天元研究生数学丛书

高等概率论

程士宏 主编

北京大学出版社

O211
C75

天元研究生数学丛书

高等概率论

程士宏　编著

北京大学出版社
北　京

EA01 / 15

图书在版编目(CIP)数据

高等概率论/程士宏编著.—北京:北京大学出版社,1996.12
(天元研究生数学丛书)
ISBN 7-301-03318-4

Ⅰ.高…　Ⅱ.程…　Ⅲ.概率论　Ⅳ.O211

书　　　名:**高等概率论**(天元研究生数学丛书)
著作责任者:程士宏　编著
责 任 编 辑:王明舟
标 准 书 号:ISBN 7-301-03318-4/O·389
出　版　者:北京大学出版社
地　　　址:北京市海淀区中关村北京大学校内　100871
电　　　话:出版部 62752015　发行部 62559712　编辑部 62752032
排　印　者:北京大学印刷厂
发　行　者:北京大学出版社
经　销　者:新华书店
850×1168毫米　32开本　11.875印张　290千字
1996年12月第一版　1996年12月第一次印刷
印　　　数:0001—3,000册
定　　　价:20.00元

《天元研究生数学丛书》编委会

名誉主编：程民德

主　　编：张恭庆

副 主 编：刘绍学

编　　委：(按姓氏笔画为序)

王仁宏　王兴华　仇庆久　龙瑞麟　叶其孝

史树中　冯克勤　刘应明　刘嘉荃　严加安

李邦河　时俭易　吴黎明　张继平　张荫南

陆善镇　陈怀惠　陈恕和　林　伟　郑忠国

贾荣庆　徐明曜　郭懋正　黄玉民　彭家贵

内 容 简 介

本书主要讲授高等概率论的基本理论和方法，特别突出离散鞅的研究成果. 全书共分五章. 内容包括：概率论基础、离散鞅、Wiener 过程、弱收敛理论、强收敛理论等. 本书旨在架设从初等概率论的大学课程到现代概率论研究之间的桥梁，为读者进行深入研究打下坚实的基础. 本书选材精练，说理清楚，推导严谨，用通俗易懂的语言介绍了近代概率论中的研究成果，使读者尽快进入前沿研究领域.

本书可作为大学数学专业高年级本科生和研究生相关选修课的教材或参考书，也可供不专门从事概率统计研究的数学工作者阅读.

此为试读，需要完整PDF请访问：www.ertongbook.com

《天元研究生数学丛书》书目

（标△号者表示待出版）

1. 复变函数论选讲	张南岳等编著
2. 近代分析引论	苏维宜编著
3. 高等概率论	程士宏编著
4. 复半单李代数引论△	孟道骥编著
5. 群表示论△	曹锡华等编著
6. 模形式讲义△	陆洪文等编著

前　言

我国实行学位制度以来，研究生教育有了很大的发展。人们逐渐认识到：拓宽研究生的知识面是时代发展的需要。许多数学硕士点和博士点都要求在研究生阶段设立专业基础课程，使得不同专业、不同专题方向的研究生能对本专题以外的重要的、带基础性的近代发展也有所了解。

开设这类研究生专业基础课程的教材，当然是要介绍该方面的基本概念和基本方法。但在涉及近代的发展上不应过于专门，要照顾到各个不同分支的需要；也不能过于拘泥在技术细节上的推导，而是要在总体上、思想方法上给读者对该学科的主要内容有一个清晰的了解。因此在编写这类教材时，在深与广、精与粗、全貌与专题等方面要掌握适度才能使大多数来自不同专题方向的学生受益。

国内过去出版的大量为本科生编写的教材，因其没有反映近代的内容，不能满足需要，就是许多为研究生编写的教材，因其过分专门而不适用。可喜的是最近几年，出现了一批经这一段教学实践检验后符合上述要求的研究生专门基础课讲义。出版《天元研究生数学丛书》就是为了推动这类教材的编写，促进我国数学研究生培养水平的提高，希望得到数学界同仁们共同的关心和支持。

1995 年 3 月于北京

序　　言

本书是在1984—1988年向北大概率统计专业硕士研究生开设的概率论基础课的讲稿的基础上编写而成的. 讲稿1987年曾油印成讲义,作为研究生教材沿用至今.

由于北大概率统计系的研究生一般都学过测度论,所以本书并不包含这部分内容,而只是在第一章用测度论的观点对概率论的基本概念作出严格的叙述. 第二章集中讨论了离散鞅. 本书的大部分内容贯穿着这样一种考虑:先得到关于鞅或鞅差的一般结论,再把这些一般结论特殊化到概率论的古典对象——独立随机变量. 这样做可能有两方面的好处:第一,可以节省篇幅;第二,可以突出离散鞅研究在60年代末70年代初那一段时间的研究成果,更新课程的内容. 这种考虑在第四章、第五章的某些地方不得不放弃. 例如,讨论(弱)不变原理和强逼近的时候,我们把重点放在透彻地理解问题的提法而不是这些结果可以推广到什么程度. 因此,在那里我们只对i.i.d.的情况进行了详尽的讨论,而对于它们的推广则吁请读者参阅书末的有关文献. 在具体内容的讲法上,我们注意从文献资料上吸收好的想法. 例如,Lindeberg-Feller中心极限定理(定理4.3.4)的证明取自于美国北卡罗来纳大学(University of North Carolina at Chapel Hill)统计系的概率论讲义. 又例如关于非负随机变量级数收敛的定理5.1.1摘自1978年Chen发表在Annals of Probability上的一篇文章,关于重对数律的新证明(定理5.3.7)摘自于1982年Ascota发表在Annals of Probability上的一篇文章,等等. 在讲授课程和编写教材的过程中,我们尽可能地将有些定理在形式上写得更一般些. 例如,Lindeberg-

Feller 定理可以写成定理 4.3.5 那种形式;Skorokhod 嵌入定理也不必加上方差有限的条件(定理 3.4.5),等等.

本书的内容每周 4 学时,一个学期讲完,时间有点紧.一般,我们能讲完不带 * 号的内容,而带 * 号的内容(4.1 小节除外)只介绍结论并解释其意义.当然,如果每周能增加 2 学时,一个学期内讲完全部内容应该是不成问题的.

感谢数学天元项目基金的资助使本书得以正式出版.在编写本书的过程中,使用过原讲义的蒋继明、祁永成等同志提出了许多宝贵意见,在此一并致谢.

在 1984 年酝酿开这个课的时候,当时北大数学系概率统计教研室的领导和同事们对我提出了两方面的要求:一是课程内容要现代化;二是要坚持严格的科学训练.但是由于个人水平的限制,究竟在多大程度上能做到这两点就很难说了.我衷心希望读者能对本书的缺点、错误提出批评意见.

程士宏

92.1.北京大学

目　录

序言……………………………………………………………（*1*）

第一章　概率论基础……………………………………………（1）

第一节　概率论的基本概念 ……………………………………（1）

第二节　距离可测空间 ………………………………………（25）

第三节　条件期望和条件概率 ………………………………（40）

第四节　距离空间的概率测度 ………………………………（61）

第二章　离散鞅论 ……………………………………………（77）

第一节　基本概念 ……………………………………………（77）

第二节　停时定理 ……………………………………………（90）

第三节　收敛定理……………………………………………（108）

第四节　鞅的不等式…………………………………………（127）

第三章　Wiener 过程 ………………………………………（149）

第一节　Wiener 过程的定义和性质 ………………………（149）

第二节*　Wiener 过程的增量 ……………………………（162）

第三节*　Wiener 过程的重对数律…………………………（176）

第四节*　Skorokhod 嵌入定理 ……………………………（187）

第四章　弱收敛理论…………………………………………（197）

第一节　距离空间概率测度的弱收敛 ………………………（198）

第二节　鞅的中心极限定理 …………………………………（218）

第三节　独立随机变量阵列的中心极限定理和弱大数律 ………（243）

第四节　随机过程的依分布收敛 ……………………………（271）

第五章　强收敛理论…………………………………………（294）

第一节　随机变量级数的收敛性 ……………………………（295）

第二节　强大数律…………………………………………………（323）
第三节　重对数律…………………………………………………（347）
第四节* Strassen 强逼近……………………………………………（362）
参考书目…………………………………………………………（367）

此为试读，需要完整PDF请访问：www.ertongbook.com

第一章　概率论基础

初等概率论是建立在排列组合和微积分等数学方法的基础上的.在那里,虽然已经接触过事件、随机变量和数学期望等基本概念,但是,对这些基本概念却始终未能给出一个明确的定义.因此,奠定概率论的严格数学基础并把它作为数学的一个分支来研究是十分必要的.1933 年 Kolmogorov 的著作《概率论基础》被公认为是概率论公理系统完成的标志.按照 Kolmogorov 公理系统,概率论是以测度论为其数学基础的.本章我们就以测度论为工具,对概率论的基本概念作一个严格叙述.在概率论及其相关领域如随机过程和数理统计中,仅仅讨论随机变量是不够的,经常还要牵涉到取值于抽象空间特别是拓扑空间的随机变量——随机元.因此,在这一章中,还包括有距离可测空间及它上面的概率测度的内容.我们还讨论了概率论的最基本的概念之一——条件期望和条件概率.其中关于正则条件分布的存在性也是对取值于完备可分距离空间的随机元来证的.

第一节　概率论的基本概念

1.1 概率空间和随机元

设 Ω 是任一集合,$\mathscr{F}$ 是 Ω 的子集组成的 σ 域,P 是 $\mathscr{F}$ 上的测度.在测度论中,$(\Omega,\mathscr{F})$称为可测空间,三位一体的$(\Omega,\mathscr{F},P)$称为测度空间.

定义 1.1.1　测度空间$(\Omega,\mathscr{F},P)$如果满足 $P(\Omega)=1$ 就称为

概率空间.

设$(\Omega,\mathscr{F},P)$是概率空间，那么$\mathscr{F}$中之集合称为事件，Ω称为必然事件，P称为概率测度. 对$A\in\mathscr{F}$，$P(A)$称为事件A发生的概率. 如果事件A发生的概率为1，我们便说A是几乎必然发生的，记成A a.s..

给定集Ω到集X的映射ξ. 对任$B\subset X$，把
$$\xi^{-1}B=\{\omega\in\Omega;\xi(\omega)\in B\}$$
叫做B在映射ξ下的完全反像. 对于X的子集所形成的集合系$\mathscr{B}$，记
$$\xi^{-1}\mathscr{B}=\{\xi^{-1}B;B\in\mathscr{B}\}.$$
当$\mathscr{B}$是一个σ域，我们称
$$\sigma(\xi)=\xi^{-1}\mathscr{B}$$
为使ξ成为到$(X,\mathscr{B})$的可测映射的最小σ域. 换言之，ξ是可测空间$(\Omega,\mathscr{F})$到可测空间$(X,\mathscr{B})$的可测变换，当且仅当$\sigma(\xi)\subset\mathscr{F}$. 如果$\xi$是测度空间$(\Omega,\mathscr{F},P)$到可测空间$(X,\mathscr{B})$的可测变换，则由

(1.1.1) $$(P\xi^{-1})(B)=P(\xi^{-1}B),\quad B\in\mathscr{B}$$

在$\mathscr{B}$上定义的测度$P\xi^{-1}$称为ξ的导出测度.

定义 1.1.2 概率空间$(\Omega,\mathscr{F},P)$到可测空间$(X,\mathscr{B})$的可测变换ξ称为定义在$(\Omega,\mathscr{F},P)$上取值于$(X,\mathscr{B})$的随机元，ξ的导出测度$P\xi^{-1}$称为它的分布.

1.2 随机变量和分布函数

以$\boldsymbol{R}$记全体实数，以$\mathscr{R}$记$\boldsymbol{R}$中形如$(-\infty,x]$之集产生的σ域，即
$$\mathscr{R}=\sigma(\{(-\infty,x];x\in\boldsymbol{R}\}).$$

定义 1.1.3 定义在概率空间$(\Omega,\mathscr{F},P)$上取值于$(\boldsymbol{R},\mathscr{R})$的随机元$\xi$称为这个概率空间上的随机变量(缩写为r.v.)；
$$F(x)=(P\xi^{-1})(-\infty,x]=P(\xi\leqslant x),\quad x\in\boldsymbol{R}$$

称为 r.v. ξ 的分布函数(缩写为 d.f.);r.v. ξ 的 d.f. 是 F,记成 $\xi \sim F$.

随机变量的 d.f. 决定了它的分布. 事实上,人们常把 $\boldsymbol{R}$ 上非降、右连续而且满足 $\lim\limits_{x\to-\infty} F(x) = 0$ 和 $\lim\limits_{x\to\infty} F(x) = 1$ 的函数叫做 $\boldsymbol{R}$ 上的 d.f.. 对于 $\boldsymbol{R}$ 上任一给定的 d.f. F,根据测度扩张定理,在 $(\boldsymbol{R},\mathscr{R})$ 上就有唯一的概率测度 μ 使得对每 $x\in \boldsymbol{R}$, $\mu(-\infty,x] = F(x)$,这个 μ 称为由 d.f. F 产生的 Lebesgue-Stieltjes 测度(简写为 L-S 测度). 不难验证,r.v. 的 d.f. 是一个我们刚才所说的 $\boldsymbol{R}$ 上的 d.f.,而 r.v. 的 d.f. 产生的 L-S 测度则正好是这个 r.v. 的分布.

按测度论的说法,r.v. 是概率空间 $(\Omega,\mathscr{F},P)$ 上取有限值的可测函数. 记

$$\boldsymbol{R}^- = \boldsymbol{R} \cup \{-\infty\} \cup \{\infty\}; \quad \mathscr{R}^- = \sigma(\{[-\infty,x]: x \in \boldsymbol{R}^-\}).$$

在概率论中,可测函数则是从概率空间到 $(\boldsymbol{R}^-,\mathscr{R}^-)$ 的随机元. 顺便说一下,讨论可测函数时要牵扯到符号 $-\infty$, ∞ 和实数 x 之间的运算,它们是这样规定的:

$$(\pm\infty) + (\pm\infty) = x + (\pm\infty) = (\pm\infty) + x = \pm\infty;$$

$$(\pm\infty)\cdot(\pm\infty) = \pm\infty; \quad (\pm\infty)\cdot(\mp\infty) = -\infty;$$

$$\frac{x}{\pm\infty} = 0; \quad x\cdot(\pm\infty) = (\pm\infty)\cdot x = \begin{cases} \pm\infty, & x > 0, \\ 0, & x = 0, \\ \mp\infty, & x < 0; \end{cases}$$

而诸如 $(\mp\infty) + (\pm\infty)$, $\dfrac{\pm\infty}{\pm\infty}$ 这类的符号则无意义. 至于 $\boldsymbol{R}^-$ 中的序,我们约定: $-\infty < x < \infty$ 对每 $x\in \boldsymbol{R}$ 成立.

关于可测函数的一个重要事实是

定理 1.1.1 设 $(X,\mathscr{B})$ 是可测空间,ξ 是从集 Ω 到集 X 的映射,则 f 是 $(\Omega,\xi^{-1}\mathscr{B})$ 上可测函数的充要条件是存在 $(X,\mathscr{B})$ 上的可测函数 g 使

$$f(\omega)=g(\xi(\omega))$$

对一切 $\omega\in\Omega$ 成立；如果 f 是有限值，上述 g 也可取成是有限值的.

1.3 期望、方差

设 ξ 是概率空间$(\Omega,\mathscr{F},P)$上的可测函数. 以 ξ^+ 和 ξ^- 记 ξ 的正部和负部，如果

$$\min\left(\int_\Omega \xi^+\,\mathrm{d}P,\int_\Omega \xi^-\,\mathrm{d}P\right)<\infty,$$

那么 ξ 的积分有意义，我们把它记作

(1.1.2) $$E\xi=\int_\Omega \xi \mathrm{d}P,$$

并称 E 为期望算子.

定义 1.1.4 概率空间$(\Omega,\mathscr{F},P)$上的可测函数 ξ 如果满足 $E|\xi|<\infty$，就说它的期望存在，并把由(1.1.2)确定的 $E\xi$ 称为它的期望. 如果 $E\xi^2<\infty$，则

$$\mathrm{var}\xi=E(\xi-E\xi)^2$$

称为 ξ 的方差.

求可测函数数学期望的有力工具是测度论中如下的积分变换公式.

定理 1.1.2 设 ξ 是测度空间$(\Omega,\mathscr{F},P)$到可测空间$(X,\mathscr{B})$的可测变换，g 是$(X,\mathscr{B})$上的可测函数，则下式在两端之一有意义时成立：

(1.1.3) $$Eg(\xi)=\int_X g\mathrm{d}P\xi^{-1}.$$

1.4 重要不等式

概率论中要用到许多测度论中的不等式，我们把一些重要的列举备查. 其中的 Jensen 不等式条件有所放宽，故给予证明.

引理 1.1.3 对 $\boldsymbol{R}$ 上连续凸函数 g，存在 $\boldsymbol{R}$ 上实值非降函数 h，使对任 $x,y\in\boldsymbol{R}$，

$$g(y)-g(x)\geqslant h(x)(y-x) \tag{1.1.4}$$

成立；又当 $x\to\infty$ 时 $g(x)$ 的极限存在，记为 $g(\infty)$；当 $x\to-\infty$ 时 $g(x)$ 的极限存在，记为 $g(-\infty)$.

证明 对任 $\lambda\in(0,1)$ 和 $u<y$，有

$$g(\lambda u+(1-\lambda)y)\leqslant\lambda g(u)+(1-\lambda)g(y).$$

若 $x<y$，取 $u<x$ 并在上式中令 $\lambda=(y-x)/(y-u)$，得

$$[g(x)-g(u)]/(x-u)\leqslant[g(y)-g(x)]/(y-x).$$

再令 $h(x)=\sup\limits_{u<x}[g(x)-g(u)]/(x-u)$，就进而得到

$$h(x)\leqslant[g(y)-g(x)]/(y-x),$$

即(1.1.4)对任何 $x<y$ 成立. 当 $y<x$ 时，由 h 的定义本身知(1.1.4)仍成立；当 $y=x$ 时，(1.1.4)无条件成立. 因此，(1.1.4)对任 $x,y\in\boldsymbol{R}$ 成立.

如对每 $x\in\boldsymbol{R}$ 均有 $h(x)\leqslant0$，则由(1.1.4)易知 g 非增，从而当 $x\to\infty$ 时 $g(x)$ 有极限. 如果存在 $x_0\in\boldsymbol{R}$ 使 $h(x_0)>0$，则由(1.1.4)又得 $\lim\limits_{x\to\infty}g(x)=\infty$. 所以当 $x\to\infty$ 时 $g(x)$ 有极限. 类似可证，当 $x\to-\infty$ 时 $g(x)$ 亦有极限. 证完.

定理 1.1.4(Jensen 不等式) 设 ξ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的可测函数，$E\xi$ 有意义，g 是 $\boldsymbol{R}$ 上连续凸函数，$Eg(\xi)$ 有意义，则

$$Eg(\xi)\geqslant g(E\xi), \tag{1.1.5}$$

其中 $g(-\infty):=\lim\limits_{x\to-\infty}g(x)$，$g(\infty):=\lim\limits_{x\to\infty}g(x)$.

证明 如果 $E|\xi|<\infty$，无妨设 ξ 是 r. v. ，这时于(1.1.4)中取 $y=\xi$，$x=E\xi$，便得

$$g(\xi)\geqslant g(E\xi)+h(E\xi)(\xi-E\xi).$$

两端再取期望即得(1.1.5).

如果 $E\xi=\infty$，可对(1.1.4)分两种情况讨论：

1） 对每 $x\in\boldsymbol{R}$，$h(x)\leqslant 0$. 此时 g 非增，故

$$g(\xi)\geqslant g(\infty)=g(E\xi).$$

上式取期望便得(1.1.5).

2） 存在 $x_0\in\boldsymbol{R}$ 使 $h(x_0)>0$. 此时 $g(\infty)=\infty$. 于是，当 $P(\xi=\infty)>0$ 时，

$$Eg(\xi)\geqslant g(\infty)P(\xi=\infty)=\infty,$$

从而(1.1.5)成立. 当 $P(\xi=\infty)=0$ 时，

$$g(\xi)-g(x_0)\geqslant h(x_0)(\xi-x_0),$$

从而取期望后仍得 $Eg(\xi)=\infty$ 并由此推知(1.1.5).

对 $E\xi=-\infty$ 的情况，类似可证(1.1.5)成立. 这样就完成了定理的证明.

定理 1.1.5(Hölder 不等式) 设 ξ,η 是概率空间$(\Omega,\mathscr{F},P)$上的可测函数，实数 p,q 满足 $1<p,q<\infty$ 和 $1/p+1/q=1$，则

(1.1.6) $$E|\xi\eta|\leqslant(E^{1/p}|\xi|^p)(E^{1/q}|\eta|^q).$$

如果 $E|\xi|^p<\infty$，$E|\eta|^q<\infty$，则(1.1.6)的等号成立当且仅当 $\xi=0$ a.s. 或 $\eta=0$ a.s. 或存在 $C>0$ 使 $|\xi|^p=C|\eta|^q$ a.s..

设 ζ 是一个可测函数，$s<t$ 是正实数. 在(1.1.6)中取 $\xi=|\zeta|^s$，$\eta=1$，$p=t/s$ 和 $q=t/(t-s)$ 便得

$$\begin{aligned}E|\zeta|^s&=E\xi\eta\leqslant(E^{1/p}|\xi|^p)(E^{1/q}|\eta|^q)\\&=E^{1/p}|\zeta|^{sp}=E^{s/t}|\zeta|^t.\end{aligned}$$

这就是所谓的矩不等式.

系 1.1.6(矩不等式) 设 $s<t$ 是正实数. 对概率空间$(\Omega,\mathscr{F},P)$上的任一可测函数 ξ，均有

(1.1.7) $$E^{1/s}|\xi|^s\leqslant E^{1/t}|\xi|^t.$$

如果 $E|\xi|^t<\infty$，上式等号成立的充要条件是 ξ 在 a.s. 意义下是一常数.

定理 1.1.7(Minkowski 不等式) 设 ξ,η 是概率空间$(\Omega,\mathscr{F},P)$上两个 r.v..

此为试读，需要完整PDF请访问：www.ertongbook.com