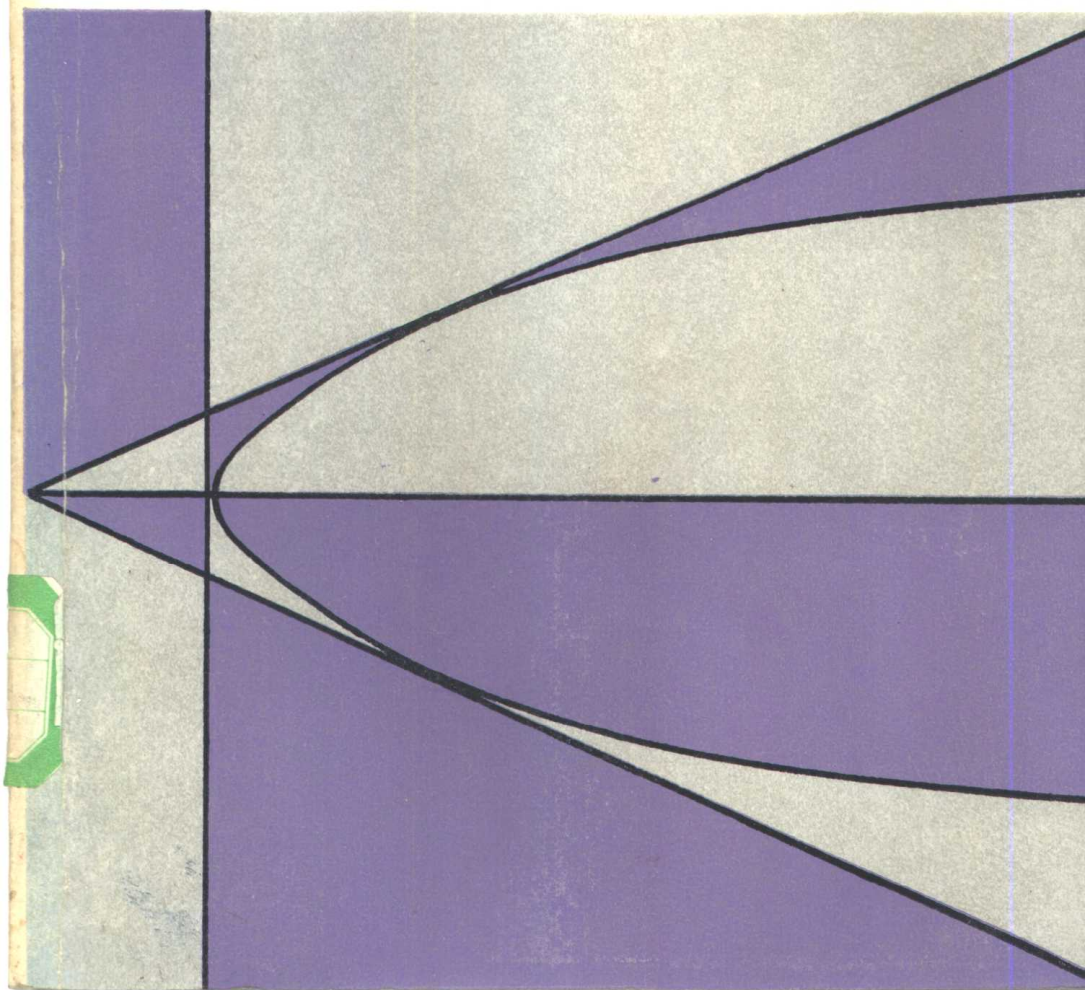


初中教师进修用书

概率论 与数理统计

吴天滨 编



初中教师进修用书

概率论与数理统计

吴天滨 编

山东教育出版社

一九八三年·济南



初中教师进修用书
概率论与数理统计

吴天滨 编

*

山东教育出版社出版
(济南经九路胜利大街)

山东省新华书店发行 山东新华印刷厂潍坊厂印刷

*

787×1092毫米 32开本 14.125印张 285千字

1983年10月第1版 1983年10月第1次印刷

印数 1—23,000

书号 7275·153 定价 1.20元



内 容 提 要

本书是初中教师进修用书之一。书中系统地讲述了概率论与数理统计的基本理论和应用,内容包括事件及其概率、随机变数及其概率分布、多元随机变数与随机变数的函数、随机变数的数字特征、极限定理、统计推断概要、方差分析和回归分析、随机过程论初步等。标*号者为选学内容。其基本部分符合高师专科教学大纲的要求,加上标*号的选学内容则符合高师本科教学大纲的要求,略去标*号的部分,不影响内容的系统性。

本书除供初中教师进修之外,也可作为高等师范院校数学专业的教学参考书。

出版说明

《初中教师进修用书》是为了适应培训教师的需要，由华东地区上海、山东、江苏、安徽、浙江、江西、福建等六省一市八家出版社协作组织编写出版的。目的是供在职初中教师业余进修，帮助他们系统地学习和掌握有关专业的基础理论、基本知识和基本技能，提高文化水平和教学能力，以便在一定时间内通过考核达到两年制高等师范专科毕业的水平。

这套用书，目前先出语文、数学两个学科，共十九种，以后将逐步扩大到其他学科。编写当中，在坚持四项基本原则，坚持思想性和科学性相统一的前提下，注意了以下几个方面：

一、根据教育部制订的高等师范专科学校教学大纲的要求，确定各册内容的深度和广度，既体现各学科知识的系统性，又力求作到简明、精炼，避免繁琐。

二、以提高教师科学文化水平为主，适当联系中学教材和教学实际，把提高知识水平和提高教学能力结合起来，达到学以致用目的。

三、从初中教师的实际水平出发，循序渐进，逐步提高要求；重视讲清学习中的难点和疑点，文字力求浅显易懂；并根据自学或函授的需要，配置必要的提示、注释、思考题和提供参考书目等学习辅助材料。

协作编写教师进修用书，尚属初次尝试。我们将在实践过程中广泛听取读者的意见和建议，努力提高书籍质量，使它更好地适合教学进修的需要。



目 录

绪论	1
第一章 事件及其概率	12
§ 1.1 试验和事件	12
§ 1.2 频率和概率	15
§ 1.3 事件间的关系和运算	19
§ 1.4 古典概型	28
§ 1.5 几何概型	40
§ 1.6 事件与概率的严格数学定义	46
§ 1.7 条件概率	53
§ 1.8 独立性	60
§ 1.9 较复杂的例题	68
本章提要	78
第二章 随机变数及其概率分布	81
§ 2.1 随机变数及其概率分布的概念	81
§ 2.2 离散型随机变数	82
§ 2.3 二项分布	85
§ 2.4 泊松分布	96
§ 2.5 连续型随机变数及其分布密度	105
§ 2.6 正态分布	110
§ 2.7 分布函数的基本性质	122
本章提要	126
第三章 多元随机变数与随机变数的函数	130



§ 3.1 多元随机变数及其分布函数	130
§ 3.2 多元离散型分布和连续型分布	133
§ 3.3 边际分布与独立性	140
§ 3.4 关于分布函数的 R-S 积分	147
§ 3.5 随机变数的函数的分布	152
§ 3.6 χ^2 分布、 t 分布和 F 分布	158
• § 3.7 分布函数的类型	168
本章提要	171
第四章 随机变数的数字特征。极限定理	176
§ 4.1 随机变数的均值	176
§ 4.2 方差与相关系数。矩的概念	184
§ 4.3 大数定律	195
• § 4.4 特征函数	201
§ 4.5 中心极限定理	208
§ 4.6 局部极限定理	216
本章提要	223
第五章 统计推断概要	228
§ 5.1 一些基本概念	228
§ 5.2 子样统计量及其分布	237
§ 5.3 参数的点估计	242
§ 5.4 参数的区间估计	252
§ 5.5 u 检验法 ($N(0,1)$ 检验法)	261
§ 5.6 t 检验法和 F 检验法	267
§ 5.7 χ^2 检验法	273
本章提要	279
第六章 方差分析和回归分析	284
§ 6.1 一元方差分析	284



§ 6.2	二元方差分析	296
§ 6.3	一元回归分析(一)	307
§ 6.4	一元回归分析(二)	317
§ 6.5	多元回归分析简介	332
*第七章	随机过程论初步	336
§ 7.1	随机过程的概念	336
§ 7.2	独立增量随机过程	340
§ 7.3	马尔可夫链	347
§ 7.4	时间连续状态离散的马尔可夫过程	357
§ 7.5	平稳随机过程	368
习题答案或提示		380
附表		423
参考书目		440



绪 论

为使初学者能够尽早对概率论与数理统计得到一个概略的了解，在叙述这门学科的系统知识之前，首先对它的研究对象、主要分支及其在自然科学、技术科学、经济与国防等方面的应用，作一个通俗简要的介绍。

一、概率论的研究对象

试验是探求科学规律的基本手段之一。所谓试验，就是在一定条件下观察发生什么结果。当观察比较粗略时，人们往往只注意决定性现象，也就是条件完全决定结果的现象；而当观察比较精细时，人们就会注意到随机现象（非决定性现象），也就是条件不能完全决定结果的现象。例如“随手掷一枚硬币”（条件），若观察较粗略，只注意硬币是否落下来，则观察到的是决定性现象，“落下”是必然事件，“不落下”是不可能事件。如果观察得精细一些，把“落下”区分为“正面向上”和“反面向上”，则观察到的是随机现象，“正面向上”和“反面向上”都是随机事件。如果观察得更精细一些，在硬币的正反面各刻一个箭头，观察落下后哪一面向上以及箭头的方向，这就是一个更复杂一点的随机现象了。实际上，随机现象在自然界和人类的一切实践活动中都是广泛存在的。例如“向目标射击一弹”，则结果可能是“命中”，也可



能是“偏高”或“偏低”，……；又如“对同一道路进行长度测量二次得值 X_1 和 X_2 ，则结果可能是“ $X_1 \leq X_2$ ”或“ $X_1 > X_2$ ”；等等。这里，对射击的例子可能有不同见解，认为神枪手命中目标是必然的。其实若目标取得充分小，神枪手也不可能枪枪命中；反之若目标充分大，劣等射手也可以弹无虚发。由上可见，当科学技术处在粗略阶段时，往往忽略随机现象；而当科学技术深刻化精密化时则不能不考虑随机现象。

在概率论产生以前，人们仅限于从决定性现象中探求科学规律，而概率论正是研究随机现象数量规律的数学分支。由于随机现象是普遍存在的，而科学技术的精密化和深刻化又必须考虑随机因素，因而概率论的方法（即“统计方法”）正日甚一日地渗入到几乎一切自然科学、技术科学、经济管理以及军事科学中去。概率论是当代发展最迅速、应用最广泛的数学分支之一。

二、随机事件的概率、随机变数和随机过程

概率论最基本的概念是随机事件及其概率。随机事件发生的可能性有大有小，比如在同一射击条件下，神枪手和初学者命中目标都是随机事件，但二者可能性的大小不同，前者大而后者小。随机事件可能性的大小可以从重复试验中显示出来（我们总假定试验是可以多次重复的，因为科学只能从重复试验中找出规律）。若随机事件 A 在 n 次独立试验中发生了 k 次，则比值 $\frac{k}{n}$ 叫做频率。经验表明，当 n 足够大



时, 频率 $\frac{k}{n}$ 常呈现稳定状态 (或叫丛集状态), 我们把频率稳定的中心叫做随机事件 A 的概率, 记为 $P(A)$; $P(A)$ 就表明随机事件 A 发生的可能性的 大小。

为了便于用数学方法研究随机现象, 很自然地可以用数来表示试验结果。这样就得到了随试验结果而变的变数, 称之为随机变数。例如向桌上掷一枚硬币, 以 ξ 表示出现正面的次数。则当出正面时, $\xi = 1$; 而当出反面时, $\xi = 0$ 。 ξ 就是一个随机变数。 ξ 取值的概率: $P\{\xi = 0\} = \frac{1}{2}$, $P\{\xi = 1\} = \frac{1}{2}$ 叫做 ξ 的概率分布。又如射击时弹着点到目标中心的距离是一个随机变数; 进行某种测量时的误差也是一个随机变数。

实际的随机现象往往是一个和时间进程相联系的过程, 每个时刻 t 的状态可用一个随机变数 $\xi(t)$ 来表示, 而一段时间 (T_1, T_2) 内的状态则需要用一族随机变数 $\{\xi(t), T_1 < t < T_2\}$ 来描写, 这就是随机过程 (或称随机函数) 的概念。例如一粒花粉在水中受水分子的推动而作随机运动 (这叫布朗运动), 它在某一段时间中的运动过程就是一个随机过程。又如飞机在飞行中由于受空气湍流等影响而产生的随机波动的过程也是一个随机过程, 等等。

关于事件、概率、随机变数和随机过程的一系列理论就是概率论的基本内容。

三、统计方法与数理统计学

作为数学学科来说, 概率论属于 “纯粹数学”, 而以概率论为基础的数理统计学则是 “应用数学” 的重要分支。二者



联系十分密切,很难严格划分界限。如前所说,概率论是在随机现象的一般数学模型的基础上研究事件、概率、随机变数和随机过程的基本规律;而数理统计学则针对实际处理随机现象的任务提出数学模型,研究其规律并提出解决问题的方法。一般教科书多是把二者结合起来叙述。

用概率论解决实际问题的方法叫做统计方法。统计方法有两个显著特点。第一个特点是由部分推断全体。被研究对象的全体在统计学中叫做母体(或称总体),从中随机抽取一部分就是子样(或称样本)。凡统计方法都是通过对子样的统计分析来推断母体的性态,否则就不能算是统计方法。例如要检验一批灯泡的耐用时间,统计方法就是抽取一个子样(比方说 10 个灯泡组成的子样)进行检验,从这 10 个灯泡的耐用时间来推断整批灯泡的情况。如果把整批灯泡挨个检验,那就不是统计方法了(虽然按照日常语言的习惯,全面检验也应是一种“统计”,但作为数学方法来说,这不叫“统计方法”)。这个例子也从一个方面表明了使用统计方法的必要性;因为对类如灯泡耐用时间这样的对象,全面检验是行不通的,全面检验就会毁掉全部灯泡。既然是由部分推断全体,那就不可能以百分之百的把握作结论。统计方法的第二个特点就是以接近于 1 的概率(例如 0.95, 0.99, 但不等于 1)保证所作结论正确。实际上这就是把概率接近于 1 的随机事件当作必然事件,这叫做“实际推断原理”。其实细想一下,我们在日常生活及生产活动所说的必然事件,往往都是可能性很大(即概率接近于 1)的事件,而不是绝对必然发生的事件。比如我们说乘车必然比步行快,其实若车子出了偶然



事故就可能比步行更慢，但车子一般不会出事故，即车子不出事故的概率通常接近于 1，因此我们把乘车比步行快当作必然事件。由此可见，概率接近于 1 的随机事件特别重要，相应地在概率论中有一套极限理论专门研究概率接近于 1 的规律。

广义的数理统计学泛指概率论在实际中的各种应用。狭义的数理统计学则指统计观察方法的拟定和统计资料的分析，主要包括以下内容：

1. 数据整理和子样统计量的研究：这是数理统计学的基础部分。

2. 统计推断理论：根据子样来推测判断母体的情况叫做统计推断，这是数理统计学的核心部分。统计推断理论包括两大方面——参数估计和假设检验。参数估计就是根据子样来估计母体的某些参数(例如平均值等)；假设检验就是针对实际问题作出假设，然后利用子样来检验这假设，以接近于 1 的概率作出正确的推断。

3. 方差分析：这是推断多个子样是否取自同一个母体的统计分析方法。

4. 回归分析：某种树的直径与高度都是随机变数，它们之间的关系是一种“统计相关”。某地区的年最高气温与降雨量，某种动物的体重与身长，人的体温与某种病的感染程度也都是统计相关的随机变数。数理统计中研究两个或多个随机变数统计相关性的理论和方法叫相关分析。在理论研究和实际应用中，只考虑一个随机变数而把其他随机变数当作普通变数看待是很有利的(例如把树的直径当作普通变数，只



考虑树高这一个随机变数),由此进行变数间关系的统计分析就叫做回归分析。包括检验相关的密切程度,找近似函数关系,通过一个或几个变数预测和控制另一个变数等等。

5. 抽样理论:研究从母体中抽取子样的方法。一个好的抽样方案一方面要求抽取的样品总数尽可能少,另一方面要求作出判断正确的概率尽可能大。抽样方法除较早的单式抽样(一次取出随机子样,据此作出判断)外,还有复式抽样(先取一个随机子样,必要时再取一个随机子样)和序贯抽样(依次取随机子样,直到作出推断为止)。

6. 质量控制:就是利用统计推断原理制定一系列公式化图表化的易于执行的方法,及时检验生产过程是否正常,及时发现出废品的倾向,以便采取措施。这对于提高产品质量显然具有重要作用。

7. 试验设计:改进生产工艺,选择优良的原料配合及生产条件等,都需要进行试验。试验设计的任务是提供一套方法,使能通过较少次数的试验而得出比较全面准确的推断。近年来推广较多的正交试验设计就是一种较好的试验设计方法。

此外,还有多元分析、极值分析、过程统计等数理统计分支,这里就不逐一介绍了。

四、有关概率论发展的一些问题

概率论的蓬勃发展是十九世纪末期以后的事情,它是与大工业生产和自然科学发展的需要相适应的。概率论的起源则早在十七世纪,当时由于对机会游戏(赌博)中随机现象的研究而提出了概率论的一些基本概念。当时研究概率论主



要用算术和排列组合的方法。后来，概率论在物理学、社会保险事业和大规模工业生产中得到了应用，由于这些实际需要的推动，概率论广泛运用了微积分、微分方程、代数和几何等数学工具，得到很大发展。但这时，概率论还不是一门成熟的数学学科，它的基本概念缺乏严格定义，事实上还停留在直观基础上。直观是知识的来源，但单靠直观绝不能发展起深刻的理论。至本世纪三十年代，概率论建立了严格的公理系统，即用集合定义事件，用测度定义概率，用可测函数定义随机变数和随机过程，用抽象积分定义数学期望（均值）等。于是概率论的发展进入了新阶段，理论上不断成熟与完善，应用上日益广泛和深化。愈来愈多地应用统计方法已成为各门自然科学和技术科学现代发展的一个特征。

这里要提一提概率论与其他数学学科的相互渗透的问题。在概率论中广泛地应用集合论、测度论、函数论、微积分、微分方程、代数、拓扑、组合论和泛函分析等数学分支的概念和结果，同时概率论也有力地推动这些数学学科的发展。另外，概率论的方法也在逐步渗透到多种数学学科中去，产生新的分支。例如微分方程是解决多种力学、电学问题的有力工具，但随着科学技术的精密化，必须考虑到随机因素的影响，比如空气中湍流的影响，随机振动的影响等，这样就提出了随机微分方程的问题。类似地也产生了随机代数方程、随机规划等数学分支。在计算数学中应用概率论的方法产生了统计试验算法（蒙特——卡洛方法），成为一种很有力的计算方法。

还应说明一个问题，即对概率论的一些概念发源于赌博

这一点应有正确的认识。概率论作为研究随机现象统计规律的一门科学，随着生产和科学技术的发展而产生和发展是必然的；但它萌芽于何时何地何种问题是有偶然性的。毫无疑问，生产实际和科学发展的需要是推动概率论发展的根本动力。

五、概率论与数理统计在科学技术和经济国防事业中的若干重要应用

如果说在二十世纪三十年代，概率论作为一门数学学科只是初具根茎枝叶的话，那么几十年来，它已成长为一株参天大树，枝多叶茂，花果累累。统计方法已经和正在渗透到各种领域，出现了许多边缘学科。这种情况的出现是必然的，因为每一门科学领域的深入研究都要出现必须分析和处理随机因素的阶段。在某些学科中，由于对象的特殊而较早地采用了统计方法，如统计物理学。到今天，几乎一切自然科学和技术科学都愈来愈多地应用了统计方法。下面我们从某些实际问题谈起，粗略地介绍概率论与数理统计的某些分支，若干边缘学科及其应用。

1. 对随机现象（特别是布朗运动，随机徘徊，无线电通讯中的随机干扰，随机服务等）的深入研究必须考虑到现象随时间进行的过程，由此发展起“随机过程论”。随机过程论是概率论中较高深的部分，主要分支有马氏过程（特点是“过去只影响现在而不直接影响将来”，因马尔可夫首先研究而叫马氏过程）、平稳过程（特点是“在长度相同的两段时间里总有相同的统计规律”）、点过程（特点是“过程在每段时间中只取非负整数值”）、鞅论，等等。



2. 在工农业生产和科学试验中, 广泛存在大批产品质量的估计、检验问题, 在生产过程中控制产品质量的问题以及试验设计问题等。这些问题都属于数理统计的应用范围, 在前面“三”中已作了大略介绍。

3. 测量是认识客观事物数量的基本手段, 而测量误差在生产和科学技术的各种领域是普遍存在的。因此, 研究误差的规律性有重大的实际意义。“误差理论”的基础就是概率论。

4. 某地年降水量、年最高气温、年平均气温、七月份降水量等等都是有统计相关性的随机变数, 而随时间变化的气象状态则是随机过程。用统计方法研究气象问题产生了“气象统计学”这一边缘学科。用气象统计方法进行气象预报叫做“统计预报”。统计预报在中、长期气象预报中作用卓著; 我国数学工作者在地震预测中应用统计预报方法也取得了较好的效果。

5. 水文学中有许多随机现象。一条河的年流量、年最高水位, 一个水库的年最大储水量等都是随机变数。在水文学中广泛应用统计方法形成了“水文统计学”。

6. 生物学和医学中有大量的随机现象。在生物学、医学中应用统计方法, 形成了“生物统计”和“医用统计”两个边缘学科。现代遗传学和基因工程理论中也广泛应用统计方法。

7. 场论在物理学中有重要地位。引力场、电场、磁场都是场论的基本研究对象。当深入到微观世界和一些复杂运动时就需要考虑随机场。量子力学、统计力学、湍流理论中都提出随机场的问题。就是说, 场中的每个点不再是一个确