验证明，由于分子数目非常大，各分子运动所具有的随机性在集体中互相抵消、互相平衡了，使得器壁所受的总压力呈现一种稳定性。分子数目越大，压力越稳定。

从上述各例可以看到，随机现象也包含着规律性，它可在相同条件下的大量重复试验或观察中呈现出来。这种规律性称为**随机现象的统计规律性**。

革命导师恩格斯指出：“在表面上是偶然性在起作用的地方，这种偶然性始终是受内部的隐蔽着的规律支配的，而问题只是在于发现这些规律。”^①

概率论就是研究随机现象统计规律的一门数学学科。

1.1.2 随机试验与事件、样本空间

^① 《马克思恩格斯选集》第四卷，第243页。

对随机现象的研究，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，我们统称为试验。例如，掷一硬币，观察哪面朝上；向一目标进行射击，观察是否命中；从一批产品中随机抽一产品，检查它是否合格；向坐标平面内任投一根针，测量此针的针尖指向与 x 轴正向之间的交角，等等；这些都是试验。仔细分析，这些试验具有如下的共同特点：

- (a) 试验可以在相同条件下重复进行；
- (b) 试验的所有可能的结果不止一个，而且是事先已知的；
- (c) 每次试验总是恰好出现这些可能结果中的一个，但究竟出现哪一个结果，试验前不能确切预言。

如掷硬币的例子，试验是可以在相同条件下重复进行的，试验的可能结果有两个，即正面和反面；每次试验必出现其中之一，但投掷之前是不可能预言正面出现还是反面出现。

我们将满足上述三个条件的试验，称为**随机试验**，简称**试验**，以字母 E 表示。

在试验中，可能发生也可能不发生的事件称为**随机事件**，简称**事件**，以字母 $A, B, C, \dots$ 表示。

例如，为检查一批产品的质量，随机地从这批产品中抽取100件来检查，结果可能是：“没有次品”，“有一件次品”，“有两件次品”， $\dots$ ，“100件都是次品”。显然，这些都是随机事件。此外，“次品数不少于5”，“次品数大于5小于10”，“次品数是偶数”，等等，也都是随机事件；不过这些事件，都是由前面那些事件组成的。例如，“次品数不小于5”这个事件，是由“有5件次品”，“有

6 件次品”， $\cdots$ ，“100 件都是次品”所组成。因此，这事件称为复合事件，而前面的那些事件称为基本事件。

一般说，随机试验的每一个可能结果，称为**基本事件**，而由多个基本事件组成的事件，称为**复合事件**。

在前面的例子中，还可以考虑两个特殊的事件：“次品数大于 100”和“次品数不大于 100”。由于 100 件产品中，次品数是不可能大于 100 的；因而，“次品数大于 100”是不可能发生的，而次品数不大于 100”是必然发生的。一般，我们将在一定的试验条件下，必然发生的事件称为**必然事件**，不可能发生的事件称为**不可能事件**。由于必然事件与不可能事件的发生与否，已失去了“不确定性”，因而本质上它们已不是随机事件；但为了以后讨论问题的方便，我们仍然把它们看作随机事件。

为了利用点集的知识来描述随机事件，我们引进样本空间的概念。

由于随机试验的所有可能的结果是已知的，从而所有的基本事件也是已知的。我们将基本事件的全体，称为**随机试验的样本空间**，记为 S ；基本事件也称为**样本点**，用 e 表示。

例如，在上述检查产品质量的试验中，若令 $e_i =$ “有 i 个次品”， $i = 0, 1, 2, \cdots, 100$ ；则

$$S = \{e_0, e_1, e_2, \cdots, e_{100}\}$$

由于随机事件是基本事件或是由基本事件组成的复合事件，故随机事件本身就是样本空间 S 的一个子集。

如在上例中，设 A 表示“次品数不少于 5”的事件，则 $A = \{e_5, e_6, \cdots, e_{100}\}$ ；设 B 表示“次品数大于 5 小于 10”的事件，则 $B = \{e_6, e_7, e_8, e_9\}$ ；设 C 表示“有 5 个次品”的

事件, 则 $C = \{e_s\}$ 。

由此可知, 引入样本空间之后, 事件便可定义为样本空间的子集; 而且, 当且仅当子集中的一个基本事件在试验中发生了, 才称事件发生。

样本空间 S 和空集 $\emptyset$ 作为 S 的子集也看作事件。由于 S 包含所有的基本事件, 故在每次试验中, 必有一个基本事件 $e \in S$ 发生, 即在试验中, 事件 S 必然发生; 因此, S 是**必然事件**。又因在 $\emptyset$ 中不包含任何一个基本事件, 故在任一次试验中, $\emptyset$ 永远不会发生; 因此, $\emptyset$ 是**不可能事件**。今后我们常用 S , $\emptyset$ 分别表示必然事件与不可能事件。

对一个随机试验, 应当弄清楚它的样本空间是什么, 这是弄清事件这一概念的基础, 下面看几个例子。

例 1 掷一均匀对称的硬币, 观察正反面出现的情况, 这是个随机试验。可能结果有两个: 正(正面朝上), 反(反面朝上)。故样本空间

$$S = \{\text{正}, \text{反}\}$$

例 2 将上述硬币掷两次, 观察正反面情况, 这也是一个随机试验。试验的可能结果有四个: (正, 正), (正, 反), (反, 正), (反, 反)。^① 故样本空间

$$S = \{(\text{正}, \text{正}), (\text{正}, \text{反}), (\text{反}, \text{正}), (\text{反}, \text{反})\}$$

在这个试验中, 若设 $A =$ “第一次出现正面”, 则 $A = \{(\text{正}, \text{正}), (\text{正}, \text{反})\}$; 若设 $B =$ “两次出现同一面”, 则 $B = \{(\text{正}, \text{正}), (\text{反}, \text{反})\}$; 若设 $C =$ “至少有一次出现正面”, 则 $C = \{(\text{正}, \text{正}), (\text{正}, \text{反}), (\text{反}, \text{正})\}$; 若设 $D =$ “第一次出现反面”, 则 $D = \{(\text{反}, \text{正}), (\text{反}, \text{反})\}$ 。

^①括号内第一个和第二个字, 分别表示第一次和第二次掷的结果。

反) }。

例 3 记录某电话交换台在一段时间内接到的呼叫次数, 这个试验的基本事件(记录结果)是一非负的整数, 由于难于规定一个呼叫次数的上界, 所以样本空间

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

在这个试验中, 若设 $A =$ “呼叫次数不超过三次”, 则 $A = \{0, 1, 2, 3\}$; 若设 $B =$ “呼叫次数大于五次”, 则 $B = \{6, 7, 8, \dots\}$ 。

例 4 从一批灯泡中抽取一只灯泡, 测试它的使用寿命。设 t 表示寿命, 则样本空间 $S = \{t: t \geq 0\}$; 基本事件 $e = t$ 。设事件 $A =$ “寿命小于 5 小时”, 则 $A = \{t: 0 \leq t < 5\}$ 。

例 5 观察某地区一昼夜最低温度 x 和最高温度 y 。设这个地区的温度不会小于 T_0 , 也不会大于 T_1 , 则样本空间 $S = \{(x, y): T_0 \leq x < y \leq T_1\}$; 基本事件 $e = (x, y)$, $T_0 \leq x < y \leq T_1$ 。设事件 $A =$ “最高温度与最低温度之差不超过 10°C ”, 则 $A = \{(x, y): y - x \leq 10, T_0 \leq x < y \leq T_1\}$ 。

例 6 一尺之棰, 折成三段, 观察其长度。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 $1 - x - y$ 。由于各段的长度必须大于 0 小于 1, 故 x, y 的可能值满足: $0 < x < 1, 0 < y < 1, 0 < 1 - x - y < 1$ 。从而, 样本空间

$$S = \{(x, y): 0 < x < 1, 0 < y < 1, 0 < x + y < 1\}$$

基本事件 $e = (x, y)$, $0 < x < 1, 0 < y < 1, 0 < x + y < 1$ 。

设事件 $A =$ “每段长都小于 $\frac{1}{2}$ ”, 则 A 可写成

$$A = \{(x, y): 0 < x < \frac{1}{2}, 0 < y < \frac{1}{2}, \frac{1}{2} < x + y < 1\}$$