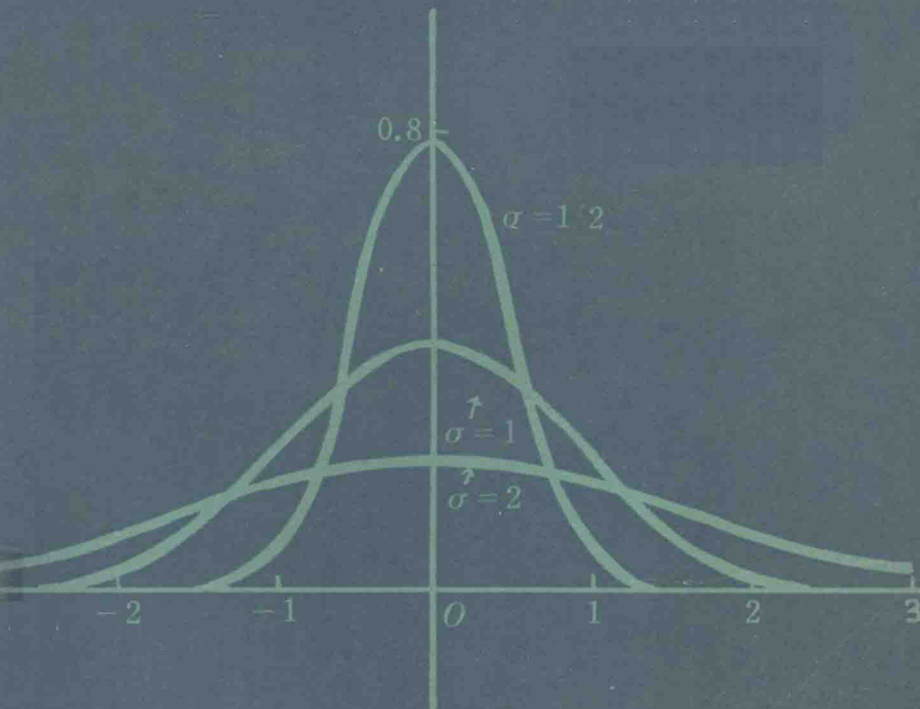
概率论与数理统计

（引论）

〔美〕E·勒克斯 著

黄志远 冯文权 胡则成译



人民教育出版社

概率论与数理统计(引论)

[美] E. 勒克斯　著
黄志远　冯文权　胡则成　译

人民教育出版社

内 容 提 要

本书叙述概率论与数理统计的基础知识，分两部分共九章．第一部分的六章叙述概率论，第二部分的三章叙述数理统计的基本概念与方法．内容简明扼要，只要求读者具有初等微积分的知识．可供理科、工科、师范等院校教学概率统计课程时参考．

概率论与数理统计(引论)

［美］E. 勒克斯 著

黄志远 冯文权 胡则成 译

*

人民教育出版社出版

新华书店北京发行所发行

人民教育出版社印刷厂印装

*

开本 850×1168 1/32 印张 6.125 字数 130,000

1982 年 2 月第 1 版 1982 年 12 月第 1 次印刷

印数 00,001—14,500

书号 13012·0700 定价 0.79 元

译者的话

概率论和数理统计是研究客观现象统计规律性的数学学科．在大量同类的客观现象中，就其个别现象来说，它的结果是不肯定的，但从整个集体现象来说，却遵守一定的规律性，这种规律性叫做统计规律性．概率论和数理统计的任务，就在于透过大量表面的偶然性去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，是概率论和数理统计研究的主要矛盾．

概率论和数理统计，同其它数学分支一样，是从实践的需要产生的．我国古代的人口调查、天文观测就应用了统计方法．系统的研究开始于 17 世纪的欧洲，当时西欧正是由封建社会向资本主义社会的过渡时期，资本主义生产方式的出现提出了精密测量、产品检查、社会保险等一系列的问题，促使人们去从事这一方面的研究．随着人们社会实践的发展，概率论和数理统计方法越来越广泛地应用于自然科学、技术科学、国民经济以及军事技术的各个部门，例如现代物理学对于微观世界的研究，无线电通信和导航，生产过程的最优控制，气象、水文和地震预报，生产企业和公用事业的计划管理，地质勘探和海洋开发以及天文学、地球物理学、化学、生物学、医学的研究等，都离不开这个方法．概率论和数理统计的发展历史，完全证明了人们的社会实践是数学发展的源泉和动力．对于古典概率论著作中有着许多赌博的语言或例子这件事，我们必须加以分析：其中一部分是以赌博作为某种物理现象的简单模型来研究的，因为这种语言生动形象，问题典型化，所以在研究和教学中都有一定价值；但是如果没有社会实践的需要，就不可能引起广泛的兴趣，概率论也不会有今天这样蓬勃的发展，因而本书

作者关于概率论起源于对赌博研究的观点，是我们所不能同意的.

这本书是作为一本简明的大学教材来写的，其特点是不需要一般概率论教材所需要的实变函数、傅立叶变换等大学高年级才学的知识作基础，只要是学过初等微积分的低年级学生，都可以通过本书掌握概率论和数理统计的基础. 因此本书适合于师范、工科院校以及理科中需要概率论和数理统计知识的非数学专业选用. 因为本书的论证严格、清楚，所以稍加补充也可作为数学专业基础课教材.

本书第一、二、三、九章由黄志远同志翻译，第四、六、七、八章由冯文权同志翻译，第五章和附录由胡则成同志翻译，全部译文由黄志远同志统一校正. 由于时间仓促、水平有限，译文难免有错，望读者批评指正.

译　者

原　序

这本书是作为大学生的一个简明的概率论和数理统计教程而写的，它的取材很受美国数学协会大学课程委员会提出的一个 2P 教程大纲的影响．因此这本书可以给 2P 教程的那些只有最低限度数学准备知识的学生阅读．

这本书要求的数学准备知识是适中的．读者应当熟悉微积分的基础知识，但是，通常在微积分教程第二或第三部分中才讲授的许多内容，我都谨慎地避开．这样当然会使得在证明定理所用的技巧上受到限制．我没有用雅可比行列式、矩母函数和特征函数，也没有用矩阵和 n 维空间的几何学，不熟悉这些内容的学生在读本书时将不会感到困难．偏微分用得很少，它只在讨论最小平方回归时出现．

由于这些限制，自然必须删去某些定理的证明而指出进一步的参考书．除此以外，所有证明都是适合中等的数学准备知识的．例如 Chi-平方分布作为独立标准正态变量平方和的分布是用数学归纳法推导的；在正文中证明了关于贝努里试验的中心极限定理，而关于独立、同分布、具有有限方差的随机变量序列的中心极限定理在正文中只作了叙述．但是在附录 C 中，我们证明了更一般形式的中心极限定理，这个证明没有用特征函数，只用了概率算子方法．尽管概率算子技巧对于这种水平的教程是适当的，我还是把它放到附录中去，因为我认为在使用本书作教材时，在时间分配上不一定允许做这样的证明．

鉴于受适中的数学准备知识的限制，不可能要求本书对这门学科作完全而严格的叙述．但我并不希望只是简单地提供一些结果或处方．我的目的是给概率论与数理统计提供一个牢固的、适当的基础．(下略)

前　　言

象掷骰子一类的赌博，早在古代就有了，但到文艺复兴时期特别盛行．这个时期的一些赌徒观察到许多不能解释的现象，于是去请教一些著名的科学家，引起了他们研究赌博的兴趣①．这些科学家的研究中应用了数学工具，但是那些工具还没有形成一种数学理论，这种研究被称为"概率论"．长时间里，概率论能不能属于物理学或数学的一部分，一直是一个疑问．这个首先由希尔伯特(D. Hilbert)明确提出的问题，在二十世纪中叶已由柯尔莫哥洛夫(A. H. Колмогоров)所解决，他奠定了概率论的严格数学基础．由于我们将采用这种体系，因此有必要先对数学理论的结构作一些解释．

数学理论要处理概念之间的关系，首先引进一些未定义的概念(基本概念)，它们的性质用某些不加证明的命题来描述，这些命题就是我们的理论的基本假设，称为"公理"．从这些公理以及由此导出的命题出发，建立一个演绎的体系．在这个体系中，引进新的概念必须给出明确而详尽的定义，然后研究它的性质；除公理外的每一个命题，必须用严格的演绎方式给出证明．因此，用来构造理论的基本概念的一切性质，都应当包含在公理中，从这种意义上可以说一组公理定义了一组基本概念，这种定义的方式表面上不同于以后引进新概念时那种明显的方式．然而从逻辑观点看来，公理起了定义相同的作用：它们阐明了基本概念的意义．形式上的不同是：所有基本概念的意义由一组公理同时规定，而以后的每个概念是明确地规定的，习惯上把前一种方式称为由公理定义或隐

---

① 参看后面所列参考文献．

式定义，以区别于后一种的显式定义．当然，所有公理必须相容，也就是说，必须保证所有公理及其推论都不包含矛盾．

从严格的逻辑观点看来，公理是可以任意选择的，但是人们很少这样去做，因为随意选择一组公理去进行构造近于一种游戏而缺乏实际意义，公理的选择常常受实验的启发，目的在于构造某些物理现象的数学模型．因此公理化系统提供了这些物理现象一种理想化的描述．基本概念对应于物理实体，公理对应于它们的性质，至少有一些命题应当对应于可观察的事实．当然，也会有一些推论无法在实践中检验——至少在现有技术条件下无法做到．

赌博现象有一种独特的性质曾引起了概率论的研究：它的不肯定性使得我们在一次特定的赌博中不能预测结果，但是，如果多次进行下去，情况就不一样了．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．其它许多现象也具有这种性质，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，某种设备零件的寿命等．这些现象在单独一次观测中其结果是不能预测的，但在多次重复观测或实验时，就会呈现某种规律性——我们称它为统计规律性．概率论的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，我们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，我们关心的是它出现“正面”还是“反面”，当然用不着去考虑那些想入非非的偶发事件(象硬币侧立着，或者掉进地上的一个洞里)．又如掷一颗骰子，有六种可能结果，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，人们还可以按是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”这一事件却只有一种方式发生．我们把前者称为复合事件，后者称为简单事件(或基本事件)．假设掷一颗骰子多次(例如 m 次)，而出现 6 点 h 次，那末 h/m 就称为 m 次

试验中“出现6点”这一事件的相对频率．假定骰子是均匀的，我们将会看到当 m 很大时 h/m 接近1/6．所谓统计规律性就表现在这种相对频率的稳定性中．我们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．

参考文献

Gerolamo Cardano(1501—1576), “De ludo aleae.” 1663年发表于Cardano选集第一卷，英译文可在Ore的书中找到．

G. Galilei(1564—1642), “Considerazione sopra il gioco dei daddi”(日期不详)Vol. 3, Opere, Firenze 1718. Vol. 14, Opere, Firenze 1855.

B. Pascal(1623—1662), “Oeuvres”, Vol. IV. Paris 1819. P. Fermat(1601—1665), “Varia opera mathematica”. Tolosac 1679.

O. Ore, “Cardano the Gambling Scholar”. Princeton Univ. Press, Princeton, New Jersey, 1953. Reprinted by Dover, New York, 1965.

D. Hilbert, Mathematical Problems (1900年巴黎举行的国际数学家会议的讲稿，英译本见 Bull. Amer. Math. Soc. 8, 437—479(1901—1902).

A. N. Kolmogorov, “Foundations of the Theory of Probability.” Springer, Berlin 1933. 英译本 Chelsea, Bronx New York, 1956.

H. Weyl, “Philosophy of Mathematics and Natural Science,” P. 27. Princeton Univ. Press, Princeton, New Jersey, 1949.

概率论历史概述可参看:

I. Todhunter, “History of the Theory of Probability.” Cambridge, 1865. Reprinted by Chelsea, Bronx, New York, 1949, 1965.

E. Czuber, “Die Entwicklung der Wahrscheinlichkeitstheorie und ihrer Anwendungen.” Jahresberichte d. Deutschen Mathematiker Vereinigung VII, No. 2, 1—127, 1899.

目　　录

第II部分 数理统计

第I部分 概 率 论

第一章 概率空间

§1.1 基本事件空间

构造随机现象的数学模型，首先是列出一切可能的结果，我们把这些结果的集合记为 Ω，叫作基本事件空间(或样本空间)，所有简单事件都是 Ω 的元素.

例 1.1.1 同时掷三个可区别的硬币，共有八个可能的结果①:

$\omega_1=$(正正正)，$\omega_2=$(正正反)，$\omega_3=$(正反正)

$\omega_4=$(正反反)，$\omega_5=$(反正正)，$\omega_6=$(反正反)

$\omega_7=$(反反正)，$\omega_8=$(反反反)

这里，基本事件空间 $\Omega=(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8)$ 由 8 个简单事件(基本事件)组成，每个简单事件只包含 Ω 中一个点. 另外还有各种复合事件，例如:

A: 三个硬币出现同一面;

B: 恰有两个硬币出现同一面;

C: 至少两个硬币出现正面;

D: 恰有两个硬币出现正面;

E: 恰有两个硬币出现反面;

F: 第一、二个硬币出现正面.

---

① "正"表示出现正面，"反"表示出现反面，三个字依次表示第一、二、三个硬币.

这些复合事件可以用它所包含的简单事件(Ω 中的点)来表示，因此复合事件就是 Ω 的子集合，例如:

$A=(\omega_1, \omega_8)$； $B=(\omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7)$；

$C=(\omega_1, \omega_2, \omega_3, \omega_5)$； $D=(\omega_2, \omega_3, \omega_5)$；

$E=(\omega_4, \omega_6, \omega_7)$； $F=(\omega_1, \omega_2)$.

对某一事件，可以定义它的逆事件. 类似地，还可以定义某两事件至少一个发生所构成的事件以及某两事件同时发生所构成的事件. 我们用集合论的符号来表示它们: 事件 A 的逆事件(余集)记为 A^c；“事件 A 和 B 至少一个发生”(A 和 B 的并集)记为 $A\cup B$；“事件 A 和 B 同时发生”(A 和 B 的交集)记为 $A\cap B$. 还可以定义两个事件的差: $A-B=A\cap B^c$①. 基本事件空间 Ω 也是一个事件，称为必然事件. 我们以 $\varnothing$ 表示不可能事件，显然 $\varnothing=\Omega^c$. 从以上例子可以看出(沿用 例 1.1.1 的 记 号): $B^c=A$, $B\cap C=D$, $D\cup E=B$, $B-D=E$, $D\cap E=\varnothing$. 注意到 D 中每一点(构成 D 的简单事件)都包含在 B 中，所以只要事件 D 发生，事件 B 必发生，在这种情况下我们就说事件 D 含于 B，并记为 $D\subset B$(或 $B\supset D$).

两个事件的并和交可以推广到任意多个事件: $\bigcup\limits_{j=1}^{n} A_j$ 表示 n 个事件 $A_1, A_2, \cdots, A_n$ 中至少一个发生，$\bigcap\limits_{j=1}^{n} A_j$ 表示 n 个事件 $A_1, A_2, \cdots, A_n$ 同时发生. 既然复合事件对应于基本事件空间的子集合，所以事件之间的关系全都可以用集合论的术语来描述，下面的表 1.1 列出了这种对应关系.

并、交、余等集合论运算服从某些类似于代数运算的规则，因

① 集合论更详细的知识可参看[1]、[2]或[3].

表1.1 事件运算与集合运算的对应

事　　件	集　　合	记　　号
简单事件	样本空间的点	$\omega, \omega_1, \omega_2, \cdots$
复合事件	样本空间的子集合	$A, B, \cdots$
简单事件ω是复合事件A的组成成分	点ω是集合A的元素	$\omega \in A$
必然事件	样本空间	Ω
不可能事件	空　集	$\varnothing$
A的逆事件	A的余集	A^c
事件“A或B”(至少一个发生)	A和B的并集	$A \cup B$
事件“A与B”(两者同时发生)	A和B的交集	$A \cap B$
事件A和B不相容(或互斥)	A和B不相交	$A \cap B = \varnothing$
事件A含于B	A是B的子集合	$A \subset B$或$B \supset A$
事件A和B等价	A和B相等	$A = B$
事件A发生但B不发生	A和B的差集	$A - B = A \cap B^c$

此也称为事件(集合)代数，现列举其中一些规则如下：

$$A \cup B = B \cup A; \quad A \cap B = B \cap A \tag{1.1.1}$$

$$(A \cup B) \cup C = A \cup (B \cup C); \quad (A \cap B) \cap C = A \cap (B \cap C) \tag{1.1.2}$$

$$\begin{aligned} (A \cup B) \cap C &= (A \cap C) \cup (B \cap C) \\ (A \cap B) \cup C &= (A \cup C) \cap (B \cup C) \end{aligned} \tag{1.1.3}$$

公式(1.1.1)、(1.1.2)和(1.1.3)分别称为交换律、结合律和分配律.

$\varnothing$和Ω有以下关系：

$$A \cup \varnothing = A, A \cap \varnothing = \varnothing, A \cup \Omega = \Omega, A \cap \Omega = A \tag{1.1.4}$$

还应当注意：

$$A \cup A = A \cap A = A \tag{1.1.5}$$

关于余集，我们有下列规则：

$$(A^c)^c=A,\ A\cup A^c=\Omega,\ A\cap A^c=\varnothing,\ A^c=\Omega-A \quad (1.1.6)$$

“含于”关系是可传的，即由 $A\subset B$ 和 $B\subset C$ 可以推出 $A\subset C$. 并、交和余三种基本运算之间还有下列关系：

$$(A\cup B)^c=A^c\cap B^c,\ (A\cap B)^c=A^c\cup B^c \quad (1.1.7)$$

类似地，对任意有限个事件有：

$$\left(\bigcup_{j=1}^{n}A_j\right)^c=\bigcap_{j=1}^{n}A_j^c,\ \left(\bigcap_{j=1}^{n}A_j\right)^c=\bigcup_{j=1}^{n}A_j^c$$

对无穷序列 $\{A_j\}$，仍有：

$$\left(\bigcup_{j=1}^{\infty}A_j\right)^c=\bigcap_{j=1}^{\infty}A_j^c,\ \left(\bigcap_{j=1}^{\infty}A_j\right)^c=\bigcup_{j=1}^{\infty}A_j^c \quad (1.1.8)$$

这些公式的证明留给读者.

一族互斥事件 $\{A_j\}$，如果它们的并是整个样本空间 Ω，则称为完备事件类，这时样本空间的每一点必属于 A_j 之一，且仅属于其中之一.

§1.2 概　　率

在前言中讲过，随机现象的模型必须反映这些现象所呈现的统计规律性. 构造这种模型，必须对每个事件指定一个数字（叫做这个事件的概率），它反映了观察到的相对频率的稳定性. 因此，为了对概率作合理的假定，我们要首先讨论相对频率的性质. 我们考察一个试验及某个（简单的或复合的）事件 A，假定在 m 次重复试验中 A 发生 h 次，显然：

$$0\leqslant h\leqslant m$$

令

$$R(A)=\frac{h}{m}$$

为m次试验中A的相对频率，则

$$0\leqslant R(A)\leqslant 1 \tag{1.2.1}$$

再考察两个互斥事件A和B(即$A\cap B=\varnothing$)，假定在m次试验中A发生h_A次，B发生h_B次. 令$C=A\cup B$代表"A或B至少一个发生"这一事件，容易看出，如果$A\cap B=\varnothing$，就有：

$$R(C)=R(A\cup B)=R(A)+R(B) \tag{1.2.2}$$

显然这个式子可以推广到任意有限个互斥事件.

假定某个试验，具有n个互不相容的结果，在m次重复试验中第j个结果观察到h_j次，于是$\sum_{j=1}^{n}h_j=m$. 用R_j表示第j个结果的相对频率，则$\sum_{j=1}^{n}R_j=1$.

因为试验有n个基本事件，我们可以构造一个包含n个点的样本空间$\Omega=(\omega_1, \omega_2, \cdots, \omega_n)$，对每个点$\omega_j$指定一个数字（它的概率)：$P(\omega_j)=p_j$，它们满足以下条件：

$$0\leqslant p_j\leqslant 1 \tag{1.2.3}$$

$$\sum_{j=1}^{n}p_j=1 \tag{1.2.4}$$

这些条件反映了相对频率的性质(1.2.1)和(1.2.2). 对任一复合事件A，我们规定它的概率为：

$$P(A)=\sum_{\omega_j\in A}p_j \tag{1.2.5}$$

即复合事件的概率等于它所包含的简单事件的概率之和.