"十二五"国家重点图书出版规划项目

电子与信息工程系列

ANALYSIS AND PROCESSING FOR STOCHASTIC PROCESS

随机过程分析与处理

● 高玉龙　陈艳平　何晨光　编著

哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

"十二五"国家重点图书出版规划项目
电子与信息工程系列

ANALYSIS AND PROCESSING FOR STOCHASTIC PROCESS

随机过程分析与处理

高玉龙　陈艳平　何晨光　编著

哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内容简介

本书从概率论和随机过程的发展历史出发，以随机过程的概率论基础、随机过程基本理论、随机过程的时域和频域分析原理、平稳随机过程通过系统、窄带随机过程、离散随机过程以及非平稳随机过程为主要内容。在介绍的过程中采用先确定后随机、先连续后离散的思路，便于读者对相关知识进行对比理解以至达到融会贯通的程度。

本书可作为通信工程、电子工程、信息工程的高年级本科生和研究生的教学参考书，也可供相关的工程技术人员参考。

图书在版编目(CIP)数据

随机过程分析与处理/高玉龙，陈艳平，何晨光编著．—哈尔滨：哈尔滨工业大学出版社，2017.3

ISBN 978-7-5603-6283-0

Ⅰ.①随… Ⅱ.①高…②陈…③何… Ⅲ.①随机过程—过程分析—研究 Ⅳ.①O211.6

中国版本图书馆 CIP 数据核字(2016)第 265134 号

电子与通信工程
图书工作室

责任编辑 李长波
封面设计 刘洪涛
出版发行 哈尔滨工业大学出版社
社　　址 哈尔滨市南岗区复华四道街 10 号　邮编 150006
传　　真 0451-86414749
网　　址 http://hitpress.hit.edu.cn
印　　刷 哈尔滨市久利印刷有限公司
开　　本 787mm×1092mm　1/16　印张 16　字数 395 千字
版　　次 2017 年 3 月第 1 版　2017 年 3 月第 1 次印刷
书　　号 ISBN 978-7-5603-6283-0
定　　价 33.00 元

“十二五”国家重点图书
电子与信息工程系列

序

FOREWORD

教材建设一直是高校教学建设和教学改革的主要内容之一。针对目前高校电子与信息工程教材存在的基础课教材偏重数学理论，而数学模型和物理模型脱节，专业课教材对最新知识增长点和研究成果跟踪较少等问题，及创新型人才的培养目标和各学科、专业课程建设全面需求，哈尔滨工业大学出版社与哈尔滨工业大学电子与信息工程学院的各位老师策划出版了电子与信息工程系列精品教材。

该系列教材是以“寓军于民，军民并举”为需求前提，以信息与通信工程学科发展为背景，以电子线路和信号处理知识为平台，以培养基础理论扎实、实践动手能力强的创新型人才为主线，将基础理论、电信技术实际发展趋势、相关科研开发的实际经验密切结合，注重理论联系实际，将学科前沿技术渗透其中，反映电子信息领域最新知识增长点和研究成果，因材施教，重点加强学生的理论基础水平及分析问题、解决问题的能力。

该系列教材具有以下特色：

(1)**强调平台化完整的知识体系。**该系列教材涵盖电子与信息工程专业技术理论基础课程，对现有课程及教学体系不断优化，形成以电子线路、信号处理、电波传播为平台课程，与专业应用课程的四个知识脉络有机结合，构成了一个通识教育和专业教育的完整教学课程体系。

(2)**物理模型和数学模型有机结合。**该系列教材侧重在经典理论与技术的基础上，将实际工程实践中的物理系统模型和算法理论模型紧密结合，加强物理概念和物理模型的建立、分析、应用，在此基础上总结牵引出相应的数学模型，以加强学生对算法理论的理解，提高实践应用能力。

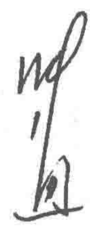
(3)**宽口径培养需求与专业特色兼备。**结合多年来有关科研项目的科研经验及丰硕成果，以及紧缺专业教学中的丰富经验，在专业课教材编写过程中，在兼顾电子与信息工程毕业生宽口径培养需求的基础上，突出军民兼用特色，在满足一

般重点院校相关专业理论技术需求的基础上，也满足军民并举特色的要求。

电子与信息工程系列教材是哈尔滨工业大学多年来从事教学科研工作的各位教授、专家们集体智慧的结晶，也是他们长期教学经验、工作成果的总结与展示。同时该系列教材的出版也得到了兄弟院校的支持，提出了许多建设性的意见。

我相信：这套教材的出版，对于推动电子与信息工程领域的教学改革、提高人才培养质量必将起到重要推动作用。

中国工程院院士
哈尔滨工业大学教授 张乃通



2010年11月于哈工大

前言

PREFACE

随机性是世间万物存在的一种形态，因此研究随机性的科学便应运而生。从静态的概率论一路发展到动态的随机过程(在电信类专业可以称之为随机信号)，随机性分析与处理的理论和方法逐步完善。本书主要介绍概率论和随机过程处理的发展历史，随机过程分析与处理的概率论基础，随机过程基本概念和理论，平稳随机过程的谱分析，随机过程通过系统，窄带随机过程，离散随机过程分析与处理和非平稳随机信号处理和分析等。在介绍随机过程分析与处理理论时，采用先确定性后随机性，先连续后离散，先理论后实验的思路，便于学生掌握和理解随机过程分析与处理的难点和重点。紧密结合电信专业的具体情况给出相应的习题和例题，使学生真正明白理论的应用价值和应用方法。另外，本书侧重在物理概念和分析方法上，把复杂的理论和数学问题与实际的应用技术问题相联系，使学生深入理解随机过程的本质，掌握随机过程处理分析的相关理论与方法。更为重要的是，通过本书内容的学习不仅仅学习一些具体的理论和方法，更应该学会用统计的、随机性观点看待物质世界和现实生活中的人和动物，逐步用这些观点完善我们的世界观和人生观，如果更进一步的学习可以上升到哲学的高度去引领我们更好地认识世界的本质和自然界的规律。

本书的第一个重要特点是不仅介绍原理和方法，还要介绍方法和原理的发展历史，使读者不仅学习知识，还要了解知识的来龙去脉，掌握科学发展的规律，从而激发读者的创新思维。第二个特点是针对电信专业设计习题和例题，使学生掌握理论的应用方法，明白理论的应用价值。第三个特点是把随机信号历史、关键技术瓶颈及其解决思路和国际前沿动态相联系，有利于引导科研创新思维并跟踪国际发展，开阔读者的视野。

本书共分 8 章：

首先为绪论。主要介绍概率论和随机通信信号处理的发展历史、研究内容、各内容之间的联系及其相应的学习方法。

第 1 章随机过程的概率论基础。主要介绍概率论的有关知识，包括概率的概念，随机变量基本概念，随机变量的数字特征，多维随机变量，随机变量函数的分布函数和概率密度函数，大数定理和中心极限定理。

第 2 章随机过程基本概念和理论。介绍随机过程的概念和定义，统计特征，随机过程的积分和微分定义，平稳随机过程，非平稳随机过程，各态历经过程，随机过程的联合分布和互相关函数，实际通信中常用的高斯随机过程。结合实际通信信号和噪声对这些概念和理论进行应用。

第 3 章平稳随机过程的功率谱。主要包括随机过程的功率谱定义，功率谱密度的性质，功率谱密度与自相关函数之间的关系，联合平稳随机过程的互功率谱密度，互谱密度和互相关函

数的关系，平稳随机过程互谱密度的性质，通信中的噪声，理想白噪声，高斯噪声，高斯型白噪声，限带高斯噪声等。

第 4 章随机过程与系统。主要包括线性系统和非线性系统的基本理论，随机过程通过线性系统的时域数字特征，系统输出过程的平稳性和遍历性分析，平稳随机过程通过线性系统的功率谱，两个平稳随机过程之和通过线性系统，白噪声通过线性系统，线性系统输出随机过程的概率分布，随机过程通过非线性系统。

第 5 章通信中的窄带随机过程。主要介绍窄带随机过程的三种表示形式，窄带随机过程对应的解析过程及其性质，窄带随机过程正交和同相分量的性质，高斯随机过程包络与相位的概率密度，高斯过程包络平方的概率密度，马尔可夫过程的概念和相关理论。

第 6 章离散时间随机过程分析与处理。主要介绍随机过程的采样定理，随机序列的时域分析(概率密度和数字特征)，随机序列的频域分析(功率谱和功率谱采样定理)，随机序列通过系统(可以包括建模)，平稳随机序列的时域模型，离散随机过程的参数估计(数字特征和功率谱估计)。

第 7 章非平稳随机信号分析与处理。主要包括非平稳信号处理和分析的基本概念、性质及非平稳的通信信号分析，主要讲述循环平稳信号分析和处理方法在通信信号处理中的应用。

本书从工科技术研究和教学人员的角度进行资料的选择和内容安排。作者在编撰过程中，参考了国内外有关随机过程分析与处理的书籍和著作，得到很多的启发，并根据自己的教学实践和学生的情况对随机信号处理理论和方法从一种学生和读者容易理解的角度进行撰写。本书绪论和第 3 章由高玉龙和何晨光撰写；第 1、2 章由高玉龙和陈艳平撰写；第 4、7 章由何晨光撰写；第 5、6 章由陈艳平撰写。

由于编者水平有限，书中难免有疏漏与不妥之处，敬请各位读者和同仁批评指正。

作　者
2017 年 1 月

目录

CONTENTS

绪　论

1. 生活中的随机过程

在人们的日常生活和科学研究中处处充满了各种确定性或随机性的现象，概率论和随机过程则是研究随机现象数量规律的数学分支。所谓的随机现象是指人们提前无法预知事情发生的结果，但这些结果也不是完全没有规律，而只是多种可能结果中的一种。例如，掷一枚硬币，不知道出现正面或反面，但肯定是二者其中之一。另外，如果测量一个物体或房间的尺寸，由于工具和环境的影响，每次测量结果可能不同，但应该在某个范围之内，也就是说测量误差应该符合某种分布。可以看出，虽然在一次随机试验中出现某种结果是有偶然性的，但如果在相同条件下大量重复的随机试验，却往往呈现出明显的数量规律性。例如，连续多次抛硬币，出现正面和反面的频率几乎相等；多次测量房间尺寸所得结果的平均值随着测量次数的增加，逐渐稳定于一个常数，并且测量值大都落在此常数的附近，测量次数越少对应的测量误差越大，也就是测量误差随测量次数的分布符合高斯分布。人们在长期实践和研究中已逐步发现了某些随机现象的规律性，经过几代人近 400 年的研究形成了今天的概率论体系，其中最著名的就是大数定律和中心极限定理。在实际中，人们往往还需要研究随着时间的变化某一特定随机现象的演变情况，此时采用古典概率理论已经无法准确描述这种现象，因此逐渐出现了随机过程的相关理论。例如，某个地区每一年的降水量由于受天气等许多随机因素的影响导致其具有随机性，因此年降水量就是一个随机过程；食堂不同时刻到达窗口需要服务的学生人数；电话交换局从确定时刻起到其后的某一时刻为止所收到的呼叫次数；微小粒子在液体中因受周围分子的随机碰撞而形成不规则的布朗运动等都是随机过程。因此，随机过程可以简单地认为是随时间变化的随机变量，通常被视为概率论的动态部分。概率论主要研究概率空间上的一个或有限多个随机变量的规律性，而随机过程则研究随时间不断变化的随机变量，内容包括不同时间随机变量的规律或不同时间随机变量之间的相互关系，是随机变量概念在二维空间的拓展和延伸。目前，随机过程论已在自然科学、工程技术科学、社会科学、军事和工农业生产中得到广泛应用，在诸如天气预报、物理学、运筹决策、经济数学、安全科学、生物、通信、自动控制及计算机科学等领域都要经常采用随机过程的理论来建立数学模型。

概率论和随机过程的发展史说明了理论与实际之间的密切关系。它们的许多研究方向来源于生产活动、科学研究和工程技术中的大量实际问题所提出的应用需求。反过来，当这些方向被深入研究形成科学理论后，又可以指导实践，进一步扩大和深化应用范围。

2. 概率论和随机过程的发展历史

(1)概率史前阶段。

随机现象存在于自然界的各个方面,在很早以前这些现象已经引起了人们的注意,例如,古希腊人从航海实践中发现了许多概率经验规律,古犹太人在纪元之初就有概率加法定律和乘法定律的应用记录。古希腊哲学家亚里士多德认为自然科学是一种可能性科学,他把事件分成3类,分别是必然发生的确定性事件、在大部分情况下发生的可能性事件以及偶然发生的不可预测或不可知事件。但是由于受到知识的限制以及结果不确定因素的影响,人们一直认为随机现象的结果都由天神决定,其规则是人类无法掌握的。在现实生活中能够促使人们思考概率的事情很多,但最终孕育概率论的却是广泛使用的骰子赌博游戏。原因大概是人们经常接触骰子赌博,比较容易引起大家的注意,另外就是如果把生活中的机遇问题精确到数量上去考虑,只有赌博这种较为简单的活动才有可能。其实很早以前赌博几乎出现在世界各地的许多地方,如埃及、印度和中国等。在赌博的过程中,人们发明了骰子(俗称"色子")等赌博性工具,总结了赌博的简单规律。后来,慢慢有人利用简单的数学思想计算赌博输赢机会的大小,以便在赌博中获胜。公元960年左右,怀特尔德大主教计算出掷3个骰子时不计次序所能出现的不同组合有56种。13世纪左右,拉丁诗歌《维图拉》指出这56种组合出现的机会不是相同的:3枚骰子点数一样,每个点数只有一种方式;2枚骰子点数一样而另一枚不一样,则有3种方式;如果3枚都不一样就有6种方式。但是这些结果并没有引起人们更多的思考,输赢概率的计算仍处于直觉的、散乱的经验水平上。一直到文艺复兴时期,人们一直没有把赌博与数学直接进行关联,甚至没有人意识到骰子点数下落频率的计算是可能的、有效的或每一面会以相同的频率出现等这些最简单的概率思想。对于概率思想出现得如此缓慢的现象,人们提出了许多原因。第一就是由于缺少完美平衡和"诚实"的骰子,因而阻碍了人们发现任何可察觉的规律。或者由于缺少适当的数学概念和符号而阻碍了数学的探索。缺乏刺激概率思想研究的经济问题。还有一个更可能的原因是"随机"概念本身与时空观念相对,长期以来,人们一直认为一系列的好运和坏运都是神授的。人们相信上帝或众神以某种预先确定的计划指导着世俗的事件,所以随机性的研究不但是不可能的,甚至是不可想象的。还有一个解释涉及道德的规范,赌博长期以来被视为一种不道德的行为,历史上充满了限制、制止赌博的各种尝试。既然赌博被视为不道德的,那么将机会性游戏作为科学研究的对象也就几乎没有可能。然而这些原因没有一个得到广泛的认可,人们对每一个猜测都提出了反驳的理由。

(2)概率萌芽阶段。

在文艺复兴前,概率还是一个非数学的概念。直到文艺复兴时期,随着阿拉伯数字和计算技术的广泛传播,简单代数和组合数学的发展,并且影响人们思想的哲学观念开始转变,此时概率的思想才开始逐渐浮出水面。现在有史可查的对于赌博问题最早加以研究的事件是从意大利开始的,那时候贵族之间盛行赌博,掷骰子是他们常用的一种赌博方式。因骰子的形状为小正方体,当它被掷到桌面上时,每个面向上的可能性是相等的,即出现1点至6点中任何一个点数的可能性是相等的。当时就有人思考不同情况出现的可能性大小以及对观察的结果如何加以归类计算等问题。例如,1477年意大利但丁的《神曲》的注释本讲到了投掷3颗骰子可能出现的各种结果。15世纪后期和16世纪早期,一些意大利数学家开始思考赌博游戏中各种存在结果的数学概率,开始从纯数学的角度研究概率,根据这些研究成果最终提出了古典概率的概念,即要把出现的结果分解为一些同等可能出现的结果,其数目与全部可能结果数之比

就是该情况的概率。此定义最初出自何人已无法考证，因为这些早期的赌博家或者学者，都没有著作流传下来。另外比较重要的事件是在这期间出现了分赌本问题，它的文字记载最早在1494年帕西奥利的一本著作中，该问题成为推动概率论研究的着力点和动力之一。所谓的分赌本问题就是，A、B二人赌博，每人的赌金是a元，每人每局获胜的概率为1/2，并且规定，谁先赢得S局，谁就获得全部的$2a$元赌金。当进行到A胜S_1局，B胜S_2局时(S_1和S_2都小于S)，赌博因故停止，问此时的$2a$元赌金怎样分配才公平合理？由于当时并没有很明确的规定公平合理的标准是什么，因此人们给出了不同的解决方法。帕西奥利本人则提出按照$S_1:S_2$分配，意大利数学家塔泰格利亚则认为很难找到数学方法进行分配，而是应该交给法官决定。虽然这样，他也给出了自己的解法，如$S_1>S_2$，则A取走自己的a元，并取走B的a元中的$(S_1-S_2)/S$，反之亦然。

现在有流传下来关于赌博问题的研究最为人知的著作是意大利的卡尔达诺(1501—1576)，他是意大利数学和医学教授，天资聪明，常常不循规蹈矩，有着丰富有趣的生活经历。他最著名的著作是1524年出版的《伟大的艺术》，其中包括了那个时代所有发展起来的代数规则，包括三次和四次方程的解法。在他一生中超过40年的时间里，卡尔达诺几乎每天都参与赌博。早年时他就认定，如果一个人赌博不是为了钱，那么就没有什么能够弥补在赌博中耗去的时间，这些时间本来是可以花在更值得做的事情上，比如学习。作为对在不合适的活动中浪费时间的补偿，他认真地分析了这种活动中的智力因素所起的作用，例如，从一副牌中抽出A的概率，同时掷两个骰子，出现点数的和为7的概率等问题。最终，在一本名叫《机遇博弈》的书中，他阐述了对这些调查、思考的结果和关于赌博实践的体会，从道德、理论和实践等方面对赌博进行了全面的探讨。这本书大约在1546年完成，但直到一百多年后的1663年才出版。这本书的内容一部分是他自己个人赌博经验的总结，如什么时候适合赌博，如何判断赌博是否公平，如何识别和防止赌博中的欺骗。在书中他认为在分牌时，得到某一张牌的机会是随着前一张牌的选走而增大的，如果掷一个骰子能掷出2、4、6，同时也能掷出1、3、5。因此，如果骰子是“诚实”的，那么下赌注就应依据这种等可能性；如果骰子不是“诚实”的，那么它就以一定的或大一点或小一点的比例离开这种等可能性。他的思想已包含了“在‘诚实’的情况下把概率定义为等可能性事件的比”的思想萌芽，即一个特殊结果的概率是所有达到这个结果可能方法的数目与所有可能结果的方法总和的比值。这是人们第一次看到关于骰子的问题由经验向理论的思想转变。从这一角度来讲，有人认为卡尔达诺可以被称为是“概率论之父”，概率论这一个数学分支应当以他作为起点。另外，他明确了胜率是有利结果与不利结果之比，由于在他的那个时代组合理论发展还很不充分，无法用现有的方法进行求解，但书中给出了计算全部结果数的一些方法。但令人遗憾的是，卡尔达诺作为经验丰富的赌博家，在其著作中没有关于实际赌博中各种结果出现频率的记载，这应该是因为那个时候人们还没有认识到频率和概率的关系，特别是频率逼近概率这个后来被伯努利称之为“笨人皆知”的事实。另外，卡尔达诺也对分赌本问题进行了研究，通过较深的推理和计算于1539年提出了按照$r_2(r_2+1):r_1(r_1+1)$的比例进行分配，其中$r_i=S-S_i$，$i=1,2$，他的解法在今天看来是不正确的，但他已经注意到$S-S_i$的作用，而不是S_1或S_2本身。

除了卡尔达诺，伟大的天文学家伽利略也对骰子问题的数学化进行了研究，在《关于骰子游戏的思想》的著作中伽利略解释了在抛掷3枚骰子时为什么会有216种同等可能的结果的问题以及3枚骰子的某些和数的出现看来似乎有同样大小的可能性，而实际上却不是同等可

能的，如和数为 10 比和数为 9 更占有优势等问题。至此，概率论思想开始萌芽、发展，随后人们逐渐开始对其进行更加详细的研究。

(3)组合概率阶段。

17 世纪中叶，法国有一位热衷于骰子游戏的贵族德·梅耳受到前述概率论思想的启蒙，经常从数学的角度提出和思考赌博中出现的需要计算可能性大小的一些问题，包括前面所说的分赌本问题，但他自己无法给出答案。于是他将这些问题寄给了当时的法国数学家帕斯卡(1623—1662)，这开启了概率论发展的历史。不知什么原因，帕斯卡没有立即回答德·梅耳的问题，而是把它交给另一位法国数学家费马(1601—1665)，他是一位业余数学研究爱好者，虽然没受过特别的数学训练，但是在数学这一领域，却取得了同时代其他数学家不可比拟的重大发现，他和笛卡尔(1596—1650)各自独立发明的解析几何学为微积分奠定了技术基础。费马在 17 世纪的数论领域里有着丰富的成果，以至于后来数论成为一个正式的抽象数学领域，这与他的贡献密不可分。作为一个谦逊朴实的人，费马很少发表文章，但是他与当时很多一流数学家不断通信，并在他的同时代人中有相当的影响力。费马的众多重要的贡献丰富了数学的很多领域，所以被称为“业余数学家之王”。1654 年 7～10 月，两位伟大的法国数学家帕斯卡和费马之间为了解决德·梅耳的问题开始了具有划时代意义的通信。他们共有 7 封信，其中费马写了 3 封，在信中他们围绕着赌博中的数学问题开始了深入细致的研究和交流。借鉴前人关于概率的相关知识，他们用计算等可能的有利和不利情况数作为计算机遇数的方法，此时他们还没有使用概率这个术语，但与前人相比，他们广泛使用组合和递推公式等初等概率的有关知识对赌博中的问题进行了研究，终于用不同的方法完整地解决了“分赌本问题”。他们给出的分赌本的计算结果中包含了二项式定理以及概率的加法和乘法定理等，更为重要的是他们将此题的解法向更一般的情况推广，从而建立了概率论的一个基本概念——赌博的值，并定义为赌注乘以胜率。另外，在这些信中他们还讨论更为复杂的赌博输光问题。这些问题后来被与牛顿齐名的荷兰科学家惠更斯获悉并进行研究，于 1657 年出版《论赌博中的机会》一书。这本书被认为是迄今为止关于概率论最早的论著，在欧洲作为概率论的标准教材长达 50 年之久，在著作中惠更斯把帕斯卡和费马所说的赌博的值正式定义为期望。在这本书中，他从公平期望的一条公理出发推出关于期望的 3 条定理，在此基础上利用递推法等解决了当时的一些赌博问题，最后提出了 11 个问题，这些定理和问题称为惠更斯的 14 个命题。

因此可以说早期概率论的真正创立者是帕斯卡、费马和惠更斯，他们主要采用组合数学的知识对概率的相关基本问题进行研究，得到了古典概率若干重要的结论，因此这一时期被称为组合概率时期。在这期间，帕斯卡和费马正确解决了“分赌本问题”，这一事件被伊夫斯称为“数学史上的一个里程碑”，成为数学概率论的起始标志。之所以不把卡尔达诺的著作作为概率论的起源的始点，是因为在卡尔达诺的著作中只有一小部分内容是处理机会计算的。就像卡尔达诺的大多数作品一样，这种处理似乎只是零碎和模糊的，混杂于卡尔达诺的个人奇闻轶事、哲学思考、大量流行的赌博者常用的欺骗策略和精明的心理应用等建议之中，并且他的著作中所阐述的数学思想对数学家和一般的赌博者几乎都没有什么影响。因为对于当时的数学家而言，概率太游戏化，而对赌博者来说，概率又太数学化。而帕斯卡和费马的通信除了正确解决了一些概率论问题和概念之外，还创造了一种研究的传统——用数学方法研究和思考机会性游戏。这种传统统治这个领域达半个多世纪的时间。所以，综合考虑所有这些因素，帕斯卡和费马关于赌博问题的解决在数学概率论历史中的标志性地位是当之无愧的。

(4)古典概率阶段。

惠更斯的《论赌博中的机会》一书的主要研究内容是针对赌博中各种情况出现概率的计算。而把概率论由局限于对赌博机遇的研究扩展到其他领域,使概率论成为数学一个分支的另一奠基人是瑞士数学家雅各布·伯努利(1654—1705)。雅各布·伯努利在前人研究的基础上,继续分析赌博中的其他问题,给出了"赌徒输光问题"的详尽解法,并证明了概率论中第一个极限定理——伯努利大数定理。伯努利大数定理用公式表示为

$$\lim_{N\to+\infty} P\left(\left|\frac{X}{N}-p\right|\geqslant\varepsilon\right)=0 \tag{1}$$

其中,X/N 表示事件发生的频率;p 表示事件发生的概率;ε 为任意小的正数。

定理表示当试验次数很大时,事件出现的频率以概率意义收敛于事件概率。这是研究等可能性事件的古典概率论中极其重要的结果,第一次在单一概率值与众多现象的统计度量之间建立了演绎关系,构成了从概率论通向更广泛应用领域的桥梁,标志了概率概念漫长形成过程的终结,也是数学概率论的开始,对概率论的发展有着不可估量的影响。因此,雅各布·伯努利被称为概率论的奠基人。大数定律证明的发现过程极其困难,为了证明大数定律的猜想,雅各布·伯努利用了二十多年时间,做了大量的试验计算,最终完成了这一定理的证明,在证明的过程中他发现了很多新方法,取得了很多新成果,并在 1713 年将这些研究成果发表在他的遗著《推测术》中。该书第 1 部分对惠更斯的《论赌博中的机会》做了详细的注解,内容总量为前者的 4 倍以上,第 2 部分较为系统地阐述了组合数学的相关内容,总结了包括莱布尼兹等一些数学家关于组合数学的研究结果。第 3 部分是利用前面两部分的内容解决赌博中的问题。《推测术》的前三部分是古典概率的系统化和深化,而第 4 部分则是雅各布·伯努利对大数定理的推导过程,也是本书的精华所在。伯努利在书的结尾对自己工作的意义进行了总结:"如果我们能把一切事件永恒地观察下去,我们终将发现世间的一切事物都受到因果律的控制,而我们也注定会在种种极其杂乱的现象中认识到某种必然。"从大数定律可以看出,虽在一次随机试验中其结果受偶然性支配难以确定,但当试验次数足够多时其结果就会呈现出某种规律性,这就是所谓的"统计规律性"。大数定律体现了事物偶然性和必然性的关系,第一次用数学的语言对此进行了描述,这一认识具有很高的哲学意义,需要说明的是他的思想也有因果论的某些局限性。

当然,现在学过概率论的人都能够简单利用切比雪夫不等式推导大数定律,但当时还没有方差的概念,而切比雪夫不等式也是在 1887 年发现并证明的。直到 1909 年博雷尔证明了强大数定律,$P(\lim\limits_{n\to+\infty}\hat{m}_n=m_x)=1$,用数学的语言解释就是时间平均以概率 1 收敛于统计均值。

此后,雅各布·伯努利的侄子同时也是《推测术》定稿人的尼古拉·伯努利对概率论的相关问题进行了研究,提出了著名的"圣彼得堡问题",提高了伯努利逆概问题的求解精度,也就是给定概率和误差时试验次数 N,用公式表示为:如果已知 ε,c,且 ε 很小,c 很大,求满足

$$P\left(\left|\frac{X}{N}-p\right|\leqslant\varepsilon\right)\geqslant\frac{c}{c+1} \tag{2}$$

的 N 的最小值。雅各布·伯努利给出的答案不够精确,甚至是非常粗糙的。尼古拉·伯努利改变了思路,他把 N 固定,估计概率 P_d,其中 $P_d=P(|X-Np|\leqslant d)$,曾经有个例子,在同样的条件下雅各布·伯努利给出的 N 为 25 550,尼古拉·伯努利计算的 N 为 17 350。后来按照棣莫弗的方法 N 为 6 600,这主要是因为 P_d 是一些二项概率之和,棣莫弗之前的人还没有找到

计算它们的方法。

棣莫弗(1667—1754)出生于法国,成名于英国,但一生困苦。棣莫弗上学期间非常喜欢数学,但由于当时教会的阻挠,只能偷偷地学习,他最感兴趣的是惠更斯《论赌博中的机会》,这给了他概率论的启蒙。1684年,棣莫弗在法国杰出的数学教育家奥扎拉姆的鼓励下学习了欧几里得的《几何原本》等数学家的重要著作,1685年由于参加震惊欧洲的宗教骚乱被捕入狱,出狱后移居伦敦,抵达伦敦不久,棣莫弗利用做家庭教师路上的间隙仔细研读了牛顿刚刚出版的《自然哲学的数学原理》,为他打下了坚实的数学基础,并开始进行学术研究。棣莫弗在雅各布·伯努利的《推测术》出版之前,就对概率论进行了广泛而深入的研究。1711年,他在英国皇家学会的《哲学学报》上发表了"论抽签的原理",该文于1718年出版时翻译成《机会的学说》,后来扩充成一本著作。当时,他在书中并没讨论上述雅各布·伯努利讨论的问题,直到1738年《机会的学说》再版时,才给出了上述问题的解决方法。

促使棣莫弗进行二项概率研究的不是伯努利的工作,而是一个偶然的事件。1721年的某天,一位名叫亚历山大·碲明的人向棣莫弗询问了一个和赌博相关的问题:A、B两人在赌场里赌博,A、B各自的获胜概率是 $p, q=1-p$,共赌 N 局,假设 X 表示A胜的局数,并且两人约定:若 $X > Np$ 则A付给赌场 $X-Np$ 元;反之则B付给赌场 $Np-X$ 元。问赌场挣钱的期望值是多少?该问题用现在的理论解决非常简单,其实质是一个二项分布的数学期望,棣莫弗假设 Np 为整数,求出的结果是 $2Npqb(N,p,Np)$。其中 $b(N,p,Np)=\mathrm{C}_i^N p^i q^{N-i}$ 是常见的二项概率。但是对具体的 N,其中的二项公式中有组合数,但在当时并没有阶乘组合数的计算方法,对于比较大的 N 计算相当困难,这就驱动棣莫弗寻找近似计算的方法。与此相关联的另一个问题是,如果随机变量 $X \sim B(N,p)$,求 X 落在二项分布中心点一定范围内的概率 $P_d = P(|X-Np| \leqslant d)$。对于 $p=1/2$ 的情形,棣莫弗进行了研究,给出了关于阶乘 $N!$ 的近似公式,但是结果还不够精确,与此同时,他把这个问题告诉了他的朋友斯特林(1692—1770),他在棣莫弗的基础上给出了关于阶乘运算的更一般结果,棣莫弗的计算公式是它的一个特例,因此阶乘公式 $N!$ 在数学上被称为斯特林公式或斯特林逼近。可以看出斯特林公式的雏形是棣莫弗最先得到的,但斯特林改进了这个公式,进行了一般化推广。1733年,棣莫弗很快利用斯特林公式计算二项概率问题,得到如下重要的公式

$$P\left(\left|\frac{X}{N}-\frac{1}{2}\right| \leqslant \frac{c}{\sqrt{N}}\right) \sim \int_{-2c}^{2c} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \mathrm{d}x \tag{3}$$

它是二项分布的中心极限定理的雏形,是概率论中第二个基本极限定理的原始表达。这一发现开辟了中心极限定理这个概率论发展的新方向。令人惊讶的是,在这个公式中首次以积分的形式出现了正态分布的概率密度函数,这也正说明了二项分布的极限分布是正态分布。但当时棣莫弗还不知道这就是正态分布的概率密度函数,只是以数学的形式体现出来。

棣莫弗的《机会的学说》在概率论发展中起着承前启后的作用,尤其在二项分布、正态分布函数和中心极限定理等方面的工作。他的发现同样包含着重要的哲学意义:在人们以为是纯粹偶然的事件中,可以寻找出其规律和必然。正如他在该书英文第3版中所指出的那样,尽管机会具有不规则性,但是由于机会无限多,随着时间的推移,不规则性与秩序相比将显得微不足道。他认为,这种秩序自然是从"固有设计中"产生而来的。

除了上述成果外,棣莫弗在《机会的学说》中得到了泊松分布的一种特殊情形,并将母函数用于对正态分布的讨论,开创了一种概率论研究的新方法。此外,他在这部著作中还对赌博

中涉及的其他概率问题进行了深入探讨。

棣莫弗发表成果 40 年后的 1770 年，拉普拉斯(1749—1827) 把 $p=1/2$ 推广到 $0<p<1$ 的情形，建立了中心极限定理较一般的形式

$$\lim_{x\to+\infty} P\left(\frac{X_n-Np}{\sqrt{Np(1-p)}}\leqslant x\right)\sim\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-t^2}{2}}\mathrm{d}t \tag{4}$$

此公式被后人称为棣莫弗－拉普拉斯极限定理，这是概率论中心极限定理的原始形式。中心极限定理随后又被其他数学家推广到其他任意分布的情形，在样本量 N 趋于无穷的时候，其极限分布都是正态的形式，这构成了数理统计学中大样本理论的基础。

除了上述成果之外，在 1812 年拉普拉斯的《概率的分析理论》一书中，他把在概率论上的发现以及前人所有成果进行归纳整理，明确地对古典概率做出了定义，并在概率论中引入了随机变量、数字特征、特征函数、拉普拉斯变换、差分方程和母函数等更有力的数学分析工具，从而实现了概率论由单纯的组合计算到分析方法的过渡，将概率论推向一个新的发展阶段。该书和雅各布·伯努利的《推测术》以及棣莫弗的《机会的学说》成为概率论发展史上的三大具有里程碑意义的著作，标志着古典概率的形成。直至拉普拉斯《概率分析理论》著作的问世，概率论从 17 世纪到 19 世纪初的经典时期才被画上了一个完整的句号。随后，另一在概率论发展史上的代表人物是法国的数学家、几何学家和物理学家泊松(1781—1840)，他推广了伯努利形式下的大数定律。伯努利大数定律证明了事件在完全相同条件下重复进行的随机试验中频率的稳定性，而泊松定理表明，当独立进行的随机试验的条件变化时，随着试验次数 N 的无限增大，在 N 次独立试验中，事件 A 的频率在各次试验中事件 A 出现概率的算术平均值处取得稳定值，也就是频率仍然具有稳定性。基于这些成果，泊松于 1838 年提出了泊松分布。另外，德国数学家高斯也对概率论的发展做出了很大贡献，他在研究天文测量数据误差处理时引入了最小二乘方法，首次给出了极大似然的思想，解决了误差概率密度分布的问题，明确证明了误差的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{5}$$

此式就是正态分布函数表达式。高斯的这项工作对后世的影响极大，而正态分布也因此被冠名为高斯分布。从此之后，概率论的研究中心则集中在推广和改进伯努利大数定律和中心极限定理两个方面。19 世纪后期由于受到概率论哲学思想的争论，概率论学科在西欧备受抨击，使概率论研究无人问津。正是以俄国数学家切比雪夫为首的圣彼得堡数学学派以及物理学的蓬勃发展挽救了岌岌可危的概率论研究。切比雪夫(1821—1894) 一生发表了 70 多篇科学论文，内容涉及数论、概率论、函数逼近论、积分学等方面。他证明了贝尔特兰公式，自然数列中素数分布的定理，大数定律的一般公式以及中心极限定理。他一开始就抓住了古典概率论中的大数定律这个具有基本意义的问题。1845 年，在其硕士论文《试论概率论的基础分析》中借助十分初等的工具——$\ln(1+x)$ 的麦克劳林展开式，对雅各布·伯努利大数定律做了精细的分析和严格的证明。1846 年，他又发表了《概率论中基本定理的初步证明》，文中给出了泊松形式的大数定律的证明。1866 年，切比雪夫发表了论文《论平均数》，进一步讨论了作为大数定律极限值的平均数问题，用他所创立的切比雪夫不等式建立了有关独立随机变量序列的大数定理，即具有相同数学期望和方差的独立随机变量序列的算术平均值依概率收敛于数学期望，即