21世纪高等学校规划教材

SUIJI XINHAO FENXI

随机信号分析

高岩 主编
吴冰 郭辉 张延良 副主编

中国电力出版社
http://jc.cepp.com.cn

21世纪高等学校规划教材

SUIJI XINHAO FENXI

随机信号分析

主　编　高　岩
副主编　吴　冰　郭　辉　张延良
编　写　钱　伟　李宝平
主　审　邱天爽

中国电力出版社
http://jc.cepp.com.cn

内 容 提 要

本书为21世纪高等学校规划教材。

全书共分为八章，主要内容包括概率论基础、随机信号、平稳随机信号及其谱分析、各态历经性与随机实验、随机信号通过线性系统、窄带随机信号、随机信号通过非线性系统、泊松过程与马尔可夫链。本书主要从工程应用的角度，讨论随机信号的基本理论和分析方法，力求联系工程实践，强调物理意义，内容全面，叙述清楚，例题与图示丰富，便于教学与自学。

本书可作为高等院校通信、电子信息类专业本科生或研究生的教材或教学参考书，也可供相关专业的师生和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/高岩主编．—北京：中国电力出版社，2009
21世纪高等学校规划教材
ISBN 978-7-5083-9296-7

Ⅰ．随…　Ⅱ．高…　Ⅲ．随机信号—信号分析—高等学校—教材　Ⅳ．TN911.6

中国版本图书馆CIP数据核字（2009）第140257号

中国电力出版社出版、发行
（北京三里河路6号　100044　http：//jc.cepp.com.cn）
航远印刷有限公司印刷
各地新华书店经售
*
2009年9月第一版　　2009年9月北京第一次印刷
787毫米×1092毫米　16开本　11.75印张　283千字
定价 **19.00** 元

前 言

随机信号分析是电子信息类专业的重要基础课程之一，是目标检测、估计、滤波等信号处理理论的基础，在通信、自动控制、雷达、图像处理、生物医学、地震信号处理、气象预报等领域有着广泛地应用。近年来，随着现代通信、信息处理和计算机科学技术的飞速发展，随机信号分析理论与应用将日益广泛和深入。

本书是作者在多年讲授随机信号分析、信号检测与估计课程的基础上，结合教学体会和相关科研工作实践编写而成，期望读者通过本书的学习掌握随机信号分析的基本理论和系统的分析方法。本书可以作为本科生与研究生的教材或教学参考书，建议教学学时数为32～56学时。

本教材突出的特点体现在以下几个方面。

（1）加强对基本概念的阐述，减少烦琐的公式推导，采用简明的讲解方式。书中对许多重要基础知识采用了定义与定理的描述形式，条理清晰，主次分明，并引入MATLAB仿真工具来帮助解释一些基本概念和方法。

（2）联系工程实践，强调物理意义。本课程的目的是以基础数学理论为工具，研究电子电气工程应用的实际问题，故本书在保证基础数学理论准确、严谨的前提下，注意解释各种数学概念与结论的物理含义，阐述数学问题所关联的工程背景。

（3）结合专业背景，突出实例的地位和重要性。书中例题大多数都来自通信工程、电子工程等学科领域经常遇到的随机信号实例，另外还特别注意对其分析方法和重要结论进行总结和归纳。

全书共分八章。第一章对概率论的基础知识进行复习与总结，同时补充随机变量的条件数学期望、特征函数等新知识点；第二章主要介绍随机信号的定义、统计特性与描述方式，讨论几种典型的随机信号实例，及其重要的高斯信号；第三章介绍平稳随机信号的定义和统计特性，重点讨论平稳信号的自相关函数和功率谱密度；第四章介绍随机信号的各态历经性与随机模拟的思想，讨论参数的实验测量方法和简单随机数的产生；第五章讨论随机信号通过线性时不变系统的分析方法，介绍噪声中的信号处理；第六章介绍希尔伯特变换与复随机信号的概念，重点讨论窄带随机信号的定义和统计特性，以及正弦信号叠加窄带高斯噪声后的统计特性；第七章介绍随机信号通过非线性系统的分析方法；第八章介绍马尔可夫链、独立增量过程和泊松过程。

本书由河南理工大学计算机科学与技术学院通信工程系的一线教师编写，高岩任主编，吴冰、郭辉、张延良任副主编。李宝平编写第一章、第四章，郭辉编写第二章、第三章，张延良编写第五章、第六章，钱伟编写第七章的第一～第三节、第八章，高岩编写第七章的第四节，吴冰编写第七章的第五节。本书由大连理工大学邱天爽教授主审，

在此表示衷心地感谢。本书的出版得到了中国电力出版社的大力支持，在此也一并表示感谢。

限于作者水平，书中难免有错误与疏漏，敬请广大读者批评指正。

作　者

2009年8月

目　　录

第一章 概率论基础

概率论是随机信号分析的理论基础。因此，本章将简单地复习与总结这些知识，更为详细的内容请大家参考有关书籍。同时也扩充一些本书后面需要用到的知识点，例如，随机变量的条件数学期望与特征函数，随机变量的MATLAB仿真等有关知识。

第一节 随机事件与概率

一、随机试验、样本空间

随机试验（Random Experiment）是概率论的基本概念，试验的结果事先不能确定地预言，但具有如下3个特性。

（1）可以在相同的条件下重复进行。

（2）每次试验的结果不止一个，但预先知道试验的所有可能的结果。

（3）每次试验前不能确定哪个结果会出现。

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（Sample Space）或基本事件空间，记为Ω。Ω中的元素e称为样本点（Sample Point）或基本事件（Elementary Event），Ω中的子集A称为事件（Event），样本空间Ω称为必然事件（Certain Event），空集$\varnothing$称为不可能事件（Impossible Event）。必然事件与不可能事件实际上是确定性现象，即它们不是随机事件，但是为了方便起见，我们把它们看成两个特殊的随机事件。

由于事件是集合，故集合的运算（并、交、差、上极限、下极限、极限等）都适用于事件。事件的关系与运算如下。

（1）包含关系与相等。“事件A发生必导致B发生”，记为$A\subset B$或$B\supset A$；$A=B\Leftrightarrow A\subset B$且$B\subset A$。

（2）和事件（并集）。“事件A与B至少有一个发生”，记为$A\cup B$。

（3）积事件（交集）。“事件A与B同时发生”，记为$A\cap B$或AB。

（4）差事件、对立事件（余事件）。“事件A发生而B不发生”，记为$A-B$称为A与B的差事件；$\Omega-B=\overline{B}$称为B的对立事件；易知$A-B=A\overline{B}$。

（5）互不相容性：$AB=\varnothing$，A、B互为对立事件$\Leftrightarrow A\cup B=\Omega$且$AB=\varnothing$。

（6）事件的运算法则。

1）交换律。$A\cup B=B\cup A$，$AB=BA$；

2）结合律。$A\cup(B\cup C)=(A\cup B)\cup C$，$(AB)C=A(BC)$；

3）分配律。$(A\cup B)C=AC\cup BC$，$(AB)\cup C=(A\cup C)(B\cup C)$；

4）摩根律。$\overline{A\cup B}=\overline{A}\,\overline{B}$，$\overline{AB}=\overline{A}\cup\overline{B}$，可推广至$\overline{\bigcup_k A_k}=\bigcap_k\overline{A_k}$，$\overline{\bigcap_k A_k}=\bigcup_k\overline{A_k}$。

定义1-1 设Ω是一个集合，事件域F是Ω的某些子集组成的集合族。如果

（1）$\Omega\in F$。

（2）若$A\in F$，则$\overline{A}=\Omega\setminus A\in F$，其中“\”表示去除。

(3) 若 $A_n \in F$，$n=1$，2，…，则 $\bigcup_{n=1}^{\infty} A_n \in F$。

则称 F 为 σ-代数（Borel 域），(Ω，F) 称为可测空间，F 中的元素称为事件。$\{A, \overline{A}, \Omega, \varnothing\}$ 为最小的 Borel 域。

由定义可知：

(1) $\varnothing \in F$。

(2) 若 A，$B \in F$，则 $A \setminus B \in F$。

(3) 若 $A_i \in F$，$i=1$，2，…，则 $\bigcup_{i=1}^{n} A_i$，$\bigcap_{i=1}^{n} A_i$，$\bigcap_{i=1}^{\infty} A_i \in F$。

二、随机事件的概率

随机事件如何描述？它发生的可能性的测度可以用概率、隶属度、证据等描述。进行随机试验时，有些是必然会发生的，称为必然事件，$P(\xi)=1$；有些是必然不会发生的，称为不可能事件，$P(\xi)=0$；但一般随机事件发生的可能性介于 0 与 1 之间，将事件 A 发生的可能性称为事件 A 的概率（Probability）。

在实际应用中，事件 A 出现的概率可由统计概率来表示。

统计概率（Statistics Probability）：事件 A 在 n 次重复试验中出现 n_A 次，则比值 $\frac{n_A}{n}$ 称为事件 A 在 n 次重复试验中出现的频率，记为 $f_n(A)$，当 $n \to \infty$ 时，频率 $f_n(A)=\frac{n_A}{n} \to P(A)$；当 n 很大时，$P(A)=P \approx f_n(A)$ 称为事件 A 的统计概率。

概率具有的一些基本特性，称为概率公理。

概率公理：任何事件 A 的概率满足以下条件。

(1) 非负性。任取事件 $A \in F$，$P(A) \geqslant 0$。

(2) 规范性。$P(\Omega)=1$。

(3) 可列可加性。对两两互不相容事件 A_1，A_2，…（当 $i \neq j$ 时，$A_i A_j=\varnothing$），有 $P\left(\bigcup_{j=1}^{\infty} A_j\right)=\sum_{j=1}^{\infty} P(A_j)$。

概率与概率公理是概率论的基石，它们源于科学实践，其正确性在于它能合理地表述与解释客观世界，并能有效地解决实际问题。

三、条件概率与乘法公式

1. 条件概率

条件概率（Conditional Probability）：设 A，B 是 Ω 中的两个事件，则 $P(B|A)=\frac{P(AB)}{P(A)}$ 称为事件 A 发生的条件下事件 B 发生的条件概率。

$P(AB)$ 是在样本空间 Ω 内，事件 AB 的概率，而 $P(B|A)$ 是在试验 B 增加了新条件 A 发生后的缩减的样本空间 Ω_A 中计算事件 B 的概率。虽然 A、B 都发生，但两者是不同的。一般来说，当 A、B 同时发生时，常用 $P(AB)$，而在有包含关系或明确的主从关系时，用 $P(B \mid A)$。如袋中有 9 个白球 1 个红球，做不放回抽样，每次任取一球，取 2 次，求：①第二次取到白球的概率；②第一次取到的是白球的条件下，第二次取到白球的概率。问题①求的就是一个积事件概率，而问题②求的就是条件概率。

2. 乘法公式

乘法公式：设 A，$B\in F$，则 $P(AB)=P(A)P(B|A)=P(B)P(A|B)$ 称为事件 A，B 的概率乘法公式。

四、全概率公式与贝叶斯（Bayes）公式

(1) 全概率公式。设 A_1，A_2，…，A_n 是 Ω 的一个划分，且 $P(A_i)>0$ ($i=1,2,\cdots,n$)，则对任何事件 $B\in F$，有 $P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)$，称为全概率公式。

(2) 贝叶斯（Bayes）公式。设 A_1，A_2，…，A_n 是 Ω 的一个划分，且 $P(A_i)>0$($i=1,2,\cdots,n$)，则对任何事件 $B\in F$，有 $P(A_j|B)=\dfrac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}$，其中 $j=1$，…，n，该公式称为贝叶斯公式或逆概率公式。

当所求的事件概率为许多因素引发的某种结果，而该结果又不能简单地看作是诸多事件之和时，可考虑用全概率公式。在对样本空间进行划分时，一定要注意它必须满足的两个条件。贝叶斯公式用于实验结果已知，追查是何种原因（情况、条件下）引发的概率。通常将各个原因事件概率 $P(A_i)$ 称为先验概率（Priori Probability），它是观察事件 B 出现与否的实验之前已存在的；将条件概率 $P(B|A_j)$ 称为转移概率（Transition Probability），它是第 j 个原因事件转移成（或引起）事件 B 的概率；而将观察到 B 出现之后，条件概率 $P(A_j|B)$ 称为后验概率（Posteriori Probability）。考虑因果推测问题：A_j 是 m 个原因事件，B 是某种结果事件，各原因的概率是先验概率，由原因引起结果的概率是转移概率，而知道结果后推测原因的概率就是后验概率。贝叶斯公式正是基于结果 B 推测某种原因 A_j 的可能性的方法。

【例 1-1】 某电子设备制造厂所用的晶体管是由三家元件制造厂提供的。根据以往的记录有以下数据，见表 1-1。

设这三家工厂的产品在仓库中均匀混合的，且无区别的标志。

(1) 在仓库中随机取一只晶体管，求它是次品的概率。

(2) 在仓库中随机取一只晶体管，若已知取到的是次品，为分析此次品出自何厂，求此次品由三家工厂生产的概率分别是多少。

表 1-1 例 1-1 数据

元件制造厂	次品率	提供晶体管的份额
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05

解 设 A 表示“取到的是一只次品”，B_i($i=1$，2，3) 表示“所取到的产品是由第 i 家工厂提供的”。易知，B_1、B_2、B_3 是样本空间 Ω 的一个划分，且有

$$P(B_1)=0.15,\quad P(B_2)=0.80,\quad P(B_3)=0.05$$
$$P(A|B_1)=0.02,\quad P(A|B_2)=0.01,\quad P(A|B_3)=0.03$$

(1) 由全概率公式：$P(A)=\sum_{i=1}^{3}P(B_i)P(A|B_i)=0.0125$

(2) 由贝叶斯公式：

$$P(B_1|A)=\frac{P(A|B_1)P(B_1)}{P(A)}=0.24,\quad P(B_2|A)=0.64,\quad P(B_3|A)=0.12$$

以上结果表明，这只次品来自第二家工厂的可能性最大。

第二节 随机变量及其分布

概率论中的样本空间与事件等都是一般意义上的集合，为了更加有效地进行讨论与分析，我们希望将原来的集合描述形式转换为实数描述形式，因为在实数域上存在着丰富的数学分析方法与结果。为此，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 1-2 设（Ω，F，P）是概率空间，$X=X(\xi)$ 是定义在 Ω 上的实函数，如果对于任意实数 x，有 $\{\xi: X(\xi)\leqslant x\}\in F$，则称 $X(\xi)$ 是 F 上的随机变量（Random Variable），简记为 X。称

$$F_X(x)=P\{X(\xi)\leqslant x\},(-\infty<x<+\infty)$$

为 X 的概率分布函数或概率累积分布函数（Probability Distribution），简称为分布函数。

分布函数完整地描述了随机变量取值的统计规律性，具有以下性质。

(1) $0\leqslant F(x)\leqslant 1,(-\infty<x<+\infty)$。

(2) 如果 $x_1<x_2$，则 $F(x_1)\leqslant F(x_2)$。

(3) $F(x)$ 为右连续，即 $F(x+0)=F(x)$。

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

(5) $P\{x_1<X\leqslant x_2\}=P\{X\leqslant x_2\}-P\{X\leqslant x_1\}=F(x_2)-F(x_1)$。

一、离散型随机变量及其概率分布

如果随机变量 X 只能取有限个或可列个可能值，则称 X 为离散型随机变量。如果 X 的一切可能值为 x_1，x_2，…，并且 X 取 x_k 的概率为 p_k，则称 $p_k=P\{X=x_k\}(k=1,2,3,\cdots)$ 为离散型随机变量 X 的概率函数（概率分布或分布律）。列成表格形式，也称为分布列（见表 1-2）。

表 1-2 离散型随机变量分布列

X	x_1	x_2	x_3	…
P	p_1	p_2	p_3	…

注 $p_i\geqslant 0$，$\sum\limits_i p_i=1$。

常见的离散型随机变量的分布有

(1) 0—1 分布，记为 $X\sim(0,1)$，概率函数

$$P\{X=k\}=p^k(1-p)^{1-k}(\text{其中 }k=0,1;0<p<1)$$

(2) 二项分布，记为 $X\sim B(n,p)$，概率函数

$$P\{X=k\}=C_n^k p^k(1-p)^{n-k}(\text{其中 }k=0,1,\cdots,n;0<p<1)$$

(3) 泊松分布，记为 $X\sim P(\lambda)$，概率函数

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}(\text{其中 }k=0,1,\cdots;\lambda>0)$$

定理 1-1 泊松定理 设 $\lambda>0$ 是一个常数，n 是任意正整数，设 $np_n=\lambda$，则对于任一固定的非负整数 k，有 $\lim\limits_{n\to\infty}C_n^k p_n^k(1-p_n)^{n-k}=\dfrac{\lambda^k e^{-\lambda}}{k!}$。

当 n 很大且 p 很小时，二项分布可以用泊松分布近似代替，即

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}(\text{其中 } \lambda = np)$$

（4）超几何分布，记为 $X \sim H(n, M, N)$，概率函数 $P\{X=k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k=0,1,\cdots,\min(n,M)$，其中 n、N、M 为正整数，且 $M \leqslant N$，$n \leqslant N$。

当 N 很大且 $p=\frac{n}{N}$ 较小时，有 $\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \approx C_n^k p^k (1-p)^{n-k}$。

（5）几何分布，记为 $X \sim G(p)$，概率函数

$$P\{X=k\} = p(1-p)^{k-1}(\text{其中 } k=0,1,\cdots;0<p<1)$$

二、连续型随机变量及其概率分布

定义 1-3 如果对于随机变量 X 的分布函数 $F(x)$，存在非负函数 $f(x)$，使对于任一实数 x，有 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$，则称 X 为连续型随机变量。函数 $f(x)$ 称为 X 的概率密度函数（Probability Density Function）。

概率密度函数具有如下性质。

（1）非负性与归一性：$f(x) \geqslant 0, \int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$；

（2）区间 A 上的概率计算公式：$P\{X \in A\} = \int_A f(x)\mathrm{d}x$。

$f(x)$ 的定义涉及求导运算，当 $F(x)$ 不连续时，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

$$u(x) = \begin{cases} 1, & x \geqslant 0 \\ 0, & x < 0 \end{cases} \tag{1-1}$$

$$\delta(x) = \frac{\mathrm{d}}{\mathrm{d}x}u(x) \tag{1-2}$$

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 $P(X=x_i)=P_i$ 的离散型随机变量，其分布函数为

$$F(x) = \sum_i P_i u(x-x_i) \quad (i \text{ 为整数}) \tag{1-3}$$

密度函数为

$$f(x) = \sum_i P_i \delta(x-x_i) \quad (i \text{ 为整数}) \tag{1-4}$$

式中，取值位置对应 $u()$ 与 $\delta()$ 自变量的偏移量，取值概率对应前面的幅值。

常见的连续型随机变量的分布如下。

（1）均匀分布（Uniform Distribution），记为 $X \sim U(a, b)$，概率密度为

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

相应的分布函数为 $F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leqslant x < b \\ 1, & x \geqslant b \end{cases}$。

（2）指数分布，记为 $X \sim E(\lambda)$，概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda}, & x \geqslant 0 \\ 0, & \text{其他}\end{cases}$$

相应的分布函数为 $F(x)=\begin{cases}1-e^{-\lambda x}, & x \geqslant 0 \\ 0, & x<0\end{cases}$。

（3）正态分布（高斯分布）（Normal Distribution），记为 $X \sim N(\mu,\sigma^2)$，概率密度为

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

相应的分布函数为 $F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$

当 $\mu=0$，$\sigma=1$ 时，即 $X \sim N(0,1)$ 时，称 X 服从标准正态分布。这时分别用 $\phi(x)$ 和 $\Phi(x)$ 表示 X 的密度函数和分布函数，即 $\phi(x)=\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$，$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$。具有性质：$\Phi(-x)=1-\Phi(x)$。

一般正态分布 $X \sim N(\mu,\sigma^2)$ 的分布函数 $F(x)$ 与标准正态分布的分布函数 $\Phi(x)$ 有关系：$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$。

第三节　多维随机变量及其分布

一、多维随机变量

定义 1-4　设 X_1，…，X_n 是定义在同一可测空间（Ω，F）上的随机变量，则称这 n 个随机变量的整体（X_1，…，X_n）为（Ω，F）上的 n 维随机变量。分别称 n 元函数

$$F_{X1X2\cdots Xn}(x_1,x_2,\cdots,x_n)=P\{X_1 \leqslant x_1,\cdots,X_n \leqslant x_n\} \tag{1-5}$$

$$f_{X1X2\cdots Xn}(x_1,x_2,\cdots,x_n)=\frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{X1X2\cdots Xn}(x_1,x_2,\cdots,x_n) \tag{1-6}$$

为 n 维随机变量（X_1，…，X_n）的联合分布函数和联合密度函数。

联合概率特性包含了分量随机变量各自的（边缘）概率特性与相互间“交叉”的概率特性，它们的基本性质与前面的相仿。

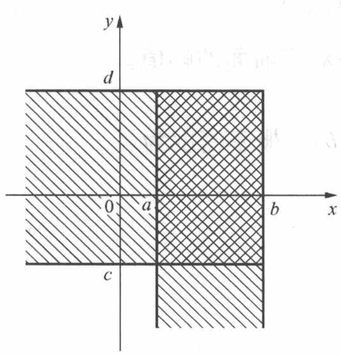

图 1-1　$P\{a<X \leqslant b;c<Y \leqslant d\}$ 的几何表示

以二维为例，分布函数 $F_{XY}(x,y)$ 反映了 $(-\infty,+\infty)$ 到 (x,y) 处的联合累积概率，它有下面的性质：

（1）$F_{XY}(x,-\infty)=0,F_{XY}(-\infty,y)=0,F_{XY}(+\infty,+\infty)=1$。

（2）$F_{XY}(x,y)$ 是 x 或 y 的单调非减函数。

（3）$P\{a<X \leqslant b;c<Y \leqslant d\}=F_{XY}(b,d)-F_{XY}(a,d)-F_{XY}(b,c)+F_{XY}(a,c)$，如图 1-1 所示。

（4）边缘概率分布为

$$F_X(x)=F_{XY}(x,+\infty),\ F_Y(y)=F_{XY}(+\infty,y)$$

二维概率密度函数反映了在（x，y）处的联合概率“强度”，它具有下面基本性质：

（1）非负与归一化：

$$f_{XY}(x,y)\geqslant 0,\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f_{XY}(x,y)\mathrm{d}x\mathrm{d}y=1$$

（2）区域概率：

$$P\{(x,y)\in D\}=\iint_D f_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{1-7}$$

（3）边缘概率密度为

$$f_X(x)=\frac{\mathrm{d}}{\mathrm{d}x}F_X(x)=\int_{-\infty}^{+\infty}f_{XY}(x,y)\mathrm{d}y,f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=\int_{-\infty}^{+\infty}f_{XY}(x,y)\mathrm{d}x \tag{1-8}$$

离散型随机变量的概率特性常用联合分布律来描述

$$P\{X=x_i,Y=y_j\}=p_{ij},\ \sum_i\sum_j P_{ij}=1$$

从中可以清楚地看到（x，y）取各离散点的概率。联合密度函数完全由如下形式的多维冲激函数组成

$$f_{XY}(x,y)=\sum_i\sum_j P_{ij}\delta(x-x_i,y-y_j) \tag{1-9}$$

其中，$\delta(x,y)=\delta(x)\delta(y)$。其联合分布函数由如下形式的多维阶跃函数组成

$$F_{XY}(x,y)=\sum_i\sum_j P_{ij}u(x-x_i,y-y_j) \tag{1-10}$$

其中，$u(x,y)=u(x)u(y)$。式（1-9）与式（1-10）中，随机变量取值位置对应 $u()$ 与 $\delta()$ 自变量的偏移量，**取值概率**对应前面的幅值。

【例 1-2】 已知随机变量（X，Y）仅在区域 $D=\{0\leqslant x\leqslant 2,|y|\leqslant 2-x\}$ 上取值，如图 1-2 所示，并且为均匀分布。求此二维均匀分布的 $f(x,y)$，$f_X(x)$ 与 $f_Y(y)$。

解 由于区域 D 如图 1-2 所示，可得面积 $S=4$，因此

$$f(x,y)=\begin{cases}1/S, & (x,y)\in D\\ 0, & \text{其他}\end{cases}=\begin{cases}1/4, & (x,y)\in D\\ 0, & \text{其他}\end{cases}$$

利用积分可求得边缘密度函数为

$$f_X(x)=\begin{cases}\displaystyle\int_{x-2}^{2-x}\frac{1}{4}\mathrm{d}y=1-x/2, & x\in[0,2]\\ 0, & \text{其他}\end{cases}$$

$$f_Y(y)=\begin{cases}\displaystyle\int_0^{2-y}\frac{1}{4}\mathrm{d}x,y\in[0,2]\\ \displaystyle\int_0^{y+2}\frac{1}{4}\mathrm{d}x,y\in[-2,0]\\ 0, \qquad \text{其他}\end{cases}=\begin{cases}\dfrac{2-|y|}{4}, & |y|\leqslant 2\\ 0, & \text{其他}\end{cases}。$$

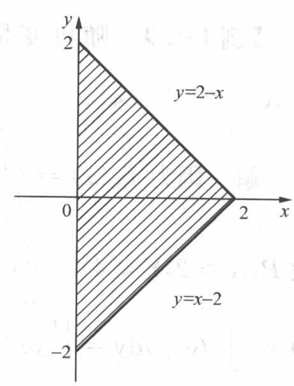

图 1-2 区域 D

二、条件随机变量

随机变量与多维随机变量的条件事件形如

$$X\in A|Y\in B$$
$$X\leqslant x|Y\in B$$
$$X_1\leqslant x_1,\cdots,X_n\leqslant x_n|Y_1\in B_1,\cdots,Y_n\in B_n$$
$$X_1\leqslant x_1,\cdots,X_n\leqslant x_n|Y_1=y_1,\cdots,Y_n=y_n$$

式中，A，B 与 B_i 都是实数集。上面最后一种是“点事件”作为条件的情形。由于随机变量取值为单点的事件概率常为 0，因此这类条件事件的概率采用极限形式来定义。以二维为例，条件概率分布与条件密度函数定义为

$$F_{X|Y}(x|y)=\lim_{\Delta y\to 0}\frac{P\{X\leqslant x,y<Y\leqslant y+\Delta y\}}{P\{y<Y\leqslant y+\Delta y\}}=\int_{-\infty}^{x}\frac{f_{XY}(u,y)}{f_Y(y)}\mathrm{d}u \tag{1-11}$$

$$f_{X|Y}(x|y)=\frac{\partial}{\partial x}F_{X|Y}(x|y)=\frac{f_{XY}(x,y)}{f_Y(y)} \tag{1-12}$$

该式在形式上与条件概率的基础定义相似。

离散型随机变量可采用条件分布率

$$P(X=x_i|Y=y_j)=\frac{p_{ij}}{P_j},P_j=\sum_i P_{ij} \tag{1-13}$$

条件概率分布和密度函数具有与普通分布和密度函数相似的性质，并且还满足

全概率公式：

$$f(x)=\int_{-\infty}^{+\infty}f(x|y)f(y)\mathrm{d}y \tag{1-14}$$

贝叶斯公式：

$$f(x|y)=\frac{f(y|x)f(x)}{\int_{-\infty}^{+\infty}f(y|x)f(x)\mathrm{d}x} \tag{1-15}$$

链式公式：

$$f(x_1x_2\cdots x_n)=f(x_1)f(x_2|x_1)f(x_3|x_2x_1)\cdots f(x_n|x_1x_2\cdots x_{n-1}) \tag{1-16}$$

对比第一节的有关公式，可以看到它们在概念与形式上的相似性。

【例 1-3】 随机变量（X，Y）的密度函数为 $f(x,y)=\begin{cases}\mathrm{e}^{-y}, & x>0,y>x\\ 0, & \text{其他}\end{cases}$，求 $P\{X>2|Y<4\}$。

解 因为 $f_X(x)=\begin{cases}\int_x^{+\infty}\mathrm{e}^{-y}\mathrm{d}y=\mathrm{e}^{-x}, & x>0\\ 0, & x\leqslant 0\end{cases}$，$f_Y(y)=\begin{cases}\int_0^y\mathrm{e}^{-y}\mathrm{d}x=y\mathrm{e}^{-y}, & y>0\\ 0, & y\leqslant 0\end{cases}$，

故 $P\{X>2,Y<4\}=\iint\limits_G f(x,y)\mathrm{d}x\mathrm{d}y=\int_2^4\int_2^y\mathrm{e}^{-y}\mathrm{d}x\mathrm{d}y=\int_2^4(y-2)\mathrm{e}^{-y}\mathrm{d}y=\mathrm{e}^{-2}-3\mathrm{e}^{-4}$，又 $P(Y<4)=\int\limits_{Y<4}f_Y(y)\mathrm{d}y=\int_0^4 y\mathrm{e}^{-y}\mathrm{d}y=1-5\mathrm{e}^{-4}$，所以 $P\{X>2|Y<4\}=(\mathrm{e}^{-2}-3\mathrm{e}^{-4})/(1-5\mathrm{e}^{-4})$。

【例 1-4】 二维均匀分布为 $f(x,y)=\begin{cases}1/4, & 0\leqslant x\leqslant 2,|y|\leqslant 2-x\\ 0, & \text{其他}\end{cases}$，求 $P\{X\leqslant 1|Y=0\}$ 与 $P\{X\leqslant 1\}$。

解 根据例 1-2 的 $f_Y(y)$ 结果，由定义有

$$f(x|y)=\frac{f(x,y)}{f(y)}=\begin{cases}\dfrac{1}{2-|y|}, & (X,Y)\in D\\ 0, & \text{其他}\end{cases}$$

于是

$$f(x|y=0)=\begin{cases}1/2, & x\in[0,2)\\ 0, & \text{其他}\end{cases}$$

即条件事件（$X|Y=0$）服从均匀分布$U(0, 2)$。实际上，任意给定$y\in[-2, 2]$，条件事件（$X|Y=y$）服从均匀分布$U(0, 2-|y|)$。于是

$$P\{X\leqslant 1|Y=0\}=\int_{-\infty}^{1}f(x|y=0)\mathrm{d}x=\int_{0}^{1}\frac{1}{2}\mathrm{d}x=\frac{1}{2}$$

而
$$P\{X\leqslant 1\}=\int_{-\infty}^{1}f(x)\mathrm{d}x=\int_{0}^{1}\left(1-\frac{x}{2}\right)\mathrm{d}x=\frac{3}{4}$$

由此可见，条件$Y=0$对$\{X\leqslant 1\}$的概率有影响，X与Y是不独立的。

三、独立性

定义 1-5 设X_1，…，X_n是n个随机变量，若任取n个实数x_1，…，x_n，有

$$F_{X1X2\cdots Xn}(x_1x_2\cdots x_n)=F_{X1}(x_1)F_{X2}(x_2)\cdots F_{Xn}(x_n)$$

则称它们相互独立。

独立性的核心在于联合事件的概率等于各自概率的积。采用密度函数，可以将独立性的定义表述如下

$$f_{X1X2\cdots Xn}(x_1x_2\cdots x_n)=f_{X1}(x_1)f_{X2}(x_2)\cdots f_{Xn}(x_n)$$

若随机变量是离散型的，也可以用分布律来表述

$$P(X_1=x_{1k1},X_2=x_{2k2},\cdots,X_n=x_{nkn})=P(X_1=x_{1k1})P(X_2=x_{2k2})\cdots P(X_n=x_{nkn})$$

式中，x_{1k1}，x_{2k2}，$\cdots x_{nkn}$分别是X_1，X_2，$\cdots X_n$的可能取值。

随机变量的独立性是原事件独立性概念的引申。还可以得出下面几点：

（1）若两组随机变量（X_1，…，X_n）与（Y_1，…，Y_m）满足

$$F(X_1,\cdots,X_n,Y_1,\cdots,Y_m)=F_{X1,\cdots,Xn}(x_1,\cdots,x_n)F_{Y1,\cdots,Ym}(y_1,\cdots,y_m)$$

则称两组变量独立，但它们各自内部不必彼此独立。

（2）在独立随机变量组之间，条件不起作用，即

$$F_{X1X2\cdots Xn|Y_1Y_2\cdots Y_m}(x_1,x_2,\cdots,x_n|y_1,y_2,\cdots,y_m)=F_{X1X2\cdots Xn}(x_1,x_2,\cdots,x_n)$$

【例 1-5】 设（X，Y）的联合密度函数为$f(x,y)=\begin{cases}Cxy,0<x<1, & 0<y<1\\ 0, & 其他\end{cases}$。试求：

（1）常数C；（2）$f_X(x)$，$f_Y(y)$；（3）X与Y是否相互独立。

解 （1）因为$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\int_0^1 Cxy\mathrm{d}x\mathrm{d}y=\frac{C}{4}$，因此$C=4$。

（2）因为$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$，当$0<y<1,0<x<1$时，$f_X(x)=\int_0^1 4xy\mathrm{d}y=2x$；当$x$，$y$为其他情况时，$f_X(x)=0$，所以$f_X(x)=\begin{cases}2x, & 0<x<1\\ 0, & 其他\end{cases}$；同理$f_Y(y)=\begin{cases}2y, & 0<y<1\\ 0, & 其他\end{cases}$。

（3）$f_X(x)f_Y(y)=\begin{cases}4xy, & 0<x<1,0<y<1\\ 0, & 其他\end{cases}$，

则有$f(x,y)=f_X(x)f_Y(y)$，因此，X与Y相互独立。

四、随机变量函数的概率分布

一个或多个随机变量的函数为

$$Y=g(X)\text{ 或 }Z=g(X_1,X_2,\cdots,X_n)$$

它们构成从原样本空间到实数域的复合映射，可见，Y或Z是新的随机变量。自变量的多样性与不确定性导致因变量的多样性与不确定性，自变量的概率特性决定因变量的概率

特性。

1. 一元函数

一元函数形如：$Y=g(X)$。确定其分布函数的基本方法是从定义出发：

$$F_Y(y)=P[g(X)\leqslant y]=\int_{g(x)\leqslant y} f_X(x)\mathrm{d}x$$

定理 1-2　设 $Y=g(X)$，若 $g(X)$ 处处可导且恒有 $g'(X)>0$ 或 $g'(X)<0$，则

$$f_Y(y)=\begin{cases} f_X[h(y)]\,|h'(y)|, & a<y<b \\ 0, & \text{其他} \end{cases} \tag{1-17}$$

式中　$a=\min\{g(-\infty),g(+\infty)\}$；$b=\max\{g(-\infty),g(+\infty)\}$；$h(y)$ 是 $g(x)$ 的反函数。

【例 1-6】　求 $Y=aX+b$ 的密度函数。

解　由于函数符合定理的条件，且反函数形式为 $x=(y-b)/a$，导函数为 $1/a$，于是

$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)$$

【例 1-7】　设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，求随机变量函数 $Y=\mathrm{e}^X$ 的概率密度。

解　由 $f_X(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$-\infty<x<+\infty$ 可知 $y=\mathrm{e}^x>0$，所以 Y 的取值区间为 $(0,+\infty)$。当 $y\leqslant 0$ 时，$f_Y(y)=0$；当 $y>0$ 时，有反函数 $x=\ln y$，从而 $f_Y(y)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}\dfrac{1}{y}=\dfrac{1}{\sqrt{2\pi}\sigma y}\mathrm{e}^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}$，由此得随机变量 Y 的概率密度为

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

2. 二元函数

二元函数表示为 $Z=\varphi(X,Y)$，(X,Y) 的联合概率密度函数为 $f(x,y)$，则确定 Z 的分布函数从定义出发 $F_Z(z)=P[\varphi(X,Y)\leqslant z]=\displaystyle\iint_{\varphi(x,y)\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y$。

(1) $Z=X+Y$ 的分布。若 (X,Y) 为离散型随机变量，联合分布律为 p_{ij}，则 Z 的概率函数为

$$P_Z(z_k)=\sum_i p(x_i,z_k-x_i)\text{ 或 }P_Z(z_k)=\sum_j p(z_k-y_j,y_j)$$

若 (X,Y) 为连续型随机变量，概率密度函数为 $f_{XY}(x,y)$，则 Z 的概率函数为

$$F_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x=\int_{-\infty}^{+\infty} f(z-y,y)\mathrm{d}y$$

(2) $Z=\dfrac{X}{Y}$ 的分布。若 (X,Y) 为连续型随机变量，概率密度函数为 $f(x,y)$，则 Z 的概率函数为

$$F_Z(z)=\int_{-\infty}^{+\infty}|y|f(yz,y)\mathrm{d}y$$

【例 1-8】　设随机变量 X 和 Y 相互独立，有 $f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}$，$f_Y(y)=\begin{cases}2y, & 0\leqslant y\leqslant 1\\ 0, & \text{其他}\end{cases}$，