（修订版）

随机过程基础

SUIJI GUOCHENG JICHU

宋占杰　王家生　王　勇　编著

天津大学出版社
TIANJIN UNIVERSITY PRESS

随机过程基础

（修订版）

宋占杰　王家生　王　勇　编著

天津大学出版社
TIANJIN UNIVERSITY PRESS

内 容 提 要

本书是为工科院校研究生及本科生学习随机过程而编写的.内容包括概率论的基本知识、随机过程的基本概念、Markov链、平稳过程和时间序列分析.本书力求贯彻选材精当而叙述详细的原则,注重说明概念的直观背景和实际意义,在基本理论和方法的阐述上力求通俗易懂、深入浅出.书中收入许多实际问题的典型例子,章末配有习题,有助于读者学习和理解全书的内容.

本书可供工科院校研究生及本科生作为教材,也可作为工程技术、管理人员自学参考用书.

图书在版编目(CIP)数据

随机过程基础/宋占杰,王家生,王勇编著.—天津:天津大学出版社,2011.3

ISBN 978-7-5618-1735-3

Ⅰ.①随…　Ⅱ.①宋…②王…　Ⅲ.①随机过程－高等学校－教材　Ⅳ.①O211.6

中国版本图书馆CIP数据核字(2003)第010529号

出版发行	天津大学出版社
出 版 人	杨欢
地　　址	天津市卫津路92号天津大学内(邮编:300072)
电　　话	发行部:022-27403647　邮购部:022-27402742
网　　址	www.tjup.com
印　　刷	廊坊市长虹印刷有限公司
经　　销	全国各地新华书店
开　　本	169mm×239mm
印　　张	12.75
字　　数	285千
版　　次	2003年3月第1版　2011年3月第2版
印　　次	2011年3月第4次
印　　数	8 001－10 000
定　　价	19.00元

第二版前言

在自然界,随机现象远多于确定性现象,因此加强随机数学的训练和培养在高层次人才能力培养方面显得十分突出.本书的初衷就是为这一目的编排和设计的,本次的修订,依然体现这一目的.

本书第一版问世7年来,历经3次重印,得到国内许多同行的大力支持和帮助,他们提出许多珍贵的修改意见,成为本书修订的重要依据.

在保留原书基本特点和优势的前提下,考虑到使用对象和层次,本书删去了有关鞅论的部分内容和附录中的部分内容,例题和习题也有一定数量的增删,增加了作者近5年来在相关方面的最新理论成果,以使内容密切地联系国际发展前沿,增加读者的学习兴趣.

本书的修订,得到了相关课题国家自然科学基金(批号:60872161)的支持,得到天津大学研究生院出版基金的支持;还得到天津大学研究生院和天津大学出版社领导及相关工作人员的热忱帮助,在此一并致以衷心的感谢.

鉴于作者水平有限,疏漏和不当之处再次恳请同行与读者指正.

联系邮箱:zhanjiesong@ tju. edu. cn

编著者

2011年元月于天津大学

第一版前言

对研究生进行数学教育的一个重要目的是培养学生应用数学方法解决实际问题的能力.而应用数学方法解决实际问题的一个主要方法是构造一个用来描述实际问题的数学模型,再利用数学的理论去分析解决实践提出来的问题.在构造数学模型时,按照实际问题条件的因果性质可分为确定性模型与随机性模型两类.我们知道,在应用中,抽象的数学模型只是当做工具使用,对同一个具体问题可以采用不同的模型来描述,数学理论的应用不依赖于事先形成的意见,它是一个有目的的技术,依赖于经验,而且随着经验的改变而改变.下面我们以群体增长问题为例给出确定性与随机性的两类模型,并加以比较.

设 $x(t)$ 表示 t 时刻群体中的细菌数,假定群体的细菌数只能增加,在 $[t,t+\Delta t]$ 内细菌的增加与时刻 t 时的细菌数成正比,即 $\Delta x(t)=\lambda x(t)\Delta t+o(\Delta t)$,其中 $\lambda>0$ 为常数.由此可得微分方程 $\mathrm{d}x(t)=\lambda x(t)\mathrm{d}t$,假定 $x(0)=x_0$,可求出 $x(t)=x_0\mathrm{e}^{\lambda t}$.这里,我们假定 $x(t)$ 是时间 t 的连续可微函数,这就是描述细菌群体的确定性模型.下面我们考虑同一问题的随机模型.用 $X(t)$ 表示 t 时刻的细菌数,$X(t)$ 是一个取非负整数值的随机变量,于是 $\{X(t),t\geqslant 0\}$ 是一**随机过程**.我们要研究 $P\{X(t)=k\}=p_k(t)$,即 t 时刻细菌数为 k 的概率.在类似的假定下,建立可列多个方程组成的差分微分方程组,求解可得 $p_k(t)$ 的表达式.在随机模型中,我们得到的 $X(t)$ 是取非负整数值的,并得到了 t 时刻群体细菌数为 k 的概率.若令 $m(t)=E[X(t)]$,经计算可得 $m(t)=x_0\mathrm{e}^{\lambda t}$.即 t 时刻群体细菌数的数学期望与确定性模型群体含量表达式是相同的,也就是说确定性模型得到的解是 t 时刻的平均群体含量.显然,随机模型给出了更多的信息,比如利用随机模型还可以求出 t 时刻群体细菌数 $X(t)$ 的方差或其他数字特征,在实际应用中这是有重要意义的.可以说在这个问题中,随机模型比确定性模型更好地描述了群体增长问题.

在很多场合,随机模型是确定性模型所无法代替的,量子力学产生的历史就是一个光辉的例证."量子"这个词是与20世纪一起出生的.1900年12月14日德国物理学会在柏林的一次例会上,普朗克第一次引进"量子"这一概念.1911年洛伦兹在第一次索尔维会议上提出了创建微观世界力学的任务,物理学家们为此进行了不断的努力.人们试图从经典力学的原则出发对量子力学给出合理的解释,但都没有获得成功.什么是经典力学的原理?经典力学的创建者之一,法国数学家拉普拉斯认为,如果把宇宙间所有物体或者质点在给定时刻的坐标和速度的精确值告诉给一个物理学家的话,那么他就能预先描述世界在任何另外一个时刻的画面,不论是未来的和过去的.经典力学是铁一般的因果律在哲学上的体现,是用确定性模型方法描述的.也就是说,宇宙间所有事件的因果都是事先完全决定了的,单方向确定了的.经典力学和决定论哲学就是这样认识

世界的,他们认为可能与存在是一致的. 但是玻尔和海森堡以及当时认为量子跃迁思想是现实存在的客观反映的人们早就怀疑这一点了. 在玻尔的思维中,物理学变成了自然哲学. 他们想到了,为了给量子力学以正确的解释必须与经典力学和决定论哲学分手. 1927年2月,他们终于用随机数学的方法刻画了测不准原则和互补原则,给出了量子力学的正确描述. 经过16年的努力,新的微观世界力学出现了. 值得指出的是,在这一过程中,洛伦兹、卢瑟福、爱因斯坦这些物理学的革新家却都是新力学的反对者. 1927年秋,第五次索尔维会议上,主席还是洛伦兹,他对量子力学的随机解释并不满意. 他说"对我来说,电子就是粒子,它在每一给定的时刻都处于空间的确定点上. ……"他干脆拒绝了量子力学. 而爱因斯坦则试图要推翻微观世界随机性描述的不可避免性,他的名言是"上帝不掷骰子". 在这次会议上,爱因斯坦对新力学不断地发难,他的近友恩菲斯特对他说:"我真为你羞愧,爱因斯坦,你反对新的量子理论的做法和你的敌人反对相对论的做法一模一样!"著名的英国作家肖伯纳有一句名言:唯一的历史教训就是忘记了历史的教训. 甚至连伟大的爱因斯坦也未能成为这一真理的例外. 量子力学的产生在科学史上是一座丰碑,这也说明了随机数学具有方法论的重要性和自然哲学上的意义.

由于现实世界几乎一切可观察现象都具有随机性,所以与确定性数学相对,随机模型方法具有越来越重要的意义. 当我们用雷达观测飞机时,要测量飞机的高低角、方位角与距离,对这三个量的量测不可避免地带有误差. 影响量测精度的因素非常多:有来源于安装雷达载体的(海岸、飞机、船等);有来自目标的(目标发出的闪烁、噪声等干扰);有来自电磁波传递的(传递误差、多路误差等);有来自周围环境的(大气折射、大气中的电磁现象等),等等. 这些因素本身在不断地变化着,而影响它们变化的又有许多因素. 在用雷达观测飞机的过程中要不间断地进行量测. 人们常说影响每次量测精度的因素太多,太复杂了,所以无法知道每次量测的精确的数值. 那么能否在将来的某一天,提高了我们的科技水平后能够精确进行量测呢? 显然,这些人试图每次进行精确量测就如同要给电子轨道以精确描述的人一样陷入了决定论的泥坑. 事实上,问题不仅在于能否精确地知道每次量测误差的大小,问题还在于对实际工作人员来说需不需要知道这些,或至少在不知道这些时能否有效地利用雷达量测结果满足实际需要. 从每次量测误差的因果关系出发,试图求出精确的测量结果,是否认随机性的客观存在,是不正确的. 恩格斯在谈到决定论时,就批评过那种企图对一个豌豆荚中有几粒豌豆也要用因果关系加以说明的人,认为这"也还是没有从神学的自然观中走出来".(《自然辩证法》)在用雷达跟踪飞机的问题中,分析多次观测的结果在表面上杂乱无章的误差中存在着一定的规律性,随机数学方法正是透过这些表面的偶然性揭示过程的内在规律性的数学工具. 用雷达跟踪飞机是如此,我们用量测仪器仪表进行任何一种量测也同样是这样.

从20世纪50年代开始,特别是近年来,随机数学方法已广泛地应用到科学和工程的各个方面. 天文学中星云密度起伏,辐射传递的研究要用到随机场的理论. 物理学中研究电子光子级联过程的起伏时要用到波利亚过程及更新过程的理论. 放射性衰变,粒

子计数器,原子核照相乳胶中的经迹理论中要用到Poisson过程的理论.研究化学反应的时变率及影响这些时变率的因素问题,自动催化反应、单分子反应、双分子反应的动力学模型要用Markov过程来描述.在生物遗传学、生态学中都要用到随机模型方法.在研究群体增长问题时就有生灭型随机模型、两性增长模型、依龄增长模型、扩散模型、群体中的竞争模型与相克模型.群体迁移模型也已成功地应用随机过程方法.各种随机服务系统,如电话通讯、机器维修、路口交通信号控制、库存问题、水库调度、港口装卸、计算机网络的优化设计都涉及排队过程的应用.在经济学中,时间序列建模与预报对期货、股票、证券及金融投资分析都有重要的意义.在金融风险分析和保险精算学中用到更新过程及鞅论.甚至于政治学、社会学、历史学等人文、社会科学也都不同程度地应用随机数学方法,如研究人口流动、人才交流、父子职业变化问题时要建立随机数学模型.历史学与随机数学方法相结合并产生了新的边缘学科——度量历史学.

对于工科院校各专业,可以说不使用随机数学方法的领域已经不存在了.随机振动的理论、极值分布的理论在机械制造、汽车减震器的设计、船舶的结构设计、建筑结构的抗震研究、航天器的随机疲劳问题等问题中都已成为必要的工具.电子电路设计、系统的自动控制理论、信号处理中的适应性滤波、决策及特征判别等方面广泛地应用平稳过程及Markov过程.在化学工程方面,Markov过程用于受扩散控制的反应动力学研究及描述长链高分子的降解和某些树脂的重排列;在电化学中电镀过程控制、极板腐蚀问题都要用到随机过程.在土木工程设计与规划问题要用到可靠性理论,建筑构件受到随机荷载的响应问题、空气净化的控制问题以及环境工程方面的大量问题都要用到随机数学方法.在计算机科学、管理科学、技术经济与系统工程中大量应用随机数学方法则已经成为近年来最为活跃的方向了.

综上所述,加强随机数学的教学已经成为面向21世纪基础数学教学内容和课程体系改革的重要方面之一了.我们知道,从地理大发现和日心说,上一世纪60年代的地球村,到90年代,人类已进入"数字地球"——人类认识地球的第三次飞跃——的时代.从信息高速公路到数字地球,人类仅仅经历了10年.从地球的资源、生态、环境到可持续发展的战略,我们正面临新的挑战.大量的数字信息需要统计,需要分析处理,随机数学方法是重要的工具之一.要培养出能够适应新世纪发展的合格人才,适应当代科技高速发展的需要,改革研究生基础数学教育加强随机数学教学是非常重要的.

本教材是在天津大学研究生院改革研究生基础数学教育的精神指导下完成的,是为非数学专业的研究生编写的.在编写时,我们针对学生在各自专业上已接近或达到学科的前沿,但现代数学基础相对较弱的具体情况,在注意数学概念的形式逻辑推演的同时,特别强调数学概念的直观背景和实际应用.对抽象的数学概念给出其实际的工程背景,对复杂的数学符号和公式用直观的语言进行描述.希望能提高学生应用现代数学工具刻画各自方向的实际问题,并利用数学的方法去分析和解决科研问题的能力.

作为随机过程的基础知识,全书取材较为全面,共分5章.第1章是概率论的基本知识,其主要内容是按照随机过程这门课程的要求,复习概率论的基本概念和有关结

论,并补充了一些必要的概率论知识.第2章主要讲述随机过程的基本概念和Poisson过程的基础知识.第3章讲述Markov链,第4章讲述平稳过程.第5章讲述时间序列分析的基本理论、方法和应用.除第1章外,每章后附有习题,完成这些习题对掌握这门课程是必不可少的.第2章的Poisson过程及第3章与第4、第5章内容相对独立,读者可根据专业的需要对内容进行适当取舍.

由于作者水平所限,书中错误和疏漏之处,尽请读者批评指正.

作者

2003年2月于天津大学

目录

第 1 章　概率论的基本知识

1.1　概率空间

概率论中的一个基本概念是随机试验，它是指其结果不能事先准确地预言，并且在相同条件下可以重复进行的那种试验.

随机试验的所有可能结果组成的集合称为这个试验的**样本空间**或**基本事件空间**，记为 Ω，Ω 中的元素 ω 称为**样本点**或**基本事件**，即 $\Omega=\{\omega$ 全体$\}$.

样本空间 Ω 的子集称为**随机事件**，简称**事件**. 称事件 A 发生，当且仅当 A 中的一个样本点出现. 样本空间 Ω 也是一个事件，称为必然事件，空集 $\varnothing$ 称为不可能事件.

由于事件是集合，因此集合的运算（并、交、差、上极限、下极限、极限等）都适用于事件.

在实际问题中，我们并不总是对所有的事件（Ω 的一切子集）都感兴趣，常常仅关心某些事件（Ω 的某些子集）及其发生的可能性大小（概率），这样便导致了事件域的概念.

1.1.1　事件域

定义 1.1.1　设 Ω 是样本空间，$\mathscr{F}$ 是由 Ω 的一些子集构成的集类（族），如果它满足：

(i) $\Omega\in\mathscr{F}$，

(ii) 若 $A\in\mathscr{F}$，则 $\bar{A}\in\mathscr{F}$，

(iii) 若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$，则 $\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{F}$，

则称 $\mathscr{F}$ 为**事件域**，$\mathscr{F}$ 中的元素称为**事件**.

由定义 1 易知：

(1) $\varnothing\in\mathscr{F}$，

(2) 若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$，则 $\bigcup\limits_{i=1}^{n}A_i\in\mathscr{F}$，$\bigcap\limits_{i=1}^{n}A_i\in\mathscr{F}$，$\bigcap\limits_{i=1}^{\infty}A_i\in\mathscr{F}$，

(3) 若 $A,B\in\mathscr{F}$，则 $A-B\in\mathscr{F}$.

1.1.2　概率

定义 1.1.2　设 Ω 是样本空间，$\mathscr{F}$ 是 Ω 的一个事件域，定义在 $\mathscr{F}$ 上的实值集函数 $P(A)$ 如果满足：

(i) $\forall A\in\mathscr{F}$，有 $P(A)\geqslant 0$，

(ii) $P(\Omega)=1$,

(iii) 若 $A_i\in\mathscr{F}, i=1,2,\cdots$, 且 $A_iA_j=\varnothing(i\neq j)$, 则

$$P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i),$$

称 $P(\cdot)$ 为 $\mathscr{F}$ **上的概率**, $P(A)$ 为**事件 A 的概率**, 并且称三元总体 $(\Omega,\mathscr{F},P)$ 为**概率空间**.

以下论及的事件都是指某一事件域中的事件.

由定义 1.1.2 易知概率 P 具有如下性质.

性质 1 $P(\varnothing)=0$. (1)

性质 2 若 $A_i\in\mathscr{F}, i=1,2,\cdots,n, A_iA_j=\varnothing(i\neq j)$, 则

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i). \tag{2}$$

性质 3 对任意事件 A, 有

$$P(A)=1-P(\bar{A}). \tag{3}$$

性质 4 若 $A\subset B$, 则

$$P(B-A)=P(B)-P(A), \tag{4}$$

$$P(A)\leqslant P(B). \tag{5}$$

性质 5 对任意 $A_i\in\mathscr{F}, i=1,2,\cdots,n$, 有

$$P(\bigcup_{i=1}^{n}A_i)\leqslant\sum_{i=1}^{n}P(A_i), \tag{6}$$

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n). \tag{7}$$

性质 6 设 $A_n\in\mathscr{F}, n=1,2,\cdots$, 若 $A_n\subset A_{n+1}$, 则

$$\lim_{n\to\infty}P(A_n)=P(\bigcup_{n=1}^{\infty}A_n); \tag{8}$$

若 $A_n\supset A_{n+1}$, 则

$$\lim_{n\to\infty}P(A_n)=P(\bigcap_{n=1}^{\infty}A_n). \tag{9}$$

1.1.3 条件概率

定义 1.1.3 设 $(\Omega,\mathscr{F},P)$ 是一个概率空间, $B\in\mathscr{F}$, 且 $P(B)>0$, 对任意 $A\in\mathscr{F}$, 记

$$P(A|B)=\frac{P(AB)}{P(B)}, \tag{10}$$

则称 $P(A|B)$ 为在事件 B 发生条件下事件 A 发生的**条件概率**.

由条件概率的定义可得到:

(1) 乘法公式 设 $A,B\in\mathscr{F}, P(B)>0$, 则

$$P(AB)=P(B)P(A|B). \tag{11}$$

一般地, 若 $A_i\in\mathscr{F}, i=1,2,\cdots,n$, 且 $P(A_1A_2\cdots A_{n-1})>0$, 则有

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}). \tag{12}$$

(2) 全概率公式 设$(\Omega,\mathscr{F},P)$是概率空间,$A\in\mathscr{F}$,$B_i\in\mathscr{F}$,$i=1,2,\cdots,n$,$B_iB_j=\varnothing(i\neq j)$,且$\bigcup\limits_{i=1}^{n}B_i=\Omega$,$P(B_i)>0$,则

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i). \tag{13}$$

(3) 逆概率(Bayes)公式 设$(\Omega,\mathscr{F},P)$是概率空间,$A\in\mathscr{F}$,$B_i\in\mathscr{F}$,$i=1,2,\cdots,n$,$B_iB_j=\varnothing(i\neq j)$,且$\bigcup\limits_{i=1}^{n}B_i=\Omega$,$P(B_i)>0$,$P(A)>0$,则

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum\limits_{j=1}^{n}P(B_j)P(A|B_j)}. \tag{14}$$

1.1.4 事件的独立性

定义1.1.4 设$(\Omega,\mathscr{F},P)$是概率空间,$A,B\in\mathscr{F}$,若

$$P(AB)=P(A)P(B), \tag{15}$$

则称事件A,B是**相互独立**的.

一般地,设$A_i\in\mathscr{F}$,$i=1,2,\cdots,n$,若对任意的$m(2\leqslant m\leqslant n)$及任意的$1\leqslant k_1<k_2<\cdots<k_m\leqslant n$,都有

$$P(A_{k_1}A_{k_2}\cdots A_{k_m})=P(A_{k_1})P(A_{k_2})\cdots P(A_{k_m}), \tag{16}$$

则称事件$A_1,A_2,\cdots,A_n$是相互独立的.

1.2 随机变量

1.2.1 随机变量

在研究随机现象时常常会遇到各种变量,这些变量的取值随着试验结果的变化而变化,这样的变量就称为随机变量.对于随机变量我们不仅关心它取什么值,而且还关心它取值的概率大小.下面我们给出随机变量的定义.

定义1.2.1 设$(\Omega,\mathscr{F},P)$是一个概率空间,X是定义在Ω上的一个实值函数$X=X(\omega)$.若对一切实数x,都有

$$\{\omega:X(\omega)\leqslant x\}\in\mathscr{F},$$

则称X是一个**随机变量**,并称

$$F(x)=P\{\omega:X(\omega)\leqslant x\},\quad -\infty<x<\infty \tag{1}$$

为X的**分布函数**.

分布函数$F(x)$具有性质:

(1) $F(x)$是$(-\infty,\infty)$上的单调非降函数,即对一切实数$x_1<x_2$,有$F(x_1)\leqslant F(x_2)$;

(2) $F(x)$ 是 $(-\infty,\infty)$ 上的右连续函数,即对一切实数 x_0,有 $\lim\limits_{x\to x_0+0}F(x)=F(x_0)$;

(3) $0\leqslant F(x)\leqslant 1$,且 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$,

$$F(\infty)=\lim_{x\to\infty}F(x)=1.$$

可以证明,定义在 $\mathbf{R}=(-\infty,\infty)$ 上的实值函数 $F(x)$,若满足上述三条性质,则必存在概率空间及其上的一个随机变量 X,其分布函数是 $F(x)$.

常见的随机变量有两种类型:离散型随机变量和连续型随机变量.

若随机变量 X 只取有限个或可列无穷多个可能值,则这样的随机变量称为离散型随机变量. 设离散型随机变量 X 的可能取值为 $x_k(k=1,2,\cdots)$,则 X 取各可能值的概率

$$P\{X=x_k\}=p_k,\ k=1,2,\cdots \tag{2}$$

称为 **X 的分布列**,其分布函数为

$$F(x)=\sum_{x_k\leqslant x}p_k. \tag{3}$$

设 $F(x)$ 是随机变量 X 的分布函数,若存在非负函数 $f(x)$,使得对任何实数 x,有

$$F(x)=\int_{-\infty}^{x}f(t)\,\mathrm{d}t, \tag{4}$$

则称 X 为**连续型随机变量**,并称 $f(x)$ 是 X 的**概率分布密度函数**,简称为 X 的**概率密度**或**分布密度**.

1.2.2 随机向量

定义 1.2.2 设 $(\Omega,\mathscr{F},P)$ 是一概率空间,$X_1(\omega),X_2(\omega),\cdots,X_n(\omega)$ 是定义在这个概率空间上的 n 个随机变量,则称

$$X(\omega)=(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$$

为 $(\Omega,\mathscr{F},P)$ 上的一个 **n 维随机向量**(或 **n 维随机变量**),简记为 $X=(X_1,X_2,\cdots,X_n)$,并称 n 元函数

$$F(x_1,\cdots,x_n)=P\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\},\ x_1,\cdots,x_n\in\mathbf{R} \tag{5}$$

为 n 维随机向量 $(X_1,X_2,\cdots,X_n)$ 的**(联合)分布函数**.

若随机向量 $X=(X_1,\cdots,X_n)$ 只取有限个或可列无穷多个向量值,则称 X 为离散型随机向量. 设 $X=(X_1,\cdots,X_n)$ 的所有可能取值为 $(x_{1i_1},\cdots,x_{ni_n})$, $i_1,i_2,\cdots,i_n=1,2,\cdots$,则称概率

$$p_{i_1i_2\cdots i_n}=P\{X_1=x_{1i_1},\cdots,X_n=x_{ni_n}\},\ i_1,\cdots,i_n=1,2,\cdots \tag{6}$$

为 $X=(X_1,\cdots,X_n)$ 的**联合分布列**.

若随机向量 $X=(X_1,\cdots,X_n)$ 的分布函数为 $F(x_1,\cdots,x_n)$,且存在非负函数 $f(x_1,\cdots,x_n)$ 使得

$$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n}f(t_1,t_2,\cdots,t_n)\,\mathrm{d}t_1\mathrm{d}t_2\cdots\mathrm{d}t_n, \tag{7}$$

则称 $X=(X_1,\cdots,X_n)$ 为连续型随机向量,并称函数 $f(x_1,\cdots,x_n)$ 为 $X=(X_1,\cdots,X_n)$ 的**联合概率分布密度函数**,简称**联合概率密度**.

1.2.3 随机变量的独立性

定义 1.2.3 设 $X_1,X_2,\cdots,X_n$ 是 n 个定义在同一概率空间 $(\Omega,\mathscr{F},P)$ 上的随机变量,若对于任意的 n 个实数 $x_1,x_2,\cdots,x_n\in\mathbf{R}$,都有

$$P\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\}=P\{X_1\leqslant x_1\}\cdots P\{X_n\leqslant x_n\},\tag{8}$$

即 $(X_1,\cdots,X_n)$ 的联合分布函数等于各随机变量分布函数的乘积

$$F(x_1,\cdots,x_n)=F_{X_1}(x_1)\cdots F_{X_n}(x_n),\tag{9}$$

则称 $X_1,\cdots,X_n$ 是**相互独立**的.

设 $(X_1,\cdots,X_n)$ 是离散型随机向量,则 $X_1,\cdots,X_n$ 相互独立的充要条件是:对任意 $i_1,i_2,\cdots,i_n=1,2,\cdots$,有

$$P\{X_1=x_{1i_1},\cdots,X_n=x_{ni_n}\}=P\{X_1=x_{1i_1}\}\cdots P\{X_n=x_{ni_n}\}.\tag{10}$$

设 $(X_1,\cdots,X_n)$ 是连续型随机向量,$f(x_1,\cdots,x_n)$ 及 $f_{X_1}(x_1),\cdots,f_{X_n}(x_n)$ 分别是 $(X_1,\cdots,X_n)$ 的联合概率密度及 $X_1,\cdots,X_n$ 的概率密度,则 $X_1,\cdots,X_n$ 相互独立的充要条件是

$$f(x_1,\cdots,x_n)=f_{X_1}(x_1)\cdots f_{X_n}(x_n).\tag{11}$$

可以证明,若随机变量 $X_1,\cdots,X_n$ 相互独立,$f_i(x)$ 为 Borel 可测函数,$i=1,2,\cdots,n$,则 $f_1(X_1),f_2(X_2),\cdots,f_n(X_n)$ 也相互独立.

1.3 随机变量的数字特征

分布函数是随机变量概率分布的完整描述,但在实际问题中求一个随机变量的分布函数往往不是一件容易的事.而有些实际问题,只要求知道描述随机变量的某些概率特征,并不需要知道其分布函数.本节介绍的随机变量的数学期望、方差等正是刻画随机变量某些概率特征的数值指标.

定义 1.3.1 设 $F(x)$ 是随机变量 X 的分布函数,若 $\int_{-\infty}^{\infty}|x|\mathrm{d}F(x)<\infty$,则称

$$EX=\int_{-\infty}^{\infty}x\mathrm{d}F(x)\tag{1}$$

为 X 的**数学期望**.

若 X 是离散型随机变量,其分布列为

$$p_k=P\{X=x_k\},\ k=1,2,\cdots,$$

且 $\sum_k|x_k|p_k<\infty$,则 X 的**数学期望**是

$$EX=\sum_k x_kp_k.\tag{2}$$

若 X 是连续型随机变量,其概率密度为 $f(x)$,且 $\int_{-\infty}^{\infty}|x|f(x)\mathrm{d}x<\infty$,则 X 的**数学期望**是

$$EX=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x.\tag{3}$$

定义 1.3.2 设 X 是随机变量，若 EX 存在，且 $E(X-EX)^2<\infty$，则称

$$DX=E(X-EX)^2 \tag{4}$$

为随机变量 X 的**方差**.

定义 1.3.3 设 X,Y 是随机变量，$EX^2<\infty$，$EY^2<\infty$，则称

$$\mathrm{cov}(X,Y)=E[(X-EX)(Y-EY)] \tag{5}$$

为 X 与 Y 的**协方差**，并且称

$$\rho_{XY}=\frac{\mathrm{cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}\quad(DX>0,DY>0) \tag{6}$$

为 X 与 Y 的**相关系数**.

若 $\rho_{XY}=0$，则称 X 与 Y **不相关**.

随机变量的数学期望和方差有如下性质：

(1) $E(aX+bY)=aEX+bEY$，其中 a,b 是常数；

(2) 若 X,Y 独立，则 $E(XY)=E(X)\cdot E(Y)$；

(3) 若 X,Y 独立，则 $D(aX+bY)=a^2DX+b^2DY$，其中 a,b 是常数；

(4) 若 n 维随机向量 $(X_1,\cdots,X_n)$ 的联合分布函数为 $F(x_1,\cdots,x_n)$，$g(x_1,\cdots,x_n)$ 是 n 元连续函数，则

$$E[g(X_1,\cdots,X_n)]=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\cdots,x_n)\,\mathrm{d}F(x_1,\cdots,x_n);$$

(5) Schwarz 不等式　若 $EX^2<\infty$，$EY^2<\infty$，则

$$[E(XY)]^2\leqslant E(X^2)E(Y^2).$$

相关系数 ρ_{XY} 有下列性质：

(1) $-1\leqslant\rho_{XY}\leqslant1$；

(2) $|\rho_{XY}|=1$ 的充要条件是 X 与 Y 以概率 1 线性相关，即存在常数 a,b，且 $a\neq0$ 使

$$P\{Y=aX+b\}=1;$$

(3) 若 X 与 Y 相互独立，则 $\rho_{XY}=0$.

1.4 条件数学期望

1.4.1 边缘分布与条件分布

定义 1.4.1 设 (X,Y) 为二维离散型随机向量，联合分布列为

$$p_{ij}=P\{X=x_i,Y=y_j\},\ i,j=1,2,\cdots,$$

则称

$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{\infty}P\{X=x_i,Y=y_j\}=\sum_{j=1}^{\infty}p_{ij} \tag{1}$$

为 (X,Y) 关于 X 的**边缘分布列**；并称

$$p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{\infty}P\{X=x_i,Y=y_j\}=\sum_{i=1}^{\infty}p_{ij} \tag{2}$$

为(X,Y)关于 Y 的**边缘分布列**.

定义 1.4.2 设(X,Y)具有联合概率密度$f(x,y)$,记

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y, \tag{3}$$

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x, \tag{4}$$

则称$f_X(x)$,$f_Y(y)$分别为(X,Y)关于 X 和 Y 的**边缘概率分布密度**.

定义 1.4.3 设(X,Y)的联合分布列为

$$p_{ij}=P\{X=x_i,Y=y_j\},\ i,j=1,2,\cdots,$$

且$P\{Y=y_j\}>0$,则称

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}}\quad(i=1,2,\cdots) \tag{5}$$

为在 $Y=y_j$ 条件下随机变量 X 的**条件分布列**.

同样,可定义

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}}\quad(j=1,2,\cdots) \tag{6}$$

为在 $X=x_i$ 条件下随机变量 Y 的**条件分布列**.

定义 1.4.4 设(X,Y)有联合概率密度$f(x,y)$和边缘概率密度$f_Y(y)$,$f(x,y)$和$f_Y(y)$连续,且$f_Y(y)>0$,则称

$$f(x|y)=\frac{f(x,y)}{f_Y(y)} \tag{7}$$

为在条件 $Y=y$ 下 X 的**条件概率密度**.

同样,可定义在条件 $X=x$ 下随机变量 Y 的条件概率密度为

$$f(y|x)=\frac{f(x,y)}{f_X(x)}. \tag{8}$$

1.4.2 条件数学期望

定义 1.4.5 设(X,Y)是二维离散型随机向量,在 $Y=y_j$ 条件下 X 的条件分布列为$P\{X=x_i|Y=y_j\}$,$i=1,2,\cdots$,如果$\sum\limits_{i=1}^{\infty}|x_i|P\{X=x_i|Y=y_j\}<\infty$,则称

$$E(X|Y=y_j)=\sum_{i=1}^{\infty}x_iP\{X=x_i|Y=y_j\},\ j=1,2,\cdots \tag{9}$$

为 X 对 Y 在 $Y=y_j$ 条件下的**条件数学期望**.

定义 1.4.6 设(X,Y)是二维连续型随机向量,在 $Y=y$ 条件下 X 的条件概率密度为$f(x|y)$,如果$\int_{-\infty}^{\infty}|x|f(x|y)\,\mathrm{d}x<\infty$,则称

$$E(X|Y=y)=\int_{-\infty}^{\infty}xf(x|y)\,\mathrm{d}x \tag{10}$$

为 X 对 Y 在 $Y=y$ 条件下的**条件数学期望**.