普通高等教育“十一五”国家级规划教材

高等学校电子信息类精品教材

随机信号分析基础

（第5版）

◆ 王永德　王 军　编著

中国工信出版集团　電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

普通高等教育“十一五”国家级规划教材
高等学校电子信息类精品教材

随机信号分析基础

(第5版)

王永德　王　军　编著

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

内容简介

本书为普通高等教育“十一五”国家级规划教材。

本书主要从工程应用的角度讨论随机信号(随机过程)的理论分析和实验研究方法。全书共10章,内容包括:随机信号两种统计特性的描述方法,重点介绍数字特征,如均值、方差、相关函数、相干函数、功率谱密度、高阶谱、谱相关理论和概率密度函数等的表述和实验测定(估计)方法;随机信号通过线性、非线性系统统计特性的变化;在通信、雷达和其他电子系统中常见的一些典型随机信号,如白噪声、窄带随机过程、高斯随机过程、马尔可夫过程等;以及在通信、雷达与模式识别系统中常用到的信号统计检测理论的基础知识。

全书是以连续时间随机信号和离散时间随机信号(随机序列)两条线展开讨论的,内容丰富、概念清楚、系统性强、理论联系实际,反映了本学科的一些新进展。书中安排了大量例题和MATLAB应用程序举例。每章末附有大量的习题供练习。附录B介绍了广泛应用的统计实验模拟方法,即蒙特卡罗模拟方法。书末给出了部分习题解答供参考。

本书可作为高等学校通信、电子信息类专业高年级本科生教材,以及研究生参考用书。也可供通信、电子及相关领域科技人员参考。

图书在版编目(CIP)数据

随机信号分析基础/王永德,王军编著. —5版. —北京:电子工业出版社,2020.3
ISBN 978-7-121-38295-6

Ⅰ.①随… Ⅱ.①王… ②王… Ⅲ.①随机信号-信号分析-高等学校-教材 Ⅳ.①TN911.6

中国版本图书馆CIP数据核字(2020)第014135号

责任编辑:韩同平
印　　刷:北京盛通商印快线网络科技有限公司
装　　订:北京盛通商印快线网络科技有限公司
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:787×1 092　1/16　印张:15　字数:432千字
版　　次:2003年7月第1版
　　　　　2020年3月第5版
印　　次:2020年3月第1次印刷
定　　价:55.90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)88254888,88258888。

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。

本书咨询联系方式:010-88254525,hantp@phei.com.cn。

前　言

本书为普通高等教育“十一五”国家级规划教材。

本书集编著者30多年教学经验，并经多次再版、修改补充后完成，是电子信息类专业本科生、研究生学习随机过程（信号）分析基本方法的教学用书。本书的核心内容、基本概念、分析方法和手段对于需要接触随机信号统计分析的其他专业的学生同样也是适用的。

“信号与系统”与“随机信号分析基础”是电子信息类专业的两门主要的专业基础课，前者主要以分析确定性的信号与系统为主要内容，后者则以分析随机信号以及与系统的相互作用为主要内容。

“随机信号分析基础”课程一般在大学本科三年级以后开设，在本课程之前，学生所接触的大多数课程都是建立在因果律或者确定性的基础上的，因而我们的思维方法也往往是这样的，对具体的函数形式、波形、必然结果感兴趣。学生初学这门课程时，往往会感到这门学科不可靠、模糊、难懂。为此，老师在讲授时有必要对本课程的特点与学习方法向学生做一些介绍，必须从它的特点出发，采用不同的学习方法才能对本课程有较好的把握。归纳起来本课程有以下3个特点：

（1）统计的概念。由于对随机过程（信号）的分析来讲，我们往往不是对一个实验结果（一个实现或一个具体的函数波形）感兴趣，而是关心大量实验结果的某些平均量（统计特性），因而随机过程（信号）的描述方式以及推演方式都应以统计特性为出发点。这样，尽管从个别的实现看不出什么规律性的东西，但从统计的角度却表现出了一定的规律性，即统计规律性。它是本门学科一个最根本的概念，从一开始就必须加以注意。

（2）模型的概念。本课程重点研究一般化（抽象化）的系统、干扰和信号，往往仅给出它们的系统函数（模型）和数学模型，而不讨论具体的系统，更不会局限于一些具体的电路系统。举出一些具体的电路系统例子，也只是用于说明一般的带普遍性的问题和处理方法。

（3）物理概念。本课程是电子信息类专业的一门专业基础课，而不是一门数学课。概率论与数理统计、随机过程理论等只是处理本门学科有关问题的一种数学工具，或者说是一种解决问题的手段。因而学习本课程除了注意处理问题的方法，更重要的是对一些数学推演的结果和结论的物理意义应有深入的理解。对一些十分复杂的数学推演的中间步骤不要死记硬背，更不必深究其数学的严密性，而要重点掌握处理问题的思路与方法。这也是将本课程命名为随机信号分析基础的原因，尽管在本书中随机信号与随机过程是同义语。

因而在学习方法上，应重点抓住上述3个特点，学习时既要理论联系实际，又要学会建立数学模型的抽象思维方法。本课程虽属基础理论性课程，但要真正理解、掌握上述3个特点，能够应用其解决实际问题，必须演算大量的习题。因而本书选编了大量的习题，除了每章指定的必做题，其他题也可根据自己的情况加以选做。

另外，利用计算机为工具，对特定随机过程采集的实验数据，或者直接由计算机模拟实际过程产生的数据进行统计分析，是研究随机过程的重要方法。因而在本书中密切结合教材内容，选编了大量基于MATLAB的典型应用程序代码举例和上机操作的习题，相信这些内容对读者更加深入地理解、掌握，以及学会用现代分析手段去分析、研究随机信号，甚至用来解决工程应用中的实际问题是会有所帮助的。

在内容安排上，本书始终以连续随机信号和离散随机信号（随机序列）两条线并行的方式展开讨论。考虑到某些学校的学生先前未学过概率论的相关知识，本书在第1章给出了概括性的介绍。对于已学过概率论的学生，教学时可以跳过此章直接从第2章开始。本教材最基本和核心的内容是第2~5章，以及第7、8章；如果学时受限，教学时可以全部或部分略去其他章节。本教材建议学时数为54~72。

不同于大多数已出版的同类教材，本教材增加了以下一些新内容，供本科课堂教学或自学时选择。其中对随机信号进行实验分析研究和计算机统计实验模拟、现代谱估值的基本方法等新内容，使本书既能成为本科生学习阶段的教材，又可作为研究生阶段深造、乃至工作中的重要参考工具。这些内容对深入理解、掌握随机信号分析的精髓，特别是利用随机信号分析的知识去解决工程实际问题是有帮助的。

(1) 鉴于信号统计检测在通信、雷达、模式识别和图像处理等领域中有重要的应用，也是随机信号分析与处理的重要内容，而很多学校在高年级并未开设专门的信号统计检测与估计的课程，因而在本书中专门开辟一章（第10章）介绍信号统计检测理论的基础知识。

(2) 功率谱估值是随机信号分析的重要内容之一。本书对此内容进行了充实细化，并单独用一章（第6章）来讨论。在介绍功率谱估值的经典法的基础上，重点增加了现代谱估值的基本方法，如参数谱估值（AR、MA和ARMA模型谱估值）、最大熵谱估值、Pisarenko谱估值等方法的介绍。由于第6章内容有一定的理论深度，本科教学时可按照各个学校和专业的需求进行适当的取舍。

(3) 除了对实际工作中记录的随机信号进行实验分析，很多时候我们还需要人为地产生各种分布和功率谱的随机过程以进行统计实验模拟，即蒙特卡罗模拟。蒙特卡罗模拟在科研工作和工程实际中有非常重要的应用，但往往难以找到一本简要介绍此内容的书籍。本书将此内容以附录的形式在书末列出，供广大师生和科技工作者参考。

在本书的修订过程中，四川大学电子信息学院夏秀渝等老师对错漏的订正和新内容的补充提出了宝贵的建议。四川大学研究生徐自励在撰写部分习题解答中做了大量工作。另外，本教材还得到四川大学电子信息学院从事过本课教学工作的师生的宝贵建议和大力支持。在此一并表示感谢。

广大读者通过书面和电子邮件方式对本书也提出了许多宝贵的意见，在此一并表示衷心的感谢，并希望读者继续给予批评和指正。

编著者

目　　录

第1章 概率论简介

本章专为未学过概率论或相关课程的读者而设置,它可以作为学习随机信号分析的基础。对于学过概率论的读者,则跳过它,直接进入第2章。

1.1 概率的基本概念

概率的概念与我们日常生活中某件事出现的机会,或说几率密切相关。把一个事件(结果)的概率同该事件出现的相对频率联系起来,是直观而易于理解的。例如,假定我们做一个实验(如一个骰子的滚动),可能有6种不同的结果$A_1,A_2,\cdots,A_6$。把实验重复N次并记录每一事件出现的次数,分别用$n_1,n_2,\cdots,n_6$表示,则它们出现的相对频率即为$n_1/N,n_2/N,\cdots,n_6/N$。在N趋于无穷大的极限情况下,这些比率就趋近于事件出现的真实概率。即

$$\lim_{N\to\infty}\frac{n_i}{N}=P(A_i),i=1,2,\cdots,6 \tag{1.1.1}$$

因此,概率是在0 ~ 1之间、并包括0和1在内的一个数,实际上,这样的概率不可能绝对准确地求得。

实验的全部可能结果的集合记为实验的样本空间。一个事件可以是样本空间的一个元素,也可以是一些可能结果的集合。两种情况中,事件出现或不出现由实验结果确定。我们用括号来表示事件,例如,$\{A\}$是样本空间的子集,其元素具有A的特性。

对任何事件$\{A\}$,都存在事件$\{非A\}$,记为$\{\overline{A}\}$事件。事件$\{A或\overline{A}\}$为全部可能结果的集合(必然事件)。事件$\{A与\overline{A}\}$是没有元素的集合(零事件)。事件$\{A或B\}$指的是,或者$\{A\}$出现,或者$\{B\}$出现,或者二者同时出现。事件$\{A与B\}$指的是$\{A\}$和$\{B\}$同时出现。以掷骰子为例,假设出现偶数点为事件$\{A\}$,则其元素为$\{2,4,6\}$,而$\{B\}$则为$\{1,3,5\}$。因而上述结论不难理解。相对频率定义的直接结果是,必然事件和零事件的概率各自为1和0。如果事件$\{A\}$和$\{B\}$互不相容(即一个出现了,另一个就不可能出现),则对事件$\{A\}$或$\{B\}$,我们得到概率$P(A或B)=P(A)+P(B)$。

假定进行两个实验,其可能结果分别记为$A_i(i=1,2,\cdots)$和$B_j(j=1,2,\cdots)$,则定义联合事件为$\{A_i和B_j\}$。像单一事件的情况那样,用一个概率与之对应,把这一联合事件的概率表示为$P(A_i,B_j)$。如果这些A_i和B_j是完备的,即不可能有其他的事件,则

$$\begin{cases}\sum\limits_i\sum\limits_j P(A_i,B_j)=1\\ \sum\limits_i P(A_i,B_j)=P(B_j),\ \sum\limits_j P(A_i,B_j)=P(A_i)\end{cases} \tag{1.1.2}$$

1.2 条件概率和统计独立

条件概率所涉及的是一事件在另一事件出现后的知识。在事件$\{A\}$出现后,事件$\{B\}$出现的概率用$P(B|A)$表示,在给定$\{A\}$时$\{B\}$的条件概率定义为

$$P(B \mid A) = P(A,B)/P(A) \tag{1.2.1}$$

这里假定 $P(A) \neq 0$。类似地,在$\{B\}$ 出现的条件下,$\{A\}$ 出现的概率为

$$P(A \mid B) = P(A,B)/P(B),\ P(B) \neq 0 \tag{1.2.2}$$

如果实验由互不相容且又完备的结果 $B_i(i = 1,2,\cdots)$ 组成,则 $\sum_i P(B_i \mid A) = 1$。

如果对于两个事件$\{A\}$ 和$\{B\}$,求得 $P(A \mid B) = P(A)$,则由条件概率的定义可以导出

$$P(A,B) = P(A)P(B) \tag{1.2.3}$$

还可以得到 $P(B \mid A) = P(B)$。式(1.2.3) 表明,其中一个事件的出现并未提供另一事件出现概率的任何信息,这样的两个事件称为是统计独立的。

若三个事件$\{A_1\}$、$\{A_2\}$ 和$\{A_3\}$ 是统计独立的,则它们必须满足下面的关系:

$$\begin{cases} P(A_1,A_2) = P(A_1)P(A_2) \\ P(A_1,A_3) = P(A_1)P(A_3) \\ P(A_2,A_3) = P(A_2)P(A_3) \\ P(A_1,A_2,A_3) = P(A_1)P(A_2)P(A_3) \end{cases} \tag{1.2.4}$$

若有三个以上事件是独立的,那么一次取二个、三个、四个事件等的概率都必须等于这些单独事件概率的乘积。

1.3 概率分布函数

我们定义随机变量或随机变数为样本空间的实值函数,即实验结果的实值函数。例如掷骰子,出现的点数是随机变量或者随机变数,点数的任意函数也是随机变量。若随机变量的取值数目(样本空间)是有限或者可数无穷的①,即称之为离散随机变量;否则是连续随机变量。

假定一个随机变量可以取6个可能值 x_i 中的任何一个,这里 $x_6 > x_5 > x_4 > x_3 > x_2 > x_1$,则相应的概率记为 $P(x_i)$ 或者 $P(X = x_i)$,如图1.1所示。这样的例子适合于研究随机变量取值小于或等于某值(比如 x_3) 的概率,在这种情况下

$$P(X \leqslant x_3) = P(x_1) + P(x_2) + P(x_3)$$

用$P(X \leqslant x)$ 定义 x 的函数称为随机变量 X 的概率分布函数(也叫分布函数或累积分布函数),图1.1也给出了前例的累积分布函数。结果$\{X \leqslant x\}$ 就是通常概率意义上的一个事件,所以累积分布函数必须满足前面所讨论的性质,特别是 $P(X < -\infty) = 0$ 和$P(X < \infty) = 1$。同样,X 落在间隔 $x_i < X \leqslant x_j$ 的概率是

$$P(X \leqslant x_j) - P(X \leqslant x_i) = P(x_i < X \leqslant x_j)$$

通常也把分布函数记为 $F_X(x)$,即有

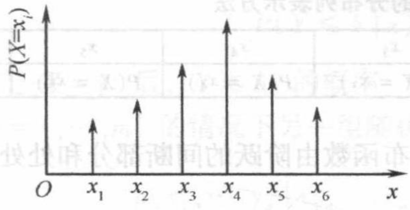
$$F_X(x) = P(X \leqslant x) \tag{1.3.1}$$

上面的讨论不难外推到两个随机变量(二元分布) 或更多随机变量(多元分布) 的情况。对于两个随机变量 X 和 Y(它们可以是连续的,也可以是离散的),下面的公式显然成立:

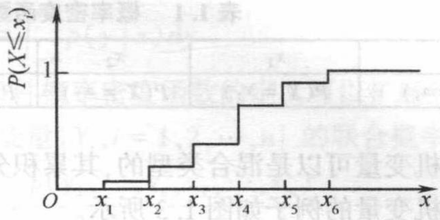
---

① 若一数集能与正整数集合一一对应关系,则该数集是可数的。

在本书中,一般用大写字母表示随机变量,用相应的小写字母表示样本空间的元素。但是用不同的符号来区分它们有时是很烦琐的,所以在以后各章中,如果文中的说明比较清楚就只用一个符号。



(a) 代表离散概率的冲激函数



(b) 累积分布函数

图 1.1　离散随机变量的概率函数

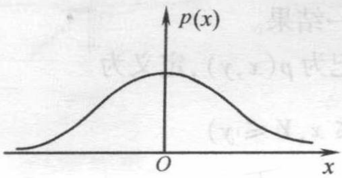
$$\begin{cases} P(X\leqslant -\infty, Y\leqslant y)=0,\ P(X\leqslant x, Y\leqslant -\infty)=0,\ P(X\leqslant \infty, Y\leqslant \infty)=1 \\ P(X\leqslant x, Y\leqslant \infty)=P(X\leqslant x),\ P(X\leqslant \infty, Y\leqslant y)=P(Y\leqslant y) \end{cases} \tag{1.3.2}$$

1.4　连续随机变量

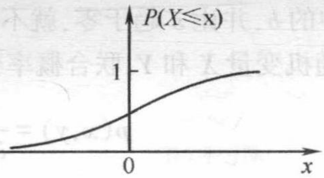
考虑一个随机变量 X,它具有图 1.2(b) 所示的连续累积分布函数,这是连续随机变量的一个例子,这种随机变量取值的数目是不可数的。例如,样本空间可以是整个实数轴。如果累积分布函数的导数存在,定义这个导数为连续随机变量的概率密度函数(或者简称密度函数)。用 $p(x)$ 表示随机变量 x 的概率密度函数,有

$$p(x)=\frac{\mathrm{d}F_X(x)}{\mathrm{d}x}=\frac{\mathrm{d}P(X\leqslant x)}{\mathrm{d}x} \tag{1.4.1}$$

注意,密度函数的定义必须包括它取值范围的说明。



(a) 概率密度函数



(b) 累积分布函数

图 1.2　连续随机变量的概率函数

如果函数 $P(X\leqslant x)$ 是定积分形式,则可以用微积分中 Leibnitz 法则来求微分。

若 $g(x)=\int_{a(x)}^{b(x)}f(t,x)\mathrm{d}t$,式中 $a(x)$, $b(x)$ 是 x 的可微函数, $f(t,x)$ 和 $\partial f(t,x)/\partial x$ 对 x 和 t 都是连续的,则

$$\frac{\mathrm{d}g(x)}{\mathrm{d}x}=\int_{a(x)}^{b(x)}\frac{\partial f(t,x)}{\partial x}\mathrm{d}t+f[b(x),x]\frac{\partial b(x)}{\partial x}-f[a(x),x]\frac{\partial a(x)}{\partial x}$$

因为累积分布函数是非降的,所以 $p(x)\geqslant 0$。图1.2(a) 为概率密度函数的一个例子。利用 δ 函数(冲激函数) 也可以把离散随机变量的概率密度函数定义为累积分布函数的导数。δ 函数在间断点出现,如图 1.1 所示,其概率密度函数可以表示为

$$p(x)=\sum_{i=1}^{6}P(x)\delta(x-x_i)$$

也可以列成一个表,用所谓分布列的形式表示,见表 1.1。

表 1.1　概率密度函数的分布列表示方法

状态 x_i	x_1	x_2	x_3	x_4	x_5	x_6
概率 $P(X=x_i)$	$P(X=x_1)$	$P(X=x_2)$	$P(X=x_3)$	$P(X=x_4)$	$P(X=x_5)$	$P(X=x_6)$

通常,随机变量可以是混合类型的,其累积分布函数由阶跃的间断部分和处处连续的部分组成,这类随机变量的例子如图 1.3 所示。

从概率密度函数的定义直接得出

$$P(X \leqslant a)=\int_{-\infty}^{a} p(x)\,\mathrm{d}x \tag{1.4.2}$$

$p(x)\,\mathrm{d}x$ 可以解释为随机变量落在 x 和 $x+\mathrm{d}x$ 之间的概率。随机变量落在区间 $a \leqslant X < b$ 的概率为

$$P(a < X \leqslant b)=\int_{a}^{b} p(x)\,\mathrm{d}x \tag{1.4.3}$$

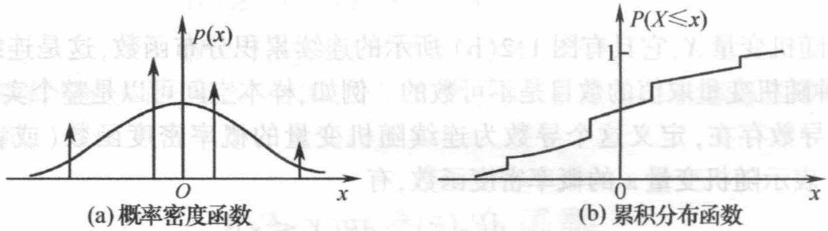


图 1.3　混合随机变量的概率函数

对连续随机变量来说,它落在一个区间的概率随这个区间的减小而趋于零。用 $a+\varepsilon$ 来代替式(1.4.3) 中的 b,并让 ε 趋于零,就不难看出这一结果。

对于两个随机变量 X 和 Y,联合概率密度函数记为 $p(x,y)$,定义为

$$p(x,y)=\frac{\partial^2}{\partial x \partial y} P(X \leqslant x, Y \leqslant y) \tag{1.4.4}$$

由此可得①

$$\begin{cases} P(X \leqslant a, Y \leqslant b)=\int_{-\infty}^{b} \mathrm{d}y \int_{-\infty}^{a} p(x,y)\,\mathrm{d}x \\ P(X \leqslant a)=\int_{\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{a} p(x,y)\,\mathrm{d}x \\ P(Y \leqslant b)=\int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{b} p(x,y)\,\mathrm{d}y \end{cases} \tag{1.4.5}$$

同理可证

$$p(x)=\int_{-\infty}^{+\infty} p(x,y)\,\mathrm{d}y \tag{1.4.6}$$

以及

$$p(y)=\int_{-\infty}^{+\infty} p(x,y)\,\mathrm{d}x \tag{1.4.7}$$

用这种方法引出的概率密度函数有时叫做边缘概率密度函数,这也就是由高维概率密度函数求低维概率密度函数通用的方法。

给定随机变量 X 后,变量 Y 的条件概率密度函数定义为

$$p(y \mid x)=p(x,y)/p(x), \qquad p(x) \neq 0 \tag{1.4.8}$$

---

① 按照普通习惯,用 $p(\cdot)$ 表示 X、Y 各自的概率密度函数,也表示 X,Y 的联合概率密度函数,即使它们有不同的函数形式,但却可用宗量来加以区分。如果这一点在文中不够清楚,则要用脚标来区别不同的函数形式。

于是 $$P(Y \leqslant b \mid x) = \int_{-\infty}^{b} p(y \mid x)\,\mathrm{d}y \tag{1.4.9}$$
应解释为给定$\{X = x\}$后，$\{Y \leqslant b\}$的概率。条件概率密度函数的定义可以扩展到给定一组随机变量$\{X_j, j = 1,\cdots,m\}$的情况下另一组随机变量$\{Y_i, i = 1,2,\cdots,n\}$的联合概率。这样

$$p(y_1,\cdots,y_n \mid x_1,\cdots,x_m) = \frac{p(y_1,\cdots,y_n,x_1,\cdots,x_m)}{p(x_1,\cdots,x_m)} \tag{1.4.10}$$

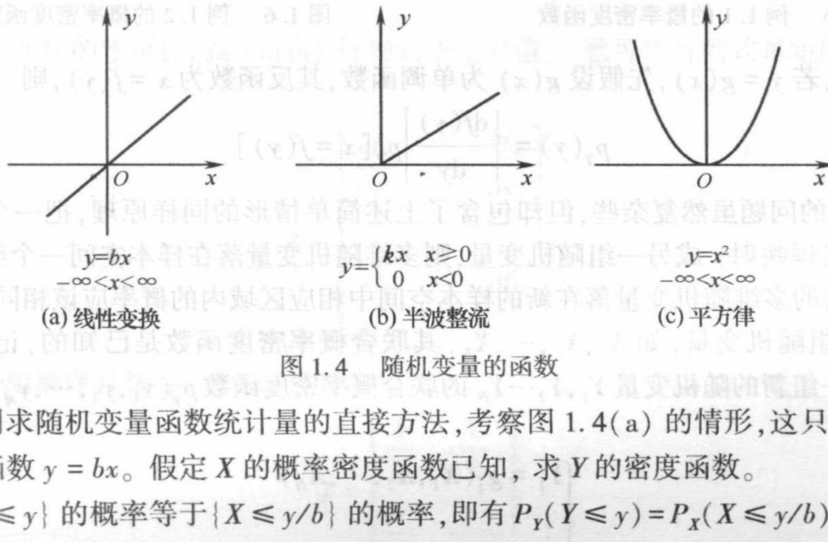
统计独立的条件同样可以用概率密度函数来表述：对$x,y,\cdots,z$的所有取值，当满足

$$p(x,y,\cdots,z) = p(x)p(y)\cdots p(z) \tag{1.4.11}$$

而且也只有满足这个条件时，随机变量$X,Y,\cdots,Z$才是统计独立的。

1.5 随机变量的函数

一个或多个随机变量的函数，经常在随机信号分析、检测理论以及同概率论和统计数学有关的其他学科中出现。一个随机变量的函数$Y = g(X)$是这样表述的：观察由实验得到的实数x，然后完成由$y = g(x)$定义的算术运算。典型例子如图1.4所示。这也可以推广到多个随机变量函数的情形。例如$Y = g(X,Z)$，即可由观察一对实数值x和z并完成$y = g(x,z)$的函数映射来表述，求和$y = x + z$便是一例。



(a) 线性变换　　(b) 半波整流　　(c) 平方律

图1.4　随机变量的函数

为了说明求随机变量函数统计量的直接方法，考察图1.4(a)的情形，这只是一个按比例变化的线性函数$y = bx$。假定X的概率密度函数已知，求Y的密度函数。

由于$\{Y \leqslant y\}$的概率等于$\{X \leqslant y/b\}$的概率，即有$P_Y(Y \leqslant y) = P_X(X \leqslant y/b)$。由概率密度函数的定义直接得到

$$p_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} P_X(X \leqslant y/b)$$

由于分布函数是非降的，所以导数不能为负，应用Leibnitz法则可以证明

$$p_Y(y) = \frac{1}{|b|} p_X(x = y/b)$$

式中，$|\cdot|$表示绝对值。Y的取值范围是X的取值范围乘以b。

例1.1　若上述情况下X的概率密度函数是指数的（换句话说，X按指数规律分布），即$p_X(x) = \mathrm{e}^{-x}(x \geqslant 0)$，则可以直接得出

$$p_Y(y) = \frac{1}{|b|}\mathrm{e}^{-y/b}, \begin{cases} 若\ b > 0, & 则\ y > 0 \\ 若\ b < 0, & 则\ y \leqslant 0 \end{cases}$$

X 的概率密度函数以及 $b=2$ 时 Y 的概率密度函数如图 1.5 所示。

例 1.2 设概率密度函数 $p_X(x)$ 和前例相同,求 $Y=X+a$ 时 Y 的概率密度函数。

解: 应用直接方法。

$$P_Y(Y\leqslant y)=P_X[X\leqslant(y-a)]$$

因为
$$p_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}P_Y(Y\leqslant y)=\frac{\mathrm{d}}{\mathrm{d}y}P_X[X\leqslant(y-a)]$$

所以
$$p_Y(y)=p_X(x=y-a)=\mathrm{e}^{-(y-a)},\quad y>a$$

其示意图如图 1.6 所示。

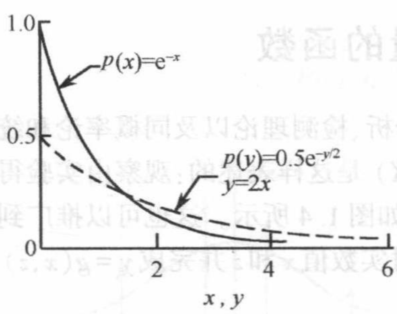


图 1.5　例 1.1 的概率密度函数

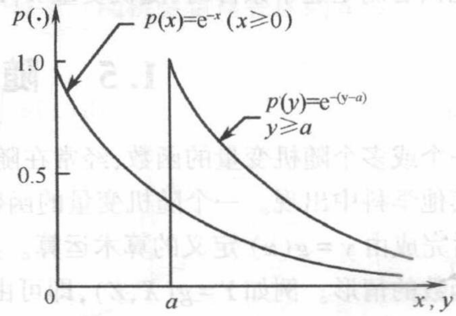


图 1.6　例 1.2 的概率密度函数

一般情况,若 $y=g(x)$,先假设 $g(x)$ 为单调函数,其反函数为 $x=f(y)$,则

$$p_Y(y)=\left|\frac{\mathrm{d}f(y)}{\mathrm{d}y}\right|p_X[x=f(y)]\tag{1.5.1}$$

下面讨论的问题虽然复杂些,但却包含了上述简单情形的同样原理,把一个或多个随机变量变换(或者说映射)成另一组随机变量,则多维随机变量落在样本空间一个给定区域内的概率,与变换后的多维随机变量落在新的样本空间中相应区域内的概率应该相同。

假定有一组随机变量,如 $X_1,X_2,\cdots,X_N$,其联合概率密度函数是已知的,记为 $p_X(x_1,x_2,\cdots,x_N)$,想求一组新的随机变量 $Y_1,Y_2\cdots Y_N$ 的联合概率密度函数 $p_Y(y_1,y_2,\cdots,y_N)$,Y 同 X 的函数关系为

$$\begin{cases}Y_1=g_1(X_1,X_2,\cdots,X_N)\\ \vdots\\ Y_N=g_N(X_1,X_2,\cdots,X_N)\end{cases}\tag{1.5.2}$$

例如,对 $N=2$ 的简单情况

$$Y_1=X_1+X_2,\qquad Y_2=X_1-X_2$$

暂时假定新变量 Y_i 的个数 N 等于旧变量 X_i 的个数,新随机变量为旧随机变量的单值连续函数,具有处处连续的偏导数,而且旧变量也可以表示为新变量的单值连续函数(反函数),即

$$\begin{cases}X_1=f_1(Y_1,Y_2,\cdots,Y_N)\\ \vdots\\ X_N=f_N(Y_1,Y_2,\cdots,Y_N)\end{cases}\tag{1.5.3}$$

用前面的例子,反函数为

$$X_1=\frac{1}{2}(Y_1+Y_2),X_2=\frac{1}{2}(Y_1-Y_2)$$

因此，$X_i(i=1,2,\cdots,N)$ 样本空间中的每一个点，对应于而且只对应于 $Y_i(i=1,2,\cdots,N)$ 样本空间中的一个点，即旧变量和新变量之间有一一对应的映射关系。

假定一个特殊样点集包含在 X 域的范围 A 内，由于 $\{X_i\}$ 和 $\{Y_i\}$ 之间的函数关系，所以范围 A 映射成 Y 域内的范围 B，正如图 1.7 所示的那样，则样点 $x_1,x_2,\cdots,x_N$ 落入 A 的概率与样点 $y_1,y_2,\cdots y_N$，落入 B 的概率相同。图中 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 分别表示多维变量 $x_1,x_2,\cdots,x_N$ 和 $y_1,y_2,\cdots,y_N$。所以

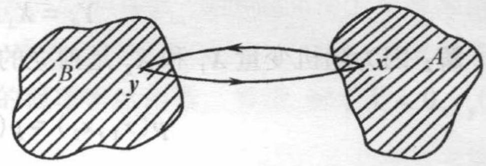


图 1.7　空间 A 一一对应地映射成空间 B 的说明

$$\underbrace{\int\cdots\int}_{A} p_X(x_1,x_2,\cdots,x_N)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_N=\underbrace{\int\cdots\int}_{B} p_Y(y_1,y_2,\cdots,y_N)\,\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_N \tag{1.5.4}$$

式中，$p_X(x_1,x_2,\cdots,x_N)$ 是 $X_i(i=1,2,\cdots,N)$ 的联合概率密度函数，是已知的；$p_Y(y_1,y_2,\cdots,y_N)$ 是 Y_i 的联合概率密度函数，是待求的。

在等式左端积分中采用多元函数积分中变换变量的标准方法，得到

$$\underbrace{\int\cdots\int}_{A} p_X(x_1,x_2,\cdots,x_N)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_N$$

$$=\underbrace{\int\cdots\int}_{B} p_X\{(x_1=f_1(y_1,y_2,\cdots,y_N),\cdots,x_N=f_N(y_1,y_2,\cdots,y_N))\}\,|J|\,\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_N$$

式中，$|J|$ 是变换的雅可比(Jacobian)行列式的绝对值。雅可比行列式是矩阵 $\boldsymbol{J}$ 的行列式，定义为

$$\boldsymbol{J}=\begin{bmatrix}\dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_N}{\partial y_1}\\ \vdots & & \vdots\\ \dfrac{\partial f_1}{\partial y_N} & \cdots & \dfrac{\partial f_N}{\partial y_N}\end{bmatrix} \tag{1.5.5}$$

因为假定积分限顺序是增加的，所以使用雅可比式的绝对值。对于前面的例子

$$\boldsymbol{J}=\begin{bmatrix}\dfrac{1}{2} & \dfrac{1}{2}\\ \dfrac{1}{2} & -\dfrac{1}{2}\end{bmatrix}$$

可得 $|J|=1/2$。

将这个积分与式(1.5.4)的积分相比较，我们可以看出新旧联合概率密度函数有下述关系：

$$p_Y(y_1,y_2,\cdots,y_N)=p_X\{[x_1=f_1(y_1,y_2,\cdots,y_N),\cdots,x_N=f_N(y_1,y_2,\cdots,y_N)]\}\,|\boldsymbol{J}| \tag{1.5.6}$$

这个公式连同 y_i 的定义域，完全确定了它的密度函数，显然一维情况的雅可比式的绝对值是 $|\mathrm{d}x/\mathrm{d}y|$。

例 1.3　假定独立随机变量 X_1 和 X_2 按正态(高斯)分布，其概率密度函数为①

$$p(x_1)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x_1^2/2},\quad p(x_2)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x_2^2/2}$$

---

①　指数函数 e^x，也写做 $\exp[x]$，本书两种表示均用到。

求 Y_1 和 Y_2 的概率密度函数,这里

$$Y_1 = X_1 + X_2 = g_1(X_1, X_2)$$
$$Y_2 = X_1 - X_2 = g_2(X_1, X_2)$$

解: 因为随机变量 X_1 和 X_2 是独立的,所以 X_1 和 X_2 的联合概率密度函数是

$$p(x_1, x_2) = p(x_1)p(x_2) = \frac{1}{2\pi}e^{-(x_1^2+x_2^2)/2}$$

反变换是

$$X_1 = \frac{1}{2}(Y_1 + Y_2) = f_1(Y_1, Y_2)$$

$$X_2 = \frac{1}{2}(Y_1 - Y_2) = f_2(Y_1, Y_2)$$

雅可比行列式为

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial Y_1} & \dfrac{\partial f_2}{\partial Y_1} \\ \dfrac{\partial f_1}{\partial Y_2} & \dfrac{\partial f_2}{\partial Y_2} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

所以 $|J| = 1/2$,且

$$p(y_1, y_2) = \frac{1}{4\pi}e^{-(y_1^2+y_2^2)/4} = \frac{1}{\sqrt{4\pi}}e^{-y_1^2/4}\frac{1}{\sqrt{4\pi}}e^{-y_2^2/4}$$

可见,在这种情况下,Y_1 和 Y_2 是统计独立的高斯随机变量。

前面涉及的是新旧随机变量数目相同的情况。当新随机变量数目少于旧随机变量数目时可以类似地处理。这里用一个实例来说明。

例 1.4 给定两个随机变量 X_1 和 X_2,其联合概率密度函数为 $p_X(x_1, x_2)$。求新随机变量 Y 的概率密度函数 $p_Y(y)$,这里

$$Y = X_1 + X_2 = g(X_1, X_2)$$

$$p_X(x_1, x_2) = \frac{1}{2\pi}e^{-(x_1^2+x_2^2)/2}$$

解: 为了使用上面的结果,改写上面的变换为

$$Y_1 = X_1 + X_2 = g_1(X_1, X_2)$$
$$Y_2 = X_2 = g_2(X_1, X_2)$$

以保持新旧变量数目相同。可得其反变换为

$$X_1 = Y_1 - Y_2 = f_1(Y_1, Y_2)$$
$$X_2 = Y_2 = f_2(Y_1, Y_2)$$

雅可比行列式为

$$J = \begin{vmatrix} \dfrac{\partial f_1}{\partial Y_1} & \dfrac{\partial f_2}{\partial Y_1} \\ \dfrac{\partial f_1}{\partial Y_2} & \dfrac{\partial f_2}{\partial Y_2} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

Y_1 和 Y_2 的联合概率密度函数为

$$p_Y(y_1, y_2) = p_X(x_1 = y_1 - y_2, x_2 = y_2)\,|J| = \frac{1}{2\pi}e^{-(y_1^2 - 2y_1y_2 + 2y_2^2)/2}$$

然后,通过求边缘分布的方法即可得 Y_1,也就是 Y 的概率密度函数,即