概率论与数理统计

王林书 鲍兰平 赵瑞清 编著

科学出版社

概率论与数理统计

王林书　鲍兰平　赵瑞清　编著

科学出版社

1999

内 容 简 介

本书是工科本科概率论与数理统计课程的教材. 主要内容包括:随机事件及其概率、随机变量及其分布、随机变量的数字特征、估计、假设检验、方差分析、回归分析、多元分析、贝叶斯统计. 本书节末附有习题.

本书可供工科大学学生、教师阅读.

图书在版编目(CIP)数据

概率论与数理统计/王林书等编著. -北京:科学出版社,1999

ISBN 7-03-007386-X

Ⅰ. 概… Ⅱ. 王… Ⅲ. ①概率论-高等学校-教材 ②数理统计-高等学校-教材 Ⅳ. O21

中国版本图书馆 CIP 数据核字(1999)第 07253 号

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

新蕾印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1999年8月第 一 版 开本:850×1168 1/32

1999年8月第一次印刷 印张:9 1/2

印数:1—6 000 字数:246 000

定价:15.00元

(如有印装质量问题,我社负责调换(北燕))

序

概率统计教本，包括国人的著作和国外介绍过来的译作，已出版了不少．这门课程的基本内容，要算是已比较定型的了．正因为如此，也就不易写出自己的特点．可是，一本没有自己特点的著作，就失掉了其存在的意义．细读王林书等同志所著的这本教材，欣喜地发现作者们确实尽力表现了自己的特点，不失为一本优秀的教材．

一是结构体系上，既遵循了世界上大多数同类教材所使用的科学体系，又在讲法上有许多创新，许多地方更为合理、好懂．例如，在讲随机事件的概率时，把概率论的公理分成五条列出，陈述了古典概型、几何概率、统计概率的具体特性，这种将抽象的公理与几种普遍模型结合起来的讲法更易于为读者所接受．

二是内容上题材丰富，深、广度适中．本教材既对本学科的基本内容做了全面、系统的论述，又在最后两章为有兴趣的或有特殊需要的读者提供了课余自学的材料．本书既有广泛的军事题材，又注意了军、地两全，兼容并包．在对基本内容如参数估计、检验、方差分析与回归分析等的论述中，作者都以不大的篇幅，对基本原理做了清楚的交代，又简明扼要地归纳出了便于掌握的操作方法，既保证了教材的科学性，又不流于纯理论的介绍，并有许多独到之处．

三是例题和习题上很有特色．本教材既收集、精选了中外教材中的优秀题目，又不满足于借鉴抄袭，且花大力气自编了大量的有趣味性、有吸引力的例题、习题，使其能紧扣课文中的原理的阐述、方法的讲解．现行教科书中的例题、习题不少有这样一个缺点，即多数是通过机械地套公式去做，缺乏启发性，也不能通过观摩这些例题或做习题，对课文中的公式、定理的用法得到深一层的训练．

本教材克服了这一缺点,这是本教材优于同类书的很显著的一点.

四是文字上,行文严谨,条理清楚,深入浅出,语言流畅,既能联系读者思想实际,步步诱导,又能联系现实客观背景,层层剖析.

总之,这是一本科学性、实用性、时代感都很强,并有许多创新的高水平教材.我们相信,今后在使用者的帮助下,本教材会进一步得到锤炼而成为更高水平的精品.

陈希孺

1999年1月8日

前　　言

本书是工科本科概率论与数理统计课程的教材. 在编写过程中,我们注意到并在一定程度上做到了以下几点:

1. 既遵循已公认的讲授模式与规矩, 充分吸收借鉴他人的经验与成果, 又大胆而谨慎地进行新探索, 力争有所创新, 有所发展.

2. 根据多年讲授过程中发现的学生们学习中的疑难、症结和规律,在结构上作了必要的调整,在内容上做了适当的取舍,在叙述上也做了有针对性的伸缩.

3. 为了更好地配合基本原理的讲述,激发学生学习的兴趣,在本书中精心编入了部分例题和习题. 其中,自编例题 7 个(占 7%),习题 79 个(占 28%). 均以" * "号标出.

4. 既有一定的军队特色,又有较强的通用性,可为军、地通用.

5. 面适当广些,内容尽量丰富些. 既可供教师选讲,又有利于学生课外阅读.

本书在编写中承蒙北京大学数学学院副院长陈家鼎教授的指导和帮助,初稿完成后他又与中国科学院应用数学所前常务副所长、中国数学会前秘书长王寿仁研究员,中国统计学会副会长、中国现场统计研究会理事长、中国科学院院士陈希孺教授,中国工程概率统计学会副会长、河北工业大学刘文教授等分别仔细审阅了全稿,提出了许多宝贵的意见和建议,谨此表示衷心感谢.

此外,本书的编写是作为中国人民解放军军械工程学院教材精品建设之一完成的,所以自始至终得到了学院各级领导及有关部门的大力支持和本教研室长期从事概率论与数理统计教学的同志们的许多帮助,在此一并致谢.

由于我们的水平所限，本书难免有缺点和错误，欢迎读者批评、指正.

编者

1997年10月

目　录

绪　论

一、概率论与数理统计的研究内容

自然界和人类社会中存在着大量的随机现象.例如,某地未来某时的天气可能是晴、多云、阴或雨,掷硬币可能掷出正面或反面,我国国民经济达到发达国家的水平所需的时间或长或短,等等.研究随机现象的规律性的理论构成了随机理论,其中包括概率论和数理统计两部分.

随机理论研究的对象是随机变量,主要研究随机变量随机取值的规律性.对于多个(有限个或无限个)随机变量,还要研究变量之间在随机取值的规律性方面的依赖关系,而不象传统数学那样只研究变量之间的对应(函数)关系.正如传统数学中整个变量部分构成函数论一样,整个随机变量部分构成了随机理论.

概率论是整个随机理论的基础,首先研究随机现象最基本的规律性,其次给出刻划随机变量的方法,然后在某些假设下进一步研究随机变量的各种规律性.数理统计则是通过观测试验数据,根据建筑在概率论基础上的原理进行推断或预测.这里说的试验是广义的,除了真正的有计划的试验外,还包括对某些并非出自试验目的活动,客观现实的存在及大自然的运行等的观测.数理统计研究的是怎样有效地收集、整理和分析带有随机性的数据以及对考察的现象做出推断或预测,直至为采取一定的决策和行动提供科学的依据和建议.

概率不仅是概率论的研究内容,也是整个概率论及数理统计最基本的工具和最基本的语词.

二、概率论与数理统计的发展历史

概率论起源于16世纪前后,成型于17世纪中叶,从博弈问题

中萌生，受最大影响于保险行业.它的发展经历了三个时期.

17世纪到18世纪末为古典概率时期.16世纪前后，赌博的盛行为研究概率问题提供了优良的模型.相当多的数学家对赌博中的问题(不是赌博本身)产生了浓厚的兴趣，因为他们注意到"其中实际上包含着很有趣、很深刻的理论基础"(惠更斯(Huygens)).法国的帕斯卡(Pascal)、费马(Fermat)、惠更斯等人研究了大量的赌金分配问题(其中最著名的是:甲、乙两人相约赌若干局，谁先赢 s 局就可得全部赌金，现中途终断，两人分别赢了 a 局和 b 局——$a,b<s$，怎样按比例分配赌金?)，逐步建立了概率、条件概率、数学期望等重要概念及其基本性质.1657年出版的惠更斯的《论赌博中的计算》是概率论发展史上的第一部专著，是概率论产生的标志之一.18世纪，由于瑞士的伯努利(Bernoulli)的奠基性工作，使概率论逐步形成了一个独立的数学分支.1713年出版的他的巨著《猜度术》(或译《推论法》)是概率论发展史上的影响深远的名著之一.伯努利是公认的概率论的主要创始人.

在这段时间里，保险行业、人口统计、大地测量、射击试验、天文观测等社会活动及自然科学既向数学提出了一系列急待解决的问题，又为概率论的产生与发展提供了可贵的素材和丰富的营养.这个时期里，概率论以计算各种古典概率为中心，以组合论为主要工具，所以这个时期也称组合概率时期.

从18世纪末到20世纪40年代为分析概率时期.整个18世纪与19世纪上半叶，概率论发展到了令人着迷的程度，一大批科学家争先恐后地跻身于这一领域.一个个强有力的新方法得到应用，尤其是把成熟的分析方法引进概率论，使其进入了一个新时期.1812年出版的拉普拉斯的著作《分析概率论》是进入这一新时期的标志.这个时期里，研究中心变成随机变量，一系列重要分布相继问世.特别是，德国的高斯(Gauss)找到了"正态分布"，他还和法国的勒让德(Legendre)、拉普拉斯(Laplace)建立了"最小二乘理论".伯努利和法国的泊松(Poisson)、德莫弗(De. Miove)、拉普拉斯的开拓性的研究，以特殊情形打开了极限理论的新篇章，这

时的西欧，特别是法国处于领先地位.

19 世纪下半叶到本世纪初的几十年里，由于种种原因，概率论在当时的西欧出现了不幸的停滞. 多亏了俄国以“三夫”为代表的一批数学家拯救了概率论于危难. 切比雪夫（Чебышев）以卓越的天才率先迈出了决定性的一步，接着，马尔科夫（Марков）、李亚普诺夫（Ляпунов）乘胜前进，再造了概率史上的辉煌.

在本世纪 20 年代以后的 20 多年里，除前苏联继续保持发展的势头外，法国、英国、德国及美国等也赶了上来，涌现了一批代表人物. 法国的莱维（Levy）、美国的费勒（Feller）、前苏联的辛钦（Хинчин）、柯尔莫哥洛夫（Колмогоров）等先驱者们在极限理论等方面的杰出贡献，使概率论获得了一次又一次突破性的进展. 所以本世纪的二三十年代，被人称为概率论的英雄时代.

伴随着概率论的深入发展，概率概念本身的刻画也经历了相当长的时期. 从起初的直观描述到以后的几何概率，再到 19 世纪拉普拉斯的古典概率，20 世纪米泽斯（Mises）的统计定义，都在不同范围内发挥了一定的效用，但也都暴露了各自的局限性. 概率概念的准确刻划问题急待解决. 1933 年，柯尔莫哥洛夫成功地建立了概率的公理化体系，终于提出了概率的科学定义. 由于他和伯恩斯坦（Бернштейн）、辛钦、斯卢茨基（Слуцкий）、莱维等的努力，使概率论中的一整套基本概念完全置于集合论、函数论、测度论等观点之下，这又使概率论发生了一次深刻变革，树起了概率史上一个新的里程碑，为现代概率论的迅速发展打下了坚实的基础.

从本世纪 50 年代之后为现代概率论时期. 在现代技术的刺激下，它的理论和应用都有显著的发展. 这时的研究内容更加丰富，特别是随机过程领域在不停地拓宽和加深着，并构成了现代概率论的核心、概率论精英们的创造性努力的主要用武之地，而其中的随机分析和鞅论更是获得了长足的发展，并形成了与传统的确定性分析法相应的随机分析法这一研究工具. 概率论在工程技术和社会经济中的应用日益广泛，出现了理论概率和应用概率比翼齐飞的局面.

数理统计的发展也经历了三个时期. 20 世纪以前为萌芽时期. 由于高斯和勒让德的最小二乘法应用于误差分析，终于使统计摆脱了萌芽时期的描述性岁月，开始了推断统计的新生. 这个时期最重要的发展，首先在于确定了这样一种观点，即数据来自服从一定概率分布的母体，而统计就是用数据去推断这个分布. 比利时的凯特房(Quetelet)第一个将法国古典概率论引入社会统计. 英国的高尔顿(Galton)和 K. 皮尔逊(Pearson)在生物统计的研究中取得了一系列重要进展，如引进了矩估计和频度分布、回归、相关、拟合度等重要概念，极大地丰富了统计内容. 德国的 F. 赫尔梅特(Helmert)发现了十分重要的 χ^2 分布.

20 世纪初到二战结束为数理统计的蓬勃发展到成熟的时期，即近代统计时期. 许多重要的基本观点、方法及主要分支都是在这个时期建立和发展起来的. 这个时期的成就，包含了至今仍在广泛使用的大多数统计方法，并占据了当今教科书的主要篇幅. 在其发展中，以费歇耳(Fisher)为代表、包含 K. 皮尔逊在内的英国学派起了主导作用. 他们构造了数理统计的基本框架并开创了一系列重要领域研究的先河. K. 皮尔逊的 χ^2 检验、戈塞特(Gosset)的 t 分布，费歇耳的极大似然估计、实验设计、方差分析、信任推断，包含 F 分布在内的种种抽样分布，奈曼(Neyman)和 S. E. 皮尔逊的假设检验、瓦尔德(Wald)的序贯概率比检验等都是这时研究出来的. 特别值得我们引以自豪的是，我国著名学者许宝騄先生在多元分析和线性模型的研究中做了奠基性工作，成为世界公认的第一流统计学家.

战后是数理统计在理论上和应用上继续大发展的时期，也即现代统计时期. 旧领域里的研究继续向纵深推进，新领域、新思想不断涌现. 尤其是，贝叶斯(Beyes)学派的崛起使数理统计别开生面，而电子计算机的出现又为数理统计的发展注入了新的活力，并对其广泛应用起了巨大的促进作用.

三、概率论与数理统计的应用

概率论与数理统计是近几十年来发展最快的数学分支，其根源就在于它的应用范围的日益扩大，以及对它的应用的认识的不断觉醒. 1943 年以前，英美运输船队在横穿大西洋运送战略物资时，频频遭受德国 U 型潜艇的攻击，损失惨重. 数学家们用概率统计分析后发现：船只被击沉的危险性大小与船队数目（批次）和船队规模（每队船数）之比成正比. 于是，英美改变了过去由各个港口分散启航的做法，先到指定海域汇合，集体通过危险海区，然后各自驶向预定港口. 结果，击沉的船数大大减少，保证了战略物资的及时供应.

二战期间，盟国数学家们又用概率统计法，由俘获的德军坦克上的号码成功地估计出了德国坦克的总数，为赢得战争的胜利起到了一定的促进作用.

以上是概率统计应用于人类生死关头的两个光辉典范. 而在平时，正如科学巨匠拉普拉斯所说："生活中最重要的问题，其中绝大多数在实质上只是概率问题."如今概率统计不仅广泛地渗透到各个学科领域，而且大量地应用于各个生产和工作部门. 它不仅渗透到许多自然科学和技术科学领域，而且渗透到一些社会科学领域，其中最突出的是经济学和军事学. 它应用于许多老学科，成功地解决了该学科的一系列重大课题，并与之相结合产生了一系列边缘学科，如统计物理、生物统计、医学统计、气象统计、地质统计……. 它还是许多新学科的基础，如信息论、控制论、弹道学、信号与系统、人工智能等. 所以，对于这些学科，掌握概率统计不仅是深入研究所必须而且是学习入门的首要条件.

除此之外，随着大规模生产企业的发展，概率统计不但用来进行产品质量的科学检验，而且更重要的是用于生产过程本身的组织上. 所以，在一些生产工作部门，尤其是质量控制和可靠性部门，概率统计中的名词术语简直成了他们的工作语言.

大哉，概率统计之为用！

第一章 随机事件的概率

第一节 随机事件及其概率

一、随机事件

在一定条件下可能发生也可能不发生的事件称为随机事件，通常用 $A,B,\cdots$ 等字母表示. 比如打靶时，“命中”、“命不中”都是随机事件；掷骰子时，“掷出 2”、“掷出奇数”等也都是随机事件. 实际上，做随机试验时，可能出现种种结果，而每一种可能出现的结果在试验结束前都是随机事件. 随机事件往往可以而且需要再分解成若干个“小”的随机事件，如“掷出奇数”可分解成“掷出 1”、“掷出 3”、“掷出 5”等. 不能或不必再分解的随机事件叫基本事件或样本点. 从集合论的角度看，基本事件或样本点相当于元素，而随机事件是基本事件或样本点的集合. 基本事件还可看作单元素的集合，所以基本事件又是特殊的随机事件.

一个随机事件发生与否全在于它所包含的诸基本事件中是否有发生者. 所有基本事件（样本点）的集合，构成必然要发生的事件，叫必然事件或样本空间，用 Ω 表示. 而不可能发生的事件，叫不可能事件，它是基本事件（样本点）的空集合，用 $\varnothing$ 表示.

二、随机事件的概率

概率是随机事件的随机性的一种度量.

随机事件 A 的概率是一个非负数，表示 A 发生的可能性的大小，即在所有可能发生的结果中，A 发生的机遇所占的“比例”，用 $P(A)$ 表示. P 是“概率”的英文（probability）字头，A 是“A 发生”的省略（以下常这样用）.

概率满足以下公理.

1. $0\leqslant P(A)\leqslant 1, P(\varnothing)=0, P(\Omega)=1.$ (1.1)

2. (古典概率)如果基本事件的总数有限,设为 n(或 n_Ω),每个基本事件的发生都是等可能的,事件 A 中有 n_A 个基本事件,则事件 A 的概率

$$P(A)=\frac{n_A}{n} \tag{1.2}$$

比如,在掷骰子时,若 A 表示掷出奇数,则 $P(A)=\frac{3}{6}=\frac{1}{2}$. 再如,从一副扑克牌中任取一张,$A$ 表示取到黑桃之事件,则 $P(A)=\frac{13}{54}$. 若任取两张,B 表示取的两张都是黑桃之事件,则 $P(B)=\frac{C_{13}^2}{C_{54}^2}$.

当基本事件有无限多时,若能合并成 n 个等可能的随机事件,A 中含有 n_A 个这种随机事件,则(1.2)式仍可用.

比如,在全体自然数中任取一个,则 P(取到被 5 整除的数)$=\frac{1}{5}$, P(取到被 2 或 3 整除的数)$=\frac{4}{6}=\frac{2}{3}$. (把自然数分成 6 部分,被 6 除分别余 0,1,2,…,5,其中余 0,2,3,4 者都可被 2 或 3 整除)

3. (几何概率)若 Ω 为某几何空间中的一个区域,在 Ω 中任何度量相等的区域里发生的可能性都相等,$A\subset\Omega$,则

$$P(A)=\frac{A\text{ 的度量}}{\Omega\text{ 的度量}} \tag{1.3}$$

比如,在平面坐标系 xoy 中的 $|x|\leqslant 1$, $|y|\leqslant 1$ 的正方形内各点等可能地投针,A 表示落入单位圆内,则 $P(A)=\frac{\pi}{4}$.

再如,某汽车站上,每隔 5 分钟开来一趟汽车,一人在一站上等候上车,A 表示他在 2 分钟内能上车,则 $P(A)=\frac{2}{5}$.

4. (统计概率)进行 m 次独立试验,若 A 发生 m_A 次(即 A 的

频数为 m_A)，则 A 的频率 $f_m(A)\triangleq\dfrac{m_A}{m}$，当 m 充分大时，将在 $P(A)$ 附近稳定摆动，m 越大这种现象越明显. 所以 $P(A)$就是 $f_n(A)$在 m 非常大时的稳定值.

传统的教科书中，通常把上面三种概率分别称为概率的古典定义、几何定义和统计定义.

5. 如果 $A_1,A_2,\cdots$是一列随机事件，其中任何两个都不同时发生，则

$$P(\{A_n\}\text{中至少有任意一个发生})=\sum_{n=1}^{\infty}P(A_n) \qquad (1.4)$$

比如，在射击试验时，直到命中为止，则

P(射击 3 次以上)$=P$(射击 3 次)$+P$(射击 4 次)$+\cdots\cdots$

三、(古典概率及几何概率)举例

古典概率计算中的大量实例属于排列组合问题. 下面的前 4 例给出了用排列组合计算的几种(不是全部)古典概率中的分子分母形式. A_m^n(或 P_m^n)和 C_m^n(或$\binom{m}{n}$))分别表示从 m 个不同元素中取 n 个元素时的所有不同的排列数和组合数，$P_n=A_n^n$(或 P_n^n).

[**例 1**](任意不重复排列数/任意可重复排列数)　从标号为 1,2,3,4,5 的五张卡片中抽取三次，每次任意抽取一张，取后放回，第一、二、三次取得的标号依次为百位、十位、个位上的数字. 求这样得到的三位数没有重复数字的概率.

解　因为每张卡片可以重复抽取，所以基本事件的总数为 5^3. 设 A 表示得到的三位数没有重复数字之事件，则 A 所包含的基本事件数为从这五个不同元素中任意取三个元素的选排列数. 即为 A_5^3. 因此，所求概率为

$$P(A)=A_5^3/5^3=60/125=0.48$$

[**例 2**](特定的不重复排列数/任意不重复排列数)　把 a,b,c,d,e 五个字母任意排列，求字母 a 和 b 排在一起的概率.

解　设事件 A 表示字母 a 和 b 排在一起. 基本事件的总数为五个不同字母的全排列数，即为 P_5. 再看事件 A 所包含的基本事