

统计数学

学数学系 編著

学技术出版社



內 容 提 要

本书系复旦大学数学系数学专业革新教材之一,内容包括随机变量与概率分布、数理统计、随机过程、信息论等四个方面的基本知识。本书可作综合大学数学专业统计数学课程的教科书,讲授92学时,亦可作高等院校有关专业和工程技术人员的参考书。

統 計 数 学

(試用本)

复旦大学数学系 編著

*

上海科学技术出版社出版

(上海瑞金二路450号)

上海市书刊出版业营业许可证出 093号

新华书店上海发行所发行 各地新华书店經售

商务印书馆上海厂印刷

*

开本850×1168 1/32 印张13 14/32 字数318,000

1960年8月第1版 1960年8月第1次印刷

印数1—12,000

統一書号: 13119 · 378

定 价: (十) 1.50 元

編輯說明

我們受到全國持續躍進的大好形勢的鼓舞和推動，積極響應了黨的號召，在兩年來教育大革命已經取得偉大成績的基礎上，掀起了一个聲勢浩大的教學改革的群眾運動。通過這個運動，我們揭露了現在教學體系、教學內容和教學方法上陳舊落后的狀況，抓住訂方案、編大綱、寫教材、搞試驗等重要環節，試圖建立一套以馬克思列寧主義、毛澤東思想為指導的、反映現代科學發展水平的、理論聯繫實際的新的教學體系和內容，以及與之相適應的教學方法，使培養人才的工作更好地貫徹黨的社會主義建設總路綫的精神。

作為這種新的探索和嘗試，我們在教學改革運動中，師生結合，提供了“關於綜合大學數學專業課程革新的建議”，編寫了一套可供綜合大學數學專業試用的基礎課程教材。全套教材包括數學分析(一)、數學分析(二)、泛函分析、高等代數、綫性規劃和計算實習、計算數學、數理邏輯與控制論、常微分方程、數學物理方程、一般力學、連續介質力學、統計數學(包括信息論)等十二種，尚有物理學一種，因力量所限，未能及時編出。

根據“關於綜合大學數學專業課程革新的建議”的精神，我們力圖使這套教材具有以下几个特點：

一、在選材上，注意克服資產階級教育思想的影響，體現為社會主義建設服務和反映現代科學發展的要求。中國數學會提出的“數學發展的方向必須以解決尖端技術和重大工程、現代物理、自動化、國民經濟和大量計算任務中的數學問題為綱”，具體地說明了社會主義建設和現代科學發展對數學的要求，我們即以此作為選材的主要標準，同時，也考慮到基礎課的某些特殊要求，適當注

意了根据理論与实践之間的直接联系与間接联系，当前需要与长远需要等关系，来确定材料的取舍和不同材料的主次安排。根据这个精神，我們精簡了原来基础課内容中一部分不必要的古典内容，添加了一部分現代材料，还增加了一些新課程。

二、在材料处理上，注意克服过去課程設置各自为政、互不联系的缺点，体现科学知識的綜合与分类的辯証統一的关系。特别是近年来，边缘学科大量出現，科学发展在原有基础上愈来愈明显地趋向新的更高级的綜合，我們想力求使这套教材适应这个趋势。

具体說来，对那些条件已經成熟的、可以綜合处理的内容，即加以統一处理，例如将泛函分析与实变函数、积分方程以及綫性代数中的部分内容統一处理，在泛函分析中加以綜合；对那些綜合趋势已經比較明显，但独立設課条件尚不成熟的，也分別情况，注意在有关課程間建立密切的有机联系，若干材料还重新另行配置，例如原来理論力学中振动理論的一部分内容，这次就移到常微分方程中去了。

三、在材料的处理与闡述上，以辯証唯物主义观点为武器，破除形而上学和唯心主义对数学教学的影响。数学的研究对象是客观世界的量的側面，所以它具有較多的抽象性，在研究方法上也較多地运用邏輯上的演繹推証。这些特点，本来應該有利于深刻地闡明問題的本質，但唯心主义者却总是加以歪曲，企图在引出抽象概念时，掩盖其实践来源，在形式論証中，避免闡述問題的本質。在这套教材中，我們力求消除这些唯心主义观点的影响。具体說来，对某些与生产实践有着更加直接联系的課程（如数学物理方程），既吸收已經严格建立数学理論的材料，也采用在实践中有广泛应用而理論上尚未成熟的材料，重新加以組織，恢复这门課程本来的生动活潑的面貌。各門課程中，对重要数学概念与問題的引进，都尽量闡明它們的直接和間接的实践来源；闡述論証过程中，

插入若干必要的描述性材料；得到的結論，也闡明它在實踐中直接和間接的作用。本學期我系幾門主要基礎課程，都初步做到了減少學時、提高質量，據了解，主要是在教學過程中初步體現了這個精神。因此，根據我們一些不成熟的經驗，要徹底解決這個問題，除在教材內容上盡量克服這些唯心主義觀點的影響以外，還要注意在教學方法中消除這些影響。

徹底實現教學改革，建立一套新的體系，是一個艱巨複雜的任務，也需要一個較為長期的時間來摸索。我們所作的一些嘗試，僅是個開端，既受到思想水平和科學水平的限制，又缺乏較充分的實踐經驗，某些課程的教材，還是在師生結合、邊學邊寫的情況下編出來的。因此不論在處理原則上或者在處理方法上都還不夠成熟。有不少問題，例如如何在教材中反映我國社會主義建設實際中所提出的數學問題、如何在有關各課程間建立更密切的有機聯系等等，在編寫過程中，也還把握不定，處理不盡適當。我們懇切地希望同志們批評和指正。

上海科學技術出版社和商務印書館上海印刷廠對這套教材的迅速出版，給了極大的支持，我們在這裡表示衷心的感謝。

復旦大學數學系

1960年5月

序

統計数学是研究自然界中随机現象的数量規律的一門学科，它在尖端技术、大型工程、現代物理、国民經济及大量計算中都有广泛的应用。但在解放前，这一学科在我国却是空白点。解放后，特别是 1958 年大跃进以来，在党的领导下，在整个飞跃发展的形势推动下，統計数学在我国也得到了迅速的发展。而且随着技术革新和技术革命运动的蓬勃开展，工农业生产水平和科学技术的迅速提高，統計数学的应用也将愈为广泛。为了使培养出来的干部能更好地符合我国社会主义建設的需要，在綜合大学数学专业的教学計劃中，必須加强統計数学的内容。为此，迫切需要有一本符合綜合大学数学专业教学計劃的統計数学教科书。

本书就是为了上述目的，根据复旦大学数学系課程革新方案中所提出的原則进行編写的。在編写过程中，我們力求貫徹辯証唯物主义的观点和理論联系实际的原則，并以社会主义建設的需要与科学技术发展的前景为取材的依据。与以往的教材相比，我們大大地增加了在国民經济及在現代物理、工程技术中有广泛应用的数理統計及随机过程内容。同时，鉴于信息論这一分支在現代科学中愈来愈重要，因此有必要也将它列入本教材中。本书中对于一些重要的基本概念和理論，尽可能地做到从实践的需要引入。在闡述一些重要理論时，力求避免純粹邏輯上的推导，而將它的本质交代清楚。我們认为，这不仅可以把概念講深講透，而且有助于讀者正确对待理論与实际的关系，树立辯証唯物主义的世界观。

本书共分四篇：第一篇是随机变量与概率分布，叙述了随机

变量、分布函数、数字特征、极限定理等统计数学中的基本概念与基础理论；第二篇是数理统计，叙述了参数估计和假设检验的基本理论，并在此基础上介绍了在国民经济中有着广泛应用的抽样理论、离差分析、回归分析、质量控制与极值分布等内容；第三篇是随机过程，介绍了马尔可夫过程和平稳过程的基本理论以及它们在统计物理、公用事业、自动控制等方面的应用；第四篇是信息论，扼要地介绍了信息论的基本概念和一些主要定理。

本书的全部内容是假定读者已经掌握了数学分析^①的基础知识，根据讲授时数为 92 学时的课程来选取的。但是，为了理论的完整性，我们亦用小字编入了一些有关的内容。

由于我们的经验和知识都很不够，有些部分还是“边干边学”写出来的，因此一定存在着不少的缺点，我们热烈地希望听到各方面的宝贵意见，并希望通过对本书的批评和讨论，有助于很快地产生一批密切结合我国实际情况、具有高度科学水平的统计数学教科书。

复旦大学数学系统计数学教材编写小组

1960 年 5 月

^① 包括实变函数及复变函数的基本知识。

目 录

序

緒論	1
----------	---

第一篇 随机变量与概率分布

第一章 基本概念	5
----------------	---

§ 1 事件和概率	5
§ 2 古典概型	11
§ 3 概率场	16
§ 4 概率的基本运算法则	23
§ 5 独立試驗序列概型	28
§ 6 例	35

第二章 随机变量及分布函数	44
---------------------	----

§ 1 随机变量	44
§ 2 多元随机变量及多元分布函数	51
§ 3 經驗分布函数与直方图	61
§ 4 随机变量函数的分布	69

第三章 随机变量的数字特征	76
---------------------	----

§ 1 数学期望	77
§ 2 方差	82
§ 3 随机变量函数的数字特征	91
§ 4 矩	96

第四章 特征函数	99
----------------	----

§ 1 特征函数的定义及性質	99
§ 2 逆轉公式及唯一性定理	102

§ 3 多元回归	215
§ 4 回归系数与相关系数的假设检验	223
第六章 质量控制	228
§ 1 问题的提出	228
§ 2 次序统计量	230
§ 3 控制图	234
§ 4 控制图的利用	238
第七章 极值分布	241
§ 1 问题的提出	241
§ 2 三种类型的原始分布	242
§ 3 第一种渐近分布	244
§ 4 第二种与第三种渐近分布	251
§ 5 最小值的分布	251

第三篇 随机过程

第一章 马尔可夫过程	255
§ 1 定义	255
§ 2 马尔可夫链	257
§ 3 时间连续状态离散的马尔可夫过程	261
§ 4 状态为连续的马尔可夫过程和扩散过程	272
第二章 平稳过程	282
§ 1 引言	283
§ 2 平稳过程和相关函数	285
§ 3 相关函数的谱分解及各态历经定理	290
§ 4 平稳随机函数的谱分解	296
§ 5 平稳过程在线性动力学系统中的应用	302
§ 6 平稳过程的线性滤波	308

第四篇 信息论

第一章 信息量的概念	320
------------------	-----

§ 1 信息量	320
§ 2 基本性质	324
§ 3 唯一性定理	329
§ 4 連續信源的熵及其基本性质	333
第二章 編碼問題	343
§ 1 最优編碼	343
§ 2 抗扰編碼	352
第三章 离散信源与通路	357
§ 1 无记忆的离散通路	357
§ 2 有限记忆的离散通路	366
§ 3 漸近等同分割性	373
§ 4 漸近等同分割性的应用	379
第四章 帶干扰通路的信息傳送問題	385
§ 1 帶干扰通路的基本定理	385
§ 2 帶干扰通路的信息傳遞	394
§ 3 連續信号的构造	400
附录一 測度与积分	403
附录二	409
参考文献	417

緒 論

在自然界中广泛地存在着一类所谓随机现象。例如我们在进行某种测量时，由于种种偶然因素的影响（如测量仪器受大气的影响，观察者生理上或心理上的变化等），不可避免地会产生测量的误差。又如在容器里盛着一定体积的气体，气体分子由于受到其它分子的冲击而产生运动速度和方向的随机变化；又如飞机在高空飞行时，由于大气中湍流等各种影响，它环绕着重心作随机摆动；与此类似地，船舶在海洋中航行时，由于受到海洋波浪的影响而产生各种各样的摇摆（纵摇、横摇）以及高低起伏的深沉现象等等。

实践证明，研究了大量的同类随机现象后，通常总揭露出一种完全确定的规律性，也就是大量随机现象所特有的一种规律性。例如由分子物理学的观点来看，气体是由无数气体分子所组成的，而这些分子是在不断运动着，且在运动过程中彼此影响着的，因而每个分子的运动轨道、速度、方向都是随机的。但是我们知道，从宏观看来，气体对器壁的压力却是稳定的，这是因为分子数足够大，因而各个分子的运动所具有的随机性在集体作用下就互相抵消了。又如在射击中，如果射击次数不大时，靶上命中点的分布是完全没有规则的，杂乱无章的，没有什么显著的规律性；当射击次数增加时，分布就开始出现一些规律性，射击次数越大，规律性越清楚，我们在本书中将证明它是服从正态分布的。综上所述，知道个别随机现象虽然是无规律的，但大量性质相同的随机现象总是有“统计”规律性的。统计学就是一门研究随机现象的数量规律

的科学。由于随机现象的普遍性，使统计数学的概念与方法具有极为普遍的意义。仅就我们自 1958 年以来所解决或接触到的实际问题来看，就有原子能、大型工程、自动控制、电子学、生物学等方面提出的随机过程与信息论的问题；气象、水文、纺织、医学等方面提出的数理统计问题。这种多方面的需要，就决定了统计数学在数学领域中所处的地位与作用以及它的广阔的发展前途。

统计数学中的一切定律（如大数定律）是宇宙中客观存在的大量随机现象统计规律性的反映。我们必须指出，随机性现象与决定性现象之间，必然性与偶然性之间并没有不可逾越的鸿沟。“那被断定为必然的东西，是由种种纯粹的偶然所构成的，而被认为是偶然的東西，則是一种有必然性隐藏在里面的形式”（恩格斯：“费尔巴哈与德国古典哲学的终结”，1959 年人民出版社出版，第 34 页）。事实上宇宙中没有那一件实际的现象，其中不带有某种程度的随机成分。因而任何决定性现象也不可避免地有随机偏差产生。在许多实际问题中，为了处理问题的方便，在所要求的精确度的范围内，可忽略掉那些造成随机偏差的次要因素，而只考虑起主要和基本作用的那些因素，这是在研究自然现象时，为现代自然科学与工程技术方面所常采用的方法。也就是说，先找出那些对现象起决定作用的、最主要和最基本的条件，然后利用数学工具（例如引出描述现象的微分方程）把它解出来，找出在这组基本条件下现象所具有的主要规律。例如过去在船舶设计时，就是根据流体力学的原理列出微分方程，求出船舶在航行中所受的阻力，它的航行规律。但是，这种做法是以牺牲精确度为代价的。在某些精确度要求较高的问题中（如火箭发射角的确定），就必须将这些随机因素考虑在内。这样一来，由于众多的因素，使得研究决定性现象的那些数学工具（如微分方程等）无法适用了，因而就必须采用统计数学的方法。

显然,随着我国社会主义建设的飞速发展及科学技术愈来愈高的要求,统计数学的应用范围也将愈来愈广泛。例如,由于考虑了海洋波浪对船舶摇摆的影响而出现了以平稳过程理论为基础的最新的“适航性理论”。在建造数百米高的电视塔时,由于要考虑到空气中湍流的影响,因而就提出了“随机微分方程”的问题。又如马尔可夫过程在统计物理学中的应用,平稳过程在无线电电子学及线性动力学系统中的应用,数理统计在气象、产品质量检查等方面的应用,以及信息论在通信技术、生物学等方面的应用……。可以这样说:几乎没有一门自然科学、没有一个工业部门不用到统计数学的方法。

可见,统计数学在高速度建设社会主义中所起的作用是巨大的,这也给我国统计数学工作者带来了极其光荣而艰巨的任务。虽然我国统计数学工作者自1958年以来已取得了巨大的成绩,但是这还远不能适应高速度进行社会主义建设所提出的要求。我们坚信,在党的领导下,高举毛泽东思想的红旗,坚持理论联系实际的原则,我国统计数学工作者一定能赶上形势,适应形势的要求,在解决实际问题上,在理论研究上,作出显著的成绩,用最短的时间攀登世界科学的高峰。

第一篇 随机变量与概率分布

第一章 基本概念

§1 事件和概率

当我们多次观察自然现象后,会发现许多事情在一定的条件下必然会发生。例如在沒有力作用的条件下,作等速直线运动的物体必然繼續作等速直线运动。又如在标准大气压下,水加热到 100°C 时,必然会沸腾等等。这种在一定的条件下,必然会发生的事情就称为必然事件;而那种在一定条件下,必然不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就必然不可能改变其等速直线运动的状况。从所举例子中可以看出,必然事件和不可能事件之间有着很紧密的联系。事实上,如果在一定的条件下,某个事情是必然事件,那么在同样的条件下,那事情的反面就必然是不可能事件;反过来也一样。

但是在自然现象中,除了上面提到的必然事件和不可能事件外,也还存在着另一类与此有本质不同的事情。这种事情在一定的条件下可能发生也可能不发生。这种事情我们称为随机事件,简称为事件。为了说明这种事件是在自然现象中广泛存在着的,我们来看下面的一些例子。

[例1] “在6到8月間黃浦江的最高水位小于5米”便是一个随机事件。因为我们无法断言在6到8月間黃浦江的最高水位小于5米还是超过5米。这种事情可能发生也可能不发生,因而是一随机事件。但是由于我们对6到8月間黃浦江的最高水位进

行的多次观察，我們还是可以断定 6 到 8 月間黃浦江最高水位小于 5 米这一事件发生的可能性的太小。

[例 2] “在一分鐘內，一个电话交換台至少接到 15 次呼喚”也是一个事件。因为在一分鐘內，可能接到不到 15 次呼喚，也可能接到超过 15 次呼喚等等。但是由于人們以往长期多次重复观察的結果，仍可以預断在一分鐘內电话交換台接到至少 15 次呼喚的可能性的太小。

[例 3] “在一定的温度下，氫分子运动速度小于 300 米/秒”也是一个随机事件，因为在一定的温度下，氫分子运动速度是变化的，它可以小于 300 米/秒，也可以大于 300 米/秒，因而无法肯定氫分子运动速度一定都小于 300 米/秒。但是观察大量的氫分子运动，按照物理学上的定律，人們仍能断言“在一定温度下，氫分子速度小于 300 米/秒”这一事件发生的可能性的太小。

[例 4] “在抽查某工厂生产的 10 件产品时，发现有一件次品”是一个事件。因为抽查的結果可能正好只发现一件次品，也可能沒有发现次品或发现的次品不止一件。但是按照人們对这个工厂过去生产情况的了解，我們还是能够預断这个事件发生的可能性的太小的。

类似的例子，还可以举出很多，但从上面所列举的一些例子中，就可以得到一些結論。

首先，事件确实是广泛存在于自然現象中，而且对这些事件的研究是非常必要的。例如对黃浦江每年最高水位的研究將有助于水坝等工程的設計，电话交換台一分鐘內接到呼喚次数的研究將有助于决定应設置多少綫路等等。

其次，事件虽然有其不确定的一面，即它在一次試驗中，可能发生也可能不发生，但是在多次和长期的观察中或在大量現象中，人們还是可以发现其中的規律性的。这就是說事件发生的可能性

究竟有多大是可以决定的。为了阐明随机现象的这种规律性,保持在一定的相同的条件下进行观察是很重要的。因为根据经验知道,在条件不同时,事件发生的可能性也会发生变化。

也许人们会误会,事件发生的可能性只是人们主观上的一种东西,其实这是完全错误的。它完全是由事件本身的客观规律性所决定的。为了更清楚的说明这一点,我们来看下面的例子。

[例5] 检查大批的产品,当被检查的产品长度介于13.60厘米到13.90厘米内,则产品为合格的,否则是次品。我们分别抽取5件、10件、60件、150件、600件、900件、1200件、1800件来检查,其情况如下表和图1-1-1所示:

总 数	5	10	60	150	600	900	1200	1800
合格产品数	5	7	53	131	548	820	1091	1681
合格品频率	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

$$\text{合格频率} = \frac{\text{合格数}}{\text{总数}}$$

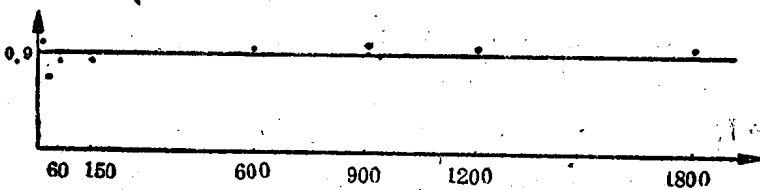


图 1-1-1

虽然抽出的产品中,次品数目是随机的,然而随着抽查的总数增多,合格的频率愈来愈趋于一个稳定值0.9,这完全是为它本身的规律性所决定的。

从上面的例子中可以清楚地看出,随着试验次数的增多,频率越来越清楚地呈现出稳定性来,亦即频率越来越接近于某个固定

的数值 p ，只是偶而产生較大的偏差。这种客观存在的稳定性就正是前面所說的“可能性”量度的客观基础。数值 p 的大小，也正是可能性大小的数量描述。我們把数 p 就称为該事件的概率。因此，事件的概率是表征事件出現可能性大小的一个量。如前所述，它是客观規律性的一种反映，只是这种規律性，跟以往經常碰到的“在一定条件下，某事件必然发生”这类規律性不同而已。这种規律性我們称之为統計規律性，它正是統計数学所要研究的对象。

上面我們討論了事件和它的概率这两个統計数学中最基本的概念。今后我們將用大写的拉丁字母 A, B, C 等等表示事件，而用 $P\{A\}, P\{B\}, P\{C\}$ 等等表示相应事件的概率。

順便指出，必然事件和不可能事件亦可以看成特殊的随机事件。必然事件是每次观察中一定发生的事件（当然都是指在一定条件下的），因而其出現的頻率為 1，所以它的概率也就是 1。必然事件通常記作 U ，所以有 $P\{U\}=1$ 。完全类似地可討論不可能事件，我們通常把它記为 V ，而有 $P\{V\}=0$ 。

上面我們只是一个个地来研究事件及其概率。但是在实际生活中，往往要求我們同时研究几个在同样条件下的事件以及他們之間的联系等等。例如在檢查某些圓柱形的产品中，要求它的长度和直徑都符合規格才算合格，这时我們要考慮“产品合格”、“产品不合格”、“直徑合格”、“直徑不合格”、“长度合格”、“长度不合格”、“直徑合格但长度不合格”等等这类事件。显然，这些事件相互之間是有联系的，从而他們的概率之間亦必然有关系。又如在电话交换机的問題中，我們常要考慮“在一分鐘內接到 1 次呼喚”、“在一分鐘內接到 2 次呼喚”以及“在一分鐘內接到不多于 5 次呼喚”、“在一分鐘內接到多于 5 次呼喚”等等的事件。显然在这些事件中也是有着联系的，从而它們的概率之間也必然有关系。总之，在考慮任何一个随机現象时，总要同时考慮与之联系的种种事件