



工业和信息化部“十二五”规划教材

随机信号分析

王新宏 马 艳◎编



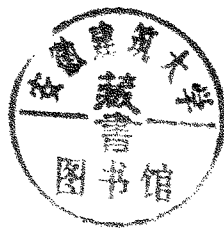
西北工业大学出版社



工业和信息化部“十二五”规划教材

随机信号分析

王新宏 马 艳 编



西北工业大学出版社

【内容简介】 本书是高等院校电子信息类专业的专业基础课教材。主要包括概率论基础,随机信号的时域、频域分析方法,随机信号通过线性系统、非线性系统的理论与分析方法,窄带随机过程的特性和分析方法,以及马尔可夫过程等,共分七章,每章都附有相应数量的习题。

本书强调对基本概念和物理含义的理解,注重基本理论、分析方法与工程实践相结合,通过结合信号处理领域的典型实例的分析,以加深对基本理论与分析方法的理解和应用。

本书可作为电子信息类专业高年级本科生和研究生教材,亦可供相关领域的科研和工程技术人员参考。

图书在版编目(CIP)数据

随机信号分析/王新宏,马艳编. —西安:西北工业大学出版社,2014.10

工业和信息化部“十二五”规划教材

ISBN 978-7-5612-4167-7

I. ①随… II. ①王… ②马… 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(2014)第230042号

出版发行:西北工业大学出版社

通信地址:西安市友谊西路127号 邮编:710072

电话:(029)88493844 88491757

网址:<http://www.nwpu.com>

印刷者:陕西宝石兰印务有限责任公司

开本:787 mm×1 092 mm 1/16

印张:14.5

字数:351千字

版次:2014年10月第1版

2014年10月第1次印刷

定价:39.00元

前 言

随机信号分析主要研究随机信号的特点和分析方法,是信号处理的基础理论之一,广泛应用于雷达、声呐、通信、语音信号处理、图像信号处理、自动控制、随机振动、气象预报、生物医学、地震信号处理等领域。随着信息技术的发展,随机信号分析的理论和应用将日益广泛和深入。

随机信号分析课程是电子信息类专业的专业基础课,是信号检测与估计、自适应滤波等信号处理类课程的基础。在学习本课程之前,学生已经系统地学习了确定信号的特性和分析方法,设置本课程的目的是使学生建立随机信号的观念,认识到绝大多数的信号处理领域研究的问题都是针对随机信号的,培养学生采用统计思维认识、研究问题的能力。在此基础上,学生能够全面、系统地掌握随机信号的基本概念、物理含义和分析方法。

随机信号分析是一门将数学与实际工程应用相结合的课程,它采用统计学的方法研究信号分析的问题。初学者常常会陷入两种状况,一种是将这门课程当作数学课,不注重对基础概念的物理含义的理解,不注重与实际问题相联系,死背概念和公式;另一种是迟迟建立不起随机信号的观念,不习惯用统计的方法来研究问题,总是用确定信号的方法来看待随机信号的问题。

本书是笔者在多年讲授随机信号分析课程及指导学生论文工作的基础上,参考了目前同类教材的长处,结合教学体会和科研经验编写的。针对初学者容易遇到的问题,本书在内容的编排和讲解上,强调对基本概念和物理含义的理解,注重基础理论、分析方法与工程实践相结合,在基本概念、分析方法的讲述中,结合对信号处理领域的典型实例的分析,加深对基础理论与分析方法的理解和应用。本书还在内容上做了相应的加深和拓展,引入了高阶累积量、高阶谱等相对较新的内容。

读者在使用本书时,也应时刻意识到这是一门专业基础课,而不是一门数学课,数学工具的运用是为了解决工程应用问题。因此,在学习时,要注重数学公式的物理含义的理解,注重数学推导过程的方法、结果和结论的理解,明白数学公式是要表达某个物理量,以及与这个物理量有关的因素,明白数学推导的目的是为了得到某个工程应用需要的结果或结论,以及为得到这个结果或结论而采用的方法。在学习时,不要对数学公式和数学推导过程死记硬背,更不必深究其数学推

导的严密性,而重在结合实际问题,加强基本概念和物理含义的理解,以及分析问题的思路和方法的掌握。

全书共分7章。第1章概率论基础,归纳了后续章节要用到的概率论的基础知识。第2章引入了随机信号和相关理论的概念,讲述了随机信号时域分析中常用的数字特征,以及工程应用中重要的平稳随机过程、各态历经过程的概念和特性。第3章讲述随机信号的频域分析方法,以及时域和频域的关系。第4章和第6章讲述了随机信号通过线性系统和非线性系统后输出端的统计特性和分析方法,第5章讲述了窄带随机信号的特性和分析方法。第7章讲述了实际中可能遇到的马尔可夫过程、独立增量过程和独立随机过程等。

读者在使用本书之前,应具有微积分、概率论、电路基础、信号与系统和数字信号处理等基础知识。

本书第2~5章由王新宏编写,第1,6,7章由马艳编写。在编写过程中,硕士生张娜、年华做了大量的文字录入、编程和图形绘制工作,在此一并表示感谢。

由于水平有限,书中难免有错误和疏漏之处,敬请读者批评指正。

编者

2014年7月

目 录

第 1 章 概率论基础	1
1.1 概率简述	1
1.2 随机变量与概率分布	3
1.3 随机变量的函数变换	14
1.4 随机变量的数字特征	20
1.5 多维正态随机变量	32
1.6 复随机变量及其统计特性	35
习题一	35
第 2 章 随机信号的时域分析	39
2.1 随机过程的基本概念	39
2.2 随机过程的统计特性	43
2.3 随机过程的微分与积分	53
2.4 平稳随机过程	56
2.5 各态历经(遍历性)过程	64
2.6 复随机过程	69
2.7 随机序列	72
2.8 高斯随机过程	76
习题二	79
第 3 章 随机信号的谱分析	83
3.1 随机过程的功率谱密度	83
3.2 功率谱密度与相关函数之间的关系	87
3.3 平稳随机过程功率谱密度的性质	93
3.4 随机序列的功率谱密度	96
3.5 白噪声	99
3.6 时变功率谱和高阶谱	102
习题三	104
第 4 章 随机信号通过线性系统	106
4.1 随机信号通过连续时间线性系统	106

4.2 随机序列通过离散时间线性系统	117
4.3 白噪声通过线性系统	120
4.4 线性系统输出端随机信号的概率分布	127
习题四	129
第 5 章 窄带随机过程	133
5.1 希尔伯特变换	133
5.2 复信号	137
5.3 窄带随机过程的统计特性	144
5.4 窄带高斯过程包络和相位的概率分布	147
习题五	154
第 6 章 随机信号通过非线性系统	156
6.1 引言	156
6.2 直接法	157
6.3 特征函数法	170
6.4 随机信号通过有记忆非线性系统	181
6.5 非线性变换后信噪比的计算	186
习题六	188
第 7 章 马尔可夫过程、独立增量过程及独立随机过程	191
7.1 马尔可夫过程	191
7.2 独立增量过程	204
7.3 独立随机过程	218
习题七	220
附录 常用傅里叶变换对	223
参考文献	225

第 1 章 概率论基础

概率论是研究随机信号分析的基础,本章简单介绍了概率论的相关知识,为后续章节的学习奠定基础。如果学过概率论课程,可以直接跳过本章。

本章的内容主要包括概率简述、条件概率与统计独立、随机变量与概率分布、随机变量的函数变换、统计平均、特征函数、多维正态随机变量、复随机变量及其统计特性。

1.1 概率简述

自然界和社会上发生的各种现象可以粗略地划分成两类。有一类现象是在一定条件下必然发生的,称之为确定性现象。微积分学、线性代数等就是研究这类现象的数学工具。另有一类现象,在相同条件下进行多次重复试验,每次试验的可能结果不止一个,而试验前不能准确地预言它的结果,即呈现出不确定性。这类现象,虽然在个别试验中呈现出不确定性,但在大量重复试验中,却呈现出某种规律性——统计规律性。这类现象称之为随机现象。概率论与数理统计就是研究随机现象统计规律性的数学学科。

1.1.1 概率

在概率论中,把具有以下 3 个特性的试验称为随机试验:①可以在相同条件下重复进行;②每次试验的可能结果不止一个,并能事先确定试验的所有可能结果;③每次试验前不能确定哪个结果会出现。

在随机试验中,对一次试验可能出现也可能不出现的事情称为随机事件(简称事件)。每个可能出现的结果都是一个随机事件,它们是这个试验的最简单的随机事件,称之为基本事件。把随机试验所有基本事件(可能结果)组成的集合,叫做试验的样本空间,记为 S 。一个事件可以是样本空间中的一个元素,也可以是一些可能结果的集合。亦即 S 中的每一事件是 S 中的子集。

在大量实践中发现,尽管随机事件的发生与否是偶然的、变化的,但每一个随机事件发生的“可能性大小”却是固定不变的。为了在数量上比较各个事件出现的可能性的,用数字 $P(A)$ 来表示事件 A 出现的可能性,并称 $P(A)$ 为事件 A 的概率。 $P(A)$ 大,则表示 A 出现的可能性大; $P(A)$ 小,则表示 A 出现的可能性小。不过,当重复试验次数不多时,事件出现的频率有明显的随机性,一组试验与另一组试验的结果会不相同。只有当试验次数增多时,事件出现的相对频率才逐渐减小随机性,趋于稳定。

可以证明:在相当广泛的条件下,当试验次数 n 无限增加时,在一定意义下,事件 A 出现的频率趋近于事件 A 出现的概率,即

$$\lim_{n \rightarrow \infty} \frac{n_A}{n} \rightarrow P(A) \quad (1.1)$$

式中, n_A 是事件 A 出现的次数。

现在给出度量事件出现可能性大小的概率的公理化定义。

定义 1.1 设 S 是随机试验的样本空间, 对于试验中的每一事件 A 赋予一个实数, 记为 $P(A)$, 如果它满足条件:

- 1) 对于每一事件 A 有 $0 \leq P(A) \leq 1$;
- 2) $P(S) = 1$;
- 3) 对于两两互不相容的事件 $A_k (k=1, 2, \dots, n)$, 有

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) \quad (1.2)$$

则称 $P(A)$ 为事件 A 的概率。

1.1.2 条件概率

1. 条件概率的定义

条件概率所涉及的是一个事件在另一个事件发生后的知识。

定义 1.2 设 A, B 为随机试验的两个事件, 且 $P(A) > 0$, 称

$$P(B | A) = \frac{P(AB)}{P(A)} \quad (1.3)$$

为事件 A 发生条件下事件 B 发生的条件概率。类似地有

$$P(A | B) = \frac{P(AB)}{P(B)} \quad (1.4)$$

为事件 B 发生条件下事件 A 发生的条件概率。

2. 概率的乘法定理

设 $P(A) > 0$, 则

$$P(AB) = P(B | A)P(A) \quad (1.5)$$

利用这个定理可以计算两个事件 A, B 同时发生的概率 $P(AB)$ 。

3. 全概率公式

设试验的样本空间为 S , A 为试验的一个事件, $B_1, B_2, \dots, B_n$ 为 S 的一个划分, 且 $P(B_i) > 0 (i=1, 2, \dots, n)$, 则

$$P(A) = P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_n)P(B_n) = \sum_{i=1}^n P(A | B_i)P(B_i) \quad (1.6)$$

全概率公式可用来解决这样一类问题: 直接计算 $P(A)$ 困难, 但能找到 S 的一个划分 $B_1, B_2, \dots, B_n$, 且能求得 $P(B_i)$ 及 $P(A | B_i) (i=1, 2, \dots, n)$, 此时就可用全概率公式求得 $P(A)$ 。

4. 贝叶斯 (Bayes) 公式

设 $B_1, B_2, \dots, B_n$ 为样本空间 S 的一个划分, 且 $P(B_i) > 0 (i=1, 2, \dots, n)$ 。对于任一事件

$A, P(A) > 0$, 由乘法定理 $P(AB_i) = P(B_i | A)P(A) = P(A | B_i)P(B_i)$ 及全概率公式

$P(A) = \sum_{i=1}^n P(A | B_i)P(B_i)$, 即得

$$P(B_i | A) = \frac{P(A | B_i)P(B_i)}{\sum_{i=1}^n P(A | B_i)P(B_i)} \quad (1.7)$$

式(1.7)称为贝叶斯公式。

贝叶斯公式可用来解决这样一类问题: 事件 A 可以在不同条件 $B_1, B_2, \dots, B_n$ 下实现, 并在已知 $P(B_i)$ 及 $P(A | B_i)$ ($i=1, 2, \dots, n$) 的情况下, 当一次试验中, 已观察到 A 发生时, 求与 A 的发生相联系的某一假设事件 B_i 发生的概率 $P(B_i | A)$ 。

通常把观察前分析计算得到的概率 $P(B_i)$ 称为先验概率, 而把已观察到 A 发生后, 按贝叶斯公式计算得到的条件概率 $P(B_i | A)$ 称为后验概率。

1.1.3 统计独立性

设 A, B 为随机试验的两个事件, 若 $P(A) > 0$, 可以定义 $P(B | A)$, 一般情况下, $P(B | A) \neq P(B)$, 即 A 的发生对于 B 发生的概率有影响。只有当 $P(B | A) = P(B)$ 时, 才可以认为这种影响不存在。此时, 由条件概率定义可导出:

$$P(AB) = P(A)P(B | A) = P(A)P(B) \quad (1.8)$$

这就是说, 其中一个事件的发生并未提供另一事件发生概率的信息, 这样两个事件称为统计独立的。式(1.8)也就是两事件相互独立的定义式。

式(1.8)还可以推广到 n 个相互独立事件的情况。设 A, B, C 是 3 个事件, 若满足

$$\left. \begin{aligned} P(AB) &= P(A)P(B) \\ P(BC) &= P(B)P(C) \\ P(AC) &= P(A)P(C) \end{aligned} \right\} \quad (1.9)$$

$$P(ABC) = P(A)P(B)P(C) \quad (1.10)$$

则称 A, B, C 为相互独立的事件。若只满足式(1.9), 而不满足式(1.10)时, 则称 3 事件 A, B, C 两两相互独立。

1.2 随机变量与概率分布

1.2.1 随机变量

一次随机试验有许多种可能的结果。有些随机试验的各种结果与某一数值相联系, 而另一些试验, 其可能的各种结果往往和数值没有直接的联系, 但可以规定一些数值来表示它的各个可能结果。如掷一枚硬币, 可规定“1”表示出现正面, “0”表示出现反面。于是, 总可以用一个变量 X 来定量地表示随机试验的结果。换言之, 随机试验的各个可能结果, 可通过 X 可能

取的数值定量地表示出来,这个变量就称作随机变量。显然随机变量可能取值的范围是由它所对应的那个随机试验来决定的,而且在试验前就已知。但在试验前无法准确预言随机变量将取什么值,只能知道以多大概率分别取这些值。对应一次试验结果,随机变量取唯一的确定值。下面给出随机变量的定义。

定义 1.3 设随机试验的样本空间 $S=\{\zeta\}$,如果对于每一个 $\zeta \in S$ 有一实数 $X(\zeta)$ 和它对应,这样就得到一个定义在 S 上的实值单值函数 $X(\zeta)$,称 $X(\zeta)$ 为随机变量。

本书将以大写英文字母 X, Y 表示随机变量,用相应的小写字母 x, y 表示随机变量的可能取值。随机变量概念的引入使概率论的研究对象由具体事件抽象为随机变量,使人们有可能利用数学分析的方法去研究随机试验。

有些随机变量,其全部可能取到的值是有限个或者可列无限多个,这种随机变量称之为离散型随机变量;否则就是连续型随机变量。连续型随机变量有无限多个可能值,且连续地占据着整个取值区间。

1.2.2 概率分布

1. 离散型随机变量的分布列

为了研究随机变量 X 的统计规律,必须要知道 X 的所有可能取值,以及取每个可能值的概率。把随机变量 X 各可能取值与其相应概率之间的各种不同形式的对应关系,统称为随机变量的分布律。

若离散型随机变量 X 的所有可能取值为 $x_i (i=1, 2, \dots)$, X 取各 x_i 相应的概率为 $p_i (i=1, 2, \dots)$, 即 $P\{X=x_i\}=p_i$, 则 X 的统计特性可用表 1.1 来表示。

表 1.1 X 的概率分布

x_1	x_2	$\dots$	x_n	$\dots$
p_1	p_2	$\dots$	p_n	$\dots$

由于 x_i 是 X 的所有可能取值,而且 X 取某一值时不能同时取另一值,有

$$p_1 + p_2 + \dots + p_n + \dots = \sum_{i=1}^{\infty} p_i = 1 \quad (1.11)$$

可见,表 1.1 把总量为“1”的概率,分配到各个取值 x_i 上,而且不同的随机变量将有着不同的概率分布,此表称之为随机变量的分布列。离散型随机变量的概率分布如图 1.1 所示。

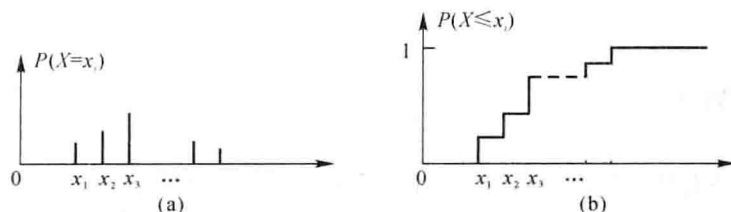


图 1.1 离散型随机变量的概率分布

2. 分布函数

用分布列描述随机变量的统计特性只适用于离散型随机变量。对于连续型随机变量,由于其可能取的值不能一一列举出来,所以不能用分布列来描述;再者,连续型随机变量通常取任一指定值的概率为零,因此转而去研究随机变量的取值落在一个区间内的概率: $P\{x_1 < X \leq x_2\}$ 。

设随机变量 X 为图 1.2 所示 x 轴上随机点的坐标,由图可得

$$P\{X \leq x_2\} = P\{X \leq x_1\} + P\{x_1 < X \leq x_2\} \quad (1.12)$$

故得

$$P\{x_1 < X \leq x_2\} = P\{X \leq x_2\} - P\{X \leq x_1\} \quad (1.13)$$

因此,只需知道 $P\{X \leq x_2\}$ 和 $P\{X \leq x_1\}$ 就可以了。现在给出随机变量分布函数的定义。

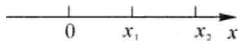


图 1.2 式(1.13)的图解说明

定义 1.4 随机变量 X 取值不超过 x 的概率 $P\{X \leq x\}$, 称为 X 的分布函数, 记为 $F_X(x)$, 即

$$F_X(x) = P\{X \leq x\} \quad (1.14)$$

式(1.14)表明,只要知道了分布函数 $F_X(x)$,就可获知 X 落在区间 $(-\infty, x]$ 内的概率。因此分布函数 $F_X(x)$ 能够完整地描述随机变量的统计特性。显然,分布函数对连续型和离散型随机变量都适用。

分布函数具有以下性质:

- 1) 单调不减性,即 $F_X(x)$ 是 x 的单调不减函数。因为 $x_2 > x_1$, 故有 $F_X(x_2) - F_X(x_1) = P\{x_1 < X \leq x_2\} \geq 0$, 即 $F_X(x_2) \geq F_X(x_1)$ 。
- 2) $0 \leq F_X(x) \leq 1$, $F_X(-\infty) = \lim_{x \rightarrow -\infty} F_X(x) = 0$, $F_X(\infty) = \lim_{x \rightarrow \infty} F_X(x) = 1$ 。
- 3) $F_X(x)$ 右连续,即 $F_X(x+0) = F_X(x)$ 。
- 4) 如果 $F_X(x)$ 在 $x = x_i$ 处不连续,则在这一点的阶跃等于 $P\{X = x_i\}$ 。

以上的讨论可以推广到两个随机变量(二维分布)或更多个随机变量(多维分布)的情况。对于两个随机变量 X 和 Y (它们可以是连续的,亦可以是离散的),它们的联合分布函数定义为

$$F_{XY}(x, y) = P\{X \leq x, Y \leq y\} \quad (1.15)$$

联合分布函数具有以下性质:

- 1) $F_{XY}(x, y)$ 是 x 或 y 的不减函数,即对于任意固定 y 值,当 $x_2 > x_1$ 时,有 $F_{XY}(x_2, y) \geq F_{XY}(x_1, y)$;对任意固定 x 值,当 $y_2 > y_1$ 时,有 $F_{XY}(x, y_2) \geq F_{XY}(x, y_1)$ 。
- 2) $0 \leq F_{XY}(x, y) \leq 1$, 且对于任意固定 y 值,有 $F_{XY}(-\infty, y) = 0$, 对于任意固定 x 值,有 $F_{XY}(x, -\infty) = 0$, 并有 $F_{XY}(-\infty, -\infty) = 0$, $F_{XY}(\infty, \infty) = 1$ 。
- 3) 二维随机变量 (X, Y) 关于 X 和关于 Y 的边缘分布函数分别为

$$F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < \infty\} = F_{XY}(x, \infty) \quad (1.16)$$

换言之, $F_{XY}(x, y)$ 中令 $y \rightarrow \infty$ 就能得到 $F_X(x)$; 同理可得 $F_{XY}(\infty, y) = F_Y(y)$ 。

$$4) P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \geq 0.$$

3. 概率密度函数

随机变量的统计特性还可利用分布函数的另一变换形式来描述, 这就是概率密度函数。

(1) 一维随机变量。若随机变量 X 的分布函数 $F_X(x)$, 存在非负的函数 $f_X(x)$, 使得对于任意实数 x , 有

$$F_X(x) = \int_{-\infty}^x f_X(u) du \quad (1.17)$$

则称 X 为连续型随机变量, 称 $f_X(x)$ 为 X 的概率密度函数, 式(1.17)也可写成

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (1.18)$$

可见, 概率密度函数是分布函数的导数。图 1.3 所示为连续随机变量 X 的概率密度函数 $f_X(x)$ 和分布函数 $F_X(x)$ 的例子。

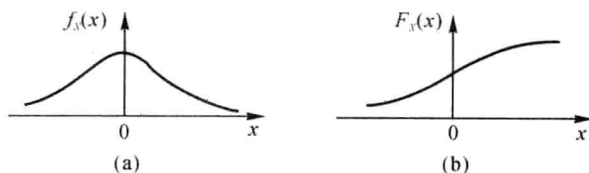


图 1.3 连续随机变量的概率分布

概率密度函数具有以下性质:

1) $f_X(x) \geq 0$, 由于 $F_X(x)$ 是单调不减函数, 所以其导数必是非负的。

$$2) \int_{-\infty}^{\infty} f_X(x) dx = 1.$$

$$3) \int_a^b f_X(x) dx = \int_{-\infty}^b f_X(x) dx - \int_{-\infty}^a f_X(x) dx = F_X(b) - F_X(a) = P\{a < X \leq b\}.$$

以上的讨论实际是在 X 为连续型随机变量的情况下进行的。利用 δ 函数, 则对离散型随机变量亦可用概率密度函数来描述。对于离散型随机变量 X , 它的分布函数可表示成

$$F_X(x) = \sum_i P_i U(x - x_i) \quad (1.19)$$

式中, $U(x)$ 是一个阶跃函数, 即 $U(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ 。离散型随机变量 X 的概率密度函数为

$$f_X(x) = \frac{dF_X(x)}{dx} = \sum_i P_i \delta(x - x_i) \quad (1.20)$$

通常, 还可能遇到混合型的随机变量, 其分布函数由阶跃的间断部分和处处连续部分组成。这类随机变量的例子如图 1.4 所示。

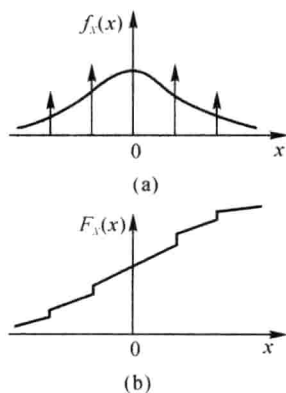


图 1.4 混合型随机变量的概率密度函数与分布函数

(2) 二维随机变量。两个随机变量 X 和 Y , 其联合概率密度函数记作 $f_{XY}(x, y)$ 。对于二维随机变量 (X, Y) 的分布函数 $F_{XY}(x, y)$, 如果存在非负的函数 $f_{XY}(x, y)$, 使得对于任意实数 x, y , 有

$$F_{XY}(x, y) = \int_{-\infty}^y \int_{-\infty}^x f_{XY}(u, v) du dv \quad (1.21)$$

则称 (X, Y) 是连续型的二维随机变量, 函数 $f_{XY}(x, y)$ 称作二维随机变量 (X, Y) 的概率密度函数。显然, 亦可定义为

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \quad (1.22)$$

联合概率密度函数具有以下性质:

$$1) f_{XY}(x, y) \geq 0;$$

$$2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = F_{XY}(\infty, \infty) = 1;$$

$$3) \begin{cases} F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^{\infty} \int_{-\infty}^x f_{XY}(x, y) dx dy \\ F_Y(y) = F_{XY}(\infty, y) = \int_{-\infty}^y \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy \end{cases};$$

$$4) \begin{cases} f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy \\ f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx \end{cases}.$$

$f_X(x)$ 和 $f_Y(y)$ 依次称作二维随机变量 (X, Y) 关于 X 和关于 Y 的边缘概率密度函数。

4. 条件分布

将随机事件条件概率的概念引入概率分布问题中, 可以得出条件分布函数和条件概率密度函数。

在条件 B 下随机变量 Y 的条件分布函数为

$$F_Y(y | B) = P\{Y \leq y | B\} = \frac{P\{Y \leq y, B\}}{P(B)} \quad (1.23)$$

如果事件 B 用随机变量 $X \leq x$ 表示, 则

$$F_Y(y | X \leq x) = \frac{P\{X \leq x, Y \leq y\}}{P\{X \leq x\}} = \frac{F_{XY}(x, y)}{F_X(x)} \quad (1.24)$$

对于事件 B 为 $X = x$ 的情况下, 对二维连续型随机变量而言, 就不能像以上所述那样, 直接由条件概率的公式引入“条件分布函数”概念, 此时要用极限方法来处理。

假设对于任意正数 ϵ , 有

$$P\{x - \epsilon < X \leq x + \epsilon\} \geq 0 \quad (1.25)$$

若极限

$$\lim_{\epsilon \rightarrow 0^+} P\{Y \leq y | x - \epsilon < X \leq x + \epsilon\} = \lim_{\epsilon \rightarrow 0^+} \frac{P\{Y \leq y, x - \epsilon < X \leq x + \epsilon\}}{P\{x - \epsilon < X \leq x + \epsilon\}} \quad (1.26)$$

存在, 则称此极限为“在条件 $X = x$ 下, Y 的条件分布函数”, 记为 $F_Y(y | X = x)$ 或 $F_Y(y | x)$ 。

可推导出:

$$\begin{aligned} F_Y(y | x) &= \lim_{\epsilon \rightarrow 0^+} \frac{P\{Y \leq y, x - \epsilon < X \leq x + \epsilon\}}{P\{x - \epsilon < X \leq x + \epsilon\}} = \lim_{\epsilon \rightarrow 0^+} \frac{F(y, x + \epsilon) - F(y, x - \epsilon)}{F_X(x + \epsilon) - F_X(x - \epsilon)} = \\ &= \frac{\lim_{\epsilon \rightarrow 0^+} \{ [F(y, x + \epsilon) - F(y, x - \epsilon)] / 2\epsilon \}}{\lim_{\epsilon \rightarrow 0^+} \{ [F_X(x + \epsilon) - F_X(x - \epsilon)] / 2\epsilon \}} = \frac{\partial F_{XY}(x, y) / \partial x}{dF_X(x) / dx} \end{aligned} \quad (1.27)$$

或

$$F_Y(y | x) = \frac{\int_{-\infty}^y f_{XY}(x, v) dv}{f_X(x)}, \quad f_X(x) \neq 0 \quad (1.28)$$

等号左、右对 y 求导, 得

$$f_Y(y | x) = \frac{f_{XY}(x, y)}{f_X(x)}, \quad f_X(x) \neq 0 \quad (1.29)$$

$f_Y(y | x)$ 称之为“在条件 $X = x$ 下, Y 的条件概率密度函数”。

条件概率分布的概念可推广到多维随机变量的情况, 用 $f_Y(y_1, y_2, \dots, y_n | x_1, x_2, \dots, x_m)$ 表示在条件 $X_1 = x_1, X_2 = x_2, \dots, X_m = x_m$ 下, $Y_1, Y_2, \dots, Y_n$ 的条件概率密度函数, 于是由上述概念可得

$$f_Y(y_1, y_2, \dots, y_n | x_1, x_2, \dots, x_m) = \frac{f_{XY}(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n)}{f_X(x_1, x_2, \dots, x_m)} \quad (1.30)$$

5. 统计独立

现在用概率密度函数给出随机变量统计独立的条件。由式(1.29), 可得

$$f_{XY}(x, y) = f_X(x) f_Y(y | x) \quad (1.31)$$

若

$$\left. \begin{aligned} f_X(x | y) &= f_X(x) \\ f_Y(y | x) &= f_Y(y) \end{aligned} \right\} \quad (1.32)$$

且对所有 x 和 y 取值都成立, 则称 X 和 Y 是相互统计独立的随机变量。于是可得 X 和 Y 统计独立的充要条件为

$$f_{XY}(x, y) = f_X(x) f_Y(y) \quad (1.33)$$

推广到多维随机变量的情况,若满足:

$$f_{X_1, X_2, \dots, X_N}(x_1, x_2, \dots, x_N) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_N}(x_N) \quad (1.34)$$

则随机变量 $X_1, X_2, \dots, X_N$ 相互统计独立。

1.2.3 几种常用的概率分布

1. 二项式分布

假设进行了 n 次独立试验,每次试验中某事件 A 可以发生或不发生。若 A 在每次试验中发生的概率是 p ,而 A 不发生的概率是 $q=1-p$ 。现在求事件 A 在这 n 次试验中发生 m 次的概率 $P_n(m)$ 。

如果事件 A 在 n 次独立试验的前 m 次出现,而其余的 $n-m$ 次不出现,依概率乘法定理可知,其概率为 $p^m q^{n-m}$ 。又由于只要求 A 在 n 次试验中出现 m 次,并未限定在前 m 次出现,实际上可以在 n 次中任意组合,这样共可以有 C_n^m 种组合方法。于是有

$$P_n(m) = C_n^m p^m q^{n-m} = \frac{n!}{m! (n-m)!} p^m q^{n-m} \quad (1.35)$$

显然,在 n 次独立试验中,事件 A 发生的次数是一个离散型随机变量,将其记为 X 。 X 可以随机地取从 0 到 n 的任一整数,对应每一整数都有一相应的概率,即

$$P\{X=m\} = P_n(m) = C_n^m p^m q^{n-m}, \quad m=0,1,\dots,n \quad (1.36)$$

因此,它的分布函数为

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \sum_{m \leq x} P_n(m), & 0 \leq x < n \\ 1, & x \geq n \end{cases} \quad (1.37)$$

由于 $P_n(m)$ 等于 $(px+q)^n$ 展开式中 x^m 项的系数,因此式(1.36)所表示的分布称作二项式分布。

图 1.5 给出二项式分布函数 $F_X(x)$ 的图形,可见它是一个阶梯形曲线,在 $0,1,\dots,n$ 处有阶跃,而且对固定的 n 及 p 而言,随着 m 的增加,阶跃高度加大,直到出现最大值。

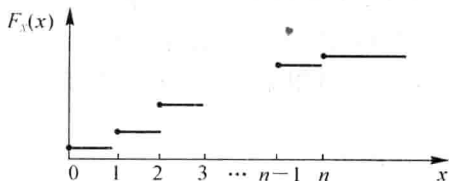


图 1.5 二项式分布函数

二项式分布的数学期望和方差分别为

$$E[X] = np, \quad D[X] = npq \quad (1.38)$$

2. 泊松(Poisson)分布

在二项式分布中,若试验次数 n 无限增大,同时 p 无限减小(即 $n \rightarrow \infty, p \rightarrow 0$),且 $np = \lambda$

不变,那么,在 n 次独立试验中事件 A 发生 m 次的概率可有一近似式,即

$$P\{X=m\}=P_n(m)=\frac{\lambda^m}{m!}e^{-\lambda} \quad (1.39)$$

满足式(1.39)的随机变量 X 为服从泊松分布的随机变量。泊松分布仅取决于参量 λ 。知道了 $P_n(m)=P\{X=m\}$ 就不难得出泊松分布函数的表达式为

$$F_X(x)=\begin{cases} 0, & x < 0 \\ \sum_{m \leq x} P_n(m), & 0 \leq x < n \\ 1, & x \geq n \end{cases} \quad (1.40)$$

图 1.6 所示为几种不同 λ 值下的分布情况。泊松分布的数学期望和方差分别为

$$E[X]=\lambda, \quad D[X]=\lambda \quad (1.41)$$

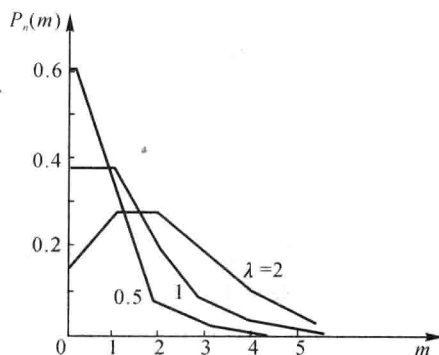


图 1.6 不同 λ 值下的泊松分布

以上所介绍的二项式分布及泊松分布都是离散型随机变量的分布函数。现在介绍几种较为重要的连续型随机变量分布函数。

3. 正态分布(高斯分布)

正态分布又称高斯分布,在概率论和统计数学中起着重要的作用,是科学技术领域中最常遇到的分布,也是无线电技术领域(包括噪声理论、信号检测理论、信息理论等)最重要的概率分布,大量的随机现象可以用正态分布描述。同时,正态分布具有许多良好的性质,许多分布可以用正态分布来近似,一些分布可以通过正态分布来导出。因此,无论在理论研究还是在实际应用中,正态分布都十分重要,而且根据正态分布各高阶矩的特点,出现了信号处理的一个重要分支——高阶累积量。

若随机变量 X 的概率密度函数为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (-\infty < x < \infty) \quad (1.42)$$

其中, $\sigma > 0$, m 与 σ 为常数,则称 X 服从正态分布,记为 $X \sim N(m, \sigma^2)$,相应的分布函数为

$$F_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^x e^{-\frac{(y-m)^2}{2\sigma^2}} dy \quad (-\infty < x < \infty) \quad (1.43)$$

特别当 $m=0, \sigma=1$ 时,称 X 服从标准正态分布,记为 $X \sim N(0,1)$,有时也称为归一化的正态分布。