高等学校教材

基础概率与数理统计

林少宫 编

高等教育出版社

高等学校教材

基础概率与数理統计

林少宫编

高等教育出版社

本书系以作者讲授概率论与数理统计课程的讲义为基础，经过修改补充而成。

内容包括概率的基本运算，重要的分布函数，常态理论，产品质量的过程控制，估计理论，假设的测（检）验，方差分析，相关分析以及保证率曲线的配制等。全书结构多从专业应用方面考虑，不讲求数学系统。书中强调了随机样本与统计总体的概念，突出了常态分布的价值，对各种方法的应用与评价也给予了适当的注意。此外，各章均附有适量的问题。

本书可作为高等工科院校这门课程的教学参考书，也可供具有高等数学知识的工程技术人员自修阅读。

本书原由人民教育出版社出版。现经上级决定，自 1965 年 1 月 1 日起，另行成立“高等教育出版社”；本书今后改用高等教育出版社名义继续印行。

基础概率与数理統计

林 少 宫 编

北京市书刊出版业营业许可証出字第 119 号

高等教育出版社出版（北京景山东街）

人 民 教 育 印 刷 厂 印 装

新 华 书 店 北 京 发 行 所 发 行

各 地 新 华 书 店 经 售

統一书号 K13010·1100 开本 850×1168 1/32 印张 9 3/16

字数 250,000 印数 16,001—22,000 定价 (6) ¥0.90

1963 年 7 月第 1 版 1965 年 4 月北京第 5 次印刷

前　　言

本书原先为水电站动力装置专业而写，后来兼顾及机器制造专业的需要而加以扩充，其后再考虑到更一般的需要，才写成現在的形式。

本书的大部分在試用的基础上曾經多次修訂，但也有若干部分未經过課堂上試用，后者如第六章和第八章的后半即是。1960 年 8 月华中工学院数学教研室的部分同志曾对原稿作了全面的討論，随之作了一次較大的修改。1962 年夏承西安交通大学数学教研室和人民教育出版社对这部稿件提出了宝貴的意見，編者对他們表示衷心的感謝，并参考他們的意見作了又一次的删改与补充，不少地方重新写过。

本书內容足够讲授 60—70 学时，但按专业的不同，可删去个别章节。例如，除水电专业外，一般可不讲第九章。若时数不够，第六章統計推断部分可只讲关于平均数的推断（这样，第五章常态曲綫的配制一节也可删去），第七章不讲，第八章相关分析与計算只讲单相关（包括最小二乘法），复相关則限于交待概念。第三章二維分布一节可合并到相关分析中簡单地提一下；泊松分布也不另立一节，而作为二項分布的极限来提。自然，附有 * 号的各节，也都可以省略。根据教学經驗，第三章不宜讲得过簡、过快。

使用本书时，最好連同各章所附問題一并考虑。一些問題起着理論补充和减省課堂上的过細推导证明的作用，例如 2.6 节問題 8, 14, 27；3.10 节問題 13, 16, 20, 23, 28；5.8 节問題 1, 10；6.10 节問題 5, 18, 19；7.4 节問題 6；8.14 节問題 8, 14 等是。若感到必要，也可选擇一些問題在正文中讲授。

最后，由于本书的撰写和修訂工作受到我們的思想水平与科学水平的限制，錯誤或不当之处，在所难免。我們誠恳地希望讀者对本书提出宝貴的意見，并将意見寄交武昌华中工学院数学教研室为感。

編　　者　1962 年 10 月

目　录

第一章　概率与統計

1.1　本学科的地位与作用

概率論与数理統計应用到人口調查、社会保險以及生物遺傳等方面已經有許多年代的历史了。近代物理化学的发展为概率論与数理統計在自然科学方面的应用开辟了广闊的园地。例如分子物理学中所研究的分子运动及其相互之間的作用，由于这些分子为数甚伙，且每个分子运动所遵循的規律又异常复杂，故不能用古典方法（例如微分方程）去处理，只有用概率論的方法才得以适当解决；又例如个别的放射性原子的寿命是很难确定的，但可用概率論的方法确定放射性原子的平均寿命。随着工业生产的标准化与自动化，概率論与数理統計又被广泛地应用到产品的质量控制上面；并且經驗证明，利用統計方法是能够有效地提高产品质量和降低成本的。近年来，計算技术的迅速发展，对于概率論与数理統計的应用提供了极有利的条件。概率論与数理統計的內容不但由于它的应用范圍的日漸扩大而更加丰富，而且，一些新的学科如运筹学、自动控制理論等，由于使用了概率論和数理統計的方法而日臻完善。

目前，概率論与数理統計已被广泛地应用到各个科学分支和各个生产部門了。

概率論与数理統計之所以能够广泛地被应用到各个方面，是因为客观世界里普遍地存在这样的一种現象，对于这种現象，我們无法利用“因果关系”加以严格控制或准确預測；这种現象是屬于偶然性质的，我們不能用一些簡单的物理定律加以概括，而須从大量观測中綜合分析，归納出一些“大量現象”的規律来。

以水文学为例，一些水文特征如洪水位、徑流量等就是如此。我們很难断定某河流明年的最高水位是什么？某河流在什么时候将会枯竭？但洪水位和徑流量并不是沒有規律的，只不过这些規律不是一种精确的物理定律，而是一种大量現象的規律。我們只能預期，比方說，在一定的自然条件下(或一定的自然环境中)，某个水位是百年一遇的，但我們不能确定哪一年遇到。

在机械制造方面也有类似的問題。我們可以估計，在某个稳定的加工过程里，零件的某个尺寸超过公差范圍的机会，比方說，不多于千分之三，但不能肯定一千件里的哪三件超过公差范圍。机械制造上的誤差，和測量上或射击上的誤差一样，都屬于偶然性质的現象。对于偶然性质的現象，須用概率論和数理統計的方法来处理。

1.2 基本概念

概率論与数理統計是一門研究偶然現象的規律性的学科。但是，現象的偶然性每每伴随着它的必然性一同出現。关于这一点，恩格斯曾經指出过，偶然性是必然性的表現形式，是必然性的补充。因此，在运用概率論与数理統計到某一个具体問題上的时候，必須注意划分有关現象的偶然性和必然性。① 例如产品的质量首先与原料和制造过程有关，若改变原料或制造方法，則因而引起的质量变异就不能算作偶然現象。然而，即使設法控制原料和制造过程不变，产品质量仍不免有所波动，这种波动就可以归結为許多不可控制的微小因素(如原料所含杂质，操作工人的疲劳，机器运轉的偶然不协調等等)所造成的偶然現象。因此，当我們使用概率論和数理統計的方法去檢查产品质量时，我們首先假定产品的原料和制造过程是不变的，换句話說，假定一組“稳定”的生产条件。

① 事实上，数理統計的一个重要作用就是測驗某种現象是否出于偶然。

概率論与数理統計的問題永远是相对于一組条件的实現而言，只有当这組条件实現了，我們才能获得有关問題的偶然性材料。对于不同的問題，这組条件有不同的形式。

例如，当我們观測河川某处的徑流量时，无疑，我們应考虑到上下游是否有建造水庫、水壩，开凿运河以及河川改道等事情发生。因为任何徑流量的偶然性規律都只能描述一定自然条件下（或自然环境中）的徑流現象。

又例如，当我們观測某放射性元素在任其自然的状态下的蜕变情况时，則必須注意到是否有外来的强烈影响（例如受高速质点的轰击），亦即注意“任其自然”的条件是否实現。

既然对偶然現象的观測总是在一組条件的实現下进行的，故每次观測可看作一个試（实）驗（事实上一个試驗就是一組条件的实現），而观測的結果就是試驗的結果。

随机試驗 对于某些事物来說，在同一組条件实現之下就必然得到同一的試驗結果。例如，水在标准大气压下加热到摄氏 100° 以上必然化为蒸汽；将一枚銅錢向上抛，它必然受地心引力的作用而下落；等等。但如果要問，旋一枚銅錢究竟出現正面抑或反面，我們就不可能这样肯定了；我們只能說，或者出現正面，或者出現反面。在机械制造方面，当我們要制造一批某种尺寸的产品时，不管机器的精密程度多高，我們总是无法預先言定某件产品的尺寸究竟偏大，抑或偏小。

如果在同一組条件（包括操作程序）的实現下，不一定得到同一的試驗結果，然而每一个可能的試驗結果都有一定的出現机会，或者說有一定的可能程度，則我們說，这个試驗是一个随机試驗。在本书中所談到的試驗都是指随机試驗而言。

事件 随机試驗的結果叫做事件。

事件可以是数量性质的，即試驗結果可直接由測量或計数而得的，例如雨量的观測，投擲骰子的点数，等等；但也可以是屬性性质的，例如

天气的風雨云晴,出生嬰孩的性别,某种东西的颜色,等等。

事件可以是单一性质的,也可以是多重性质的。前者例如我們只考慮出生嬰孩的性别而說某某嬰孩是男性;后者例如我們同时考慮出生嬰孩的性别和体重而說某某嬰孩是女性、体重6磅,等等。

事件可以是簡单的,也可以是复合的。前者指不能再行分拆的事件;后者則指由簡单事件复合而成的事件。例如当我們考虑一批圓柱体的半徑时,"某件圓柱体的半徑是2厘米"为一簡单事件,而"某件圓柱体的半徑在2到2.1厘米之間"則为一复合事件,因后者由圓柱体半徑为2厘米,2.01厘米,2.02厘米,……等一系列事件复合而成。又当我們观察某服务站的排队人数时,"有4人在排队"为一簡单事件,而"有不多于4人在排队"則为一复合事件。

前面說,随机試驗的結果叫做事件,这只是一种簡单的籠統的說法。因为同一試驗結果可以有不同的說法。例如擲骰子出現5点,旣可以說成出現奇数点,也可以說成不少于3点;"3人排队"也可以說成"排队人数不多于5人";等等。确切地說,事件是随机試驗的所有可能結果里面的一个集合。例如"有不多于4人在排队"是0, 1, 2, 3或4个人在排队等5个可能結果的集合。这里每一可能結果相当于一个簡单事件。注意,只要集合中的任意一个結果出現,代表此集合的事件即发生。例如"3人在排队"含意着"有不多于4人在排队"这一事件的发生。

事件是概率論与数理統計中的最基本的概念。这个概念是与一个随机試驗以及它的所有可能的結果联系在一起的。只有明确了随机試驗及其全部可能結果,我們才能确定事件的可能程度(以后我們說确定事件的概率)。

熟悉点集的数学理論的讀者,可以把事件比作点的集合,集合中的点則相当于一个不能再行分拆的試驗結果或簡单事件。

随机变量 表示随机試驗的結果的一个数量叫做随机变量。随机

变量取什么値是不能在試驗前得知的;它决定于試驗的結果(故事实上是試驗結果的函数)。例如，在一定的气候和測量技术的条件下，重复測量地球和某星球的距离所得到的測量値是一个随机变量；在某一小时內,电話总机所接到的电話呼喚次数是一个随机变量;等等。如果試驗結果不是直接由測量或計数而得,則随机变量取什么値,可按照适当的惯例。例如在产品的质量檢查中,記廢品为1,成品为0。这样,随机变量取什么値就可視出現廢品或出現成品而定了。

不言而喻,随机变量取什么値也都有一定的机会或可能程度,而每次試驗事实上就是对随机变量的一次观測。

1.3 統計总体与随机样本

在統計学中，把我們准备加以观測的一个滿足指定条件的元素或个体的集合叫做統計总体。例如在稳定的生产条件下的产品，在正常情况下发射的炮彈,都是統計总体的例子。然而,我們感兴趣的并不是这些元素的本身,而是这些元素的某种性质。因此,統計总体又可理解为表征这些元素的某种性质(例如产品的某个尺寸，炮彈的射程)的数値的一个集合。統計总体(以下或簡称总体)的內容究竟是什么，不但要看所指的条件是什么，而且要看所考虑的是这些元素的什么性质。例如，我們旣可以观測某車間某班所生产的一批圓柱体的半徑的尺寸大小,也可以只观測这些半徑是否落在某个尺寸范圍以內。此外,我們旣可以观測元素(或个体)的一种性质,例如20—25岁的藏族男子的体高;也可以同时观測多种性质,例如体高与体重。

若总体中的元素有限,則称总体是有限的。例如,某地区在最近过去10年內的每年夏季最高溫度构成一有限总体。若总体中的元素是无穷的,則称总体是无穷的。例如,某河流在将来无限多年份里的年平均徑流量构成一无穷总体。至于旋一枚銅錢的試驗，則可把总体看作出現正面和出現反面两个元素各重复无穷多次的假想无穷总体。統計

总体形形色色，不一而足。下面让我們再考虑一些例子：

(1) 一批圆柱体的椭圆度；

(2) 一顆骰子每投擲两次所得点数的平均数；

(3) 某牌电池的放电时数；

(4) 某某地区地下水位的深度；

(5) 某厂产鋼的硬度、韌度、含碳、含硫及含磷量。

这些統計总体到底是无穷的抑或是有限的，还不是完全明确的。例如，总体(3)和(5)在一定时期内是有限的；但若把某牌某厂看作一定的生产条件，則任何一个时期的产品都不过是这个生产条件下所生产的产品的一部分，因而总体是无穷的。至于总体(4)，由于每点都有一个地下水位的深度，所以是无穷的。

現在我們从一个总体中抽取一些元素作为代表这个总体的样本。如果我們抽取元素的方法是使总体中的每一元素都有同等的机会被抽取，且每次抽取时总体中的元素成分不改变，那末我們所得的样本是一个**簡单随机样本**。取得簡单随机样本的过程或手續叫做**簡单随机抽样**。显然，簡单随机抽样就是重复地进行同一随机試驗。因此当我們从一个总体中抽取一个随机样本时，我們是在重复观測同一个随机变量。随机变量的具体意义就是一个統計总体。

我們說，簡单随机抽样就是重复进行同一随机試驗。关于这一点，可以作进一步的解釋。所謂重复一个試驗，是指每次試驗都在同一組条件下进行，因而每次試驗得到什么結果，其可能程度都是固定不变的。对于有限总体，簡单随机抽样意味着每次抽出一个元素后，放还再抽。若不放还，总体的成份将有所改变，那末再抽时，出現各种結果的可能程度就相对地改变了。虽然“不放还”的抽样方法(仍然叫做随机抽样，但不叫做簡单随机抽样)有它的实际意义，但对于有限总体而言，这样的抽样方法不能算重复同一随机試驗。至于无穷总体則沒有区分“放还”或“不放还”的必要。因为即使抽取后不再放还，有限次的抽取总不致改变总体的成份。

为了保证抽样的随机性，往往需要采取一些具体措施(例如使用随机数表以

保证总体中每一元素有同等被抽取的机会，舍弃不合手續的观测值，等等），以防止抽取时的人为偏見。虽然如此，完全随机的样本实际上还是很少的，因为每一次抽取都要直接或简接通过人的判断来执行；也就是說，随机抽样毕竟是一种理想的情况。况且，有时考虑到具体情况（例如經济原則），也可能有意識地不采取随机抽样的方法。但一般說来，随机样本（对于有限总体說包括“放还”和“不放还”的随机抽样）最足以代表統計总体。在概率論与数理統計中，随机抽样的概念是最基本而重要的。在本书中，我們不准备讲用其他抽样方法得来的样本，而只考虑随机样本，特别是簡单随机样本。除非特别声明，下面所談到的随机样本都是指簡单随机样本而言，并簡称为样本。

数理統計的中心問題就是如何根据样本探求有关总体的种种知識，以及利用样本的随机性测（檢）驗关于总体的种种假設。

1.4 頻率的稳定性·概率

就个别的試驗言，我們很难預料其結果。但当我們进行大量的重复試驗时，情形就根本改观。我們将会发現一种“大量現象”的規律。

以某种商品的生产廢品率为例。可把一定时期內在稳定生产条件下所生产的商品的全体看作一統計总体。現从中任取 n 件，由于生产条件稳定，每取一件时，出現廢品或成品的可能程度都可看作是不变的；也就是說，这 n 次抽取是一个 n 次重复試驗。所抽取的 n 件商品构成一个容量（或项数）为 n 的随机样本。若記廢品为 1，成品为 0，则对应着出現廢品和出現成品这两个事件，随机变量 ξ 分别取値 1 和 0。設在 n 次試驗中，有 m 次得廢品，則比率 $\frac{m}{n}$，$0\leqslant m\leqslant n$，叫做事件“$\xi=1$”（即出現廢品）的相对頻率，或簡称頻率。完全可以想像，对于不同的随机样本，事件“$\xi=1$”的頻率可以不同，并且，随着样本容量 n 的改变，这个頻率也会起一定的波动。但是，只要样本是随机的，我們将会发現，随着 n 的增大，頻率 $\frac{m}{n}$ 将圍繞着某一确定的常数 p 作平均幅度愈来愈小的波动。这就是所謂**頻率的稳定性**。如表(1.4.1)所示，这个常数 p

"似乎"是 6%。

表 (1.4.1) 对于三种不同的样本容量 (n) 各次抽样得到的廢品次数 (m) 和频率 $\left(\frac{m}{n}\right)$

抽样序号	$n=25$		$n=250$		$n=2,500$	
	m	$100\times\frac{m}{n}$	m	$100\times\frac{m}{n}$	m	$100\times\frac{m}{n}$
1	1	4	12	4.8	157	6.28
2	4	16	14	5.6	152	6.08
3	0	0	17	6.8	157	6.28
4	0	0	11	4.4	136	5.44
5	1	4	22	8.8	152	6.08
6	1	4	9	3.6	135	5.40
7	2	8	15	6.0	143	5.72
8	0	0	14	5.6	160	6 40
9	1	4	21	8.4	149	5.96
10	1	4	8	3.2	153	6.12

可以理解，事件的可能程度和它的出現頻率是有一定的联系的；事件的可能程度愈大，它的出現頻率平均地說也愈大。因此，事件"$\xi=1$"的可能程度可以很自然地用数 p 来表示，并称数 p 为事件"$\xi=1$"的概率。这样一来，当 n 很大时，就能由頻率得到概率的近似值。

可注意到，以上我們并沒有直接定义概率，而只是說，当試驗次数很大时，頻率接近于概率。問題还不在于概率的严格定义是什么①，而在于怎样得到一个事件的概率。从頻率的稳定性看概率，則概率建立在多次重复試驗的基础上。就是說，如果我們不能明确規定一个随机試驗并且能无限地(至少在概念上如此)重复这个試驗，則无概率之可言。例如，我們将不談"孔子到过某地方"的概率，也不談"屈原生于公

① 近年来，由于現代数学的发展，人們更傾向于公理化的概率定义，把概率看作是与事件相結合的一个滿足某些公理的数。但是，概率的完整的、严格的公理化处理是比較艰深的，在一个初步的教材里，我們不准备这样做。

元前若干年"的概率，等等。

概率是表示事件的可能程度的一个数。如果我們能用别的方法确定这个数，則暫时撇开頻率而談概率，也未尝不可。在下一章里，我們将看到，利用概率的古典定义确定事件的概率，也不是无益的。但我們会終于看到，从頻率的角度考虑概率有更为普遍而現实的意义。

1.5 問題

1. 我們說"天有不測之風云"与天气可以預报有无矛盾？在机械制造上能不能做到产品的尺寸只会偏大不会偏小？如果能够，是否与我們所說的"无論机床的精密程度多高总是无法預言某件产品偏大或偏小"有矛盾？

2. 两人向同一目标射击，只考虑射中或射不中。問对于这个射击試驗，可以列举哪些簡单事件，哪些复合事件？

3. 为什么說随机变量是一个函数，这个函数有什么規律？

4. 怎样能把連年測量的年徑流量看作随机抽样？一年內的月徑流量呢？

5. 从頻率的穩定性看，概率有什么平均意义？我們說某个水位百年一遇，是不是說一百年內不可能遇到两次这样的水位？

6. 能不能說当試驗次数 n 趋于无穷大时，頻率 $\frac{m}{n}$ 趋于概率 p？試以旋銅錢为例說明你的想法。

第二章　概率的基本运算

为了探求事件的概率而进行大量的实际观測，这只是一方面，虽然也是主要的一方面。另一方面，对事件的概率进行邏輯分析与計算也是很重要的事。后者往往能把問題簡化和精确化。概率論与数理統計的問題需要我們很好地从这两方面辯证地求得解决。

这一章里，我們討論古典概率計算的一些基本概念和方法。

2.1　概率的古典定义

設相对于一个試驗，有也只有 n 个可能发生的情况，并且每个情况都是**等可能的**，其中恰恰有 m 个情况具有性质 A，則謂 A 的概率为 $\frac{m}{n}$，記作

$$P\{A\}=\frac{m}{n}.$$

这就是概率的古典定义。

这个定义有很大的局限性，因为它假定等可能的情况为已知，实則我們怎样知道一个情况和另一个情况是同等可能的呢？在具体应用上，我們只好这样做，就是如果沒有任何理由认为一个情况比另一个情况更容易出現，則姑且把这些情况看做是等可能的。

利用古典定义直接計算概率时，需要把所有等可能情况罗列出来。

例 1. 任从随机数表讀两个个位数，求这两个数之和等于 3 的概率。

解：由于每次讀数为 0, 1, 2, ……, 9 都是等可能的，故第一次讀数共有 10 个等可能情况，而其中的每一个情况又可能配合着第二次讀数的 10 个等可能情况之一，因而两次讀数共有 $10\times10=100$ 个等可能的