随机过程

王志刚　编著

中国科学技术大学出版社

随机过程

王志刚　编著

中国科学技术大学出版社

内 容 简 介

本书主要介绍了随机过程的基本理论、基本方法和应用背景等，主要内容包括概率论基本知识、随机过程的基本概念和基本类型、泊松过程和更新过程、马尔可夫链、连续时间的马尔可夫链、平稳随机过程、平稳随机过程的谱分析、平稳时间序列、平稳时间序列的统计分析、随机积分和随机微分方程等.在选材上强调实用性，配有大量的应用实例，每章后附有一定数量的习题，附录给出了习题答案，可供读者选用、参考，期望帮助读者加深对基本概念和基本方法的理解和掌握.

图书在版编目(CIP)数据

随机过程/王志刚编著.—合肥：中国科学技术大学出版社，2018.9
ISBN 978-7-312-04547-9

Ⅰ.随…　Ⅱ.王…　Ⅲ.随机过程　Ⅳ.O211.6

中国版本图书馆 CIP 数据核字(2018)第 196131 号

出版　中国科学技术大学出版社
安徽省合肥市金寨路 96 号，230026
http://press.ustc.edu.cn
https://zgkxjsdxcbs.tmall.com
印刷　安徽省瑞隆印务有限公司
发行　中国科学技术大学出版社
经销　全国新华书店
开本　787 mm×1092 mm　1/16
印张　15.25
字数　380 千
版次　2018 年 9 月第 1 版
印次　2018 年 9 月第 1 次印刷
定价　38.00 元

前　　言

随机过程理论在物理、生物、工程、经济、管理等方面都有着重要的应用，现已成为科技工作者必须掌握的一个理论工具. 随机过程研究的是客观世界中随机现象演变过程的统计规律性，本书是编著者及其团队在多年讲授“随机过程”课程和科研实践的基础上编著而成的，希望能系统讲述随机过程的基本原理、基本方法和应用.

本书是在 2009 年出版的《应用随机过程》一书的基础上修订而成的，增加了随机积分和随机微分方程部分，在时间序列的统计分析章节中增加了利用统计软件 Eviews 9.0 对时间序列进行分析预测的内容，增加了鞅过程的介绍，对基本概念和基本原理的撰写更注重精炼和准确.

在撰写过程中力求反映以下几个特点：

(1) 强调基本概念和基本方法的理解与阐述. 着眼于引发读者兴趣，使读者领悟随机过程理论的思想，感受其魅力与威力，着重揭示基本概念的来源与实际背景，典型随机模型的提炼、特性的刻画、应用背景及其发展的踪迹. 较全面地介绍了现代科学技术中常见的几种重要的随机过程，如泊松过程和更新过程、马尔可夫过程、平稳随机过程、时间序列分析等，对每种过程都做了详细的分析，读者可根据专业的需要适当取舍. 每章都配有例题和习题，帮助读者理解和掌握基本概念与基本方法.

(2) 强调知识的灵活运用. 本书不仅介绍了随机过程的基本概念和基本方法，还强调了随机过程理论的应用，例如线性时不变系统、市场占有率分析、网页搜索排序、教学效果评价、$M/M/s$ 排队系统分析、机床维修、天气预报系统、股票走势预测等，展示出随机过程理论强大的生命力和广阔的应用前景.

由于笔者水平有限，书中一定还有不少不足和错误，恳请读者批评指正.

作　者

2018 年 1 月于海南大学

目　　录

试读结束：需要全本请在线购买： www.ertongbook.com

第 1 章 概率论基础知识

概率论是随机过程的基础.在传统的概率论中,由于各种原因,往往借助于直观理解来说明一些基本概念,这对于简单随机现象似乎无懈可击,但对于一些复杂随机现象就难以令人信服了.随着随机数学理论的不断完善,随机过程越来越成为现代概率论的一个重要分支和发展方向.为了更好地学习随机过程,我们必须对基础概率论的理论有一个比较深入和全面的了解.本章将系统介绍概率论基础知识,包括概率空间、随机变量及其分布、数学期望的若干性质、特征函数和母函数、随机变量列的收敛性及其相互关系、条件数学期望等.

1.1 概 率 空 间

概率论是研究随机现象统计规律的一门数学分科.由于随机现象的普遍性,概率论具有极其广泛的应用.随机试验是概率论的基本概念之一,随机试验所有可能结果组成的集合,称为这个试验的样本空间,记为 Ω. Ω 的元素 ω 称为样本点,Ω 的子集 A 称为随机事件,样本空间 Ω 也称为必然事件,空集 $\varnothing$ 称为不可能事件.

定义 1.1 设 Ω 是一个集合,$\mathscr{F}$ 是 Ω 的某些子集组成的集合簇(或称集类),如果:

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

(2) 若 $A \in \mathscr{F}$,则 $\bar{A} = \Omega \backslash A \in \mathscr{F}$;(取余集封闭)

(3) 若 $A_n \in \mathscr{F}, n = 1,2,\cdots$,则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$,(可列并封闭)

则称 $\mathscr{F}$ 为 σ-代数、Borel 域或事件域,$(\Omega,\mathscr{F})$ 称为可测空间,$\mathscr{F}$ 中的元素称为事件. 由定义可以得到:

(4) $\varnothing \in \mathscr{F}$.

(5) 若 $A, B \in \mathscr{F}$,则 $A \backslash B \in \mathscr{F}$.(取差集封闭)

(6) 若 $A_n \in \mathscr{F}, n = 1,2,\cdots$,则 $\bigcap\limits_{i=1}^{n} A_i, \bigcup\limits_{i=1}^{n} A_i, \bigcap\limits_{i=1}^{\infty} A_i \in \mathscr{F}$.(有限交、有限并、可列交封闭)

定义 1.2 设 $(\Omega,\mathscr{F})$ 为可测空间,$P(\cdot)$ 是定义在 $\mathscr{F}$ 上的实值函数,如果:

(1) 对任意 $A \in \mathscr{F}, 0 \leqslant P(A) \leqslant 1$;(非负性)

(2) $P(\Omega) = 1$;(正规性)

(3) 对两两互不相容事件 $A_1, A_2, \cdots$(当 $i \neq j$ 时,$A_i \cap A_j = \varnothing$),有 $P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$,(可列可加性)

则称 P 是 $(\Omega,\mathscr{F})$ 上的概率,$(\Omega,\mathscr{F},P)$ 称为概率空间,$P(A)$ 为事件 A 的概率.同时,由定义

我们可以得到：

(4) 如果 $A,B\in\mathscr{F}$，$A\subset B$，则 $P(B\backslash A)=P(B)-P(A)$.（可减性）

一事件列 $\{A_n,n\geqslant 1\}$ 称为单调增列，若 $A_n\subset A_{n+1}$，$n\geqslant 1$；称为单调减列，若 $A_n\supset A_{n+1}$，$n\geqslant 1$. 显然，如果 $\{A_n,n\geqslant 1\}$ 为单调增列，则 $\lim\limits_{n\to\infty}A_n=\bigcup\limits_{i=1}^{\infty}A_i$；如果 $\{A_n,n\geqslant 1\}$ 为单调减列，则 $\lim\limits_{n\to\infty}A_n=\bigcap\limits_{i=1}^{\infty}A_i$.

设 $\{A_n\}$ 为一集列，称 $\overline{\lim\limits_{x\to\infty}}A_n=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k$ 和 $\underline{\lim\limits_{n\to\infty}}A_n=\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{k=n}^{\infty}A_k$ 分别为集列 $\{A_n\}$ 的上极限和下极限. 容易看出，集列的上极限是由属于 $\{A_n\}$ 中无限多个集的那种元素的全体所组成的集合，集列的下极限是由不属于有限个集的那种元素构成的集合.

显然，$\{\bigcup\limits_{k=n}^{\infty}A_k\}$ 与 $\{\bigcap\limits_{k=n}^{\infty}A_k\}$ 分别为降列和升列，因此集列的上极限和下极限可以分别写为

$$\overline{\lim_{n\to\infty}}A_n=\lim_{n\to\infty}\bigcup_{k=n}^{\infty}A_k \quad \text{和} \quad \underline{\lim_{n\to\infty}}A_n=\lim_{n\to\infty}\bigcap_{k=n}^{\infty}A_k. \tag{1.1}$$

例如，设某人在反复地投掷硬币，观察硬币朝上的面是正面或反面. 记 A_n 为第 n 次投掷的是“正面”的事件，则

$$\overline{\lim_{n\to\infty}}A_n=\{\text{有无限多次投掷结果是“正面”}\},$$

$$\underline{\lim_{n\to\infty}}A_n=\{\text{除有限多次外，投掷结果都是“正面”}\}.$$

(5)（次可加性）若 $A_n\in\mathscr{F}$，$n=1,2,\cdots$，则 $P(\bigcup\limits_{i=1}^{\infty}A_i)\leqslant\sum\limits_{i=1}^{\infty}P(A_i)$.

(6)（概率的连续性）若 $\{A_n,n\geqslant 1\}$ 是递增的事件列，则

$$\lim_{n\to\infty}P(A_n)=P(\lim_{n\to\infty}A_n)=P(\bigcup_{n=1}^{\infty}A_n)\quad\text{（下连续性）；} \tag{1.2}$$

若 $\{A_n,n\geqslant 1\}$ 是递减的事件列，则

$$\lim_{n\to\infty}P(A_n)=P(\lim_{n\to\infty}A_n)=P(\bigcap_{n=1}^{\infty}A_n)\quad\text{（上连续性）.} \tag{1.3}$$

定义 1.3 设 $(\Omega,\mathscr{F},P)$ 为概率空间，$B\in\mathscr{F}$，且 $P(B)>0$，对任意 $A\in\mathscr{F}$，记

$$P(A\mid B)=\frac{P(AB)}{P(B)},$$

称 $P(A|B)$ 为事件 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 $A,B\in\mathscr{F}$，则

$$P(AB)=P(B)P(A\mid B).$$

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

$$P(A_1A_2\cdots A_{n-1})=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_n\mid A_1A_2\cdots A_{n-1}). \tag{1.4}$$

(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

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A\mid B_i). \tag{1.5}$$

(3) **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

$$P(B_i\mid A)=\frac{P(B_i)P(A\mid B_i)}{\sum\limits_{i=1}^{n}P(B_i)P(A\mid B_i)}. \tag{1.6}$$

一般地，若 $A_1,A_2,\cdots,A_n\in\mathscr{F}$，有 $P(\bigcap\limits_{i=1}^{n}A_i)=\prod\limits_{i=1}^{n}P(A_i)$，则称$\mathscr{F}$为独立事件族.

1.2 随机变量及其分布

随机变量是概率论的主要研究对象之一，随机变量的统计规律用分布函数来描述.

定义 1.4 设$(\Omega,\mathscr{F},P)$为概率空间，$X=X(\omega)$是定义在Ω上的实值函数，如果对于任意实数 x，有

$$X^{-1}((-\infty,x])=\{\omega:X(\omega)\leqslant x\}\in\mathscr{F}, \tag{1.7}$$

则称 $X(\omega)$为$\mathscr{F}$上的**随机变量**(random variable)，简记为 r. v. X.

值得注意的是，条件$\{\omega:X(\omega)\leqslant x\}\in\mathscr{F}$保证了 $P(X\leqslant x)$ 总有意义，通常$\mathscr{F}$是包含全体$\{X\leqslant x\}$的最小代数，并且对任意 $x\in\mathbf{R}$，有

$$\{X>x\}=\Omega-\{X\leqslant x\}\in\mathscr{F},$$

$$\{X=x\}=\{X\leqslant x\}-\bigcup_{k=1}^{\infty}\left\{X\leqslant x-\frac{1}{k}\right\}\in\mathscr{F},$$

$$\{X<x\}=\{X\leqslant x\}-\{X=x\}\in\mathscr{F}.$$

并有$\{X\geqslant x\}$，$\{x_1<X<x_2\}$，$\{x_1\leqslant X\leqslant x_2\}\in\mathscr{F}$.

随机变量实质上是$(\Omega,\mathscr{F})$到$(\mathbf{R},\mathscr{B}(\mathbf{R}))$上的可测映射(函数)，记 $\sigma(X)=\{X^{-1}(B)\mid B\in\mathscr{B}(\mathbf{R})\}\subset\mathscr{F}$，称 $\sigma(X)$为随机变量 X 所生成的σ 域. 称

$$F(x)=P(X\leqslant x)=P(\omega:X(\omega)\leqslant x)=P(X\in(-\infty,x])=P(X^{-1}(-\infty,x]) \tag{1.8}$$

为随机变量 X 的分布函数(distribution function)，简记为 d. f..

由定义可知，分布函数有如下性质：

(1) $F(x)$为不降函数，即当 $x_1<x_2$ 时，有 $F(x_1)\leqslant F(x_2)$.

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

(3) $F(x)$是右连续的，即 $F(x+0)=F(x)$.

可以证明，定义在 $\mathbf{R}$ 上的实值函数 $F(x)$，若满足上述三条性质，则其必能作为某个概率空间$(\Omega,\mathscr{F},P)$上某个随机变量的分布函数.

推广到多维情形，类似可得到：

定义 1.5 设$(\Omega,\mathscr{F},P)$为概率空间，$X=X(\omega)=(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$是定义在$\Omega$上的$n$ 维空间 $\mathbf{R}^n$ 中取值的向量实值函数. 如对于任意 $x=(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$，有$\{\omega:X_1(\omega)\leqslant x_1,X_2(\omega)\leqslant x_2,\cdots,X_n(\omega)\leqslant x_n\}\in\mathscr{F}$，则称 $X=X(\omega)$为 n 维随机变量，称

$$F(x)=F(x_1,x_2,\cdots,x_n)=P(\omega:X_1(\omega)\leqslant x_1,X_2(\omega)\leqslant x_2,\cdots,X_n(\omega)\leqslant x_n) \tag{1.9}$$

为 $X=X(\omega)=(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$ 的联合分布函数.

随机变量有两种类型:离散型随机变量和连续型随机变量.离散型随机变量的概率分布用概率分布列来描述:$p_k=P(X=x_k),k=1,2,\cdots$,其分布函数为 $F(x)=\sum\limits_{x_k\leqslant x}p_k$;连续型随机变量的概率分布用概率密度函数 $f(x)$ 来描述,其分布函数为 $F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t$.

类似可定义 n 维随机变量 $X=(X_1,X_2,\cdots,X_n)$ 的联合分布列和联合分布函数如下.

对于离散型随机变量 $X=(X_1,X_2,\cdots,X_n)$,联合分布列为

$$p_{x_1x_2\cdots x_n}=P(X_1=x_1,X_2=x_2,\cdots,X_n=x_n),$$

其中,$x_i\in I_i$,I_i 为离散集,$i=1,2,\cdots,n$;X 的联合分布函数为

$$F(y_1,y_2,\cdots,y_n)=\sum_{x_i\leqslant y_i,\ i=1,2,\cdots,n}p_{x_1,\cdots,x_n}\quad(y_1,y_2,\cdots,y_n)\in\mathbf{R}^n.$$

对于连续型随机变量 $X=(X_1,X_2,\cdots,X_n)$,如果存在 $\mathbf{R}^n$ 上的非负函数 $f(x_1,x_2,\cdots,x_n)$,对于任意 $(y_1,y_2,\cdots,y_n)\in\mathbf{R}^n$,有 $X=(X_1,X_2,\cdots,X_n)$ 的联合分布函数

$$F(y_1,y_2,\cdots,y_n)=\int_{-\infty}^{y_1}\int_{-\infty}^{y_2}\cdots\int_{-\infty}^{y_n}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n,$$

称 $f(x_1,x_2,\cdots,x_n)$ 为 X 的联合密度函数.

下面我们列出一些常见随机变量的分布.

(1) **离散均匀分布**　如果分布列为

$$p_k=\frac{1}{n},\quad k=1,2,\cdots,n,$$

则称之为离散均匀分布.

(2) **二项分布**　如果对固定 n 和 $0<p<1$,分布列为

$$p_k=C_n^kp^k(1-p)^{n-k},\quad k=0,1,\cdots,n,$$

则称之为参数为 n 和 p 的二项分布,记为 $X\sim B(n,p)$.

(3) **几何分布**　如果分布列为

$$p_k=p(1-p)^{k-1},\quad k=1,2,\cdots,$$

则称之为几何分布.

(4) **泊松分布**　如果分布列为

$$p_k=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\quad k=0,1,2,\cdots;\lambda>0,$$

则称之为参数为 λ 的 Poisson 分布,记为 $X\sim P(\lambda)$.

(5) **连续型均匀分布**　如果密度函数

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

则称之为区间 $[a,b](a<b)$ 上的均匀分布,记为 $X\sim U[a,b]$.

(6) **正态分布**　如果密度函数

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},\quad x\in\mathbf{R},$$

则称之为参数为 μ 和 σ^2 的正态分布,也称 Gauss 分布,记为 $X\sim N(\mu,\sigma^2)$.

(7) **Γ 分布**　如果密度函数

试读结束:需要全本请在线购买: www.ertongbook.com

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

其中，$\alpha>-1,\lambda>0$，则称之为参数为 α 和 λ 的 Γ 分布. 这里 Γ 函数定义为

$$\Gamma(x)=\int_0^{+\infty}y^{x-1}\mathrm{e}^{-y}\mathrm{d}y,\quad x>0.$$

(8) **指数分布**　如果密度函数

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

则称之为参数为 λ 的指数分布，记为 $X\sim E(\lambda)$.

(9) $\boldsymbol{\chi^2}$ **分布**　如果在 Γ 分布中取 $\alpha=(n-2)/2$，n 为正整数，$\lambda=1/2$，这时密度函数为

$$f(x)=\frac{1}{2^{n/2}\Gamma(n/2)}x^{\frac{n-2}{2}}\mathrm{e}^{-\frac{x}{2}},\quad x>0,$$

则称之为自由度为 n 的 χ^2 分布，记为 $X\sim\chi^2(n)$.

(10) **n 维正态分布**　设 $C=(c_{ij})$ 是 n 阶正定对称矩阵，μ 为 n 维实值列向量，随机向量 $(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 的联合概率密度为

$$f(x)=\frac{1}{(2\pi)^{n/2}(\det C)^{1/2}}\exp\left\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}C^{-1}(x-\mu)\right\},$$

其中，$x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 是实数列向量，$\mu=(\mu_1,\mu_2,\cdots,\mu_n)^{\mathrm{T}}=(EX_1,EX_2,\cdots,EX_n)^{\mathrm{T}}$ 为均值列向量，$C=(c_{ij})$ 为协方差矩阵，$c_{ij}=E[(X_i-\mu_i)(X_j-\mu_j)](i,j=1,2,\cdots,n)$，则称随机向量 $(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 服从 n 维正态分布，记为 $(X_1,X_2,\cdots,X_n)^{\mathrm{T}}\sim N(\mu,C)$.

1.3　数学期望及其性质

设 $X=X(\cdot)$ 是定义在概率空间 $(\Omega,\mathscr{F},P)$ 上的 r.v.，如果 $\int_\Omega|X|\mathrm{d}P<\infty$，就称 r.v. X 的数学期望或均值存在(或称 r.v. X 是可积的)，记为 EX. 有下列定义：

$$EX=\int_\Omega X\mathrm{d}P. \tag{1.10}$$

利用积分变换，也可写成 $EX=\int_{-\infty}^{+\infty}x\mathrm{d}F(x)$.

设 $g(x)$ 是 $\mathbf{R}$ 上的 Borel 可测函数，如果 r.v. $g(X)$ 的数学期望存在，即 $E|g(X)|<\infty$，由积分变换可知

$$Eg(X)=\int_\Omega g(X)\mathrm{d}P=\int_{-\infty}^{+\infty}g(x)\mathrm{d}F(x). \tag{1.11}$$

设 k 是正整数，若 r.v. X^k 的数学期望存在，就称它为 x 的 k 阶原点矩，记为 α_k，即

$$\alpha_k=EX^k=\int_{-\infty}^{+\infty}x^k\mathrm{d}F(x). \tag{1.12}$$

设 k 是正整数，若 r.v. $|X|^k$ 的数学期望存在，就称它为 x 的 k 阶绝对原点矩，记为 β_k，即

$$\beta_k=E|X|^k=\int_{-\infty}^{+\infty}|x|^k\mathrm{d}F(x). \tag{1.13}$$

类似地,X 的 k 阶中心矩 μ_k 和 k 阶绝对中心矩 ν_k 分别定义为

$$\mu_k = E(X - EX)^k = \int_{-\infty}^{+\infty} (x - \alpha_1)^k \mathrm{d}F(x),$$

$$\nu_k = E|X - EX|^k = \int_{-\infty}^{+\infty} |x - \alpha_1|^k \mathrm{d}F(x).$$

我们称二阶中心矩为方差,记为 $\mathrm{Var}X$ 或 DX,显然有

$$\mathrm{Var}X = \mu_2 = \nu_2 = \alpha_2 - \alpha_1^2.$$

关于数学期望,容易验证下列性质:

(1) 若 r.v. X,r.v. Y 的期望 EX 和 EY 存在,则对任意实数 α,β,$E(\alpha X+\beta Y)$ 也存在,且 $E(\alpha X+\beta Y)=\alpha EX+\beta EY$.

(2) 设 $A\in\mathscr{F}$,用 I_A 表示集 A 的示性函数,若 EX 存在,则 $E(XI_A)$ 也存在,且

$$E(XI_A) = \int_A X\mathrm{d}P.$$

(3) 若 $\{A_k\}$ 是 Ω 的一个划分,即 $A_i\cap A_j=\varnothing(i\neq j)$,且 $\Omega=\bigcup\limits_i A_i$,则

$$EX = \int_\Omega X\mathrm{d}P = \sum_i \int_{A_i} X\mathrm{d}P.$$

关于矩的存在性,有如下的必要条件和充分条件.

定理 1.1　设对 r.v. X 存在 $p>0$,使 $E|X|^p<\infty$,则有

$$\lim_{x\to\infty} x^p P(|X| \geqslant x) = 0. \tag{1.14}$$

证明　设 X 的 d.f. 为 $F(x)$,由 $E|X|^p<\infty$,因有

$$\lim_{x\to\infty} \int_{|t|\geqslant x} |t|^p \mathrm{d}F(t) = 0,$$

且

$$x^p P(|X| \geqslant x) \leqslant \int_{|t|\geqslant x} |t|^p \mathrm{d}F(t),$$

故有

$$\lim_{x\to\infty} x^p P(|X| \geqslant x) = 0.$$

定理 1.2　设对 r.v. $X\geqslant 0$(a.s.)(almost surely 的缩写),它的分布函数为 $F(x)$,那么 $EX<\infty$ 的充要条件是

$$\int_0^\infty (1 - F(x))\mathrm{d}x < \infty, \tag{1.15}$$

此时 $EX=\int_0^\infty (1-F(x))\mathrm{d}x$.

证明　对于非负随机变量 X,由 Fubini 定理,有

$$\begin{aligned} EX &= \int_\Omega X(\omega)\mathrm{d}P(\omega) = \int_\Omega \int_0^{+\infty} I_{(x\leqslant X(\omega))}\mathrm{d}x\mathrm{d}P(\omega) \\ &= \int_0^{+\infty}\int_\Omega I_{(x\leqslant X(\omega))}\mathrm{d}P(\omega)\mathrm{d}x = \int_0^{+\infty} P(X \geqslant x)\mathrm{d}x \\ &= \int_0^\infty (1 - F(x))\mathrm{d}x. \end{aligned}$$

推论 1.1　$E|X|<\infty$ 的充要条件是 $\int_{-\infty}^0 F(x)\mathrm{d}x$ 与 $\int_0^{+\infty}(1-F(x))\mathrm{d}x$ 均有限,这时有

$$EX = \int_0^{+\infty} (1 - F(x))\mathrm{d}x - \int_{-\infty}^0 F(x)\mathrm{d}x. \tag{1.16}$$

推论 1.2　对于 $0 < p < +\infty$，$E|X|^p < +\infty$ 的充要条件是 $\sum_{n=1}^{+\infty} P(|X| \geqslant n^{1/p}) < +\infty$，也等价于 $\sum_{n=1}^{+\infty} n^{p-1} P(|X| \geqslant n) < +\infty$.

证明　由 Fubini 定理，有

$$\begin{aligned} E|X|^p &= \int_0^{+\infty} P(|X|^p \geqslant x)\mathrm{d}x = \int_0^{+\infty}\int_\Omega I_{(|X|^p \geqslant x)}\mathrm{d}P\mathrm{d}x \\ &= \int_\Omega \int_0^{+\infty} p x^{p-1} I_{(|X| \geqslant x)}\mathrm{d}x\mathrm{d}P \\ &= p\int_0^{+\infty} x^{p-1} P(|X| \geqslant x)\mathrm{d}x. \end{aligned}$$

因此，$E|X|^p < +\infty$ 当且仅当 $\int_0^{+\infty} P(|X|^p \geqslant x)\mathrm{d}x < +\infty$，也等价于 $\int_0^{+\infty} x^{p-1} P(|X| \geqslant x)\mathrm{d}x < +\infty$，而这两个积分收敛分别等价于上述两个级数收敛.

1.4　特征函数和母函数

特征函数是研究随机变量分布又一个很重要的工具，用特征函数求分布律比直接求分布律容易得多，而且特征函数有良好的分析性质.

设 X, Y 是实随机变量，定义复随机变量 $Z = X + \mathrm{i}Y$，其中 $\mathrm{i} = \sqrt{-1}$，数学期望为 $E(Z) = E(X) + \mathrm{i}E(Y)$. 特别地，对于复数 $\mathrm{e}^{\mathrm{i}tX} = \cos(tX) + i\sin(tX)$，有

$$\begin{aligned} E(\mathrm{e}^{\mathrm{i}tX}) &= E(\cos tX) + \mathrm{i}E\sin(tX) \\ &= \int_{-\infty}^{+\infty} \cos(tx)\mathrm{d}F(x) + \mathrm{i}\int_{-\infty}^{+\infty} \sin(tx)\mathrm{d}F(x) \\ &= \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}tx}\mathrm{d}F(x), \quad -\infty < t < +\infty. \end{aligned}$$

由于对任意实数 t，$\cos(tx)$ 和 $\sin(tx)$ 都是关于 x 的有界连续函数，因此，对于任意随机变量 X，$E(\mathrm{e}^{\mathrm{i}tX})$ 总是存在的，而且是关于实变量 t 的函数.

定义 1.6　设 X 是 n 维随机变量（随机向量），分布函数为 $F(x)$，则称

$$g(t) = E(\mathrm{e}^{\mathrm{i}tX}) = \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}tx}\mathrm{d}F(x), \quad -\infty < t < +\infty \tag{1.17}$$

为 X 的特征函数（characteristic function），简记为 c.f..

从本质上看，特征函数是实变量 t 的复值函数，随机变量的特征函数一定是存在的. 从定义，我们能够看出特征函数有如下性质：

(1) $g(0) = 1$.

(2)（有界性）$|g(t)| \leqslant 1$.

(3)（共轭对称性）$g(-t) = \overline{g(t)}$.

(4)（非负定性）对于任意正整数 n 及任意实数 $t_1, t_2, \cdots, t_n$ 和复数 $z_1, z_2, \cdots, z_n$，有

$$\sum_{k,l=1}^{n} g(t_k - t_l) z_k \overline{z_l} \geqslant 0.$$

(5)（一致连续性）$g(t)$为 $\mathbf{R}^n$ 上的一致连续函数.

(6) 有限多个独立随机变量和的特征函数等于各自特征函数的乘积，即如随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，则 $X=X_1+X_2+\cdots+X_n$ 的特征函数为

$$g(t)=g_1(t)g_2(t)\cdots g_n(t), \tag{1.18}$$

其中，$g_i(t)$为随机变量 X_i 的特征函数.

(7)（特征函数与矩的关系）若随机变量 X 的 n 阶矩 EX^n 存在，则 X 的特征函数 $g(t)$ 可微分 n 次，且当 $k\leqslant n$ 时，有

$$g^{(k)}(0)=\mathrm{i}^k EX^k. \tag{1.19}$$

(8) 如随机变量 X 的特征函数为 $g_X(t)$，则 $Y=aX+b\,(a,b\in\mathbf{R})$ 的特征函数为

$$g_Y(t)=\mathrm{e}^{\mathrm{i}tb}g_X(at). \tag{1.20}$$

证明　这里只证明(6)，(7)和(8)，其余的显然成立.

(6) 由于 $X_1,X_2,\cdots,X_n$ 相互独立，其函数 $\mathrm{e}^{\mathrm{i}t_kX_k}$，$k=1,2,\cdots,n$ 也相互独立，从而

$$\begin{aligned} g(t)=g(t_1,t_2,\cdots,t_n) &= E\left[\mathrm{e}^{\mathrm{i}(t_1X_1+t_2X_2+\cdots+t_nX_n)}\right] \\ &= \prod_{k=1}^{n}E(\mathrm{e}^{\mathrm{i}t_kX_k})=\prod_{k=1}^{n}g_{X_k}(t_k). \end{aligned}$$

取 $t_k=t,k=1,2,\cdots,n$，有 $X=X_1+X_2+\cdots+X_n$ 的特征函数为

$$\begin{aligned} \prod_{k=1}^{n}g_{X_k}(t_k) &= g(t,t,\cdots,t) \\ &= E\left[\mathrm{e}^{\mathrm{i}(tX_1+tX_2+\cdots+tX_n)}\right]=E\left[\mathrm{e}^{\mathrm{i}t(X_1+X_2+\cdots+X_n)}\right]=g_{\sum\limits_{k=1}^{n}X_k}(t). \end{aligned}$$

特别地，若 $X_1,X_2,\cdots,X_n$ 独立同分布，特征函数为 $g(t)$，则 $X=X_1+X_2+\cdots+X_n$ 的特征函数为

$$g_X(t)=[g(t)]^n. \tag{1.21}$$

(7) 仅对 X 是连续型的情形进行证明. 设 X 的密度函数为 $f(x)$，有

$$\frac{\mathrm{d}^k[\mathrm{e}^{\mathrm{i}tx}f(x)]}{\mathrm{d}t^k}=\mathrm{i}^k x^k \mathrm{e}^{\mathrm{i}tx}f(x),$$

且

$$\int_{-\infty}^{+\infty}|\mathrm{i}^k x^k\mathrm{e}^{\mathrm{i}tx}f(x)|\,\mathrm{d}x=\int_{-\infty}^{+\infty}|x^k|f(x)\mathrm{d}x=E(|X|^k)<+\infty.$$

对 $g(t)=\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x$ 两边求导，得

$$g^{(k)}(t)=\mathrm{i}^k\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}x^kf(x)\mathrm{d}x=\mathrm{i}^kE(X^k\mathrm{e}^{\mathrm{i}tX}).$$

令 $t=0$，得 $g^{(k)}(0)=\mathrm{i}^kE(X^k)$.

(8) $g_Y(t)=E[\mathrm{e}^{\mathrm{i}t(aX+b)}]=E[\mathrm{e}^{\mathrm{i}t(aX)}\cdot\mathrm{e}^{\mathrm{i}tb}]=\mathrm{e}^{\mathrm{i}tb}E[\mathrm{e}^{\mathrm{i}t(aX)}]=\mathrm{e}^{\mathrm{i}tb}g_X(at)$.

定理 1.3　（Bocher 定理）$\mathbf{R}^n$ 上的函数 $g(t)$是某个随机变量的特征函数当且仅当 $g(t)$ 连续非负定且 $g(0)=1$.

由随机变量的分布函数可以唯一确定特征函数，下面的定理说明分布函数 $F(x)$能唯一地用特征函数 $g(t)$表示，从而建立了分布函数与特征函数的一一对应关系.

定理 1.4　（反演公式）设 $F(x)$是随机变量 X 的分布函数，相应的特征函数为 $g(t)$，则对任意两点 $x_1,x_2,-\infty<x_1<x_2<+\infty$，有：

(1) 若 x_1,x_2 为 $F(x)$的连续点，则有

$$F(x_2)-F(x_1)=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\mathrm{e}^{-\mathrm{i}tx_1}-\mathrm{e}^{-\mathrm{i}tx_2}}{\mathrm{i}t}g(t)\mathrm{d}t. \tag{1.22}$$

(2) 若 x_1,x_2 不是 $F(x)$的连续点,则有

$$\frac{F(x_2)+F(x_2-0)}{2}-\frac{F(x_1)+F(x_1-0)}{2}=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\mathrm{e}^{-\mathrm{i}tx_1}-\mathrm{e}^{-\mathrm{i}tx_2}}{\mathrm{i}t}g(t)\mathrm{d}t.$$

定理 1.4 说明,x_1,x_2 为 $F(x)$的连续点时,$F(x_2)-F(x_1)$的值完全由特征函数 $g(t)$ 给出,也就是可利用特征函数计算下面的概率:

$$P(x_1<X\leqslant x_2)=F(x_2)-F(x_1).$$

进一步,若 x_2 为 $F(x)$的连续点,令 $x_1\to-\infty$,可以得到

$$F(x)=\lim_{x_1\to-\infty}\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-\mathrm{i}tx_1}-\mathrm{e}^{-\mathrm{i}tx_2}}{\mathrm{i}t}g(t)\mathrm{d}t.$$

对于 $F(x)$的不连续点,重新定义 $\hat{F}(x)=\dfrac{F(x)+F(x-0)}{2}$,则 $\hat{F}(x)$完全由特征函数确定.

推论 1.3 （唯一性定理）分布函数 $F(x_1)$和 $F(x_2)$恒等的充分必要条件是它们的特征函数 $g_1(t)$和 $g_2(t)$恒等.

证明 若 $F_1(x)=F_2(x)$,则显然有 $g_1(t)=g_2(t)$.反过来,若 $g_1(t)=g_2(t)$,令 A 为 $F(x_1)$和 $F(x_2)$的连续点所组成的集合.若 $x\in A$,由式(1.22)可知 $F_1(x)=F_2(x)$.若 $x\notin A$,取单调下降列 $x_n\in A$,且 $\lim\limits_{n\to\infty}x_n=x$,由分布函数的右连续性,得

$$F_1(x)=\lim_{n\to\infty}F_1(x_n)=\lim_{n\to\infty}F_2(x_n)=F_2(x).$$

推论 1.4 若随机变量 X 的特征函数 $g(t)$在 $\mathbf{R}$ 上绝对可积,则 X 为连续型随机变量,其概率密度 $f(x)$和特征函数分别为

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-\mathrm{i}tx}g(t)\mathrm{d}t,\quad g(t)=\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x. \tag{1.23}$$

也就是说,特征函数 $g(t)$是概率密度 $f(x)$的 Fourier-Stieltjes 变换,而概率密度 $f(x)$是特征函数 $g(t)$的 Fourier-Stieltjes 逆变换.

推论 1.5 若离散型随机变量 X 的分布律 $p_k=P(X=k)$, $k=0,\pm1,\pm2,\cdots$,则有

$$p_k=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-\mathrm{i}tx}g(t)\mathrm{d}t,\quad g(t)=\sum_{k=-\infty}^{+\infty}p_k\mathrm{e}^{\mathrm{i}tx}. \tag{1.24}$$

具有相同特征函数的两个分布函数是恒等的.由此还可推导出一个事实:一个随机变量是对称的,当且仅当它的特征函数是实的.事实上,由 X 的对称性知 X 和 $-X$ 有相同的分布函数,根据定义 $g(t)=E\mathrm{e}^{\mathrm{i}tX}=E\mathrm{e}^{-\mathrm{i}tX}=g(-t)=\overline{g(t)}$,也就是说 $g(t)$是实的;反之,从 $g(t)=E\mathrm{e}^{\mathrm{i}tX}=\overline{g(t)}=g(-t)=E\mathrm{e}^{-\mathrm{i}tX}$ 知 X 和 $-X$ 有相同的特征函数,因此,它们的分布函数相等,这就说明 X 是对称的.

例 1.1 设 X 服从 $B(n,p)$,求 X 的特征函数 $g(t)$及 EX,EX^2,DX.

解 X 的分布列为 $P(X=k)=C_n^kp^kq^{n-k}$, $q=1-p$, $k=0,1,2,\cdots,n$,则有

$$g(t)=\sum_{k=0}^{n}\mathrm{e}^{\mathrm{i}tk}C_n^kp^kq^{n-k}=\sum_{k=0}^{n}C_n^k(p\mathrm{e}^{\mathrm{i}t})^kq^{n-k}=(p\mathrm{e}^{\mathrm{i}t}+q)^n,$$

因此

$$EX=-\mathrm{i}g'(0)=-\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}(p\mathrm{e}^{\mathrm{i}t}+q)\big|_{t=0}=np,$$

$$EX^2=(-\mathrm{i})^2g''(0)=(-\mathrm{i})^2\frac{\mathrm{d}^2}{\mathrm{d}t^2}(p\mathrm{e}^{\mathrm{i}t}+q)\big|_{t=0}=npq+n^2p^2,$$

故

$$DX = EX^2 - (EX)^2 = npq.$$

例 1.2　设 $X \sim N(0,1)$，求 X 的特征函数 $g(t)$.

解　由特征函数的定义知

$$g(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{itx - x^2/2} dx.$$

由于 $| ix e^{itx-x^2/2} | = | x e^{-x^2/2} |$，且 $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} | x | e^{-x^2/2} dx < \infty$，可对上式两边分别求导，得

$$g'(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} ix e^{itx-x^2/2} dx = \frac{i}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{itx}(-de^{-x^2/2})$$

$$= -\frac{i}{\sqrt{2\pi}} e^{itx-x^2/2}\Big|_{-\infty}^{+\infty} - \frac{t}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{itx-x^2/2} dx = -tg(t).$$

于是得到微分方程 $g'(t) + tg(t) = 0$. 这是变量可分离型方程，有

$$\frac{dg(t)}{g(t)} = -t dt.$$

对两边分别积分得 $\ln g(t) = -t^2/2 + c$，得方程的通解为 $g(t) = e^{-t^2/2+c}$. 由于 $g(0) = 1$，因此，$c = 0$. 于是 X 的特征函数为 $g(t) = e^{-t^2/2}$.

若 $X \sim N(0,1)$，令 $Y = \sigma X + \mu$，则 $Y \sim N(\mu, \sigma^2)$，由式(1.20)可得，Y 的特征函数为

$$g_Y(t) = e^{i\mu t} g_X(\sigma t) = e^{i\mu t - \sigma^2 t^2/2}.$$

例 1.3　设 X, Y 相互独立，$X \sim B(n,p)$，$Y \sim B(m,p)$，证明：$X + Y \sim B(n+m, p)$.

证明　X, Y 的特征函数分别为

$$g_X(t) = (q + pe^{it})^n, \quad g_Y(t) = (q + pe^{it})^m \quad (q = 1 - p).$$

$X + Y$ 的特征函数为

$$g_{X+Y}(t) = g_X(t) g_Y(t) = (q + pe^{it})^{n+m} \quad (q = 1 - p),$$

即 $X + Y$ 的特征函数是服从参数为 $n + m, p$ 的二项分布的特征函数，由唯一性定理有

$$X + Y \sim B(n + m, p).$$

附表 1 给出了常用分布的均值、方差和特征函数.

例 1.4　设随机变量 X 的概率密度为 $f(x) = \frac{1}{2}\cos x, -\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2}$，求 EX 和 DX.

解　由于概率密度为偶函数，于是 X 的特征函数为

$$g(t) = 2\int_0^{\frac{\pi}{2}} \frac{1}{2}\cos x \cos tx dx = \frac{1}{2}\int_0^{\frac{\pi}{2}} [\cos(t+1)x + \cos(t-1)x] dx$$

$$= \frac{1}{2}\left\{\frac{1}{t+1}\sin\left[(t+1)\frac{\pi}{2}\right] + \frac{1}{t-1}\sin\left[(t-1)\frac{\pi}{2}\right]\right\}.$$

由于 $g'(0) = 0, g''(0) = 2 - \frac{\pi^2}{4}$，因此

$$EX = i^{-1} g'(0) = 0,$$

$$DX = EX^2 + (EX)^2 = i^{-2} g''(0) = \frac{\pi^2}{4} - 2.$$

例 1.5　设随机变量 X 在 $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ 上均匀分布，且 $Y = \cos X$，利用特征函数求 Y 的概率密度.

试读结束：需要全本请在线购买： www.ertongbook.com