普通高等学校教材

SUIJI GUOCHENG JICHU

随机过程基础

主　编 陈家清 赵华玲

武汉理工大学出版社
WUTP Wuhan University of Technology Press

随机过程基础

主　编　陈家清　赵华玲

武汉理工大学出版社
·武　汉·

内容提要

本书是为统计学以及相关专业本科生编写的随机过程基础知识入门教材，主要内容包括：概率论基础知识、随机过程的概念和基本类型、平稳过程、Poisson过程、更新过程、马尔可夫链、随机过程分析等。

本书尽可能简化复杂的抽象证明或推导，重点讲述Poisson过程、更新过程、马尔可夫链、随机过程分析等内容。叙述通俗、简明且例题较多，并给出了章节习题的大部分参考答案，便于读者自学参考。

本书也可作为广大从事以统计学为基础工具的相关技术研究人员的参考或者自学书籍。

图书在版编目(CIP)数据

随机过程基础/陈家清，赵华玲主编．—武汉：武汉理工大学出版社，2013.11
ISBN 978-7-5629-4185-9

Ⅰ.①随…
Ⅱ.①陈…　②赵…
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(2013)第230242号

项目负责人：陈军东　　**责任编辑**：陈军东
责任校对：陈　平　　**装帧设计**：董君承
出版发行：武汉理工大学出版社
社　　址：武汉市洪山区珞狮路122号
邮　　编：430070
网　　址：http://www.techbook.com.cn
经　　销：各地新华书店
印　　刷：武汉兴和彩色印务有限公司
开　　本：787×1092　1/16
印　　张：10.25
字　　数：230千字
版　　次：2014年3月第1版
印　　次：2014年3月第1次印刷
定　　价：23.00元
凡购本书，如有缺页、倒页、脱页等印装质量问题，请向出版社发行部调换。
本社购书热线电话：027-87523148　87664138　87515798　87165708(传真)

前　言

随机过程理论是现代概率论中的一个重要分支，在生物物理、系统工程、管理和经济等领域都有着广泛的应用。目前国内各大高校大部分在统计、管理等相关专业开设随机过程课程。编者结合多年的教学经验，根据教学课时的安排及本科专业学生掌握统计专业基础知识的程度，在自编讲义的基础上，编写了这本教材作为本科统计学以及相关专业的随机过程入门教程。

本书大致分为四个部分：概率论基础知识以及随机过程的最基本内容（第 1 章、第 2 章、第 3 章），Poisson 过程及更新过程（第 4 章、第 5 章），马尔可夫链（第 6 章），随机过程分析（第 7 章）。作为随机过程的入门知识介绍，我们根据本科教学和应用的需要特增加了第 7 章，目的是为了让学生在学习了高等数学或者数学分析知识的基础上，进一步掌握随机过程分析的相关知识。考虑到本科统计以及相关专业学生掌握数学知识的程度不足，比如没有学习测度论知识等，因而，书中没有介绍比较复杂的随机过程，比如鞅论、布朗运动等。

本书的编写得到了武汉理工大学唐湘晋、梅顺治等多位老师的鼓励和支持，感谢他们提出的宝贵修改意见。陈志强、金倩聿、刘金等研究生对本书的编写也提供了诸多帮助，在此表示由衷的谢意！

由于编者水平有限，书中的缺点在所难免，敬请读者批评指正。

编　者

2013 年 9 月

目　　录

第 1 章　概率论基础知识

随机过程以概率论作为其主要的基础. 为此,我们首先对在本书中经常用到的概率论的基础知识作简要的介绍.

§1.1　概 率 空 间

随机试验是概率论的一个基本概念. 一个试验(或观察),若它的结果预先无法确定,但具有如下 3 个性质,则称之为**随机试验**,简称为试验:

(1) 可以在相同的条件下重复进行;

(2) 试验的可能结果不止一个,且所有可能结果是已知的;

(3) 每次实验前不能确定哪个结果会出现.

所有试验的可能结果组成的集合,称为**样本空间**,记作 Ω. Ω 中的元素则称为**样本点**,用 ω 表示. 由 Ω 的某些样本点构成的子集合,称为**事件**,常用大写字母 A,B,C 等表示. 由 Ω 中的若干子集构成的集合称为**集类**,用花体字母 $\mathscr{A},\mathscr{B},\mathscr{F}$ 等表示. Ω 中的样本点 ω 也称为**基本事件**,样本空间 Ω 称为**必然事件**,空集 $\varnothing$ 称为**不可能事件**.

由于并不是在所有的 Ω 的子集上都能方便地定义概率,一般只限制在满足一定条件的集类上研究概率性质,为此引入 σ 代数的概念:

定义 1.1　设随机试验 E 的样本空间为 Ω, $\mathscr{F}$ 是 Ω 的子集组成的集合族,如果满足:

(1) $\Omega \in \mathscr{F}$;

(2) 若 $A \in \mathscr{F}$,则 $\overline{A} \in \mathscr{F}$;

(3) 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 Ω 的一个 σ 代数(σ-域), $\mathscr{F}$ 中的集合称为事件, $(\Omega,\mathscr{F})$ 称为可测空间.

可测空间有如下性质:

(1) $\varnothing \in \mathscr{F}$;

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

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

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

定义 1.2　设 $(\Omega,\mathscr{F})$ 是一可测空间,定义在 $\mathscr{F}$ 上的实值集函数 $P(\cdot)$,若满足:

(1) 非负性：对 $\forall A \in \mathscr{F}, 0 \leqslant P(A) \leqslant 1$；

(2) 规一性：$P(\Omega) = 1$；

(3) 完全可加性：对 $\forall A_i \in \mathscr{F}\,(i = 1,2,\cdots)$，其中 $A_i \cap A_j = \varnothing\ (i \neq j)$，有

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

则称 P 是 $(\Omega,\mathscr{F})$ 上的概率测度，简称为概率，$(\Omega,\mathscr{F},P)$ 称为概率空间，$P(A)$ 表示事件 A 的概率.

由定义易知，事件的概率具有如下性质：

(1) $P(\varnothing) = 0$；

(2) 有限可加性：若 $A_i \in \mathscr{F}\,(i = 1,2,\cdots,n)$，$A_i \cap A_j = \varnothing\ (i \neq j)$，则

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

(3) 单调性：若 $A, B \in \mathscr{F}$，且 $B \subset A$，则 $P(A-B) = P(A) - P(B)$，且 $P(A) \geqslant P(B)$；

(4) 上连续性：若 $A_i \in \mathscr{F}$，且 $A_1 \supset A_2 \supset \cdots$，且 $\bigcap_{n=1}^{\infty} A_n = A$，则 $\lim_{n\to\infty} P(A_n) = P(A)$；

(5) 下连续性：若 $A_i \in \mathscr{F}$，且 $A_1 \subset A_2 \subset \cdots$，且 $\bigcup_{n=1}^{\infty} A_n = A$，则 $\lim_{n\to\infty} P(A_n) = P(A)$；

(6) 多除少补原理：设 $A_i \in F(i = 1,2,\cdots,n)$，则有

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

(7) 次可加性：若 $A_i \in F(i = 1,2,\cdots,n)$，则 $P(\bigcup_{i=1}^{n} A_i) \leqslant \sum_{i=1}^{n} P(A_i)$.

§1.2 随机变量和分布函数

1.2.1 随机变量及分布函数定义

定义 1.3 设 $(\Omega,\mathscr{F},P)$ 是概率空间，$X(\omega)$ 是定义在 Ω 上的单值实函数，若对于任意实数 $x \in \mathbf{R}$，有 $\{\omega: X(\omega) \leqslant x\} \in \mathscr{F}$，则称 $X(\omega)$ 是 $\mathscr{F}$ 上的**随机变量**. 函数 $F(x) = P\{\omega: X(\omega) \leqslant x\}, x \in \mathbf{R}$ 称为随机变量 X 的**分布函数**.

若随机变量 X 可能取值的全体是一可数集或有限集，则称 X 是离散型随机变量. 其概率分布用分布列描述：

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

分布函数 $F(x)=\sum\limits_{x_k\leqslant x}p_k$.

对随机变量 X 的分布函数 $F(x)$，若存在一非负函数 $f(x)$，对 $\forall x\in\mathbf{R}$，有

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

则称 X 为连续型随机变量，$f(x)$ 为随机变量 X 的概率密度函数.

若 $f(x)$ 连续，则

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x}=f(x)$$

即

$$\lim_{h\to 0}\frac{P\{x<X\leqslant x+h\}}{h}=f(x)$$

或

$$P\{x<X\leqslant x+h\}=f(x)h+o(h)$$

以上关系是以后用所谓"微元法"求概率密度函数的依据. 为求随机变量 X 的概率密度函数，先求 X 落在一个小区域$(x,x+h]$上的概率 $P(x<X\leqslant x+h)$，然后令 $h\to 0$，求其极限

$$\lim_{h\to 0}\frac{P\{x<X\leqslant x+h\}}{h}$$

即得 $f(x)$ 分布函数具有如下性质：

(1)$F(x)$ 是单调不减函数；

(2)$0\leqslant F(x)\leqslant 1$，$\lim\limits_{x\to-\infty}F(x)=0$，$\lim\limits_{x\to+\infty}F(x)=1$；

(3)$F(x)$ 是右连续函数，即对 $\forall x\in\mathbf{R}$，$F(x+0)=F(x)$.

1.2.2　常见分布函数

下面是常见随机变量的分布函数：

二项分布：若随机变量 X 的分布列为

$$P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k}\quad(k=0,1,\cdots,n)$$

其中 $0<p<1$，则称 X 服从参数为 n,p 的二项分布，记为 $X\sim\mathrm{B}(k;n,p)$ 或$X\sim\mathrm{B}(n,p)$.

泊松(Poisson)分布：如果随机变量 X 的分布列为

$$P\{X=k\}=\frac{\lambda^k\mathrm{e}^{-\lambda}}{k!}\ (k=0,1,2,\cdots)$$

其中 $\lambda>0$，则称 X 服从参数为λ 的泊松分布，记为 $X\sim\mathrm{P}(\lambda)$.

二点分布：如果随机变量 X 的分布列为

$$P\{X=k\}=p^kq^{1-k}\quad(k=0,1)$$

即

$$P\{X=1\}=p, P\{X=0\}=q=1-p$$

其中 $0<p<1$，则称 X 服从二点分布.

正态分布：若随机变量 X 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\quad(-\infty<x<+\infty)$$

其中 μ 和 $\sigma>0$ 是常数，则称随机变量 X 服从参数为 μ 和 σ 的正态分布(或高斯分布)，记为 $X\sim \mathrm{N}(\mu,\sigma^2)$. 称参数 $\mu=0,\sigma^2=1$ 的正态分布为标准正态分布，记为 $X\sim \mathrm{N}(0,1)$.

对数正态分布：若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}\mathrm{e}^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

其中 $\sigma>0$，μ 为常数，则称 X 服从对数正态分布. 因 X 的对数 $\ln X$ 服从 $\mathrm{N}(\mu,\sigma^2)$ 分布而得名.

均匀分布：若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & \text{其他}\end{cases}$$

则称 X 在区间 $[a,b]$ 上服从均匀分布. 记作 $X\sim \mathrm{U}(a,b)$.

韦布尔分布：若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{m}{\beta}(x-\alpha)^{m-1}\mathrm{e}^{-\frac{(x-\alpha)^m}{\beta}}, & x>\alpha\\ 0, & x\leqslant\alpha\end{cases}$$

则称 X 服从韦布尔分布. 其中 $m>0$ 称为形状参数，α 称为位置参数，$\beta>0$ 称为尺度参数.

Γ 分布：若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

其中 $\alpha>0,\lambda>0$，则称 X 服从 Γ 分布，记为 $X\sim\Gamma(\alpha,\lambda)$.

指数分布：在 Γ 分布中令 $\alpha=1,\lambda>0$，得到指数分布的密度函数为

$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

χ^2 分布：在 Γ 分布中令 $\alpha=n/2$，n 为正整数，$\lambda=\dfrac{1}{2}$，得到自由度为 n 的 χ^2 分布的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})}x^{n/2-1}\mathrm{e}^{-\frac{x}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

在二项分布中，当 n 大，p 小，而乘积 np 大小适中时，有如下近似：

定理 1.1　对于二项分布 $\mathrm{B}(k;n,p)$，若 $\lim\limits_{n\to\infty}np=\lambda$，则

$$\lim_{n\to\infty}\mathrm{C}_n^k p^k q^{n-k}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$$

由此可知，当 n 充分大，p 较小时，二项分布可用泊松分布近似表示.

1.2.3　联合分布函数

定义 1.4　如果 X 和 Y 是定义在同一概率空间 $(\Omega,\mathscr{F},P)$ 上的两个随机变量，称 (X,Y) 为二维随机向量，对 $\forall(x,y)\in\mathbf{R}^2$，

$$F(x,y)=P\{\omega:X(\omega)\leqslant x,Y(\omega)\leqslant y\}$$

称为 (X,Y) 的联合分布函数.

由 (X,Y) 的分布可确定 X,Y 各自的分布，反之不行. 如“(X,Y) 服从二维正态分布”与“X,Y 是两个正态分布随机变量”是完全不同的概念.

联合分布函数具有如下性质：

(1) 对于 x 和 y，$F(x,y)$ 是单调非减函数；

(2) $F(x,y)$ 对每个变元右连续；

(3) $\lim\limits_{x\to-\infty}F(x,y)=0$，$\lim\limits_{y\to-\infty}F(x,y)=0$，$\lim\limits_{\substack{y\to+\infty\\ x\to+\infty}}F(x,y)=1$；

(4) 对 $\forall x_1<x_2,y_1<y_2$，有

$$F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\geqslant 0.$$

定义 1.5　设 (X,Y) 是二维随机向量，对 $\forall(x,y)\in\mathbf{R}^2$ 有

$$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\}$$

成立，称 X 与 Y 相互独立.

定义 1.6　$(\Omega,\mathscr{F},P)$ 是概率空间，$\boldsymbol{X}=\boldsymbol{X}(\omega)=(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$ 是定义在 Ω 上取值于 n 维空间 $\mathbf{R}^n$ 的向量函数. 如果对于任意 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$，$\{\omega:X_1(\omega)\leqslant x_1,X_2(\omega)\leqslant x_2,\cdots,X_n(\omega)\leqslant x_n\}\in\mathscr{F}$，则称 $\boldsymbol{X}=\boldsymbol{X}(\omega)$ 为 **n 维随机变量**或 **n 维随机向量**. 称 $F(\boldsymbol{x})=F(x_1,x_2,\cdots,x_n)=P\{\omega:X_1(\omega)\leqslant x_1,X_2(\omega)\leqslant x_2,\cdots,X_n(\omega)\leqslant x_n\}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$ 为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的**联合分布函数**.

定义 1.7　$(X_1,X_2,\cdots,X_n)$ 是 n 维随机变量，若对 $\forall(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$，有

$$F(x_1,x_2,\cdots,x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

成立，则称随机变量 $X_1,X_2,\cdots,X_n$ 相互独立.

定理 1.2　若随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,则其中任意 $k(2 \leqslant k \leqslant n)$ 个随机变量 $X_{i_1}, X_{i_2}, \cdots, X_{i_k}$ 也相互独立.

§1.3　随机变量的数字特征和特征函数

1.3.1　数字特征

定义 1.8　随机变量 X 的分布函数为 $F(x)$,若$\int_{-\infty}^{+\infty} |x| \mathrm{d}F(x) < +\infty$,则

$$E(X) = \int_{-\infty}^{+\infty} x \mathrm{d}F(x)$$

称为随机变量 X 的数学期望或均值.

若 X 是连续型随机变量,则

$$E(X) = \int_{-\infty}^{+\infty} xF'(x)\mathrm{d}x = \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$$

若 X 是离散型随机变量,则

$$E(X) = \sum_{i=1}^{\infty} x_i P\{X = x_i\}$$

随机变量 X 的数学期望是对 X 所有可能取值的加权平均.

定义 1.9　随机变量 X 的分布函数为 $F(x)$,则称

$$D(X) = E\{[X - E(X)]^2\} = E(X^2) - [E(X)]^2$$

为随机变量 X 的方差,有时记 $D(X) = \mathrm{Var}X = \sigma_X^2$,称 $\sigma_X = \sqrt{D(X)}$ 为 X 的标准差.

随机变量的方差反映其取值偏离均值的程度.

定理 1.3　(切比雪夫不等式) 设随机变量 X 的期望为 $E(X)$,方差为 $D(X)$,则对任给的 $\varepsilon > 0$,有

$$P\{|X - E(X)| \geqslant \varepsilon\} \leqslant \frac{D(X)}{\varepsilon^2}$$

或

$$P\{|X - E(X)| < \varepsilon\} \geqslant 1 - \frac{D(X)}{\varepsilon^2}$$

切比雪夫不等式只利用均值及方差就描述了随机变量的变化情况,不管 X 的分布是什么,X 落在 $(E(X) - \varepsilon, E(X) + \varepsilon)$ 中的概率不小于 $1 - \frac{D(X)}{\varepsilon^2}$. 因此它在理论研究及实际应用中很有价值.

随机变量的数学期望和方差具有如下性质:

(1) $E(C) = C, D(C) = 0, E(CX) = CE(X), D(CX) = C^2 D(X)$;

(2) $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$;

(3) 若 X 和 Y 相互独立,则 $E(XY) = E(X)E(Y)$;

(4) $D(X) = 0$ 的充要条件是 $P\{X = E(X)\} = 1$;

(5) 若 X 为非负整数值的随机变量,则

$$E(X) = \sum_{i=1}^{\infty} P\{X \geqslant i\}$$

(6) 若 X 为非负值的连续型随机变量,则

$$E(X) = \int_0^{+\infty} [1 - F(x)] \mathrm{d}x$$

定义 1.10　对于两个随机变量 X、Y,称

$$\mathrm{Cov}(X,Y) = E\{[X - E(X)][Y - E(Y)]\} = E(XY) - E(X)E(Y)$$

为(X,Y) 的**协方差**.

若 X,Y 独立,则 $E(XY) = E(X)E(Y)$,从而得 $\mathrm{Cov}(X,Y) = 0$. 于是,若 $\mathrm{Cov}(X,Y) \neq 0$,则 X,Y 不独立. 因此 $\mathrm{Cov}(X,Y) \neq 0$ 刻画了 X,Y 取值存在某种统计意义上的线性相关关系.

定义 1.11　X,Y 是随机变量,若 $0 < D(X) = \sigma_X^2 < +\infty$, $0 < D(Y) = \sigma_Y^2 < +\infty$,称

$$r_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

为(X,Y) 的**相关系数**.

r_{XY} 刻画了 X,Y 之间线性关系的密切程度,若 $r_{XY} = 0$,称 X,Y 不相关.

性质:

(1) $D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D(X_i) + 2\sum_{\substack{i,j=1 \\ i<j}}^{n} \mathrm{Cov}(X_i, X_j)$;

(2) 若 X 和 Y 相互独立,则 $\mathrm{Cov}(X,Y) = 0$,即 X,Y 不相关;

(3) (Schwarz 不等式) 若 $E(X^2) < +\infty$, $E(Y^2) < +\infty$,则

$$(E \mid XY \mid)^2 \leqslant E(X^2)E(Y^2)$$

特别地,

$$\mid \mathrm{Cov}(X,Y) \mid^2 \leqslant \sigma_X^2 \sigma_Y^2;\ \mid r_{XY} \mid \leqslant 1$$

(4) $\mid r_{XY} \mid = 1$ 的充要条件是 X 与 Y 以概率 1 线性相关,即存在常数 a 和 b,使得

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

定义 1.12　称 $E(X^k) = \int_{-\infty}^{+\infty} x^k \mathrm{d}F(x)$, $k \geqslant 1$ 为随机变量 X 的 k 阶矩.

在求某些随机变量的各阶矩时,有时用矩母函数会很方便.

定义 1.13　随机变量 X 的分布函数为 $F(x)$,如果积分

$$\psi(t) = E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tx} dF(x)$$

存在,则称其为 X 的矩母函数.

显然若 X 的 k 阶矩存在,则 $E(X^k) = \psi^{(k)}(0)$. 矩母函数由此得名.

1.3.2　特征函数

当矩母函数存在时,能够用其刻画随机变量的概率分布,但有时随机变量的矩母函数不一定存在,在这种情况下,我们可以采用如下定义的特征函数.

定义 1.14　设随机变量 X 的分布函数为

$$\varphi(t) = E(e^{itx}) = \int_{-\infty}^{+\infty} e^{itx} dF(x), -\infty < t < +\infty$$

称为 X 的特征函数.

特征函数 $\varphi(t)$ 是实变量 t 的复值函数,因为 $|e^{itx}| = 1$,所以随机变量的特征函数一定存在. 显然特征函数只与分布函数有关,因此也称其为分布函数的特征函数:

若 X 为离散型随机变量,其分布列为 $p_k = P\{X = x_k\}$ $(k = 1,2,\cdots)$,则其特征函数为

$$\varphi(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k, -\infty < t < +\infty$$

若 X 为连续型随机变量,其概率密度为 $f(x)$,则其特征函数为

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx$$

特征函数有如下基本性质:

(1) 有界性: $|\varphi(t)| \leqslant |\varphi(0)| = 1, -\infty < t < +\infty$;

(2) 一致连续性: $\varphi(t)$ 在整个实数轴 $(-\infty, +\infty)$ 上一致连续;即对 $\forall \varepsilon > 0, \exists \delta > 0$,使当 $|h| < \delta$ 时,对 t 一致地有

$$|\varphi(t+h) - \varphi(t)| < \varepsilon$$

(3) 非负定性:对于任意 $n \geqslant 1$,任意实数 $t_1, t_2, \cdots, t_n$ 和任意复数 $z_1, z_2, \cdots, z_n$ 有

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \varphi(t_j - t_k) z_j \bar{z}_k \geqslant 0$$

(4) 共轭对称性: $\varphi(-t) = \overline{\varphi(t)}$;

(5) 设 $Y = aX + b$,其中 a、b 为任意实数, $\varphi_X(t)$、$\varphi_Y(t)$ 分布是 X、Y 的特征函数,有

$$\varphi_Y(t) = a^{itb} \cdot \varphi_X(at)$$

(6) 设 X 和 Y 为独立随机变量,则

$$\varphi_{X+Y}(t) = \varphi_X(t) \varphi_Y(t)$$

一般地,如果 $X_1, X_2, \cdots, X_n$ 相互独立,则

$$\varphi_{X_1+X_2+\cdots+X_n}(t) = \varphi_{X_1}(t) \varphi_{X_2}(t) \cdots \varphi_{X_n}(t)$$

(7) 如果 X 有 n 阶原点矩，则它的特征函数 $\varphi(t)$ 有 n 阶导数，且

$$E(X^k) = (-\mathrm{i})^k \varphi^{(k)}(0) \ (k = 1,2,\cdots,n)$$

(8) **唯一性定理**　若 X 的特征函数为 $\varphi(t)$，则 X 的分布函数 $F(x)$ 由 $\varphi(t)$ 唯一确定.

对于 n 维随机变量也可以定义特征函数：

定义 1.15　设 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 是 n 维随机变量，$\boldsymbol{t} = (t_1, t_2, \cdots, t_n) \in \mathbf{R}^n$，称

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \varphi_{\boldsymbol{X}}(t_1, t_2, \cdots, t_n) = E\mathrm{e}^{\mathrm{i}t'x} = E[\exp(\mathrm{i}\sum_{k=1}^{n} t_k X_k)]$$

为 $\boldsymbol{X}$ 的特征函数.

特征函数是研究概率论最重要的分析工具之一，用特征函数作为随机变量的研究工具相比用分布函数有许多方便之处. 例如，独立随机变量之和的概率分布是各被加项分布的卷积，而独立随机变量之和的特征函数则是各被加项特征函数的普通乘积. 又如，通过随机变量的分布函数求它的数字特征，一般需要进行积分运算，而通过特征函数求数字特征，一般只需要微商运算.

例 1.1　设 $X \sim \mathrm{N}(\mu, \sigma^2)$，求 X 的特征函数.

解　先求 $X^* = \dfrac{X-\mu}{\sigma}$ 的特征函数 $\varphi^*(t)$. 因为 $X^* \sim \mathrm{N}(0,1)$，故

$$\varphi^*(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}tx - \frac{x^2}{2}} \mathrm{d}x$$

等式两边进行求导得：

$$\varphi'^*(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{i}x\mathrm{e}^{\mathrm{i}tx - \frac{x^2}{2}} \mathrm{d}x$$

对右边进行分部积分，令 $u = \mathrm{e}^{\mathrm{i}tx}$，$\mathrm{d}v = x\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x = -\mathrm{d}\mathrm{e}^{-\frac{x^2}{2}}$，则 $v = -\mathrm{e}^{-\frac{x^2}{2}}$，有

$$\begin{aligned}\varphi'^*(t) &= -\frac{\mathrm{i}}{\sqrt{2\pi}}\mathrm{e}^{-\mathrm{i}tx - \frac{x^2}{2}}\Bigg|_{-\infty}^{+\infty} - \frac{t}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}tx - \frac{x^2}{2}} \mathrm{d}x \\ &= -\frac{t}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}tx - \frac{x^2}{2}} \mathrm{d}x = -t\varphi^*(t)\end{aligned}$$

于是得微分方程

$$\varphi'^*(t) + t\varphi^*(t) = 0$$

通解为

$$\varphi^*(t) = c\mathrm{e}^{-\frac{t^2}{2}}$$

因为 $\varphi^*(0) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x = 1$，得 $c = 1$. 所以 $X^* = \dfrac{X-\mu}{\sigma}$ 的特征函数为 $\varphi^*(t) = \mathrm{e}^{-\frac{t^2}{2}}$.

而 $X=\sigma X^{*}+\mu$. 由特征函数的性质(5)，X 的特征函数为

$$\varphi(t)=\mathrm{e}^{\mathrm{i}\mu t}\varphi^{*}(\sigma t)=\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2 t^2}{2}}$$

根据特征函数的定义式可求出常用分布的特征函数。现将这些常用分布的特征函数列出，如表 1.1 所示。

表 1.1 常用分布的特征函数表

分布名称	分布律或概率密度	特征函数
单点分布	$p_a=1$	$\mathrm{e}^{\mathrm{i}at}$
两点分布	$p_1=p,p_0=q,q=1-p$	$p\mathrm{e}^{\mathrm{i}t}+q$
二项分布 $\mathrm{B}(n,p)$	$p_k=\mathrm{C}_n^k p^k q^{n-k},k=1,2,\cdots,n,q=1-p$	$(p\mathrm{e}^{\mathrm{i}t}+q)^n$
泊松分布 $\mathrm{P}(\lambda)$	$p_k=\dfrac{\lambda^k\mathrm{e}^{-\lambda}}{k!}(\lambda>0),k=1,2,\cdots$	$\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}t}-1)}$
几何分布	$p_k=pq^{k-1}(0<p<1,q=1-p),k=1,2,\cdots$	$p\mathrm{e}^{\mathrm{i}t}(1-q\mathrm{e}^{\mathrm{i}t})^{-1}$
均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}$	$\dfrac{\mathrm{e}^{\mathrm{i}bt}-\mathrm{e}^{\mathrm{i}at}}{\mathrm{i}t(b-a)}$
指数分布	$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0\\ 0, & x\leqslant 0\end{cases}\quad(\lambda>0)$	$(1-\dfrac{\mathrm{i}t}{\lambda})^{-1}$
正态分布 $\mathrm{N}(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\quad(\sigma>0)$	$\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2t^2}{2}}$
$\chi^2(n)$ 分布	$f(x)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})}x^{n/2-1}\mathrm{e}^{-\frac{x}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$	$(1-2\mathrm{i}t)^{-n/2}$
$\Gamma(\alpha,\beta)$ 分布	$f(x)=\begin{cases}\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}, & x>0\\ 0, & x\leqslant 0\end{cases}$	$(1-\dfrac{\mathrm{i}t}{\beta})^{-\alpha}$

例 1.2 设 $X\sim\mathrm{N}(\mu,\sigma^2)$，利用 X 的特征函数求 $E(X)$ 和 $D(X)$.

解 已知 X 的特征函数 $\varphi(t)=\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2t^2}{2}}$，对它求导得

$$\varphi'(t)=\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2t^2}{2}}(\mathrm{i}\mu-\sigma^2t)$$

$$\varphi''(t)=\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2t^2}{2}}(\mathrm{i}\mu-\sigma^2t)^2+\mathrm{e}^{\mathrm{i}\mu t-\frac{\sigma^2t^2}{2}}(-\sigma^2)$$

再由特征函数性质(7)得

$$E(X)=(-\mathrm{i})\varphi'(0)=(-\mathrm{i})\mathrm{i}\mu=\mu$$

$$E(X^2)=(-\mathrm{i})^2\varphi''(0)=-(\mathrm{i}^2\mu^2-\sigma^2)=\mu^2+\sigma^2$$

故

$$D(X)=E(X^2)-[E(X)]^2=\sigma^2$$

例 1.3　随机变量 X 在 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 上服从均匀分布，$Y=\sin X$，试利用特征函数求 Y 的概率密度函数 $f(y)$.

解　X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\pi}, & x\in[-\dfrac{\pi}{2},\dfrac{\pi}{2}]\\ 0, & \text{其他}\end{cases}$$

Y 的特征函数为

$$\varphi_Y(t)=E(e^{itY})=E(e^{it\sin X})=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}e^{it\sin x}\frac{1}{\pi}dx$$

令 $u=\sin x$，$du=\cos x dx=\sqrt{1-u^2}dx$，得

$$\varphi_Y(t)=\frac{1}{\pi}\int_{-1}^{1}e^{itu}\frac{1}{\sqrt{1-u^2}}du$$

利用特征函数与分布函数一一对应的唯一性定理，知随机变量 Y 的概率密度为

$$f(y)=\begin{cases}\dfrac{1}{\pi}\dfrac{1}{\sqrt{1-y^2}}, & -1<y<1\\ 0, & \text{其他}\end{cases}$$

例 1.4　设随机变量 X 服从自由度为 n 的 χ^2 分布.

(1) 利用特征函数求 $E(X)$ 和 $D(X)$.

(2) 若 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，且 X,Y 相互独立，利用特征函数证明 $X+Y\sim\chi^2(n_1+n_2)$，即 χ^2 分布具有可加性.

解　(1) 因为 $X\sim\chi^2(n)$，故 X 的特征函数 $\varphi_X(t)=(1-2it)^{-n/2}$，

$\varphi'_X(0)=ni$

$\varphi''_X(0)=-n(n+2)$

$E(X)=(-i)\varphi'_X(0)=n$

$E(X^2)=(-i)^2\varphi''_X(0)=n(n+2)$

$D(X)=E(X^2)-[E(X)]^2=2n$

(2) 令 $Z=X+Y$，则

$$\begin{aligned}\varphi_Z(t)&=E[e^{it(X+Y)}]=E(e^{itX}e^{itY})=\varphi_X(t)\varphi_Y(t)\\&=(1-2it)^{-n_1/2}\cdot(1-2it)^{-n_2/2}\\&=(1-2it)^{-\frac{n_1+n_2}{2}}\end{aligned}$$

即 $Z=X+Y\sim\chi^2(n_1+n_2)$