工业和信息产业科技与教育专著出版资金
cbjj.phei.com.cn

工业和信息产业科技与教育专著出版资金资助出版

应用随机过程

Applied Stochastic Processes

◆ 李晓峰　唐　斌　舒　畅　等编著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY　http://www.phei.com.cn

工业和信息产业科技与教育专著出版资金资助出版

应用随机过程

李晓峰　唐斌　舒畅　傅志中　周宁　编著

電子工業出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书主要讨论随机过程的基础理论和应用方法。全书共七章，内容包括：概率论基础，随机过程基础，泊松过程及其推广，马尔可夫过程，二阶矩过程及其均方分析，平稳过程，以及高阶统计量与非平稳过程等。

本书强调随机过程的基础理论、物理意义与应用方法，注重理论联系实际，力求从概念的物理背景、理论的逻辑推导与应用的典型例子三个方面加以阐述。内容全面，叙述清楚，例题与图示丰富，便于教学与自学。

本书以初等概率论、高等数学的基本知识为基础，可以作为理工科高等院校有关专业研究生与高年级本科生的教学用书，也可供有关领域的师生、科研和工程技术人员参考。

图书在版编目（CIP）数据

应用随机过程 / 李晓峰等编著. —北京：电子工业出版社，2013.8
ISBN 978-7-121-20687-0

Ⅰ. ①应… Ⅱ. ①李… Ⅲ. ①随机过程 Ⅳ. ①O211.6

中国版本图书馆 CIP 数据核字（2013）第 128912 号

责任编辑：韩同平　　特约编辑：张庆杰
印　　刷：北京市李史山胶印厂
装　　订：北京市李史山胶印厂
出版发行：电子工业出版社
　　　　　北京市海淀区万寿路 173 信箱　　邮编：100036
开　　本：787×1092　1/16　印张：12　字数：350 千字
印　　次：2013 年 8 月第 1 次印刷
印　　数：2 000 册　　定价：39.90 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

前　言

本书是作者在多年来从事研究生“随机过程及其应用”课程的教学与研究的经验基础上，经对所用讲义进行反复修改与补充编写而成的。它主要讨论随机过程的基础理论和应用方法，可以作为研究生与高年级本科生的教材或教学参考书。本书内容包括：概率论基础，随机过程基础，泊松过程及其推广，马尔可夫过程，二阶矩过程及其均方分析，平稳过程，以及高阶统计量与非平稳过程。

随机过程是一门应用性很强的学科。它在各个科技领域的应用中得到充分的发展，形成了许多富有启发性的模型与典型的方法。单纯地照搬数学理论难以理解其精髓，也无法有效地解决具体问题。因此，本书强调理论联系实际，力求从概念的物理背景、理论的逻辑推导与应用的典型例子三个方面加以阐述。了解直观背景与体会物理意义，有助于读者掌握数学概念的本质，理解其理论体系，进而形成正确的思路与方法。

本书考虑到理工科专业的学生的数学背景，编写上以微积分、线性代数与初等概率论为基础，采用循序渐进、简明易懂的方式。书中第 1 章专门用于复习与总结概率论的基础知识，可作为有用的“在线”资料；同时，它也扩充一些重要的知识点。为了方便施教与自学，书中既设计了大量的举例，又绘制了丰富的图示，还在各章末配备了充分的习题。本书建议的学时数为 40～60 学时，教师可以根据具体的教学需要，适当选择其中一些内容组织教学。

本书编写过程中，李乐民院士给予了指导。本书得到电子科技大学研究生院、通信与信息工程学院、电子工程学院等师生的大力支持与帮助。作者也参考了许多文献书籍，从中获得了不少有益的启示。本书的出版得到电子工业出版社的大力支持，韩同平编辑为本书花费了不少精力。作者在此一并向他（她）们表示衷心的感谢。

本书由李晓峰、唐斌、舒畅、傅志中、周宁编著。全书由李晓峰统编定稿。限于作者水平，书中谬误与疏漏在所难免，恳请读者批评指正。

本书得到电子科技大学“十二五”规划研究生教材建设资助出版。

编著者
于电子科技大学
（xfli@uestc.edu.cn）

反侵权盗版声明

举报电话：（010）88254396；（010）88258888
传　　真：（010）88254397
E-mail：dbqq@phei.com.cn
通信地址：北京市海淀区万寿路 173 信箱
　　　　　电子工业出版社总编办公室
邮　　编：100036

目　录

第 1 章　概率论基础

概率论的基本知识读者已经学习过了，它们与本书的内容密切关联。本章将简明地复习与总结这些知识，同时，也会扩充与加深一些理论知识，比如，事件域、R-S 积分、条件数学期望与特征函数等。

1.1　概 率 空 间

随机性的研究起源于 17 世纪的赌博与机会游戏，这些研究结果渐渐地形成了一个有趣且深刻的理论。随机现象既表现出个别时的不确定性，又呈现出大数量时的有规律性。这种特性在自然界与社会生活中普遍存在。现代概率论形成于 20 世纪 30 年代，主要归功于柯尔莫格洛夫(A. Kolmogoroff)和列维(Levy)建立了概率论、集合与实变函数理论之间的紧密联系。在今天的信息社会里，信息的度量建立在消息的概率的基础上，使得概率论及其相关理论得到广泛的研究与应用。

1.1.1　概率

在概率论中，对随机现象做出的观察与科学实验被抽象为**随机试验**(Random Experiment)，它具有下述特性：(1)可以在相同条件下重复进行；(2)全部的可能结果是事先知道的；(3)每次试验的结果不可预知。

随机试验的全部可能结果构成的集合称为**样本空间**(Sample space)，记为Ω。Ω的元素是单个可能结果，称为**样本点**(Sample point)，记为ξ_i，$\xi_i\in\Omega$。**事件**(Event)是试验中“人们感兴趣的结果”构成的集合，是Ω的子集，常用大写字母A、B、C等表示。

由Ω的若干子集构成的集合称为**集类**，通常用花体字母$\mathcal{A}$、$\mathcal{B}$、$\mathcal{C}$等表示。

深入研究发现，并不是在所有的Ω子集上都能够方便地定义概率，一般只在满足一定条件的集类上研究概率及其性质，为此引入了σ域的概念。

定义 1.1　设$\mathcal{F}$是由样本空间Ω的子集构成的非空集类，它满足

（1）若$A\in\mathcal{F}$，则$\overline{A}\in\mathcal{F}$；

（2）若$A_n\in\mathcal{F}$，$n=1,2,3,\cdots$，则$\bigcup\limits_{n=1}^{\infty}A_n\in\mathcal{F}$。

则称它为σ**域**(或σ代数)，称$(\Omega,\mathcal{F})$为**可测空间**。

例1.1　集类$\mathcal{F}_0=(\varnothing,\Omega)$与$\mathcal{F}_A=(\varnothing,A,\overline{A},\Omega)$是$\sigma$域；而$\mathcal{F}_0=(\varnothing,A,\Omega)$与$\mathcal{F}_A=(A,\overline{A})$不是。

容易验证，σ域包含空集与全集，并对可列次交、并、差等运算是封闭的。即，

性质 1　σ域$\mathcal{F}$的基本性质：

（1）$\Omega\in\mathcal{F}$，$\varnothing\in\mathcal{F}$；

（2）$\forall A,B\in\mathcal{F}$，则$A\cap B\in\mathcal{F}$，$A-B\in\mathcal{F}$；

（3）$\forall A_n\in\mathcal{F}$，$n=1,2,3,\cdots$，则$\bigcap\limits_{n=1}^{\infty}A_n\in\mathcal{F}$。

定义 1.2 设$(\Omega,\mathcal{F})$为可测空间，P 为定义在$\mathcal{F}$上的实值函数，若满足

（1）非负性：$\forall A\in\mathcal{F}$，$P(A)\geqslant 0$；

（2）归一性：$P(\Omega)=1$；

（3）可列可加性：若 $A_i\in\mathcal{F}$，$i=1,2,3,\cdots$，且 $\forall i\neq j$，$A_iA_j=\varnothing$，则

$$P\left(\bigcup_i A_i\right)=\sum_i P(A_i)$$

则称 P 为$(\Omega,\mathcal{F})$上的一个**概率测度(Probability measure)**，简称概率。称(Ω,F,P)为**概率空间(Probability space)**，称$\mathcal{F}$为**事件域**。若 $A\in\mathcal{F}$，称 A 为**随机事件(Random event)**，简称事件；称 $P(A)$ 为**事件 A 的概率**。

事件的概率用于刻画事件出现可能性的大小。在实际问题中，事件 A 出现的可能性直观地由其**相对频率**(Relative frequency)来度量，认为

$$P(A)\approx\frac{\text{试验中}A\text{出现的次数}}{\text{总试验次数}}=\frac{n_A}{n}\qquad(n\text{很大})\tag{1.1.1}$$

容易看出，概率的定义与相对频率的客观特性相吻合。

性质 2 事件概率的基本性质

（1）$P(\varnothing)=0$

（2）$0\leqslant P(A)\leqslant 1$

（3）$P(\overline{A})=1-P(A)$

（4）$P(A\cup B)=P(A)+P(B)-P(AB)$

（5）$P(A-B)=P(A\cap\overline{B})=P(A)-P(AB)$

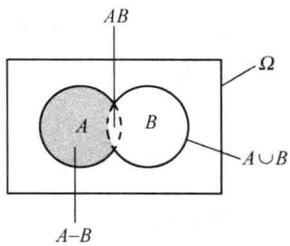

图 1.1.1 文氏图

借助集合的文氏图(Venn diagram)可以形象化地理解这些性质(见图 1.1.1)。下面不加证明地再给出概率的两个重要性质。

性质 3(Jordan 公式) 对 $A_i\in\mathcal{F}$，$i=1,2,3,\cdots$，有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n+1}P(A_1A_2\cdots A_n)\tag{1.1.2}$$

给定事件序列 $\{A_i,i=1,2,3,\cdots\}$，(1) 若 $A_i\subset A_{i+1},i\geqslant 1$，称为**单调增序列**，定义 $\lim\limits_{i\to\infty}A_i=\bigcup\limits_{n=1}^{\infty}A_n$；(2) 若 $A_i\supset A_{i+1}$，$i\geqslant 1$，称为**单调减序列**，定义 $\lim\limits_{i\to\infty}A_i=\bigcap\limits_{n=1}^{\infty}A_n$。

性质 4(连续性定理) 若 $\{A_i,i=1,2,3,\cdots\}$ 是单调增序列(或减序列)，则

$$\lim_{i\to\infty}P(A_i)=P\left(\lim_{i\to\infty}A_i\right)\tag{1.1.3}$$

基于概率论解决实际问题的基本思路是：首先为问题设计合适的随机试验模型，建立样本空间 Ω；再围绕感兴趣的事件确定事件域$\mathcal{F}$；而后合理地设定其中一些基础事件的概率；最后，由这些概率分析出我们感兴趣的事件的概率特性，从而解决所关心的问题。

例 1.2 分析掷均匀骰子问题。

解： 掷骰子试验的结果是：1 至 6 面，我们用 1 至 6 个数字表示。因此，

（1）$\Omega=\{1,2,3,4,5,6\}$

（2）$\mathcal{F}=\{\varnothing,\Omega,\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},$

$\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{1,6\},\{2,3\},\{2,4\},\cdots$

$\{1,2,3\},\{1,2,4\},\{1,2,5\},\{1,2,6\},\{1,3,4\},\{1,3,5\},\cdots$

$$\{1,2,3,4\},\{1,2,3,5\},\{1,2,3,6\},\{1,2,4,5\},\{1,2,4,6\},\cdots$$
$$\{1,2,3,4,5\},\{1,2,3,4,6\},\{1,2,3,5,6\},\{1,2,4,5,6\},\cdots\}$$

显然，其中的事件是样本的各种组合。

（3）由骰子的均匀特性可得，每个面的基本概率为$1/6$；而且，$P\{\varnothing\}=0$，$P\{\Omega\}=1$。进而，$\forall A\in\mathcal{F}$，$P(A)=k/6$，$k\in[0,6]$为事件A包含的样本点数。

在确定事件域时，通常最关心的是包含感兴趣事件的最小σ域。设$\mathcal{A}$是所有感兴趣事件构成的集类，一切包含$\mathcal{A}$的σ域的交，称为$\mathcal{A}$生成的σ域，记为$\sigma(\mathcal{A})$，它就是包含$\mathcal{A}$的最小σ域。例如，如果上述例题中我们只关注骰子 1 号面出现的事件，即$\mathcal{A}=\{\{1\}\}$，于是，$\sigma(\mathcal{A})=\{\varnothing,\Omega,\{1\},\{2,3,4,5,6\}\}$。

当样本空间为实数区间时，博雷尔域是研究中常用的事件域，一维**博雷尔(Borel)域**定义为：包含R(实数集)上所有形如$(-\infty,a]$的最小σ域，记为$\mathcal{B}=\sigma\left((-\infty,a],\forall a\in R\right)$。

例 1.3 分析测量某随机电压的问题，假定该电压在$[0,1]$(单位：伏特)上是等可能取值的。

解：本测量可视为一个随机试验，结果为$\xi\in[0,1]$。深入的分析可见，区间$[0,1]$上存在某些点集不宜作为事件。而且，由区间无限稠密性可知，$\forall\xi\in\Omega$，$P\{\xi\}=0$，这无法有效地进行描述。但是，如果选取$\{\xi\leqslant u\}$型子集为基础事件族(其中u为参变量)，由它们的交、并等运算可以构造出所有有用的事件。于是，设：

（1）$\Omega=[0,1]$；

（2）$\mathcal{F}=\mathcal{B}[0,1]=\sigma\left([0,u],\forall u\in[0,1]\right)$，称为局限在$[0,1]$上的 Borel 域。

（3）定义概率测度：考虑 Borel 域上的基础事件族$\{0\leqslant\xi\leqslant u\}$，其中参变量$u\in[0,1]$。由于电压取值$\xi$是等可能的，因此

$$P\{0\leqslant\xi\leqslant u\}=P\{\xi\in[0,u]\}=u \tag{1.1.4}$$

而其他各种事件的概率都可以由此计算。比如

$$\begin{aligned}P\{\xi\in(u,u+\Delta u]\}&=P\{\{0\leqslant\xi\leqslant u+\Delta u\}-\{0\leqslant\xi\leqslant u\}\}\\&=P\{0\leqslant\xi\leqslant u+\Delta u\}-P\{0\leqslant\xi\leqslant u\}\\&=(u+\Delta u)-u=\Delta u\end{aligned}$$

1.1.2 条件概率与独立性

定义 1.3 给定概率空间$(\Omega,\mathcal{F},P)$与事件A,B，若$P(A)>0$，可以定义**条件事件**和**条件概率**(Conditional probability)，如下

$$B\,|\,A=\text{事件 }A\text{ 发生条件下的事件 }B$$

$$P(B|A)=\frac{P(AB)}{P(A)},\quad P(A)>0 \tag{1.1.5}$$

由条件概率的定义可以得到乘法公式并推广为链式法则，如下：

$$P(AB)=P(A)P(B|A)\qquad P(A)>0 \tag{1.1.6}$$

$$P\left(A_1A_2\cdots A_n\right)=P\left(A_1\right)P\left(A_2|A_1\right)P\left(A_3|A_1A_2\right)\cdots P\left(A_n|A_1A_2\cdots A_{n-1}\right) \tag{1.1.7}$$

其中，$P\left(A_1A_2\cdots A_{n-1}\right)>0$。

事件组$A_i\in F,i=1,2,3,\cdots,n$，若满足：

$$(1)\ \forall i \neq j,\ A_iA_j = \varnothing; \qquad (2)\ \bigcup_{i=1}^{n} A_i = \Omega$$

则称该事件组为样本空间的一个**完备事件组或分割(Partition)**。完备事件组是既彼此互斥又可以完整地拼成Ω的事件组。若$A_i(i=1,2,\cdots,n)$是完备组，任取事件$B \in \mathcal{F}$，有

（1）**全概率公式**(见图 1.1.2)

$$P(B) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i) \tag{1.1.8}$$

全概率公式表明："全部"的概率是各个"部分"的概率按比例构成的。

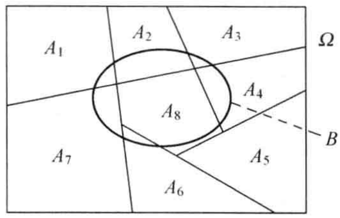

图 1.1.2　全概率公式

（2）**贝叶斯(Bayes)公式**

$$P(A_k|B) = \frac{P(A_kB)}{P(B)} = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)} \qquad k=1,2,\cdots,n \tag{1.1.9}$$

贝叶斯公式在研究因果推测、信息传输与检测等问题中有着重要的应用。通常$P(A_i)$称为**先验**(A priori)**概率**，$P(B|A_k)$称为**转移**(Transition)**概率**，$P(A_k|B)$称为**后验**(A posteriori)**概率**。考虑一下因果推测问题：A_i是m个原因事件，B是某种结果事件，贝叶斯公式正是基于结果B推测某种起因A_k的可能性的方法。

独立(Independence)的概念用于描述事件的发生不依赖于条件的特性，即，$P(B|A) = P(B)$。因此事件A与B独立等价地定义为

$$P(AB) = P(A)P(B)$$

多个事件彼此独立的定义为：

定义 1.4　设$A_i \in \mathcal{F}$，$i=1,2,\cdots,n$，若对于任意$m(0 \leqslant m \leqslant n)$与$m$个任意整数$1 \leqslant k_1 < k_2 \cdots < k_m \leqslant n$，满足

$$P(A_{k_1}A_{k_2}\cdots A_{k_m}) = P(A_{k_1})P(A_{k_2})\cdots P(A_{k_m}) \tag{1.1.10}$$

则称$A_1, A_2, \cdots, A_n$彼此独立。

可见多个事件的独立要求它们两两独立，三三独立，以及任意$m(\leqslant n)$个都独立。

例 1.4　盒中有形状相同、编号为$1 \sim N$的小球各 1 只，每次随机取出 1 个不再放回。对于$k \in [1, N]$，求：

（1）在第k次时(首次)摸到 1 号球的概率？

（2）前k次能摸到 1 号球的概率？

解："摸球问题"中最基本的事件概率是："从M个球中摸到 1 个的概率"为$1/M$。我们依据具体问题将有关事件进行恰当的分解处理。

（1）令B_k为第k次$(1 \leqslant k \leqslant N)$摸到 1 号球的事件，$X_k$为第$k$次首次摸到 1 号球的事件。它们的取值为 1(成立)或 0(不成立)。显然

$$X_k = \overline{B_1}\,\overline{B_2}\,\overline{B_3}\cdots\overline{B_{k-1}}B_k$$

应用式(1.1.7)

$$P(X_k) = P(\overline{B_1})P(\overline{B_2}|\overline{B_1})P(\overline{B_3}|\overline{B_1}\,\overline{B_2})\cdots$$

注意其中各项事件及其概率的含义，比如$\overline{B_3}|\overline{B_1}\,\overline{B_2}$指前两次未摸到的条件下，第 3 次也未摸到，这时第三次摸球时盒中球数为$N-2$，其 1 号球还在其中。于是

$$P(\overline{B_3}|\overline{B_1}\,\overline{B_2}) = \frac{N-3}{N-2}$$

由此类推得到上式左边的各种概率，于是，

$$P(X_k)=\frac{N-1}{N}\frac{N-2}{N-1}\cdots\frac{N-(k-1)}{N-(k-2)}\frac{1}{N-(k-1)}=\frac{1}{N}$$

（2）令前 k 次能摸到 1 号球的事件记为 Y_k

$$Y_k=X_1\cup X_2\cdots\cup X_k$$

1 号球只有一个，因此，$X_1\cdots X_k$ 是彼此互斥的。于是，

$$P(Y_k)=P(X_1)+\cdots+P(X_k)=k/N$$

在具体问题中求解事件概率有许多技巧与方法。基本的方法是：首先建立起最基本事件的概率，以便计算其他基本事件的概率；而后，将问题所关心的事件分解为互斥的基本事件之“或事件”，或者分解为独立的基本事件之“与事件”；或者分解为链式法则的形式；还可以将其表示为基本事件的条件事件。

例 1.5 有 N 个格子排为一列，将一只小球随机地放入其中任一格子。对于 $k\in[1,N]$ 求：

（1）小球放入第 k 号格子的概率？

（2）前 k 个格子中有小球的概率？

解： 因为是等概的，显然

$$P(\text{小球放入任一格子})=1/N$$

$$P(\text{小球放入任意}k\text{个格子})=k/N$$

其实这两个例题的数学本质是一样的。在例 1.4 中虽然有多个球，但如果将注意力集中在 1 号球上，其他球只是“摆设”而已。有关概率论的应用题有时候令人困惑，但如果选好了观察的“视角”，问题的本质可能显得很简单。

1.2 随机变量与典型分布

1.2.1 随机变量

设计一个从样本空间向实数域的映射，将样本点映射为实数值，将事件映射为实数集合，这便产生了随机变量。

定义 1.5 若 $(\Omega,\mathcal{F},P)$ 是一个概率空间，$X(\xi)$ 是定义在 Ω 上的单值实函数，如果对 $\forall x\in R$（实数集），有 $\{X\leqslant x\}\in\mathcal{F}$，则称 X 为 $(\Omega,\mathcal{F},P)$ 上的一个**随机变量**(Random variable)。并称 $F(x)=P\{X\leqslant x\}$ 为 X 的**分布函数**(Distribution function)，或**累积分布函数**(Cumulative distribution function)。

需要注意的是，定义中的 $\{X\leqslant x\}$ 是 $\{\xi:X(\xi)\leqslant x\}$ 的缩写，它是样本空间上样本点的集合，而非普通实数集合。定义关心 $X(\xi)$ 的取值及其取值的概率。定义以 $\{X\leqslant x\}$ 为基础事件簇，通过分布函数来描述 X 的概率特性。

分布函数具有如下基本性质：

（1）$F(-\infty)=0$，$F(+\infty)=1$；

（2）$F(x)$ 是右连续的非降函数，即

$$F(x_1)\leqslant F(x_2)\ \ (\text{当 } x_1<x_2 \text{ 时})；\ \ F(x^+)=F(x)$$

（3）概率计算公式

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1),\qquad P(X=x)=F(x)-F(x^-)$$

其中，$F(x^-)$ 与 $F(x^+)$ 分别表示 $F(x)$ 在 x 处的左、右极限。

分布函数可能含有间断点，但它们必定是跳跃型的。X 可分为三种类型：

（1）**连续型**：$F(x)$ 是连续的。这时，$P(X=x)=0$；

（2）**离散型**：$F(x)$ 仅含有跳跃型间断点。这些间断点就是 X 的全部可能取值，记为 $\{x_i\}$；相应的概率记为 $\{p_i\}$，即 $P(X=x_i)=p_i=F(x_i)-F(x_i^-)$，并有 $\sum_i p_i=1, p_i \geqslant 0$。$\{p_i\}$ 称为 X 的**分布律**(或**分布列**)(Distribution law)。显然

$$F(x)=P(X \leqslant x)=\sum_{x_i \leqslant x} p_i$$

（3）**混合型**：这时 $F(x)$ 既有连续的部分也有间断点，是上面两种形式的组合。

例 1.6　示性函数：给定概率空间 $(\Omega,\mathcal{F},P)$，对于某事件 $A \in \mathcal{F}$，定义函数

$$I_A(\xi)=\begin{cases}1, & \xi \in A \\ 0, & \xi \notin A\end{cases} \tag{1.2.1}$$

容易验证，$I_A(\xi)$（简记为 I_A）是一个随机变量，而且是二值的，并有 $P(I_A=1)=P(A)$，$P(I_A=0)=P(\overline{A})$。显然，I_A 指示着事件 A 发生与否。

定义 1.6　若存在非负函数 $f(x)$，对 $\forall x \in R$，有

$$F(x)=\int_{-\infty}^{x} f(u)\mathrm{d}u \tag{1.2.2}$$

则称 $f(x)$ 为 X 的**概率密度函数**(Probability density function)，简称密度函数。

利用概率密度函数可方便地计算任意区间 A 上的概率：$P\{X \in A\}=\int_A f(x)\mathrm{d}x$。

当 $F(x)$ 连续时，易见 $f(x)=\dfrac{\mathrm{d}}{\mathrm{d}x}F(x)$；而当 $F(x)$ 不连续时，可以引入阶跃函数 $u(x)$ 与冲激函数 $\delta(x)$ 来表示 $F(x)$ 和 $f(x)$。定义

$$u(x)=\begin{cases}1 & x \geqslant 0 \\ 0 & x<0\end{cases} \quad \text{和} \quad \delta(x)=\frac{\mathrm{d}}{\mathrm{d}x}u(x) \tag{1.2.3}$$

这样，即使在 $F(x)$ 的间断处，仍可认为其(广义)导数存在，于是，密度函数存在。极端的情况是，分布律为 $P(X=x_i)=p_i$ 的离散随机变量，其分布与密度函数可表示为：

$$F(x)=\sum_i p_i u(x-x_i) \quad (i\text{为整数}) \tag{1.2.4}$$

$$f(x)=\sum_i p_i \delta(x-x_i) \quad (i\text{为整数}) \tag{1.2.5}$$

例 1.7　均匀骰子实验。定义随机变量 X 为骰子顶面的编号，取值为 $\{1,2,3,4,5,6\}$。显然 X 是离散型的，其概率特性通常用分布律描述最为方便，即

$$P(X=i)=1/6 \quad i=1,2,\cdots,6$$

但如果需要分布与密度函数，可由上面公式得到

$$F(x)=\sum_{i=1}^{6}\frac{1}{6}u(x-i) \quad \text{或} \quad f(x)=\sum_{i=1}^{6}\frac{1}{6}\delta(x-i)$$

如图 1.2.1 所示。

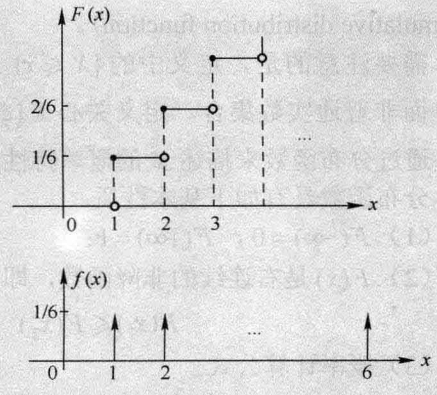

图 1.2.1　例 1.7 的图

例 1.8 随机变量 X 有 30%的可能取 0，70%的可能按 λ 的指数分布特性取正值。求其分布与密度函数。

解： 由于指数分布的分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

因此 $$F_X(x)=P\{X\leqslant x\}=\begin{cases}0, & x<0\\ 0.3, & x=0\\ 0.3+0.7(1-\mathrm{e}^{-\lambda x}), & x>0\end{cases}$$

$$=0.3u(x)+0.7(1-\mathrm{e}^{-\lambda x})u(x)$$

$$f_X(x)=0.3\delta(x)+0.7\lambda\mathrm{e}^{-\lambda x}u(x)$$

1.2.2 典型分布

在应用与理论研究中，人们发现大量的随机变量具有一些特定的分布与密度函数。本书附录 A 中列出了一些常见的分布与密度函数，以及它们的主要统计特性。下面通过几个例子说明几种重要分布的某些特性。

例 1.9 讨论指数分布。指数分布常用于描述一些具有随机性的等待时间。比如，在公交车站等车的时间，排队等候服务的时间，电话交换机或服务器等待呼叫的时间，设备工作到出现故障的时间等。假定一台 PC 机的使用寿命服从 $\lambda=1/3$ 年的指数分布。求该 PC 机可以无故障地使用两年以上的概率？如果到两年时还没有坏，它再使用两年以上的概率是多少？

解： 由于 X 为指数分布，于是，$\forall t\geqslant 0$，有

$$P\{X>t\}=1-F(t)=\mathrm{e}^{-\lambda t}$$

因此，该 PC 机无故障使用两年的概率为

$$P\{X>2\}=\mathrm{e}^{-2/3}=0.5134$$

又 $$P\{X>t+\tau|X>t\}=P\{X>t+\tau\}/P\{X>t\}=\mathrm{e}^{-\lambda(t+\tau)}/\mathrm{e}^{-\lambda t}=\mathrm{e}^{-\lambda\tau}$$

可见，再使用两年的概率为

$$P\{X>4|X>2\}=0.5134$$

看起来使用过的旧设备，其寿命与新设备一样，这源于假定设备的寿命服从简单的指数分布。而指数分布具有**无记忆性**，由上例可知：

$$P\{X>t+\tau|X>t\}=P\{X>\tau\}$$

几何分布是离散形式的指数分布。考虑一个独立试验序列，每个试验结果是“成功”或“失败”，分别记为“1”或“0”，成功的概率为 $0\leqslant p\leqslant 1$。假定以 X 记直至首次出现成功所需进行的试验次数，那么，X 的概率分布律为

$$P\{X=k\}=p(1-p)^{k-1}\qquad(k=0,1,2,\cdots)$$

这种分布称为**几何分布**。

例 1.10 讨论二项分布的计算。假定元件次品率为 0.02，求 10000 件中次品的个数小于等于 240 件的概率。

解： 这类问题中，可认为各元件的好坏是彼此独立的，且次品概率同为 0.02。令 X 为总的次品数目，则 $X\sim B(10000,0.02)$。于是

$$P\{X \leqslant 240\} = \sum_{k=0}^{240} \binom{10000}{k} 0.02^k 0.98^{10000-k}$$

由于总数n很大，计算该式是困难的。为此我们可以利用棣莫弗-拉普拉斯中心极限定理：$n \to \infty$时，$P_n(k) \to N(np, npq)$。通常$npq \geqslant 10$就行了。因此近似有，$P_n(k) \to N(200,196)$，所以

$$P\{X \leqslant 240\} \approx \Phi\left(\frac{240-200}{14}\right) = 0.9979$$

其中，$\Phi(x)$为标准正态分布函数。

运用上述近似公式时应该满足$npq \gg 1$的条件。但有一些时候，该条件并不成立。比如，如果次品率为10^{-5}，求 10000 件中无次品的概率？注意到，$npq \approx 0.1$。这时可改用泊松分布来近似处理这种大数量“稀有”(p 很小)事件的情形：$P_n(k) \to \mathcal{P}(np)$，即参数$\lambda = np$的泊松分布。于是$P_n(k) \to \mathcal{P}(0.1)$，有

$$P\{X=0\} = (1-10^{-5})^{10000} \approx \mathrm{e}^{-\lambda}\frac{\lambda^0}{0!} = \mathrm{e}^{-0.1} = 0.9048$$

二项分布的近似计算方法可归纳如下(参见 1.7.4 节)：$X \sim B(n,p)$在n很大时，可以依条件近似为两种分布：

（1）$npq \gg 1$时：
$$X \to N(np, npq) \text{ 或 } P_n(k) \approx \frac{1}{\sqrt{2\pi npq}}\mathrm{e}^{-\frac{(k-np)^2}{2npq}} \tag{1.2.6}$$

（2）p很小时：
$$X \to \mathcal{P}(np) \text{ 或 } P_n(k) \approx \frac{(np)^k}{k!}\mathrm{e}^{-np} \tag{1.2.7}$$

例 1.11 讨论随机点的均匀分布。假定有n个随机点独立且均匀地落在t轴的$(0,T]$区间中，如图 1.2.2 所示。第k点所落的位置记为T_k，显然，T_k服从$(0,T]$的均匀分布。我们做下面几点分析：

（1）$\forall t \in (0,T]$，由于实数是无限稠密的，T_k正好落在t上的概率为

$$P\{T_k = t\} = 0$$

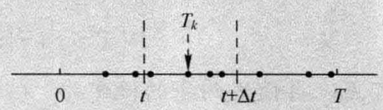

图 1.2.2　例 1.11 的图

（2）任意给定某个子区间$(t, t+\Delta t] \subset (0,T]$，$T_k$落在该子区间的概率为

$$P\{T_k \in (t, t+\Delta t]\} = \Delta t / T$$

（3）n个点各自独立，其中有k个落在上述子区间$(t, t+\Delta t]$的概率服从二项式分布

$$P_n(k) = \binom{n}{k} p^k q^{n-k}$$

其中$p = \Delta t / T$，$q = 1-p$。

（4）设想无穷多个点落入$(0,+\infty)$区间，并维持点的密度为λ的情形(即$T \to \infty$，$n \to \infty$，但$n/T = \lambda$)：考察任一区间$(t, t+\Delta t]$上有k个点的概率。由于$p = \Delta t / T \to 0$，$np = n\Delta t / T = \lambda \Delta t$，利用式(1.2.7)，得

$$P_n(k) = \mathrm{e}^{-\lambda \Delta t}\frac{(\lambda \Delta t)^k}{k!} \tag{1.2.8}$$

可见，它是参数为$\lambda\Delta t$的泊松分布。