高等学校教材

概率论基础

(第二版)

复旦大学　李贤平

高等教育出版社

高等学校教材

概率论基础

第二版

复旦大学　李贤平

高等教育出版社

（京）112号

内 容 提 要

本书内容分为事件与概率、条件概率与统计独立性、随机变量与分布函数、数字特征与特征函数、极限定理等5章，全书及各章都有内容小结，可供高等学校理科数学专业、数理统计与概率论专业作为教材使用.

图书在版编目（CIP）数据

概率论基础/李贤平编著. —2版. —北京：高等教育出版社，1997

ISBN 7-04-005921-5

Ⅰ. 概… Ⅱ. 李… Ⅲ. 概率论 Ⅳ. 0211

中国版本图书馆CIP数据核字（96）第21708号

高等教育出版社出版

北京沙滩后街55号

邮政编码:100009 传真:64014048 电话:64054588

新华书店总店北京发行所发行

北京联华印刷厂印装

*

开本 850×1168 1/32 印张 10.875 字数 280000

1997年5月第1版

1997年4月第2版 1997年4月第1次印刷

印数 0001—1712

定价 10.60元

再版前言

本书第一版受到出乎意料的欢迎：各兄弟院校广泛采用；发行20余万册；获第一届国家级优秀教材奖．这应归功于时代的需要，也说明作者试图写出一本既有前苏联教本的系统严谨，又有美国书籍的生动活泼，体现我国教学经验，理论与应用兼备的教材的目标得到一定程度的实现．

广大师生的热情反映和意见，写作时就存在的不少遗憾，陆续发现的一些不足甚至错误，18年来在教学和科研中的些许心得，这些成了第二版修改的依据．

本书的结构受到一致的肯定，这次未加更改，只是重写了过于薄弱的数字特征部分，因而增加一节，并恢复了初版时因篇幅考虑而删去的全书小结．

增写了不少典型应用事例，补写上几个影响学科初期发展的名例，改写了一些平淡无味的实例，这些使本书的理论与应用更为均衡，更为尊重历史事实，也更准确地反映概率论这一学科目前在整个自然科学和社会科学中的重要地位．相信也增添了本书的趣味性．

一种分布就是一个数学模型．初版重视分布及它们之间联系的传统得到发扬．正式引入负二项分布与埃尔兰分布，使两个等待分布序列的对比更为明显．改动了一些例题与习题．这样一来，本书把分布按重要性分成三类：第一类三大分布有专门的节加以介绍；第二类包括十几种重要分布，在正文中指明其背景、性质以及它与其他分布的关系；第三类则多数在习题中出现．最后在附录一中汇总．关于多元分布，新版也较前重视．

几个定理的证明被局部更动，多数为改善，少数为改正．增

添了唯一的一个定理（移植自浅野、江島两位先生与作者合著的一本日文书），用极其简练的办法证明了重要的多元中心极限定理.

习题是本书的重要组成部分. 这次略作调整，基本题约占四分之三，与正文紧密配合，标有星号的题目对正文作了补充，双星号的是难题. 为适应不同需要，配题较多，教学中选用约半数即可.

以上是关系全局的一些较大变动，其他增删所在多是. 定理、公式、图表的编号作了统一编排. 关键词附上英文，重要数学家的名字也都标出原文，译名有些更动. 希望通过这些修改能使原书的质量有所提高. 预料书中仍会存在不少缺点与错误，欢迎广大师生批评、指正.

修改后的第二版篇幅略有增加，如为学时所限，建议舍去相对独立的熵与信息一节和相当专深的最后两节.

在我的概率论生涯中，得到郑绍濂、吴立德、陶宗英、汪嘉冈、何声武、卞国瑞、徐家鹄等师友的许多帮助，高尚华为本书的前后两版付出大量心血，缪铨生教授主持了第二版的审稿，在此一并致谢.

李贤平

1996 年中秋

目　　录

第一章　事件与概率

§1. 随机现象与统计规律性

一、随机现象

概率论（probability theory）是研究随机现象的数量规律的数学分支．为了说明什么是随机现象，让我们先来看一个例子．设某车间有 200 台车床，由于经常需要检修、测量、调换刀具、变换位置等种种原因，因此，即使在生产期间，各台车床还是时常需要停车．若每台车床有百分之六十的时间在开动，而每台车床开动时需要耗电 1 千瓦，问应供给这个车间多少电力才能保证此车间正常生产？

类似的例子在许多实际问题中出现，解决这类问题当然具有重要意义．

显然，若供给这个车间 200 千瓦的电能则此车间便能正常生产．但这样做不合算，因为每台车床的开工率只有 60%，也就是说，平均起来这个车间中同时在工作的车床只有 120 台，供给它们 200 千瓦的电能太多．供给 120 千瓦的电能行吗？这又太少些，因为有时工作的车床数会超过 120 台，若只供给 120 千瓦，则这时会出现因电力不足而使车床无法正常运转，那么到底供给多少才能既保证生产正常进行而又节约电力呢？

用概率论的方法能给这个问题以完满的解决，现在我们仅把结果列出，详细的解法以后再讲．计算表明，只要供给这个车间 141 千瓦的电就够了，虽然在这时也可能碰到因电力不足而不能正常生产的情况．但这种机会很少，它小于 0.1%，即在 8 小时工

作中一般只有半分钟会碰到这种情况，这显然影响不大. 但节约出来的 59 千瓦电能却能用于许多别的用途.

解决这个问题的关键在于要计算出某时刻同时工作着的车床数，但是由于某台车床在某时刻是否开工很难预先确定，这受到许多偶然因素的影响，即工作着的车床数是一个受许多偶然因素影响的量.

原来，在自然界和人类社会中存在着两类不同的现象.

当我们多次观察自然现象和社会现象后，会发现许多事情在一定的条件下必然会发生. 例如在没有外力作用的条件下，作等速直线运动的物体必然继续作等速直线运动；又如在海边生活时，水加热到 100℃时必然会沸腾等等. 这种在一定条件下，必然会发生的事情称为**必然事件**. 反之，那种在一定条件下，必然不会发生的事情就称为**不可能事件**. 例如在不受外力作用的条件下，作等速直线运动的物体改变其等速直线运动状态是不可能的.

从所举例子中看出，必然事件和不可能事件，虽然形式相反，但是两者的实质是相同的. 必然事件的反面就是不可能事件，而不可能事件的反面就是必然事件.

所有这种现象我们称之为**决定性现象**，它广泛地存在于自然现象和社会现象中.

但是在自然现象和社会现象中也还广泛存在着与决定性现象有着本质区别的另一类现象，上述车间供电问题就是一例.

类似的例子还可以举出很多，例如用同一仪器多次测量同一物体的重量，所得结果彼此总是略有差异，这是由于诸如测量仪器受大气影响，观察者生理上或心理上的变化等等偶然因素引起的. 同样地，同一门炮向同一目标发射多发同种炮弹，弹落点也不一样，因为炮弹制造时种种偶然因素对炮弹质量有影响，此外，炮筒位置的误差，天气条件的微小变化等等都影响弹落点. 再如从某生产线上用同一种工艺生产出来的灯泡的寿命也有差异等等. 总之，所举这些现象的一个共同的特点是：在基本条件不变

的情况下，一系列试验或观察会得到不同的结果．换句话说，就个别的试验或观察而言，它会时而出现这种结果，时而出现那种结果，呈现出一种偶然性．这种现象称为**随机现象**．对于随机现象通常关心的是在试验或观察中某个结果是否出现，这些结果称为**随机事件**，简称**事件**（event）．例如过马路交叉口时可能遇上各种颜色的交通指挥灯，这是一个随机现象，而“遇到红灯”则是一个随机事件．以后我们用 A，B，C 等大写拉丁字母表示随机事件．

二、频率稳定性

正如恩格斯所指出的，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．

人们经过长期的实践发现，虽然个别随机事件在某次试验或观察中可以出现也可以不出现，但在大量试验中它却呈现出明显的规律性——频率稳定性．

对于随机事件 A，若在 N 次试验中出现了 n 次，则称

$$F_N(A)=\frac{n}{N}$$

为随机事件 A 在 N 次试验中出现的**频率**．

下面是关于频率稳定性的几个有名例子．援引这类例子是因为它们不但具有一定的权威性，而且都是可以反复验证的．

在掷一枚硬币时，既可能出现正面，也可能出现反面，预先作出确定的判断是不可能的，但是假如硬币均匀，直观上出现正面与出现反面的机会应该相等，即在大量试验中出现正面的频率，应接近于50%，为了验证这点，历史上曾有不少人做过这个试验，其结果如下页所示[①]．

又如，在英语中某些字母出现的频率远远高于另外一些字母．

① 引自格涅坚科．概率论教程．高等教育出版社，第44页．

实　验　者	掷硬币次数	出现正面次数	频　　率
蒲　　丰	4040	2048	0.5069
皮　尔　逊	12000	6019	0.5016
皮　尔　逊	24000	12012	0.5005

在进行了更深入的研究之后，人们还发现各个字母被使用的频率相当稳定. 例如，下面就是英文字母使用频率的一份统计表[①]. 其他各种文字也都有着类似的规律.

字母	空格	E	T	O	A	N	I	R	S
频率	0.2	0.105	0.072	0.0654	0.063	0.059	0.055	0.054	0.052
字母	H	D	L	C	F	U	M	P	Y
频率	0.047	0.035	0.029	0.023	0.0225	0.0225	0.021	0.0175	0.012
字母	W	G	B	V	K	X	J	Q	Z
频率	0.012	0.011	0.0105	0.008	0.003	0.002	0.001	0.001	0.001

近年来对汉语的统计研究有了很大的发展. 关于汉字的使用频率已有初步统计资料，对汉语常用词也作了一些统计研究. 特别是结合汉字输入方案等的研制，正在对汉字的结构作深入的统计分析. 这些研究对实现汉字信息处理自动化无疑具有重要的意义.

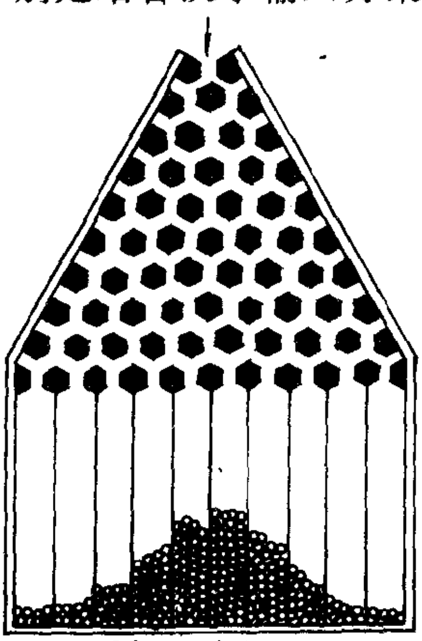
图 1　高尔顿板

另一个验证频率稳定性的著名试验是由英国生物统计学家高尔顿(Galton)设计的. 它的试验模型如图 1 所示.

自上端放入一小球，任其自由下落，在下落过程中当小球碰到钉子时，从左边落下与从右边落下的机会相等. 碰到下一排钉子时又是如此. 最后落入底板中的某一格子. 因此，任意放入一球，则此球落入

① 引自 L. Brillouin. Science and Information Theory. New York，1956.

哪一个格子，预先难以确定．但是实验证明，如放入大量小球，则其最后所呈现的曲线，几乎总是一样的．也就是说，小球落入各个格子的频率十分稳定．这个试验模型称为**高尔顿板**．试验中呈现出来的规律性，在学习第五章极限定理之后，就会有更深刻的理解．

同样，如果多次测量同一物体，其结果虽略有差异，但当测量次数增加时，就会越来越清楚地呈现出一些规律性：测量值的平均值在某固定常数附近波动，诸测量值在此常数两旁的分布呈现某种对称性．又如在射击的例子中，当射击次数不多时，炮弹的弹落点似乎是前后左右杂乱无章，看不出什么明显的规律；但当射击次数增加时，弹落点的分布就会呈现出一定的规律性：如弹落点关于目标的分布略呈对称性，偏离目标远的弹落点比偏离目标近的弹落点少等等．其他如灯泡寿命等，在进行多次观察或试验后，也都可以发现类似的规律性．

日常生活中也不乏有趣的例子，例如衣服和用具总在同样部位以相似的方式破损，下雨时地面各处总是差不多同时淋湿等等．读者只要多注意观察，就不难发现许多关于频率稳定性的有说服力的实例．

上述种种事实表明，随机现象有其偶然性的一面，也有其必然性的一面．这种必然性表现为大量试验中随机事件出现的频率的稳定性，即一个随机事件出现的频率常在某个固定的常数附近摆动，这种规律性我们称之为**统计规律性**．频率的稳定性说明随机事件发生的可能性大小是随机事件本身固有的、不随人们意志而改变的一种客观属性，因此可以对它进行度量．

对于一个随机事件 A，用一个数 $P(A)$ 来表示该事件发生的可能性大小，这个数 $P(A)$ 就称为随机事件 A 的**概率**(probability)．因此概率度量了随机事件发生的可能性的大小．

对于随机现象，只讨论它可能出现什么结果，价值不大，而指出各种结果出现的可能性的大小则具有很大意义．有了概率的

概念就使我们能对随机现象进行定量研究，由此建立了一个新的数学分支——概率论.

三、频率与概率

既然概率 $P(A)$度量了随机事件 A 发生的可能性大小，可以预料，在 N 次重复试验中，若 $P(A)$较大,则频率 $F_N(A)=\frac{n}{N}$也较大. 反之若 $P(A)$很小，则 $F_N(A)$ 也很小，而且概率 $P(A)$应与频率有许多相似的性质. 以下我们先对频率的性质进行一番考察.

首先，频率具有非负性

$$F_N(A)\geqslant 0 \tag{1.1.1}$$

其次，对于必然发生的事件，在 N 次试验中应出现 N 次. 若以 Ω 记必然事件，则应有

$$F_N(\Omega)=1 \tag{1.1.2}$$

还有，若 A 及 B 是两个不会同时发生的随机事件，以 $A+B$ 表示 A 或 B 至少出现其一这个事件，则应有

$$F_N(A+B)=F_N(A)+F_N(B) \tag{1.1.3}$$

这个性质称为**频率的可加性**.

当然还可以列出频率的许多性质，但上述三个性质是最基本的. 例如，“不可能事件在 N 次试验中出现的频率为 0”，“任何随机事件在 N 次试验中出现的频率不大于 1”，“对于有限个两两不会同时发生的随机事件也有频率可加性”，这些性质都可以由 (1.1.1)，(1.1.2) 及 (1.1.3) 推出.

最后，根据上述频率稳定性的讨论似乎可以提出这样的猜想，即当 N 足够大时 $F_N(A)$与 $P(A)$应充分接近. 这一点有很大的启发性，在历史上它一直是概率论研究的一个重大课题. 以后我们将会看到，在很一般的条件下，这个结论的确成立，但同时还须对问题的提法进一步明确化.

频率与概率的上述关系有时还提供了求某事件概率的一种手段，即当 N 足够大时，用它的频率来作为概率的近似值. 以后我们将会看到，这种做法大有好处.

四、概率论简史

概率论是一门研究随机现象数量规律的学科，一般把1654年作为概率论诞生的一年. 这年中，法国数学家巴斯卡与费马就机会博弈中的一些问题作了通信讨论；后来惠更斯也加入研究. 在这些研究中建立了概率论的一些基本概念，如事件、概率、数学期望等.

其后，随着生产实践的发展，特别是在射击、保险、测量等工作中提出的一些概率问题，促使人们在概率论的极限定理等方面进行深入研究，起初主要对伯努利试验概型进行研究，而后则推广到更为一般的场合. 这个时期先后对概率论作出重要贡献的数学家有伯努利、拉普拉斯、泊松和高斯.

极限定理的研究在18世纪和19世纪整整200年中成了概率论研究的中心课题. 在本世纪初，由于新的更有力的数学方法的引入，这些古典问题得到了较好解决，这当中俄罗斯彼得堡学派起了主导作用.

虽然概率论历史悠久，但是它的严格的数学基础的建立以及理论研究和实际应用的极大发展却主要是本世纪的事情.

由于物理学（如统计物理）、生物学以及工程技术（如自动电话、无线电技术）发展的推动，概率论得到了飞速的发展. 理论课题不断扩大与深入，概率论的思想渗入各个学科成了近代科学发展的明显的特征之一.

由于各个数学分支的发展与互相渗透，概率论的严格的数学基础被建立起来，古典问题得到了解决，新的概念和工具不断出现，概率论成了数学的一个活跃分支.

概率论大大地扩大了它的应用范围. 特别在最近几十年中，概

率论的方法被引入各个工程技术学科和社会学科. 目前，概率论在近代物理，无线电与自动控制，工厂产品的质量管理，医药和农业试验，金融保险业等等方面都找到了重要应用，这些实际需要也有力地推动了概率论的新发展，有些还形成了边缘学科（如信息论、排队论、可靠性理论等）.

在这个时期内，由于生物学和农业试验的推动，数理统计学（mathematical statistics）也获得了很大发展，它以概率论为理论基础，又为概率论应用提供了有力的工具，两者互相推动，迅速发展. 而概率论本身的研究则转入以随机过程为中心课题，取得了许多理论上和应用上都有重要价值的结果.

§2. 样本空间与事件

一、样本空间

从本节开始，我们将逐步引进概率论的基本概念. 样本空间与事件是最基本的两个概念.

对随机现象的研究必然要联系到对客观的事物进行“调查”、“观察”或“实验”，以后我们统称之为**试验**（trial），并假定这种“试验”可以在相同条件下重复进行.

我们感兴趣的是试验的结果. 例如掷一次硬币，我们关心的是出现正面或出现反面，这是两个可能出现的结果. 假如我们考察的是掷两次硬币的试验，则可能出现的结果有（正，正），（正，反），（反，正），（反，反）四种；如果掷三次硬币，则结果还要复杂，但还是可以把它们描述出来. 总之，为了研究随机试验，首先需要知道这个试验可能出现的结果. 这些结果称为**样本点**，一般用 ω 表示. 样本点全体构成**样本空间**（sample space），用 Ω 表示. 在具体问题中，给定样本空间是描述随机现象的第一步.

下面举一些例子.

［例 1］ 在研究英文字母使用情况时，把样本空间选为 $\Omega=$

{空格，A，B，…，Z}是适宜的，这个样本空间只有有限个样本点，是比较简单的样本空间.

[例 2] 观察一小时中落在地球上某一区域的粒子数，可能的结果一定是非负整数，而且很难指定一个数作为它的上界，这样，可以把样本空间取为 $\Omega=\{0, 1, 2, \cdots\}$. 这个样本空间含有无穷多个样本点，但这些样本点可以依照某种次序排列出来，以后我们将称它的点数为可列个.

[例 3] 讨论某地区的气温时，我们自然把样本空间取为 $\Omega=(-\infty, \infty)$，或 $\Omega=[a, b]$. 这个样本空间包含有无穷多个样本点，它们充满一个区间，不是一个可列集.

[例 4] 考察地震震源时，可以把样本点取为 (x, y, z)，其中 x 表示震源的经度，y 表示纬度，z 表示深度. 这时，样本空间是三维空间中某一区域.

从以上例子可以看出，随着问题不同，样本空间可以相当简单，也可以相当复杂.

在今后讨论中，经常把样本空间认为是预先给定的. 当然对于一个实际问题或一个随机现象，如何用一个恰当的样本空间来描述它也很值得研究. 但是在概率论的研究中，一般都假定样本空间是给定的. 这是必要的抽象，这种抽象使我们能更好地把握住随机现象的本质，而且得到的结果能广泛地应用. 事实上，一个样本空间可以概括各种实际内容很不相同的问题：例如只包含两个样本点的样本空间既能作为掷硬币出现正、反面的模型，也能用于产品检验中出现“好品”及“废品”，又能用于气象中“下雨”与“不下雨”，以及公用事业排队现象中“有人排队”与“无人排队”等等. 尽管问题的实际内容如此不同，但有时却能归结为相同的概率模型. 我们后面常以摸球等作为例子也是由于这个原因，它能使问题的本质更为突出.

二、事件

有了样本空间的概念，就可以定义事件．我们还是从考察一个例子开始．

［例 5］　口袋中装有 4 只白球和 2 只黑球，我们考虑依次从中摸出两球所可能出现的事件．若对球进行编号，4 只白球分别编为 1，2，3，4 号，2 只黑球编为 5，6 号．如果用数对 (i, j) 表示第一次摸得 i 号球，第二次摸得 j 号球，则可能出现的结果是

$$\begin{array}{l}(1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\\(2, 1), (2, 3), (2, 4), (2, 5), (2, 6)\\(3, 1), (3, 2), (3, 4), (3, 5), (3, 6)\\(4, 1), (4, 2), (4, 3), (4, 5), (4, 6)\\(5, 1), (5, 2), (5, 3), (5, 4), (5, 6)\\(6, 1), (6, 2), (6, 3), (6, 4), (6, 5)\end{array} \qquad (*)$$

把这 30 个结果作为样本点，则构成了样本空间．在这个问题中，这些样本点是我们感兴趣的事件；但是我们也可以研究下面另外一些事件：

A：第一次摸出黑球；

B：第二次摸出黑球；

C：第一次及第二次都摸出黑球．

后面这些事件与前面那些事件的不同处在于这些事件是可以分解的，例如为了 A 出现必须而且只须下列样本点之一出现：

$$\begin{array}{l}(5, 1), (5, 2), (5, 3), (5, 4), (5, 6)\\(6, 1), (6, 2), (6, 3), (6, 4), (6, 5)\end{array}$$

前面的 30 个事件由单个样本点构成；后面这三个事件，每一个事件都是由若干个样本点构成的，总之，它们都是样本点的某个集合．

所谓给定一个点的集合 S，是指对于任何一个点 ω，都可以确定它是不是属于 S．如果是，则记为 $\omega \in S$；如果不是，