教育部高等农林院校理科基础课程
教学指导委员会推荐示范教材

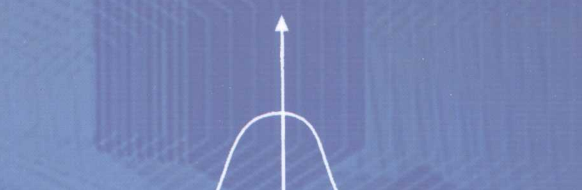
概率论与数理统计

S

吴 坚 张录达 主编

Probability and Statistics

Probability and Statistics
Probability and Statistics



中国农业大学出版社
ZHONGGUONONGYEDAXUE CHUBANSHE

教育部高等农林院校理科基础课程

教学指导委员会推荐示范教材

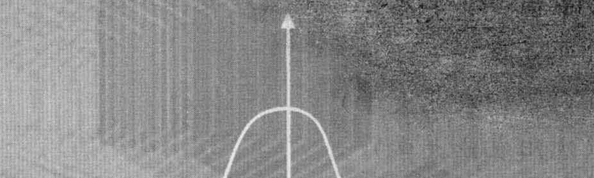
概率论与数理统计

吴坚 张录达 主编

Probability and Statistics

Probability and Statistics

Probability and Statistics



中国农业大学出版社

ZHONGGUONONGYEDAXUE CHUBANSHE

图书在版编目(CIP)数据

概率论与数理统计/吴坚,张录达主编. —北京:中国农业大学出版社,2009,12
ISBN 978-7-81117-938-5

Ⅰ.①概… Ⅱ.①吴…②张… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材
Ⅳ.①O21

中国版本图书馆 CIP 数据核字(2009)第 224500 号

书　　名 概率论与数理统计

作　　者 吴　坚　张录达　主编

策划编辑	张秀环　董夫才	**责任编辑**	洪重光
封面设计	郑　川		
出版发行	中国农业大学出版社		
社　　址	北京市海淀区圆明园西路 2 号	**邮政编码**	100193
电　　话	发行部 010-62731190,2620	读者服务部	010-62732336
	编辑部 010-62732617,2618	出　版　部	010-62733440
网　　址	http://www.cau.edu.cn/caup	**e-mail**	cbsszs @ cau.edu.cn
经　　销	新华书店		
印　　刷	北京时代华都印刷有限公司		
版　　次	2009 年 12 月第 1 版　　2009 年 12 月第 1 次印刷		
规　　格	787×1 092　　16 开本　　19 印张　　430 千字		
定　　价	29.00 元		

主　编　吴　坚（安徽农业大学）
张录达（中国农业大学）

副主编　徐凤琴（北京林业大学）
张长勤（安徽农业大学）
姚贵平（内蒙古农业大学）
吴清太（南京农业大学）
鲁春铭（沈阳农业大学）
吕金凤（河北科技师范学院）
左振钊（河北北方学院）

编　者　（按姓氏笔画排列）
左振钊（河北北方学院）
许海洋（青岛农业大学）
孙　燕（内蒙古民族大学）
吕金凤（河北科技师范学院）
吴　坚（安徽农业大学）
吴清太（南京农业大学）
张长勤（安徽农业大学）
张好治（青岛农业大学）
张录达（中国农业大学）
杨晓霞（北京林业大学）
姚贵平（内蒙古农业大学）
赵培玉（沈阳农业大学）
徐凤琴（北京林业大学）
高瑞平（河北科技师范学院）
鲁春铭（沈阳农业大学）

教育部高等农林院校理科基础课程教学指导委员会
推荐示范教材编审指导委员会

主　任　江树人

副主任　杜忠复　程备久

委　员（以姓氏笔画为序）

王来生　王国栋　方炎明　李宝华　张文杰　张良云
杨婉身　吴　坚　林家栋　陈长水　周训芳　周志强
高孟宁　戚大伟　梁保松　曹　阳　焦群英　傅承新

教育部高等农林院校理科基础课程教学指导委员会
推荐数学类示范教材编审指导委员会

主　任　高孟宁

委　员（以姓氏笔画为序）

王来生　石　峰　卢恩双　吴　坚　杜忠复　张良云
杜晓林　孟　军　房少梅　梁保松　惠淑荣

出版说明

在教育部高教司农林医药处的关怀指导下，由教育部高等农林院校理科基础课程教学指导委员会(以下简称“基础课教指委”)推荐的本科农林类专业数学、物理、化学基础课程系列示范性教材现在与广大师生见面了.这是近些年全国高等农林院校为贯彻落实“质量工程”有关精神，广大一线教师深化改革，积极探索加强基础、注重应用、提高能力、培养高素质本科人才的立项研究成果，是具体体现“基础课教指委”组织编制的相关课程教学基本要求的物化成果.其目的在于引导深化高等农林教育教学改革，推动各农林院校紧密联系教学实际和培养人才需求，创建具有特色的数理化精品课程和精品教材，大力提高教学质量.

课程教学基本要求是高等学校制定相应课程教学计划和教学大纲的基本依据，也是规范教学和检查教学质量的依据，同时还是编写课程教材的依据.“基础课教指委”在教育部高教司农林医药处的统一部署下，经过批准立项，于 2007 年年底开始组织农林院校有关数学、物理、化学基础课程专家成立专题研究组，研究编制农林类专业相关基础课程的教学基本要求，经过多次研讨和广泛征求全国农林院校一线教师意见，于 2009 年 4 月完成教学基本要求的编制工作，由“基础课教指委”审定并报教育部农林医药处审批.

为了配合农林类专业数理化基础课程教学基本要求的试行，“基础课教指委”统一规划了名为“教育部高等农林院校理科基础课程教学指导委员会推荐示范教材”(以下简称“推荐示范教材”).“推荐示范教材”由“基础课教指委”统一组织编写出版，不仅确保教材的高质量，同时也使其具有比较鲜明的特色.

一、“推荐示范教材”与教学基本要求并行　教育部专门立项研究制定农林类专业理科基础课程教学基本要求，旨在总结农林类专业理科基础课程教育教学改革经验，规范农林类专业理科基础课程教学工作，全面提高教育教学质量.此次农林类专业数理化基础课程教学基本要求的研制，是迄今为止参与院校和教师最多、研讨最为深入、时间最长的一次教学研讨过程，使教学基本要求的制定具有扎实的基础，使其具有很强的针对性和指导性.通过“推荐示范教材”的使用推动教学基本要求的试行，既体现了“基础课教指委”对推行教学基本要求的决心，又体现了对“推荐示范教材”的重视.

二、规范课程教学与突出农林特色兼备 长期以来各高等农林院校数理化基础课程在教学计划安排和教学内容上存在着较大的趋同性和盲目性，课程定位不准，教学不够规范，必须科学地制定课程教学基本要求. 同时由于农林学科的特点和专业培养目标、培养规格的不同，对相关数理化基础课程要求必须突出农林类专业特色. 这次编制的相关课程教学基本要求最大限度地体现了各校在此方面的探索成果，“推荐示范教材”比较充分反映了农林类专业教学改革的新成果.

三、教材内容拓展与考研统一要求接轨 2008 年教育部实行了农学门类硕士研究生统一入学考试制度. 这一制度的实行，促使农林类专业理科基础课程教学要求作必要的调整. “推荐示范教材”充分考虑了这一点，各门相关课程教材在内容上和深度上都密切配合这一考试制度的实行.

四、多种辅助教材与课程基本教材相配 为便于导教导学导考，我们以提供整体解决方案的模式，不仅提供课程主教材，还将逐步提供教学辅导书和教学课件等辅助教材，以丰富的教学资源充分满足教师和学生的需求，提高教学效果.

乘着即将编制国家级“十二五”规划教材建设项目之机，“基础课教指委”计划将“推荐示范教材”整体运行，以教材的高质量和新型高效的运行模式，力推本套教材列入“十二五”国家级规划教材项目.

“推荐示范教材”的编写和出版是一种尝试，赢得了许多院校和老师的参与和支持. 在此，我们衷心地感谢积极参与的广大教师，同时真诚地希望有更多的读者参与到“推荐示范教材”的进一步建设中，为推进农林类专业理科基础课程教学改革，培养适应经济社会发展需要的基础扎实、能力强、素质高的专门人才做出更大贡献.

中国农业大学出版社

2009 年 8 月

内容提要

本书是教育部高等农林院校理科基础课程教学指导委员会推荐示范教材，是教育部教学研究立项项目成果．本教材突出随机数学思想，注重概率论与数理统计的通用知识和应用性，讲授的内容包括：随机事件与概率、条件概率与独立性、一维随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、大数定律和中心极限定理、数理统计的一些基本概念、参数估计、假设检验和方差分析与回归分析．

本书可作为高等农林院校农林类各专业的本科生教材，也可作为非数学类各专业该课程的参考教材以及科技人员的参考用书．

前言

本书是根据教育部高等农林院校理科基础课程教学指导委员会(以下简称"基础课教指委")2008 年 11 月北京会议的精神,按照"基础课教指委"2009 年 4 月组织讨论制定的《普通高等学校农林类专业数理化基础课程教学基本要求》编写的.本书是"基础课教指委"首次在全国高等农林院校中推荐使用的示范教材,也是教育部组织的教学研究立项课题成果.

本书适用学时为 48～56 学时.为了该课程教学内容的系统性和高等农林院校的教学实际,也为了适应读者对该课程更多知识的需求和自学需要,本书适当增加了若干带有"*"号的章节,可供教师在教学中选用和读者选学.

概率论与数理统计是高等农林院校本科数学教育中的一门主要课程,它的理论和方法是学习和从事其他学科研究的重要基础,并在农林、经济、管理、金融、工程技术等诸多学科领域中有着广泛的应用.本课程属于随机数学范畴,讲授研究随机现象规律性的概率论基础知识和以处理统计试验数据为主的数理统计基本理论和方法.虽然该课程着眼于应用,但对于该课程的基本概念、基本理论和基本方法的认识也很重要,在讲授和学习中应力求做到理论与实际的结合.本书的例题与习题较为丰富,教师和学生可适当选用.

本书由吴坚教授(安徽农业大学)、张录达教授(中国农业大学)担任主编,徐凤琴(北京林业大学)、张长勤(安徽农业大学)、姚贵平(内蒙古农业大学)、吴清太(南京农业大学)、鲁春铭(沈阳农业大学)、吕金凤(河北科技师范学院)和左振钊(河北北方学院)担任副主编,参与该教材编写的其他人员有许海洋(青岛农业大学)、孙燕(内蒙古民族大学)、张好治(青岛农业大学)、杨晓霞(北京林业大学)、高瑞平(河北科技师范学院)和赵培玉(沈阳农业大学).

全书由安徽农业大学吴坚教授统稿.

编者十分感谢"基础课教指委"和中国农业大学出版社在本教材出版方面给予的大力支持,并感谢安徽农业大学研究生王洁、史婕在原稿打字方面所做的工作.

囿于学识,书中难免存在不妥之处,期待广大读者和教师指正.

编者

2009 年 11 月

目录 CONTENTS

Chapter 1 第 1 章

随机事件与概率

Random Event and Probability

本章主要介绍概率论的研究对象——随机事件，概率的统计定义与公理化定义，古典概型、几何概型以及概率的基本性质与简单计算.

1.1 概率论的现实背景

在自然界和各项科学研究中，有许多现象的规律是以下列形式表达的："只要条件 C 实现，则事件 A 必然会发生(或必然不发生)."例如："如果平面图形是三角形(条件 C 实现)，那么这个图形的内角和是 $180°$(事件 A 必然发生)"；又如"纯种紫花碗豆的后代开紫花"这一现象是一定会发生的. 而"在 101 325 Pa 的大气压下，水加热至 100 ℃不沸腾"和"同性电荷相互吸引"等现象是必然不会发生的. 以下就称在条件 C 下必然发生的事件 A 为条件 C 下的必然事件(或简称必然事件)，称在条件 C 下必然不发生的事件为条件 C 下的不可能事件(或简称不可能事件). 必然事件的反面是不可能事件，而不可能事件的反面就是必然事件. 它们虽然形式相反，但两者的实质是相同的. 所有这类现象我们称之为**确定性现象**(**或决定性现象、必然现象**)，它们广泛地存在于自然和社会现象中，许多数学分支研究的是确定性现象的数量规律.

除确定性现象外，在自然界和社会现象中还大量存在与确定性现象有着本质区别的另一类现象——随机现象：在条件 C 下事件 A 可能发生也可能不发生. 它是事先不能预知结果的，即在相同条件下重复观察或进行试验时，每次得到的结果未必相同，或即使知道它过去的状态，也不能肯定它将来的发展状态. 例如某地今年冬暖而导致明年小麦赤霉病流行严重等现象. 我们以后就将在条件 C 下事件 A 可能发生也可能不发生的事件称为条件 C 下的随机事件(或简称为随机事件). 下面我们举一些例子.

例 1.1.1 从某工厂生产的某种产品中任意抽取的一件产品可能是合格品，也可能是不合格品.

在此例中，条件 C 是"从某工厂生产的某种产品中任意抽取一件产品"，事件 A 是"抽取

的是合格品"，显然 A 是条件 C 下的随机事件.

许多实际问题与此例问题同类。例如，重复抛掷一枚质地均匀的硬币国徽面的朝上（约定为出现"正面"）或朝下（约定为出现"反面"）；种子发芽试验中一粒种子的"发芽"或"不发芽"；一次射击的"中"或"不中"等现象.

同样，将"从某工厂的某种产品中任意抽取 n 件产品"看作条件 C，则"其中恰有 $k(0\leqslant k\leqslant n)$ 件合格品"是在条件 C 下的随机事件.

例 1.1.2 对于一个服务系统来说，有多少顾客要求服务是具有偶然性的，因此"在时间间隔 $(a,a+t]$ 内有 k 个顾客到来"是"该服务机构在时间 $(a,a+t]$ 内工作"条件下的随机事件.

例 1.1.3 射击中，"弹着点距目标的偏差为 x"及"偏差在 x_1 与 x_2 之间"都是"进行一次射击"条件下的随机事件；如果在目标所在的平面上建立坐标系（不妨设目标位于坐标原点），则"弹着点 (ξ,η) 位于区域 $\{(x,y)\mid x_1\leqslant x\leqslant x_2, y_1\leqslant y\leqslant y_2\}$ 内"是随机事件.

例 1.1.4 植物病害流行研究中，以气传方式从菌源中心传播出的"一个病害孢子在时刻 t 的位置 (ξ,η,ζ) 位于区域

$$\{(x,y,z)\mid x_1\leqslant x\leqslant x_2, y_1\leqslant y\leqslant y_2, z_1\leqslant z\leqslant z_2\}$$

内"是随机事件.

例 1.1.5 种群动态研究中"某个野外动物种群进入某个划定区域的数量为 k"是随机事件.

总之，以上所列举现象的一个共同特点是：在基本条件不变的情况下，一系列的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，它会时而发生这种结果，时而发生那种结果，呈现出一种偶然性，这种现象就是随机现象. 在随机现象中通常关心的是在试验或观察中某个结果是否发生，这些结果就是随机事件. 以后我们用大写字母 $A,B,C,\cdots$ 表示随机事件.

另外，我们注意到随机事件的这种不确定性，不是事件本身不明确，而是发生的条件不充分，使得在条件与事件之间不能出现确定性的因果关系，从而在事件的发生与否上表现出不确定的性质，这种不确定性就是**随机性**. 在客观世界中还存在另一些不确定现象，例如**模糊现象**，它所反映的事物在概念本身就是模糊的（即一个对象是否符合这个概念也难以确定，也就是由于外延模糊而带来的不确定性）. 例如天气预报中的"多云"，医疗诊断中的"食欲不振"、"头痛"，日常生活中的"美与丑"、"冷与热"等. 这种由客观事物差异的中间过渡的"不分明性"而产生的不确定性人们称之为模糊性. 当然，客观世界中还有一些其他不确定现象，有些现象还可以是几种不确定性共存的情况，这些都是非确定性现象，而概率论研究和处理的对象是随机现象①.

如何去研究随机事件？一种自然的想法就是像我们研究必然事件那样，去进一步寻求随机事件发生的条件. 但是只要对上面举出的一些例子分析一下，就会发现这种做法几乎是不可能的，而且是不必要的. 例如对于例 1.1.2，要想事先找出"在时间 $(a,a+t]$ 内有 k 个顾

① 近些年来，一些学者将概率论方法适当改变用于人工智能、决策、不确定性推理等方面，取得了不少成果.

客到来”这一随机事件发生的条件是不可能的. 另外，服务机构所关心的是大致有多少顾客，从而可以设置相适应的服务机构，因而寻求事件确切发生的条件也是不必要的. 这就产生了这个随机事件出现的“可能性”的另一自然想法.

再如，人们对于工厂生产好坏的评判标准之一是它的产品合格率，即产品中合格品的个数与产品总数的比值；对一种药物的疗效的评判标准之一用治愈率，即使用过该药的患者中治愈人数与总人数之比；对于射手的射击技术的评判标准之一是他的命中率，即射击命中次数与射击总数之比. 合格率、治愈率及命中率都是人们所关心的随机事件发生的可能性的一个数量化描述.

从这些例子可以概括出频数与频率的概念如下：若在条件 C 的 n 次实现下，事件 A 发生的次数是 n_A，称它为 A 在条件 C 的 n 次实现下的频数；而称 $f_n(A)=\frac{n_A}{n}$ 为事件 A 在条件 C 的 n 次实现下的频率(简称频率)，它反映了事件 A 在条件 C 下发生的可能性. 当我们进一步考察 A 的频率 $f_n(A)$ 时，发现 $f_n(A)$ 对于条件 C 的不同 n 次实现一般来说是不同的. 但大量的实践表明，对于固定的随机事件 A 来说，当 n(条件 C 实现的次数)较大时，$f_n(A)$ 有经常接近于一个常数的趋势，并且当 n 越大时，接近的程度也就越显著，接近的次数也越经常，即在大量的试验或观察下呈现出明显的规律性——频率稳定性.

例如，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. 但若硬币均匀，直观上出现正面与出现反面的机会应该相等，即“出现正面”的可能性应为 1/2. 下面让我们进行抛掷硬币的试验来验证这种频率稳定性.

将一枚均匀硬币掷 50 次，结果如表 1.1.1 所示. 第一列是抛掷的次数，记作 n，第二列是抛掷的结果，记作 H(正面向上)和 T(反面向上). 到任何指定的一次抛掷，出现 H 的总数列在第三列，并记作 n_H，最后，出现 H 的频率，用 $f_n(H)=\frac{n_H}{n}$ 表示，列在第四列. $f_n(H)$ 的图形如图 1.1.1 所示.

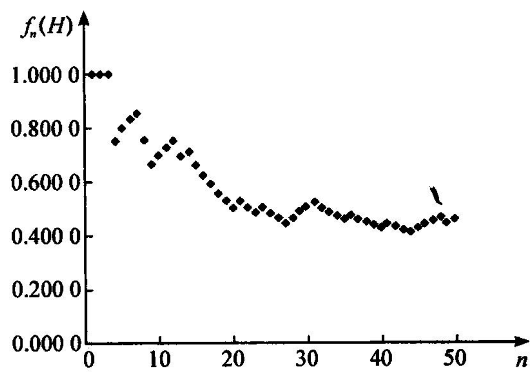


图 1.1.1

此实验已由多人做过，历史上，著名统计学家蒲丰(Comte de Buffon)和皮尔逊(Karl Pearson)曾进行过大量抛掷硬币的试验，所得结果如表 1.1.2 所示.

从这些结果看，“出现正面”的频率总是在 1/2 附近波动，且抛掷次数越多，一般来说越接近于 1/2，这个常数 1/2 客观地反映了“出现正面”这一随机事件发生的可能性大小. 读者可以思考，H 出现的频率是如何变得稳定并接近于 1/2 的.

同样，若多次测量同一个物体，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边的分布呈现某种对称性. 其他的随机现象，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.

表 1.1.1

n	H 或 T	n_H	$f_n(H)$	n	H 或 T	n_H	$f_n(H)$
1	H	1	1.000 0	26	T	12	0.461 5
2	H	2	1.000 0	27	T	12	0.444 4
3	H	3	1.000 0	28	H	13	0.464 3
4	T	3	0.750 0	29	T	14	0.482 8
5	H	4	0.800 0	30	H	15	0.500 0
6	H	5	0.833 3	31	H	16	0.516 1
7	H	6	0.857 1	32	H	16	0.500 0
8	T	6	0.750 0	33	T	16	0.484 8
9	T	6	0.666 7	34	T	16	0.470 6
10	H	7	0.700 0	35	T	16	0.457 1
11	H	8	0.727 3	36	H	17	0.472 2
12	H	9	0.750 0	37	T	17	0.459 5
13	T	9	0.692 3	38	T	17	0.447 4
14	H	10	0.714 3	39	T	17	0.435 9
15	T	10	0.666 7	40	T	17	0.425 0
16	T	10	0.625 0	41	H	18	0.439 0
17	T	10	0.588 2	42	T	18	0.428 6
18	T	10	0.555 6	43	T	18	0.418 6
19	T	10	0.526 3	44	T	18	0.409 1
20	T	10	0.500 0	45	H	19	0.422 2
21	H	11	0.523 8	46	H	20	0.434 8
22	T	11	0.500 0	47	H	21	0.446 8
23	T	11	0.478 3	48	H	22	0.458 3
24	H	12	0.500 0	49	T	22	0.449 0
25	T	12	0.480 0	50	H	23	0.460 0

表 1.1.2

试验者	掷硬币次数	“出现正面”次数	频率 $f_n(H)$
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

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观察或试验中随机事件发生的频率的稳定性，即一个随机事件发生的频率常在某个定值附近摆动，而且，试验次数越多，一般摆动越小. 这种规律性我们称之为统计规律性. 频率稳定的这种统计规律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，是不随人们意志而改变的一种客观属性，因此可以对它度量.

在寻求随机现象的规律时，常常采用模拟的方法. 例如，代替掷一个均匀骰子，我们可以从一个恰好装有 6 个球（这 6 个球的大小、形状完全相同，分别标上 1，2，3，4，5，6 的号码）的袋中任意摸一个球. 类似地，代替从一副纸牌中摸出一张牌，我们可以从一个装有