随机信号分析

概率论概要

主讲: 赵国亮

内蒙古大学电子信息工程学院

September 29, 2020

目录

目录

参考书籍

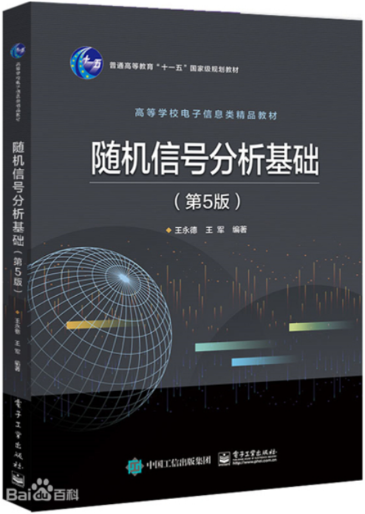
1) Umberto Spagnolini, 《工程中的统计信号处理-Statistical Signal Processing in Engineering》, 2017,

2) Statistical Signal Processing.

3) Chonavel, T. Statistical Signal Processing: Modelling and Estimation, 2002,


普通高等教育"十一五"国家级规划教材
随机信号分析
常建平 李海林 编著
科学出版社
www.sciencep.com
普通高等教育"十一五"国家级规划教材
随机信号分析
常建平 李海林 编著
科学出版社
www.sciencep.com


新坐标大学本科电子信息类专业系列教材
随机信号分析与处理
(第2版)
罗鹏飞 张文明 编著
清华大学出版社


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
电子工业出版社
Baidu百科

左边需要 gitee 的账号，登录后可以；右边 github，直接可以下载。

第 1 章概率论基础 1 教案-概率论部分

随机事件及其概率

自然界与人类社会的众多现象大致可分为两类，分别称为确定性现象与随机现象。所谓确定性现象，即在一定条件下必然会出现某一结果 (或发生某一事件) 的现象。

例 .1

标准大气压下，纯净水在一个大气压下加热至 100 摄氏度时，必然沸腾；物体以 10 米 / 秒的速度做匀速直线运动 1 分钟，其走过的路程必为 600 米。这类确定性现象由确定的规律所控制，从数量的角度来研究，从而产生了量与量之间确定的函数关系。

随机现象

所谓随机现象，即在一定条件下可能出现不同结果 (或发生不同事件)，且不能准确预言究竟出现哪一种结果的现象。

例 .2

相同条件下掷一枚硬币，可能正面向上，也可能反面向上，且在未掷之前无法准确预言究竟哪一面向上；二元数字通信系统发送的信号可能是 1，也可能是 0, 接收机在接收之前无法准确预言接收结果是信号 1 还是信号 0.

这一类现象广泛存在于自然界与社会活动中，而概率论正是探索研究这类随机现象的客观规律的一门学科。

本节首先介绍随机事件及其概率，并在此基础上分析随机变量的分布和数字特征。

观察、研究随机现象的手段与过程称为试验。当试验满足下述条件时，称之为随机试验，简称试验，记为 E. 随机试验具有如下特征：

(1) 试验可在相同条件下重复进行 (可重复性)；
(2) 试验可能出现的结果不止一个，并明确知道所有可能的结果；
(3) 每次试验总是恰好出现这些可能结果中的一个，但是在一次试验之前不能准确预言哪一种结果会出现 (结果出现的随机性)。

如：掷一颗骰子并观察出现的点数，从一批产品中任意抽取若干件产品，观察其中的次品数等都是随机试验。概率论所要研究的是随机试验中出现的各种情况。

定义 .1 基本事件

某一随机试验中可能出现的每一结果称为该试验的一个基本事件(样本点)，记为 e. 所有基本事件构成的集合称为该试验的样本空间，记为 Ω, 由样本空间 Ω 中的若干基本事件构成的子集合称为该试验中的随机事件，简称为事件，记为 A, B, C。当属于事件 A 的某一基本事件发生时，称事件 A 发生。

我们在研究随机现象时，不仅需要知道可能会出现哪些事件，更重要且更具实践意义的是研究各种事件发生可能性的大小并加以度量。

定义 .2

(事件的概率) 把刻画事件 A 发生可能性大小的数量指标称为事件 A 的概率，记为 $P(A)$.

计算 P(A) 的三种主要方法。

定义 .3 概率的统计定义

在观察某一随机事件 A 的随机试验中，随着随机试验次数 n 的增大，事件 A 发生的频率 p(A) 会越来越稳定地在某一常数 p 附近，对于具体给定的实验，这些值在 p 附近摆动，这时就以常数 p 作为事件 A 的概率，称之为统计概率，即 $P(A) = p$.

古典概率是一类特殊的随机试验——古典概型中随机事件的概率。具有以下特征的随机试验称为古典概型：

(1) 每次试验的样本空间 Ω 只包含有限个基本事件，记为如 $\omega_1, \omega_2, \cdots, \omega_n$;

(2) 各个基本事件出现的可能性相同，即基本事件的出现具有等可能性。

定义 .4 古典概型

对古典概型中的任一随机事件 A, 以 $p(A)=\frac{A中的基本事件数}{\Omega中的基本事件数}$ 作为事件 A 的概率，称为古典概率。

具有以下特征的随机试验称为几何概型：
(1) 随机试验可归结为在一个可度量的几何图形 Ω 中随机投点 (或取点)，以 $m(\Omega)$ 表示 Ω 的度量 (如长度、面积、体积等)，而事件 A 是指所投点 (取点) 落在 (取自) Ω 中的可度量图形 A 中；
(2) 事件 A 的概率与 A 的度量 $m(\Omega)$ 成正比，而与 A 在 Ω 中的位置无关。同时，几何概型是一类特殊的随机试验结果——使用几何方法度量随机事件发生的概率。

定义 .5 几何概型

对几何概型中的任一随机事件 A, 以

$$p(A) = \frac{m(A)}{m(\Omega)} \qquad (1)$$

作为事件 A 的概率，称为几何概型。

需要指出的是，随着概率论这门学科研究的深入和发展，产生了对随机事件概率高度科学概括的公理化定义。

定义 .6 概率的公理化定义

设随机试验 E 的样本空间为 n, 对于随机试验 E 的每一随机事件 A, 都赋予唯一确定的实数 P(A)，其中满足下列条件的集合函数 P$(\cdot)$ 称为事件 A 的概率：

(1) 非负性：对每一个事件 A $\subset \Omega$, 都有 P(A) ≥ 0;

(2) 规范性：P$(\Omega) = 1$;

(3) 可列可加性：对任意互不相容的事件 $A_1, A_2, \cdots$, 有 $P(\sum_{i=1}^{\infty} A) = \sum_{i=1}^{\infty} P(A)$.

定义 .7 条件概率

在概率论公理化结构中，称三元总体 (Ω, F, P) 为概率空间，其中 Ω 为样本空间，F 为事件域 (事件的全体)，P 为概率。

在统计信号处理中，我们还经常用到条件概率。

定义 .8 条件概率的度量

设 A 和 B 为任意两个随机事件，且 $P(A|B) > 0$，称

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{2}$$

为事件 B 发生条件下事件 A 发生的条件概率，也称 A 对 B 的概率。

贝叶斯公式 1 设某随机试验的样本空间 Ω 中的事件 $A_1, A_2, \cdots$ (有限个或可列个) A_i 构成一个完备事件组，且 $P(A_i) > 0\,(i = 1, 2, \cdots)$，则对任一事件 $A \subset \Omega, B \subset \Omega, P(B) > 0$, 有如下的条件概率:

$$P(A_m|B) = \frac{P(A_m)\,P(B|A_m)}{\sum P(A_i)\,P(B|A_i)}, \quad m = 1, 2, \cdots \tag{5}$$

称为贝叶斯公式。

最后，我们将简要介绍事件和试验的独立性。

定义 .9 事件独立

如果随机事件 A 与 B 满足关系

$$P(A|B) = p(B), \tag{6}$$

则称事件 A 与 B 相互独立，简称 A 与 B 独立。

独立事件具有如下的重要性质：若事件 $A_1, A_2, \cdots, A_n$ 相互独立, 则有

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

事件独立性

定义 .10 事件独立

一个试验重复进行 n 次，如果在每次试验中，任意事件出现的概率与其他各次试验结果无关，则称这 n 次试验是独立的，或称这 n 次试验是 n 次重复独立试验。

目录

在本部分，我们将引入随机变量的概念来表达随机事件。

定义 .11 随机变量

某随机试验的样本空间为 $\Omega = \{\omega\}$，对于每一个样本点 $\omega \in \Omega$ 都有唯一的实数 $X(\omega)$ 与之对应，这样就得到一个定义在 Ω 上的单值实函数 $X = X(\omega)$, 如果对任意实数 x, " $X(\omega) \le x$" 都是一个随机事件并有其确定的概率，则称 $X = X(\omega)$ 为随机变量。

随机变量的记法

随机变量常用大写拉丁字母 X, Y, Z 或希腊字母 ξ, η, ζ 等表示，随机变量的取值常用 x, y, z, a, b, c 等表示。

两类随机变量: 离散型随机变量和连续型随机变量

离散型随机变量

如果随机变量 X 的全部可能取值为有限个或可列个，则称 X 为离散型随机变量。

离散型随机变量的分布律

设离散型随机变量 X 所有可能的取值为 $x_i(i=1,2,\cdots)$, X 取各个可能值的概率，即事件 $X=x_i$ 的概率为

$$P(X=x_i)=p_i, i=1,2,\cdots \tag{7}$$

我们称(7)式为离散型随机变量 X 的概率分布或分布律。显然分布律也可以用表格的形式表示。

定义 .12 离散型随机变量的分布函数

设 X 是一个随机变量，对任意的实数 $x(-\infty < x < +\infty)$, 令

$$F(x) = P(X \leqslant x), \tag{8}$$

则称 $F(x)$ 为随机变量 X 的概率分布函数，简称为分布函数。它的定义域是 $(-\infty, +\infty)$，值域是 $[0, 1]$.

定义 .13 连续型随机变量的分布函数

如果对于随机变量 X 的分布函数 F(x)，存在一个非负可积函数 f(x) $(-\infty < +\infty)$，有

$$F(x) = \int_{-\infty}^{x} f(t)dt, \tag{9}$$

则称 X 为连续型随机变量，函数 f(x) 为其概率密度函数。

定义 .14 0-1 分布

假设随机试验只有两个可能结果 A 与 $\bar{A}$ 时，随机变量

$$X = \begin{cases} 0 & \bar{A}出现 \\ 1 & A出现 \end{cases} \tag{10}$$

X 表示在试验中事件 A 出现的次数，并设 $P(A) = p\,(0 < p < 1)$，则 X 的概率分布为

$$P(X = k) = p^k(1-p)^{1-k}, \quad k = 0, 1 \tag{11}$$

这时称 X 服从参数为 p 的 0-1 分布，记为 $X \sim B(1, p)$.

定义 .15 泊松分布

如果随机变量 X 的分布为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}, \tag{12}$$

其中 $\lambda>0$ 为常数，则称 X 服从参数为入的泊松分布，记为 $X\sim P(\lambda)$.

对于连续型随机变量，常见的分布有均匀分布和正态分布等。

定义 .16 均匀分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & \text{其他}\end{cases} \tag{13}$$

其中 a 和 b 为常数且 $a<b$, 则称 X 在区间 $[a,b]$ 上服从均匀分布，记为 $X\sim U[a,b]$.

正态分布

定义 .17 高斯分布

如果连续型随机变量 X 的概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x - m_X)^2}{2\sigma_X^2}\right], \quad -\infty < x < +\infty, \tag{14}$$

其中 m_X 和 σ_X 为常数, 且 $\sigma_X > 0$, 则称 X 服从参数为 m_X 和 σ_X 的高斯分布，记为 $X \sim N\left(m_X, \sigma_X^2\right)$.

瑞利分布

定义 .18 瑞利分布

如果连续型随机变量 X 的概率密度函数为

$$f(x,\sigma) = \begin{cases} \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), & x \geqslant 0 \\ 0, & x < 0 \end{cases}, \tag{15}$$

其中 $\sigma > 0$, 则称 X 服从参数为 σ 的瑞利分布。

瑞利分布的密度函数

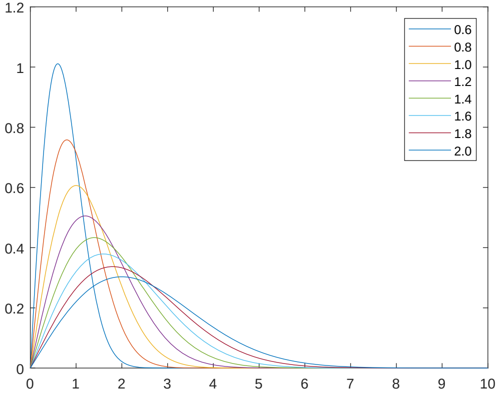


图 1: 瑞利分布的密度函数

卡方分布

定义 .19 卡方分布

如果连续型随机变量 X 的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right), & x > 0 \\ 0, & x \leqslant 0 \end{cases} \tag{16}$$

其中 $n \in \mathbb{Z}^+$，则称 X 服从自由度为 n 的 χ^2 分布。Γ 函数由以下积分为实数 $x > 0$：

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt. \tag{17}$$

Γ 函数和 factorial 函数有联系. 对于 $n \in \mathbb{Z}^+$:

$$\Gamma(n+1) = \text{factorial}\ (n) = \text{prod}(1:n) \tag{18}$$

$$\Gamma(n-1) = \frac{\Gamma(n)}{n-1} \tag{19}$$

定义 .20 莱斯分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{x}{\sigma^2}\exp\left[-\frac{(x^2+v^2)}{2\sigma^2}\right]I_0\left(\frac{xv}{\sigma^2}\right), & x>0\\ 0, & x\leqslant 0\end{cases}\tag{20}$$

其中 σ 为常数且 $\sigma>0$, $I_0(z)$ 是零阶第一类贝塞尔 (Bessel) 函数，则称 X 服从莱斯分布。

① 当 $v=0$ 时，莱斯分布退化为瑞利分布。

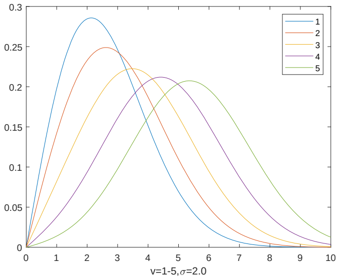


图 2: 莱斯分布在不同 v 下的密度函数

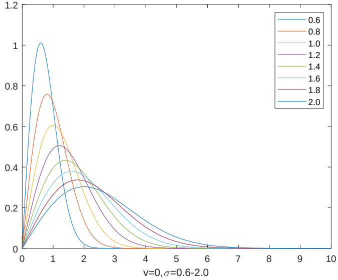


图 3: 莱斯分布在不同 σ 下的密度函数

联合分布

联合分布函数 $F_{XY}(x, y)$ 具有以下基本性质:
1° $F_{XY}(x, y)$ 分别对 x, y 单调不减。
2° $F_{XY}(x, y)$ 对每个变量均满足右连续。
3° $0 \leqslant F_{XY}(x, y) \leqslant 1$, 且 $F_{XY}(x, -\infty) = 0, F_{XY}(-\infty, y) = 0$ 和 $F_{XY}(+\infty, +\infty) = 1$.
4° 若任意四个实数 a_1, a_2, b_1, b_2, 满足 $a_1 \leqslant a_2, b_1 \leqslant b_2$, 则

$$
\begin{aligned}
P\{a_1 < X \leqslant a_2, b_1 < Y \leqslant b_2\} = F_{XY}(a_2, b_2) + F_{MY}(a_1, b_1) \\
- F_{XY}(a_1, b_2) - F_{XY}(a_2, b_1).
\end{aligned}
\tag{21}
$$

如图 4 所示。

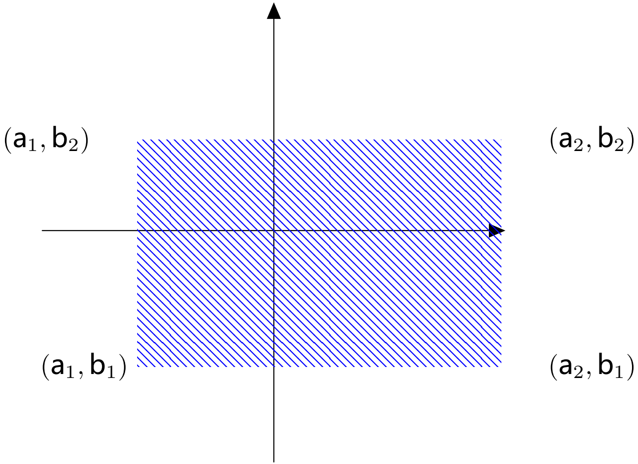


图 4: 联合分布函数

在几何上, $P(x, y) \in D$ 表示曲面 $f_X(x, y)$ 与 D 间所围的柱体体积, 如图 5 所示。

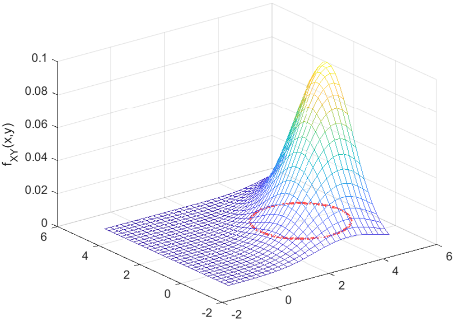


图 5: 矩形区域上的联合分布函数值

例 .1

设二维随机变量 (X, Y) 的概率密度

$$f(x,y)=\begin{cases} e^{-(x+y)}, & 0<x<+\infty, 0<y<+\infty \\ 0, & \text{其他} \end{cases} \tag{22}$$

求:

① 求分布函数 $F_X(x,y)$。

② (X, Y) 落在如图 6 所示的三角形域 G 内的概率.

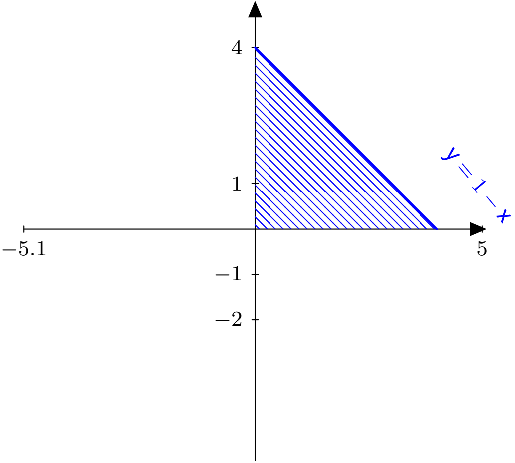


图 6: 三角积分区域 $G: x>0, y>0, x+y<1$

解: ① 分布函数

$$
\begin{aligned}
F_{XY}(x,y) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv \\
&= \left\{\int_{0}^{y}\int_{0}^{x} f(u,v)dudv, \quad 0<x<+\infty, 0<y<+\infty\right. \\
&= \begin{cases} \left(1-e^{-\alpha}\right)\left(1-e^{-y}\right), & 0<x<+\infty, 0<y<+\infty \\ 0, & \text{其他} \end{cases}
\end{aligned} \tag{23}
$$

② (X,Y) 落在三角形填 G 内的概率

$$
\begin{aligned}
P\{(x,y)\in G\} &= \iint_{\sigma} f(x,y)dxdy = \int_{0}^{1}\int_{0}^{1-y} e^{-(x+y)}dxdy \\
&= \int_{0}^{1} e^{-y}\left[\int_{0}^{1-y} e^{-x}dx\right]dy = \int_{0}^{1} e^{-y}\cdot\left(1-e^{-1+y}\right)dy \\
&= \int_{0}^{1}\left(e^{-y}-e^{-1}\right)dy = 1-2e^{-1} = 0.2642.
\end{aligned} \tag{24}
$$

3) 联合概率密度的定义

定义 .21 联合概率密度

若 $F_{XY}(x, y)$ 存在二阶偏导数, 则称

$$f_{XY}(x, y) = \frac{\partial F_{XY}(x, y)}{\partial x \partial y} \tag{25}$$

为二维随机变量 (X, Y) 的联合概率密度。

联合概率密度的基本性质

$$1^\circ\ f_{XY}(x,y) \geqslant 0 \tag{26}$$

$$2^\circ\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)dxdy = 1 \tag{27}$$

$$3^\circ\ \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,v)dudv = F_{XY}(x,y) \tag{28}$$

$$4^\circ\ P\{(x,y)\in D\} = \iint_{(u,v)\in D} f_{XY}(u,v)dudv \tag{29}$$

离散型二维随机变量

(4) 离散型二维随机变量

若二维随机变量 (X, Y) 的所有可能取值是有限对或可列无限多对, 则称 (X, Y) 的所有可能取值 (x_i, y_j) $(i, j = 1, 2, \cdots)$, 记为

$$P\{X = x_i, Y = y,\} = p_{ij}, i, j = 1, 2, \cdots \tag{30}$$

根据概率的性质，有

$$1^\circ \; p_{ij} \geqslant 0 \tag{31}$$

$$2^\circ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1 \tag{32}$$

则称 p_{ij} 为二维离散型随机变量 (X, Y) 的分布律, 或称为随机变量 X 和 Y 的联合分布律。

利用阶跃函数 $U(x)$ 与冲激函数 $\delta(x)$, 离散型二维随机变量的联合分布函数可表示为

$$
\begin{aligned}
F_{XY}(x,y) &= P\{X \leqslant x, Y \leqslant y\} \\
&= \sum_i \sum_j P\left\{X = x_i, Y = y_j\right\} U\left(x - x_i\right) U\left(y - y_j\right) \\
&= \sum_i \sum_j p_{ij} U\left(x - x_i\right) U\left(y - y_j\right). \qquad (33)
\end{aligned}
$$

离散型二维随机变量的联合概率密度可表示为

$$
\begin{aligned}
f_{XY}(x,y) &= \sum_i \sum_j P\left\{X = x_i, Y = y_j\right\} \delta\left(x - x_i\right) \delta\left(y - y_j\right) \\
&= \sum_i \sum_j p_{ij} \delta\left(x - x_i\right) \delta\left(y - y_j\right). \qquad (34)
\end{aligned}
$$

2. 二维随机变量的边缘分布和条件分布

边缘分布函数

二维随机变量 (X, Y) 作为一个整体，它具有分布函数 $F_{XY}(x, y)$; 而 X 和 Y 也都是随机变量，各自的分布函数为 $F_X(x)$ 和 $F_Y(y)$，联合分布函数 $F_{XY}(x, y)$ 具有如下关系:

$$F_X(x) = F_X(x, \infty), \tag{35}$$

$$F_Y(y) = F_{XY}(\infty, y). \tag{36}$$

则称 $F_X(x)$ 和 $F_Y(y)$ 分别为 (X, Y) 关于 X 和 Y 的边缘分布函数，简称 X 和 Y 的边缘分布函数。

边缘分布律

$$F_X(x) = P\{X \leq x\} = F_{XY}(x, \infty) = \sum_i \sum_{j=1}^{\infty} p_{ij} U(x - x_i) = \sum_{x_i \leqslant x} \sum_{j=1}^{\infty} p_{ij}. \tag{37}$$

可得

$$p_i = P\{X = x_i\} = \sum_{j=1}^{\infty} p_{ij}, i = 1, 2, \cdots, \tag{38}$$

称之为 (X, Y) 关于 X 的边缘分布律。

(X, Y) 关于 Y 的边缘分布律

$$p_j = P\{Y = y_j\} = \sum_{i=1}^{\infty} p_{ij}, j = 1, 2, \ldots \tag{39}$$

二维随机变量中各随机变量本身的统计特征: 边缘分布函数、边缘概率密度和边缘分布律。

二维离散随机变量边缘分布图

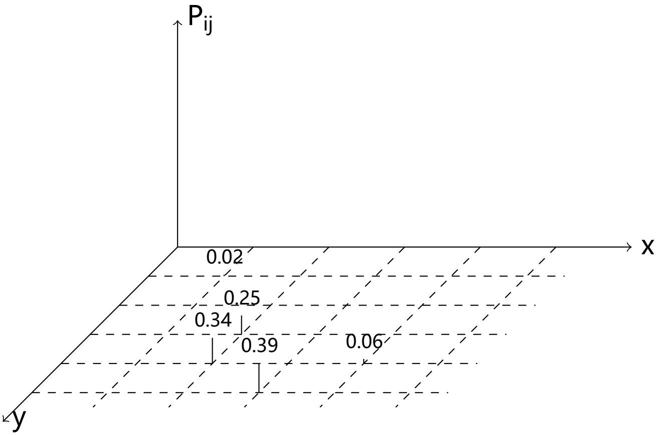


图 7: 离散型二维随机变量的分布律

(X,Y) 关于 X 的边缘分布

对于连续型随机变量 (X, Y), 有

$$F_X(x) = P\{X \le x\} = F_{XY}(x, \infty) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{XY}(u, y) dydu.$$

对边缘分布函数 $F_X(x) = P\{X \le x\} =? = F_{XY}(x, \infty) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{XY}(u, y) dydu$ 求导，得

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy. \tag{40}$$

称之为 (X, Y) 关于 X 的边缘概率密度。

对于连续型随机变量 (X, Y), 关于 Y，有

$$F_Y(y) = P\{Y \le y\} = F_{XY}(\infty, y) = \int_{-\infty}^{y} \int_{-\infty}^{+\infty} f_{XY}(u, y) dydu.$$

同理, (X, Y) 关于 Y 的边缘概率密度为

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx. \tag{41}$$

边缘分布函数——例 4

例 .2

(续) ① 求边缘分布函数 $F_X(x)$ 和 $F_Y(y)$。② 求边缘概率密度。

解: ① 已知联合分布函数 $F_{XY}(x,y)$ 的表达式, 则边缘分布函数

$$F_X(x) = F_X(x,\infty) = \begin{cases} (1-e^{-x})(1-e^{-\infty}), & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}$$
$$= \begin{cases} 1-e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \tag{42}$$

同理

$$F_Y(y) = F_{XY}(\infty,y) = \begin{cases} 1-e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \tag{43}$$

(1) 边缘概率密度

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dy = \begin{cases} \int_0^{+\infty} e^{-(x+y)}dy, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}$$
$$= \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \tag{44}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dx = \begin{cases} e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \tag{45}$$

定义 .22 条件分布函数和条件概率密度

对于连续型二维随机变量 (X,Y)，有

$$F_Y(y|X=x)=\int_{-\infty}^{y}\frac{f_{XY}(x,v)}{f_X(x)}dv, \tag{46}$$

$$f_Y(y|X=x)=\frac{f_{XY}(x,y)}{f_X(x)}. \tag{47}$$

分别称为给定 $X=x$ 的条件下, Y 的条件分布函数和条件概率密度, 可以简写为 $F_Y(y|x), f_Y(y|x)$.

2) 推导过程

前面引入了条件概率的概念，即在给定事件 B 的条件下，事件 A 发生的条件概率。

$$P(A|B) = \frac{P(AB)}{P(B)}, \tag{48}$$

把这个概念引入到随机变量的理论中。

对于连续型二维随机变量 (X, Y)，若令 $A = \{Y \leqslant y\}, B = \{X \leqslant x\}$，则称 $P(A|B) = P\{Y \leqslant y|B\} = F_Y(y|B)$ 为给定 B 条件下的 Y 的分布函数。

由以上讨论，可得条件分布函数 $F_Y(y|B)$，联合分布函数 $F_{XY}(x,y)$ 及边缘分布函数 $F_X(x)$ 三者之间的关系为

$$
\begin{aligned}
F_Y(y|X \leqslant x) = P(Y \leq y|X \leqslant x) = P(A|B) &= \frac{P(AB)}{P(B)} \\
&= \frac{P\{X \leqslant x, Y \leqslant y\}}{P\{X \leqslant x\}} \\
&= \frac{F_{XY}(x,y)}{F_X(x)}. \qquad (49)
\end{aligned}
$$

若上式对 y 的导数存在, 则有

$$f_Y(y|X \leqslant x) = \frac{\partial}{\partial y} F_Y(y|X \leqslant x) = \frac{\partial}{\partial y}\left[\frac{F_{XY}(x,y)}{F_X(x)}\right] = \frac{\partial F_{XY}(x,y)/\partial y}{F_X(x)}$$
$$= \frac{\int_{-\infty}^{x} f_{XY}(u,y)du}{\int_{-\infty}^{x} f_X(u)du}. \qquad (50)$$

若令 $B = \{X = x\}$, 需要特殊处理，代入式 (49)

$$F_Y(y|X = x) = \lim_{\Delta x \to 0} F_Y(y|x < X \leq \Delta x) = P\{?\}$$
$$= \lim_{\Delta x \to 0} \frac{P\{x < X \leq \Delta x, Y \leq y\}}{P\{x < X \leq \Delta x\}}$$
$$= \lim_{\Delta x \to 0} \frac{P\{X \leqslant x + \Delta x, Y \leqslant y\} - P\{X \leqslant x, Y \leqslant y\}}{P\{X \leqslant x + \Delta x\} - P\{X \leqslant x\}}$$

$$
\begin{aligned}
F_Y(y|X=x) &= \lim_{\Delta x\to 0}\frac{[F_{XY}(x+\Delta x,y)-F_{XY}(x,y)]/\Delta x}{[F_X(x+\Delta x)-F_X(x)]/\Delta x}\\
&= \frac{\partial F_{XY}(x,y)/\partial x}{\partial F_X(x)/\partial x}=\frac{\int_{-\infty}^{y} f_{XY}(x,v)dv}{f_X(x)}.
\end{aligned} \tag{51}
$$

得条件概率密度

$$
\begin{aligned}
f_Y(y|X=x) &= \frac{\partial F_Y(y|X=x)}{\partial y}\\
&= \frac{f_{XY}(x,y)}{f_X(x)},\ f_X(x)\neq 0.
\end{aligned} \tag{52}
$$

例 .3

(续) ⑤ 求条件分布函数 $F_X(x|y)$ 和 $F_Y(y|x)$。⑥ 求条件概率密度 $f_X(x|y)$。

解: ⑤ 条件分布函数

$$F_X(x|y) = \frac{\int_{-\infty}^{x} f_X(x,y)dx}{\int_{-\infty}^{x} f_Y(y)} = \begin{cases} \dfrac{\int_0^x e^{-(x+y)}dx}{\int_0^y e^{-y}dy} & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}$$

$$= \begin{cases} 1 - e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \tag{53}$$

同理

$$F_Y(y|x) = \begin{cases} 1 - e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \tag{54}$$

⑥ 条件概率密度

$$\begin{aligned} f_X(x|y) = \frac{f_X(x,y)}{f_Y(y)} &= \begin{cases} \dfrac{e^{-(x+y)}}{e^{-y}}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} . \end{aligned} \tag{55}$$

条件分布函数 $F_Y(y|B)$ 是求在 B 发生的条件下，事件 $\{Y(\zeta) \leqslant y\}$ 发生的概率，$\zeta \in B$. 换句话说，它是求在新的样本空间上，事件 $\{Y(\zeta) \leqslant y\}$ 发生的概率。

而无条件的分布函数 $F_Y(y)$，则是在 Ω 上求 $\{Y(\xi) \leqslant y\}$ 事件发生的概率 $\zeta \in \Omega$。因此，除了样本空间缩小成 Ω_B 以外，条件分布函数的性质与一般分布函数的性质完全相同。

条件分布函数性质:

1° $F_Y(\infty|B) = 1$, $F_Y(-\infty|B) = 0, 0 \leqslant F_Y(y|B) \leqslant 1.$

2° $F_Y(y_2|B) - F_Y(y_1|B) = P\{y_1 < Y \leqslant y_2|B\}.$

条件密度函数

条件密度函数性质：

1° $f_Y(y|B) \geqslant 0.$

2° $\int_{-\infty}^{\infty} f_Y(y|B)dy = F_Y(\infty|B) - F_Y(-\infty|B) = 1.$

3° $F_Y(y|B) = \int_{-\infty}^{y} f_Y(v|B)dv.$

4) 离散型随机变量的条件分布律

对于离散型随机变量 X 和 Y, 其在 $X=x$ 的条件下, $Y=y$ 的条件概率可直接定义为

$$P\{Y=y|X=x\}=\frac{P\{X=x,Y=y\}}{P\{X=x\}},P(X=x)>0. \tag{56}$$

4) 二维离散型随机变量的条件分布律

二维随机变量 (X,Y), 对于固定的 j, 若 $P\{Y=y_j\}>0$, 则

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}},i=1,2,\cdots \tag{57}$$

称为在 $Y=y$ 条件下随机变量 X 的条件分布律。

条件分布律

对于固定的 i, 若 $P\{X=x_i\}>0$, 则

$$P\left\{Y=y_j|X=x_i\right\}=\frac{P\left(X=x_i,Y=y_j\right)}{P\left\{X=x_i\right\}}=\frac{p_{ij}}{p_{i\cdot}},j=1,2,\cdots \quad (58)$$

称为在 $X=x_1$ 条件下随机变量 Y 的条件分布律。

3. 随机变量的统计独立

前面引入了事件独立的概念, 现在把它引入到随机变量中来。
X, Y 是两个随机变量, 若对任意实数 x 和 y, 有

$$P\{X<x,Y<y\}=P\{(X<x)\cap(Y<y)\}=P\{X<x\}\cdot P\{Y<y\}, \quad (59)$$

则称随机变量 X, Y 相互独立。

3. 随机变量相互独立的条件

对于二维随机变量 (X, Y), X 与 Y 相互独立的条件为

$$F_{XY}(x, y) = F_X(x) \cdot F_Y(y). \tag{60}$$

或

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y). \tag{61}$$

把上两式代入条件分布函数和概率密度的定义 (式 (1-73)) 可得

$$\begin{cases} F_Y(y|x) = F_Y(y) \\ f_Y(y|x) = f_Y(y) \end{cases} \tag{62}$$

同理可得

$$\begin{cases} F_X(x|y) = F_X(x) \\ f_X(x|y) = f_X(x) \end{cases} . \tag{63}$$

说明: 当 X 和 Y 相互独立时, X 在 $Y=y$ 的条件下的分布与 X 的无条件分布相同, 或者 Y 在 $X=x$ 的条件下的分布与 Y 的无条件分布相同。也就是说, 随机变量 X 的统计特征与随机变量 Y 的统计特征无关。

离散型随机变量 X 和 Y 独立的条件: 对所有 i,j, 均有 $p_{ij}=p_{i\cdot}p_{\cdot j}$ $(i,j=1,2,\cdots)$, 即

$$P\left\{X=x_i,Y=y_j\right\}=P\left\{X=x_i\right\}\cdot P\left\{Y=y_j\right\}. \tag{64}$$

独立性判定的三种方法

例 .4

(续) ⑦ X 和 Y 是否统计独立?

解: ⑦ 独立的判定可由下满的四种方法判定 (只需满足 $1-4$ 中的任一条件即可推出统计独立).

1. 分布函数的判定方法

由已知条件和求得的结论可知

$$F_{XY}(x,y)=\begin{cases}(1-e^{-x})(1-e^{-y}), & 0<x,y<+\infty\\ 0, & \text{其他}\end{cases}$$

$$F_X(x)=\begin{cases}1-e^{-x}, & 0<x<+\infty, 0,\\ \text{其他} & \end{cases} \tag{65}$$

$$F_Y(y)=\begin{cases}1-e^{-y}, & 0<y<+\infty\\ 0, & \text{其他}\end{cases} \tag{66}$$

所以 $F_{XY}(x,y)=F_X(x)\cdot F_Y(y)$ 成立, 则 X 和 Y 统计独立。

2. 条件分布函数的判定方法

由已知条件和求得的结论可知

$$F_X(x|y) = \begin{cases} 1-e^{-x}, & 0 < x < +\infty \\ 0, & 其他 \end{cases}, \tag{67}$$

$$F_X(x) = \begin{cases} 1-e^{-x}, & 0 < x < +\infty \\ 0, & 其他 \end{cases} \tag{68}$$

所以 $F_X(x|y) = F_X(x)$ 成立, 则 X 和 Y 统计独立。

2. 密度函数的判定

3. 密度函数的判定方法

由已知条件和求得的结论, 可知

$$f_{XY}(x,y)=\begin{cases} e^{-(x+y)}, & 0<x<+\infty, 0<y<+\infty \\ 0, & \text{其他} \end{cases} \tag{69}$$

$$f_X(x)=\begin{cases} e^{-x}, & 0<x<+\infty \\ 0, & \text{其他} \end{cases}, \tag{70}$$

$$f_Y(y)=\begin{cases} e^{-y}, & 0<y<+\infty \\ 0, & \text{其他} \end{cases}, \tag{71}$$

所以 $f_{XY}(x,y)=f_X(x)\cdots f_Y(y)$ 成立, 则 X 和 Y 统计独立。

4. 条件密度函数的判定方法

由已知条件和求得的结论可知

$$f_X(x|y) = \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}, \tag{72}$$

$$f_X(x) = \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \tag{73}$$

所以 $f_X(x|y) = f_X(x)$ 成立, 则 X 和 Y 统计独立。

目录

随机矢量

在实际问题中，一些随机现象常常需要两个或两个以上的随机变量来描述，例如高频信号的中心频率、带宽、振幅和初相位等等。

多个随机变量可以组成一个随机矢量，先引入多维随机变量的矢量的概念。

定义 .23 随机矢量

设某随机试验的样本空间为 $\Omega = \{w\}$, 对于每一个样本点 $\omega \in \Omega$, $X_1 = X_1(\omega), X_2 = X_2(\omega), \cdots, X_n = X_n(\omega)$ 是定义在同一个样本空间 Ω 上的 n 个随机变量，则 $(X_1, X_2, \cdots, X_n)$ 称为 n 维随机变量，其矢量形式 $\mathbf{X} = [X_1, X_2, \cdots, X_n]^T$ 也称为随机矢量。

类比方法来推广

二维随机变量的情况的很多结果都可以推广到多维随机变量上。

定义 .24 二维联合分布

设 (X, Y) 是二维随机变量，对于任意实数 x 和 y, 二元函数

$$F_{XY}(x, y) = P(X \leqslant x, Y \leqslant y), \tag{74}$$

称为二维随机变量 (X, Y) 的分布函数，或称随机变量 X 和 Y 的联合分布函数，简称联合分布。

边缘分布函数的概念

二维随机变量 (X,Y) 作为一个整体，具有分布函数 $F(x,y)$. 而 X 和 Y 都是随机变量，我们也可以对其中任何一个随机变量单独进行研究，即求随机变量 X 或 Y 的分布，这就是二维随机变量的边缘分布。

二维随机变量的边缘分布

二维随机变量的边缘分布与二维变量的分布函数具有如下关系：

$$F_X(x) = P(X \leqslant x, Y < +\infty) = F_{XY}(x, +\infty). \tag{75}$$

$$F_Y(y) = P(X < +\infty, Y \leqslant y) = F_{XY}(+\infty, y). \tag{76}$$

二维离散型和连续型随机变量

定义 .25 二维离散型随机变量

如果二维随机变量 (X, Y) 的所有可能取值都是有限对或可列无限多对，并且以确定的概率取各个不同的数对，则称 (X, Y) 为二维离散型随机变量.

定义 .26 联合概率分布

若 (X, Y) 是一个二维离散型随机变量，它的一切可能取值为 (x_i, y_j) $(i, j = 1, 2, \cdots)$, 则

$$P(X = x_i, Y = y_i) = p_{ij}, \quad i, j = 1, 2, \cdots \tag{77}$$

称为 (X, Y) 的联合概率分布。

先来回忆下二维随机变量 (X, Y) 情形下的定义：

定义 .27 联合概率密度函数

设二维随机变量 (X, Y) 的分布函数是 $F(x, y)$, 如果存在非负函数 $f(x, y)$ 使得对任意实数 x, y 有

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv, \tag{78}$$

则称 (X, Y) 是二维连续型随机变量，$f(x, y)$ 为二维连续型随机变量 (X, Y) 的联合概率密度函数。

利用二维随机变最的联合分布函数与边缘分布函数，引入随机变量独立的概念

定义 .28 独立

设 $F(x,y)$ 是二维随机变量 (X,Y) 的联合分布函数，$F_X(x)$ 和 $F_Y(y)$ 是其边缘分布函数，若对任意实数 x 和 y, 有

$$F(x,y) = F_X(x)F_Y(y), \tag{79}$$

则称随机变量 X 和 Y 是统计独立的。

此外，对于两个随机变量，我们也可以讨论它们的条件分布。对二维离散随机变量，由条件概率的公式可得

$$P(X = x_i | Y = y_i) = \frac{P(X = x_i, Y = y_i)}{P(Y = y_i)}. \tag{80}$$

上式被称作在 $Y = y_i$ 条件下 $X = x$ 的条件概率，其中 $P(Y = y_i)$ 表示边缘分布函数。

由于连续型随机变最取任何数值的概率都是零，所以不能像离散型随机变量那样直接利用条件概率公式给出连续型随机变量的条件概率密度。

采用极限的办法来解决：设 y 是定值，对任 $\Delta y > 0$, $P(y - \Delta y < Y \leqslant y + \Delta y) > 0$, 若对任意实数 X, 极限 $\lim_{\Delta y \to 0} P(X \leqslant x | y - \Delta y < Y \leqslant y + \Delta y)$ 存在，则称此极限为在 $Y = y$ 条件下 X 的条件分布函数，记为 $P(X \leqslant x | Y = y)$.

n 维随机变量中的边缘分布函数和边缘概率密度

边缘分布函数和边缘概率密度

n 维随机变量中的任意 m 个分量的联合分布函数 ($m<n$), 都称为 n 维随机变量的 m 维边缘分布函数。

由 n 维随机变量的联合分布函数 $F_X(x_1, x_2, \cdots, x_n)$, 可以得到它任意 m 个分量的边缘分布函数。

$$
\begin{aligned}
&F_X(x_1, x_2, \cdots, x_m) = F_X(x_1, x_2, \cdots, x_m, \infty, \cdots, \infty).\\
&F_X(x_i) = F_X(\infty, \cdots, \infty, x_i, \infty, \cdots, \infty).
\end{aligned} \tag{81}
$$

m 维边缘概率密度

n 维随机变量中的任意 m $(m < n)$ 个分量的概率密度, 都称为 n 维随机变量的 m 维边缘概率密度。

由 n 维随机变量的联合概率密度 $f_X(x_1, x_2, \cdots, x_n)$, 可以得到它任意 m 个分量的边缘概率密度。如

$$f_X(x_1, x_2, \cdots, x_m) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \cdots, x_m, x_{m+1}, \cdots, x_n) dx_{m+1} dx_{m+2} \cdots dx_n. \tag{82}$$

$$f_X(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n. \tag{83}$$

4. 条件概率密度

n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 在给定 $X_1 = x_1$ 的条件下, 其余 $n-1$ 个分量 $(X_2, X_3, \cdots, X_n)$ 的条件概率密度为

$$f_X(x_2, \cdots, x_n | x_1) = \frac{f_X(x_1, x_2, \cdots, x_n)}{f_X(x_1)}. \tag{84}$$

也可以将多个随机变量固定为条件, 如在 $X_1 = x_1$ 和 $X_2 = x_2$ 条件下, 随机变量 $(X_3, X_4, \cdots, X_n)$ 的条件概率密度为

$$f_X(x_3, \cdots, x_n | x_1, x_2) = \frac{f_X(x_1, x_2, \cdots, x_n)}{f_X(x_1, x_2)}. \tag{85}$$

在 $X_1, X_2, \cdots, X_{n-1}$ 固定的条件下，随机变量 X_n 的条件概率密度为

$$f_X(x_n|x_1, x_2, \cdots, x_{n-1}) = \frac{f_X(x_1, x_2, \cdots, x_n)}{f_X(x_1, x_2, \cdots, x_{n-1})}. \tag{86}$$

n 维随机变量联合概率密度的递推公式

$$f_X(x_1, \cdots, x_n) = f_X(x_n|x_1, \cdots, x_{n-1}) f_X(x_{n-1}|x_1, \cdots, x_{n-2}) \cdots f_X(x_2|x_1) f_X(x_1). \tag{87}$$

使用归纳法推导，具体步骤如下:

二维情况 $f_X(x_1, x_2) = f_X(x_2|x_1) f_X(x_1)$.

三维情况 $f_X(x_1, x_2, x_3) = f_X(x_3|x_1, x_2) f_X(x_1, x_2)$.

可得到 (X_1, X_2, X_3) 三维联合概率密度的递推关系

$$f_X(x_1, x_2, x_3) = f_X(x_3|x_1, x_2)\, f_X(x_2|x_1)\, f_X(x_1)\,. \tag{88}$$

由数学归纳法, 可推出 n 维随机变量联合概率密度的递推关系

$$f_X(x_1, x_2, x_3) = f_X(x_3|x_1, x_2)\, f_X(x_2|x_1)\, f_X(x_1)\,. \tag{89}$$

当 $X_1, X_2, \cdots, X_n$ 相互独立时, 可进一步得到

$$f_X(x_1, \cdots, x_n) = f_X(x_n)\, f_X(x_{n-1}) \cdots f_X(x_2)\, f_X(x_1)\,. \tag{90}$$

例 .1

四维随机变量 (X_1, X_2, X_3, X_4) 中各随机变量相互独立, 且都服从 $(0,1)$ 上的均匀分布。求:

① 四维随机变量的联合概率密度 $f_X(x_1, x_2, x_3, x_4)$。
② 边缘概率密度 $f_X(x_1, x_2)$。
③ 条件概率密度 $f_X(x_3|x_1, x_2)$ 和 $f_X(x_3, x_4|x_1, x_2)$。

解: ① X_i 服从 $(0,1)$ 上的均匀分布, 则 X_i 的概率密度为

$$f_X(x_i) = \begin{cases} 1, & 0 < x_i < 1 \\ 0, & \text{其他} \end{cases}, i = 1, 2, 3, 4. \tag{91}$$

随机变量 X_1, X_2, X_3, X_4 相互独立, 则四维随机变量的联合概率密度为

$$f_X(x_1, x_2, x_3, x_4) = f_X(x_1) \cdot f_X(x_2) \cdot f_X(x_3) \cdot f_X(x_4) = \begin{cases} 1, & 0 < x_i < 1, i = 1, 2, 3, 4 \\ 0, & \text{其他} \end{cases} \tag{92}$$

② 同理可知 X_1, X_2 的联合概率密度为

$$f_X(x_1, x_2) = f_X(x_1) \cdot f_X(x_2) = \begin{cases} 1, & 0 < x_1, x_2 < 1 \\ 0, & \text{其他} \end{cases} \tag{93}$$

③ 因为随机变量 X_1, X_2, X_3, X_4 相互独立, 所以条件概率密度

$$f_X(x_3|x_1,x_2) = f_X(x_3) = \begin{cases} 1, & 0 < x_3 < 1 \\ 0, & \text{其他} \end{cases} \tag{94}$$

$$\begin{aligned} f_X(x_3,x_4|x_1,x_2) &= f_X(x_3,x_4) = f_X(x_3)\cdot f_X(x_4) \\ &= \begin{cases} 1, & 0 < x_3, x_4 < 1 \\ 0, & \text{其他} \end{cases} \end{aligned}$$

目录

定理 .1 1 维随机变量函数的分布

设 X 是一个连续型随机变量，其概率密度函数为 $f(x), -\infty < X < +\infty$, $g(x)$ 处处可导且有 $g'(x) > 0$ 或恒有 $g'(x) < 0$, 则 $Y = g(X)$ 也是一个连续型随机变量，且其概率密度函数为

$$f(y) = \begin{cases} f(h(y))\,|h'(y)|, & a < y < \beta \\ 0, & \text{其他} \end{cases} \tag{95}$$

其中，$\alpha = \min\{g(-\infty), g(+\infty)\}$, $\beta = \max\{g(-\infty), g(+\infty)\}$, $h(y)$ 是 $g(x)$ 的反函数。

特别地，若随机变量 $X \sim N(\mu, \sigma^2)$, 令随机变量 $Y = \frac{X-\mu}{\sigma}$，则 $Y \sim N(0, 1)$.

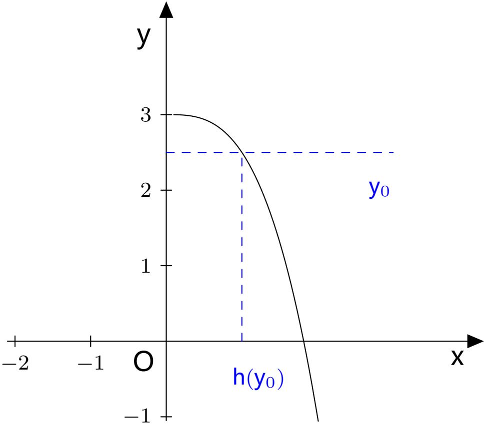


图 8: 单调递减函数

解:

- 当 $y \leq \alpha$, $F_Y(y) = P\{Y \leq y\} = P\{\emptyset\} = 0.$
- 当 $y \geq \beta$, $F_Y(y) = P\{Y \leq y\} = P\{\Omega\} = 1.$
- 当 $\alpha \leq y < \beta$,
  $F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = P\{X \leq h(y)\} = \int_{-\infty}^{h(y)} f_X(x)dx.$

概率密度函数

$$f_Y(y) = F_Y'(y) = \begin{cases} f_X(h(y))h'(y), & \alpha \leq y < \beta \\ 0, & \text{其他} \end{cases},$$

对于 $g'(x) < 0$ 的情况,

$$f_Y(y) = F'_Y(y) \begin{cases} f_X(h(y))(-h'(y)), & \alpha \le y < \beta \\ 0, & \text{其他} \end{cases},$$

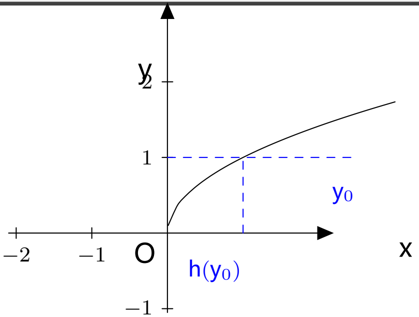


图 9: 单调递增函数

对于 $g'(x) > 0$ 的情况,

$$f_Y(y) = F_Y'(y) \begin{cases} f_X(h(y))(h'(y)), & \alpha \le y < \beta \\ 0, & \text{其他} \end{cases},$$

$$f_Y(y) = F_Y'(y) \begin{cases} f_X(h(y))|h'(y)|, & \alpha \le y < \beta \\ 0, & \text{其他} \end{cases}.$$

例 .1

设 X 的概率密度为

$$\mathrm{f}(\mathrm{x}) = \begin{cases} \frac{2\mathrm{x}}{\pi^2}, & 0 < \mathrm{x} < \pi \\ 0, & \mathrm{x} \end{cases}$$

求 $\mathrm{Y} = \sin \mathrm{X}$ 的概率密度.

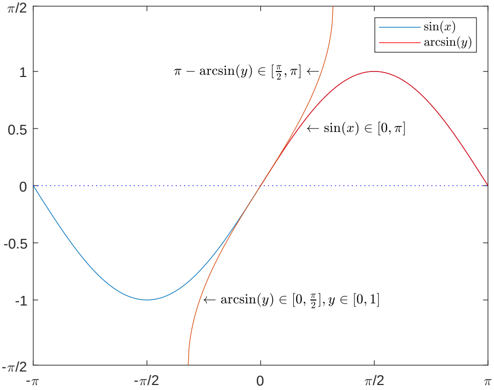


图 10: sin(x) 及其反函数的取值区间.

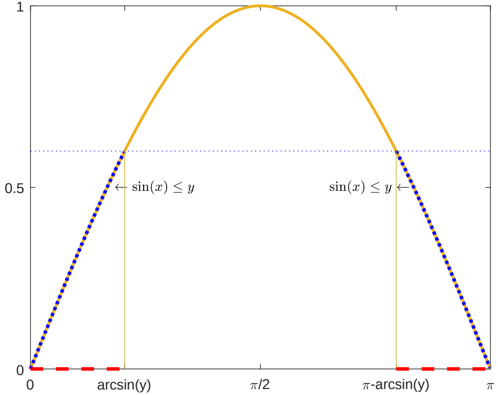


图 11: $\sin(x) \le y$ 的取值范围.

解: $F_Y(y) = P\{Y \le y\} = P\{\sin X \le y\}$.

- 当 $y < 0$ 时, $F_Y(y) = 0$.
- 当 $0 \le y \le 1$ 时: $F_Y(y) = P\{\sin X \le y\} = P\{0 \le X \le \arcsin y\} + P\{\pi - \arcsin y \le X \le \pi\}$,
  $(x = \arcsin y \in [0, \frac{\pi}{2}], 0 \le y \le 1;$
  $x = \pi - \arcsin y \in [\frac{\pi}{2}, \pi], 0 \le y \le 1.$ 如图10或者11).

$$F_Y(y) = \int_0^{\arcsin y} \frac{2x}{\pi^2} dx + \int_{\pi - \arcsin y}^{\pi} \frac{2x}{\pi^2} dx.$$

- 当 $1 < y$ 时: $F_Y(y) = 1$.

Y 的概率密度 f(y) 为:

- 当 $y \le 0$ 时, $f(y) = [F_Y(y)]' = (0)' = 0.$

Y 的概率密度 f(y) 为:

- 当 $y \leq 0$ 时, $f(y) = [F_Y(y)]' = (0)' = 0.$
- 当 $0 < y < 1$ 时,

$f(y) = [F_Y(y)]' = \left(\int_0^{\arcsin y} \frac{2x}{\pi^2} dx + \int_{\pi - \arcsin y}^{\pi} \frac{2x}{\pi^2} dx\right)',$

即

$$
\begin{aligned}
f(y) &= \left(\frac{x^2}{\pi^2}\Big|_0^{\arcsin y} + \frac{x^2}{\pi^2}\Big|_{\pi - \arcsin y}^{\pi} \right)' \\
&= \left(\frac{(\arcsin y)^2}{\pi^2} + 1 - \frac{(\pi - \arcsin y)^2}{\pi^2} \right)' \\
&= \left(\frac{(\arcsin y)^2}{\pi^2} + 1 - \frac{\pi^2 - 2\pi \arcsin y + (\arcsin y)^2}{\pi^2} \right)' \\
&= \left(\frac{2 \arcsin y}{\pi} \right)' = \frac{2}{\pi\sqrt{1 - y^2}}.
\end{aligned}
$$

当 $1 \le y$ 时, $f(y) = [F_Y(y)]' = (1)' = 0$.
Y 的概率密度 $f(y)$ 为:

$$f(y) = \begin{cases} 0, & y \ge 1 \text{或} y \le 0, \\ \frac{2}{\pi\sqrt{1-y^2}}, & 0 < y < 1. \end{cases}$$

定理 .2 n 维随机变量函数的分布

设 $(X_1, X_2, \cdots, X_n)$ 具有概率密度函数 $f(x_1, x_2, \cdots, x_n)$ 的续型 n 维随机变量。

(1) $y_1 = g_1(x_1, x_2, \cdots, x_n), \cdots, y_n = g_n(x_1, x_2, \cdots, x_n)$ 分别是 n 维实数空间到自身的一对一的映射，即存在定义在该变换值域上的逆变换：

$$x_1 = h_1(y_1, y_2, \cdots, y_n), \cdots, x_n = h_n(y_1, y_2, \cdots, y_n).$$

(2) 变换和它的逆变换都是连续的；

(3) 偏导数 $\frac{\partial h_i}{\partial y_j}\ (i = 1, 2, \cdots, n; j = 1, 2, \cdots, n)$ 存在且连续;

定理 .3 n 维随机变量

(4) 逆变换的雅可比行列式

$$J(y_1, y_2, \cdots, y_n) = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \frac{\partial h_n}{\partial y_2} & \cdots & \frac{\partial h_n}{\partial y_n} \end{vmatrix} \neq 0,$$

则 $(Y_1, \cdots, Y_n) = (g_1(X_1, X_2, \cdots, X_n), \cdots, g_n(X_1, X_2, \cdots, X_n))$
的联合概率密度函数函数

$$f(y_1, y_2, \cdots, y_n) = f(h_1(y_1, y_2, \cdots, y_n), \cdots, h_n(y_1, y_2, \cdots, y_n)|J|. \qquad \text{dotted}(96)$$

上节讨论了随机变量的概念及其分布。但实际工作中, 还经常遇到求随机变量函数分布的问题。

例 .2

电子系统中, 在 t 时刻一个概率密度为 $f(x)$ 的随机变量 X 通过一个非线性放大器。

$$Y = \begin{cases} X^{1/n}, & X \geqslant 0 \\ -|X|^{1/n}, & X < 0 \end{cases} \quad n \text{ 为正整数} \tag{97}$$

如何求出输出随机变量 Y 的概率密度呢?

显然, 若能找到求随机变量 X 的函数 $Y = g(X)$ 的概率密度的方法, 就能解决上述的实际问题。

1. 一维随机变量函数的分布

对于一维随机变量函数的分布, 分成两种情况来讨论。

1. 单值变换

随机变量 X 和 Y 存在单调函数关系 $Y = g(X)$, 存在唯一反函数 $X = h(Y)$, 即若有一个 X 出现, 则必有一个与其对应的 Y 出现。若 X 位于 $(x_0, x_0 + dx)$ 区间内, 则 Y 必位于 $(y_0, y_0 + dy)$ 区间内。因此, X 落在区间 $(x_0, x_0 + dx)$ 内的概率等于 Y 落在区间 $(y_0, y_0 + dy)$ 的概率, 有

$$P\{x_0 < X \leqslant x_0 + dx\} = P\{y_0 < Y \leqslant y_0 + dy\}. \tag{98}$$

可得

$$f_Y(y)dy = f_X(x)dx, \tag{99}$$

所以

$$f_Y(y) = f_X(x)\frac{dx}{dy} = [h'(y)] \cdot f_X[h(y)]. \tag{100}$$

由于概率密度不可能取负值, 所以 dx/dy 应取绝对值, 即

$$f_Y(y) = |h'(y)| \cdot f_X[h(y)]. \tag{101}$$

这样, 不论 $h(y)$ 是单调增函数 ($h'(y) > 0$), 还是单调减函数 ($h'(y) < 0$), 上式均成立。

例 .3

随机变量 X 和 Y 之间成线性关系: $Y = X + 5$。已知随机变量 X 服从标准高斯分布。求随机变量 Y 的概率密度。

解: 随机变量 X 和 Y 之间存在唯一的反函数, 其表达式为 $X = h(Y) = Y - 5$, 则 $x = h(y) = y - 5$, 所以 $|h'(y)| = 1$。由单值变换公式可得

$$f_Y(y) = |h'(y)|\, f_X[h(y)] = 1 \cdot f_X(y-5) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-5)^2}{2}}. \quad (102)$$

可见, x 服从高斯分布, 其线性函数 $y = x + 5$ 也服从高斯分布.
思考: 若随机变量 X 服从高斯分布, 其线性函数 $Y = ax + b\ (a \neq 0)$ 是否服从高斯分布?

多值变换

随机变量 X 和 Y 存在非单调函数关系 $Y=g(X)$, 反函数 $X=h(Y)$ 不唯一, 如果 Y 值可能对应着两个 X 值, $X_1=h_1(Y)$ 和 $X_2=h_2(Y)$。

当 X 位于 (x_1, x_1+dx_1) 内或位于 (x_2, x_2+dx_2) 内时, 两事件中只要有一个发生, 则 Y 位于 (y_0, y_0+dy) 内的事件就发生。因此, 根据和事件概率的求法，可得

$$f_Y(y)dy = f_X(x_1)\,dx_1 + f_X(x_2)\,dx_2. \tag{103}$$

将 x_1 用 $h_1(y)$ 代入, x_2 用 $h_2(y)$ 代入, 可得

$$f_Y(y) = |h_1'(y)|\, f_X[h_1(y)] + |h_2'(y)|\, f_X[h_2(y)]\,. \tag{104}$$

更复杂的情况: 一个 Y 值对应多个值

此时, 将上式作进一步推广, 由概率可加性可得

$$f_Y(y)dy = f_X(x_1)\, dx_1 + f_X(x_2)\, dx_2 + f_X(x_3)\, dx_3 + \cdots \tag{105}$$

则

$$\begin{aligned} f_Y(y) = {} & |h_1'(y)|\, f_X[h_1(y)] + |h_2'(y)|\, f_X[h_2(y)] \\ & + |h_3'(y)|\, f_X[h_3(y)] + \cdots \end{aligned} \tag{106}$$

例 .4

已知随机变量 X 服从标准高斯分布，求随机变量 $Y = X^2$ 的概率。

解：随机变量 X 和 Y 之间的反函数关系为

$$X = h(Y) = \pm\sqrt{Y}. \tag{107}$$

其反函数导数的绝对值为

$$|h_1'(y)| = |h_2'(y)| = \left|\frac{dx}{dy}\right| = \frac{1}{2\sqrt{y}}. \tag{108}$$

① 当 $y<0$ 时，$\{X^2\leqslant y\}$ 为不可能事件，所以 $P\{X^2\leqslant y\}=0$，得 $F_Y(y)=0$，因此
② 当 $y<0$ 时，其概率密度 $f(y)=0$。
③ 当 $y>0$ 时，反函数为 $X=h(Y)=\pm\sqrt{Y}$ 是双值变换。已知变量 x 服从标准高斯分布，则

$$\begin{aligned} f_Y(y) &= |h_1'(y)|\, f_X[h_1(y)] + |h_2'(y)|\, f_X[h_2(y)] \\ &= \frac{1}{2\sqrt{y}}\left(\frac{1}{\sqrt{2\pi}}e^{-(-\sqrt{y})^2/2}+\frac{1}{\sqrt{2\pi}}e^{-(\sqrt{y})^2/2}\right) \\ &= \frac{1}{2\sqrt{y}}\cdot\frac{2}{\sqrt{2\pi}}e^{-x/2}=\frac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2}. \end{aligned} \tag{109}$$

综合 ① ② 可得

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, & y > 0 \\ 0, & y < 0 \end{cases} \tag{110}$$

称之为 χ^2 分布。说明一个高斯变量经过平方变换以后, 其概率密度为 χ^2 分布, 如图 **??** 所示。

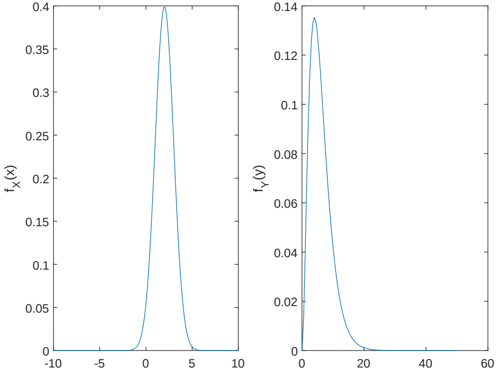


图 12: 输入、输出随机变量的概率密度: mu $= 2, \sigma = 1,$ n $= 6$

例 .5

设随机变量 $X \sim U[-1,2]$, 求随机变量函数 $Y = X^2$ 的概率密度.

解: X 的概率密度

$$f_X(x) = \begin{cases} \frac{1}{3}, & -1 \le x \le 2 \\ 0, & 其他 \end{cases},$$

由 $Y = X^2 \in [0,4]$, 用分布函数求 $Y = X^2$ 的概率密度 $f_Y(y)$.

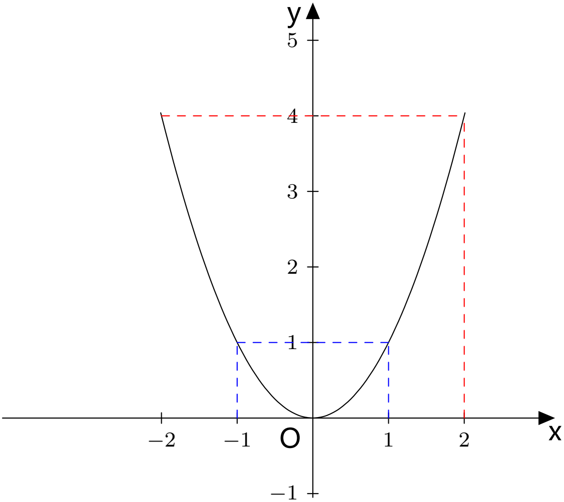


图 13: 随机变量函数关系 $Y = X^2$

1 当 $y < 0$, $F_Y(y) = P\{Y \le y\} = P\{\emptyset\} = 0$.

1. 当 $y < 0$, $F_Y(y) = P\{Y \leq y\} = P\{\emptyset\} = 0$.
2. 当 $y \geq 4$, $F_Y(y) = P\{Y \leq y\} = P\{\Omega\} = 1$.

1. 当 $y < 0$, $F_Y(y) = P\{Y \le y\} = P\{\emptyset\} = 0$.
2. 当 $y \ge 4$, $F_Y(y) = P\{Y \le y\} = P\{\Omega\} = 1$.
3. 当 $0 \le y < 4$, $x \in [-1, 2], 0 \le |x| \le 2$.

$$F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\}$$
$$= P\{-\sqrt{y} \le X \le \sqrt{y}\}$$

1 当 $y < 0$, $F_Y(y) = P\{Y \le y\} = P\{\emptyset\} = 0$.

2 当 $y \ge 4$, $F_Y(y) = P\{Y \le y\} = P\{\Omega\} = 1$.

3 当 $0 \le y < 4$, $x \in [-1, 2], 0 \le |x| \le 2$.

$$F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\}$$
$$= P\{-\sqrt{y} \le X \le \sqrt{y}\}$$

1 当 $0 \le y < 1$ 时, $0 \le \sqrt{y} < 1$, $-1 < -\sqrt{y} \le 0$, 且 $-1 \le x \le 2$.

$$F_Y(y) = P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{3} dx$$
$$= \frac{2}{3}\sqrt{y}, 0 \le y < 1.$$

1 当 $y<0, F_Y(y)=P\{Y\le y\}=P\{\emptyset\}=0.$

2 当 $y\ge 4, F_Y(y)=P\{Y\le y\}=P\{\Omega\}=1.$

3 当 $0\le y<4, x\in[-1,2], 0\le |x|\le 2.$

$$F_Y(y)=P\{Y\le y\}=P\{X^2\le y\}$$
$$=P\{-\sqrt{y}\le X\le \sqrt{y}\}$$

1 当 $0\le y<1$ 时, $0\le\sqrt{y}<1, -1<-\sqrt{y}\le 0$, 且 $-1\le x\le 2.$

$$F_Y(y)=P\{-\sqrt{y}\le X\le\sqrt{y}\}=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{3}dx$$
$$=\frac{2}{3}\sqrt{y}, 0\le y<1.$$

2 当 $1\le y<4$ 时, $1\le\sqrt{y}<2, -2\le-\sqrt{y}<-1$, 且 $-1\le x\le 2.$

$$F_Y(y)=P\{\{-\sqrt{y}\le X\le\sqrt{y}\}\cap\{-1\le X\le 2\}\}$$
$$=P\{-1\le X\le\sqrt{y}\}=\int_{-1}^{\sqrt{y}}\frac{1}{3}dx=\frac{\sqrt{y}+1}{3}.$$

分布函数

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{2\sqrt{y}}{3}, & 0 \le y < 1 \\ \frac{\sqrt{y}+1}{3}, & 1 \le y < 4 \\ 1, & y \ge 4 \end{cases},$$

对 y 求导, 得概率密度函数

$$f_Y(y) = \begin{cases} \frac{1}{3\sqrt{y}}, & 0 < y < 1 \\ \frac{1}{6\sqrt{y}}, & 1 \le y < 4 \\ 0, & \text{其他} \end{cases},$$

二维随机变量函数的分布

求解二维问题所采用的方法基本上和一维情况相似, 仅仅是稍微复杂一些。

已知二维随机变量 (X_1, X_2) 的联合概率密度为 $f_X(x_1, x_2)$, 要求新的二维随机变量 (Y_1, Y_2) 的联合概率密度 Y_1, Y_2, 其中 Y_1, Y_2 分别为 (X_1, X_2) 的函数, 则

$$\begin{cases} Y_1 = g_1(X_1, X_2). \\ Y_2 = g_2(X_1, X_2). \end{cases} \tag{111}$$

函数 $g_1(\cdot), g_2(\cdot)$ 可以是单值变换, 也可以是多值变换。

1. 单值变换

若解出的反函数

$$\begin{cases} X_1 = h_1(Y_1, Y_2) \\ X_2 = h_2(Y_1, Y_2) \end{cases} . \tag{112}$$

是唯一的, 则称二维随机变量 (X_1, X_2) 与 (Y_1, Y_2) 之间是单值的函数变换, 简称单值变换与一维随机变量类似, 单值变换是一一对应的变换。换句话说, 当随机点落入 x_1Ox_2 平面时, 在 y_1Oy_2 平面内有且仅有一个随机点与其对应, 反之亦然。

假设 $dS_{x_1x_2}$ 是 x_1Ox_2 平面内的一个任意闭域, $dS_{y_1y_2}$ 是它在 y_1Oy_2 平面中的映射, (X_1, X_2) 点落入 $dS_{x_1x_2}$ 区间的概率 $f_X(x_1, x_2)\,dS_{x_1x_2}$ 等于它的映射, 如图 14 所示, (Y_1, Y_2) 落入 $dS_{y_1y_2}$ 区间的概率 $f_Y(y_1, y_2)\,dS_{y_1y_2}$。

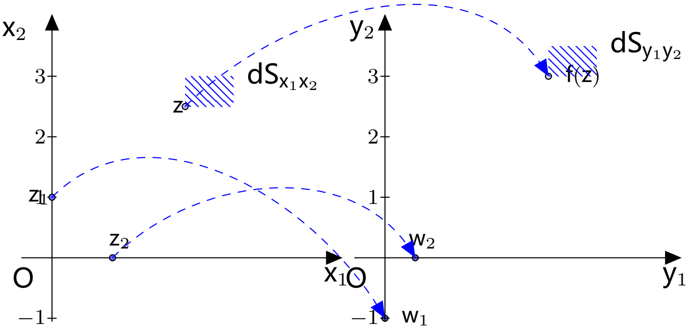


图 14: 单值映射下，函数变换对应的区间变换

雅可比变换

所以, 新的二维随机变量 (Y_1, Y_2) 的概率密度为

$$f_Y(y_1, y_2) = f_X(x_1, x_2) \cdot \left| \frac{dS_{x_1 x_2}}{dS_{y_1 y_2}} \right|. \tag{113}$$

坐标转换中 $dS_{x_1 x_2}$ 和 $dS_{y_1 y_2}$ 之间的变换称为雅可比变换。可得雅可比行列式为

$$J = \frac{dS_{x_1 x_2}}{dS_{y_1 y_z}} = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix}. \tag{114}$$

于是

$$\begin{aligned} f_Y(y_1, y_2) &= f_X(x_1, x_2) \left| \frac{dS_{x_1 x_2}}{dS_{y_1 x_2}} \right| = |J| f_X(x_1, x_2) \\ &= |J| f_X[h_1(y_1, y_2), h_2(y_1, y_2)]. \end{aligned} \tag{115}$$

例 .6

已知二维随机变量 (X_1, X_2) 具有联合概率密度

$$f_X(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)}, & x_1 > 0, x_2 > 0 \\ 0, & \text{其他} \end{cases} \tag{116}$$

新的二维随机变量 (Y_1, Y_2) 是 (X_1, X_2) 的函数, 满足关系

$$Y_1 = \frac{X_1 + X_2}{2}, \quad Y_2 = \frac{X_1 - X_2}{2}. \tag{117}$$

求: ① 二维随机变量 (Y_1, Y_2) 的联合概率密度 $f_{XY}(y_1, y_2)$. ② 边缘密度 $f_Y(y_1)$ 和 $f_Y(y_2)$, 说明 Y_1 与 Y_2 是否相互独立。

解: ① 由函数关系, 可以找出唯一的反函数

$$\begin{cases} x_1 = h_1(y_1, y_2) = y_1 + y_2 \\ x_2 = h_2(y_1, y_2) = y_1 - y_2 \end{cases} \tag{118}$$

则其雅可比行列式为

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2. \tag{119}$$

可得

$$
\begin{aligned}
f_Y(y_1, y_2) &= |J| f_X[h_1(y_1, y_2), h_2(y_1, y_2)] \\
&= 2 f_X[y_1 + y_2, y_1 - y_2] \\
&= \begin{cases} 2e^{-2y_1}, & y_1 > |y_2| \geqslant 0 \\ 0, & \text{其他} \end{cases}
\end{aligned} \tag{120}
$$

其中, 根据 (X_1, X_2) 与 (Y_1, Y_2) 的函数关系, 将 (X_1, X_2) 的值域映射到 y_1Oy_z 平面, 找出 (Y_1, Y_2) 的值域, 如图 15 所示。

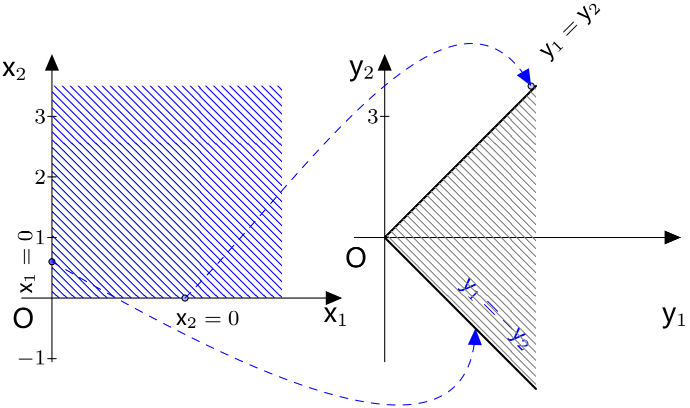


图 15: 例 6 的区间变换

关于 Y_1 边缘分布

(Y_1, Y_2) 的值域满足当 $y_1 > 0$ 时，

$$f_Y(y_1) = \int_{-y_1}^{y_1} f_{XY}(y_1, y_2)\, dy_2 = \int_{-y_1}^{y_1} 2e^{-2y_1} dy_2 = 4y_1 e^{-2y_1}. \tag{121}$$

关于 Y_2 边缘分布

当 $-\infty < y_2 < +\infty$ 时，

$$f_Y(y_2) = \begin{cases} \int_{y_2}^{\infty} 2e^{-2y_1} dy_1 = e^{-2y_2}, & y_2 \geqslant 0 \\ \int_{-y_2}^{\infty} 2e^{-2y_1} dy_1 = e^{2y_2}, & y_2 < 0 \end{cases} = e^{-2|y_2|}. \tag{122}$$

由于

$$
\begin{aligned}
f_Y(y_1) \cdot f_Y(y_2) &= 4y_1 e^{-2(y_1+|y_2|)} \\
&\neq f_Y(y_1, y_2) \\
&= \begin{cases} 2e^{-2y_1}, & y_1 > |y_2| \geqslant 0 \\ 0, & \text{其他} \end{cases},
\end{aligned}
\tag{123}
$$

所以 Y_1 与 Y_2 不是相互独立的。

多值变换

若从 $g_1(\cdot), g_2(\cdot)$ 中解出的 X_1 和 X_2 不是唯一的, 如解出两对反函数

$$\begin{cases} X_{a_1} = h_{a_1}(Y_1, Y_2) \\ X_{a_2} = h_{a_2}(Y_1, Y_2) \end{cases}, \begin{cases} X_{b_1} = h_{b_1}(Y_1, Y_2) \\ X_{b_2} = h_{b_2}(Y_1, Y_2) \end{cases}. \tag{124}$$

这种一对 (Y_1, Y_2) 有几对 (X_1, X_2) 与其对应的函数变换, 称之为多值变换。

二维随机变量的多值变换函数的概率密度

用概率的加法定理，求二维随机变量的多值变换函数的概率密度，即

$$f_Y(y_1,y_2) = |J_a| f_X[h_{a_1}(y_1,y_2), h_{a_2}(y_1,y_2)] + |J_b| f_X\left[h_{b_1}(y_1,y_2), h_{b_2}(y_1,y_2)\right], \tag{125}$$

其中，雅可比矩阵

$$J_a = \frac{\partial(x_{a_1}, x_{a_2})}{\partial(y_1,y_2)} = \begin{vmatrix} \frac{\partial h_{a_1}}{\partial y_1} & \frac{\partial h_{a_1}}{\partial y_2} \\ \frac{\partial h_{a_2}}{\partial y_1} & \frac{\partial h_{a_2}}{\partial y_2} \end{vmatrix}, \tag{126}$$

$$J_b = \frac{\partial(x_{b_1}, x_{b_2})}{\partial(y_1,y_2)} = \begin{vmatrix} \frac{\partial h_{b_1}}{\partial y_1} & \frac{\partial h_{b_1}}{\partial y_2} \\ \frac{\partial h_{b_2}}{\partial y_1} & \frac{\partial h_{b_2}}{\partial y_2} \end{vmatrix}. \tag{127}$$

n 维随机变量函数的分布

可用归纳法扩展到 n 维变换的情况。若 n 维随机变量 $(X_1, \cdots, X_n)$ 与其函数 $(Y_1, \cdots, Y_n)$ 间的变换是单值的，即有唯一反函数

$$\begin{cases} X_1 = h_1\left(Y_1, Y_2, \cdots, Y_n\right) \\ X_2 = h_2\left(Y_1, Y_2, \cdots, Y_n\right) \\ \vdots \\ X_n = h_n\left(Y_1, Y_2, \cdots, Y_n\right) \end{cases} . \tag{128}$$

则

$$
\begin{aligned}
f_Y(y_1, \cdots, y_n) &= |J| \cdot f_X(x_1, \cdots, x_n) \\
&= |J| \cdot f_X[h_1(y_1, \\
&\qquad \cdots, y_n, \cdots, h_n(y_1, \cdots, y_n).
\end{aligned} \tag{129}
$$

雅可比行列式为

$$
J = \frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)} = \begin{vmatrix} \frac{\partial h_1(y_1, \cdots, y_n)}{\partial y_1} & \cdots & \frac{\partial h_1(y_1, \cdots, y_n)}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial h_n(y_1, \cdots, y_n)}{\partial y_1} & \cdots & \frac{\partial h_n(y_1, \cdots, y_n)}{\partial y_n} \end{vmatrix} .. \tag{130}
$$

例 .7

已知 n 维随机变量 $(X_1,\cdots,X_n)$ 的联合概率密度 $f_X(x_1,\cdots,x_n)$，求随机变量 Y 的概率密度 $f_Y(y)$，其中 $Y=\sum_{i=1}^n X_i$.

解: 要想进行雅可比变换, 必须保证变换的维数相同, 因此必须构造新的 n 维随机变量 $(Y_1,\cdots,Y_n)$, 使满足

$$Y_1=X_1,Y_2=X_2,\cdots,Y_{n-1}=X_{n-1},Y_n=\sum_{i=1}^n X_i=Y. \tag{131}$$

解出反函数

$$X_1=Y_1,X_2=Y_2,\ldots,X_{n-1}=Y_{n-1},X_n=Y_n-\sum_{i=1}^{n-1} Y_i. \tag{132}$$

则 n 维雅可比行列式为

$$J = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{vmatrix} = 1. \tag{133}$$

所以 n 维随机变量 $(Y_1, \cdots, Y_n)$ 的联合概率密度为

$$f_Y(y_1, \cdots, y_a) = f_X\left(y_1, \cdots, y_{n-1}, y_n - \sum_{i=1}^{n-1} y_i\right). \tag{134}$$

其边缘分布即随机变量 Y 的概率密度 $f_Y(y)$ 为

$$f_Y(y) = f_Y(y_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X\left(y_1, \cdots, y_{n-1}, y_n - \sum_{i=1}^{n-1} y_i\right) dy_1 \cdots dy_{n-1}. \tag{135}$$

定理 .4

设二维连续型随机变量 (X,Y) 的概率密度为 $f(x,y)$, $Z=g(X,Y)$ 关于随机变量 X 或 Y 在区间 $I_i(i=1,2,\cdots,n)$ 上为严格单调函数, 且 $\frac{\partial Z}{\partial X}$ 或 $\frac{\partial Z}{\partial Y}$ 处处存在, 其在区间 I_i 上的反函数为 $y=y_i(x,z)$ 或 $x=x_i(y,z)$, 若 $I_i\,(i=1,2,\cdots,n)$ 互不相交且 $\bigcup_{i=1}^{n} I_i I_i=(-\infty,+\infty)$, 则随机变量 Z 的概率密度为

$$f_Z(z)=\sum_{i=1}^{n}\int_{-\infty}^{+\infty} f\left(x,y_i(x,z)\right)\left|\frac{\partial y_i}{\partial z}\right| dx. \tag{136}$$

或

$$f_Z(z)=\sum_{i=1}^{n}\int_{-\infty}^{+\infty} f\left(x_i(y,z),y\right)\left|\frac{\partial x_i}{\partial z}\right| dy. \tag{137}$$

证明 若 $Z=g(X,Y)$ 关于随机变量 Y 在区间 $I_i(i=1,2,\cdots,n)$ 上为严格单调函数, 则随机变量 (X,Z) 的联合分布为

$$
\begin{aligned}
F(x,z) =& P\{X\leqslant x, Z\leqslant z\} = P\left\{\bigcup_{i=1}(X\leqslant x, g(X,Y)\leqslant z, Y\in I_i)\right\} \\
&= \sum_{i=1}^{n} P\left\{X\leqslant x, g(X,Y)\leqslant z, Y\in I_i\right\} \\
&= \sum_{i=1}^{n} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f\left(x, y_i(x,z)\right) \left|\frac{\partial y_i}{\partial z}\right| dz \\
&= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} \left[\sum_{i=1}^{n} f\left(x, y_i(x,z)\right) \left|\frac{\partial y_i}{\partial z}\right|\right] dz.
\end{aligned} \tag{138}
$$

所以 (X, Z) 的联合概率密度为

$$\sum_{i=1}^{n} f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right|.$$

Z 的概率密度为

$$f_Z(z) = \sum_{i=1}^{n} \int_{-\infty}^{+\infty} f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right| dx. \tag{139}$$

同理可证当 $Z = g(X, Y)$ 关于随机变量 X 在区间 I_i $(i = 1, 2, \cdots, n)$ 上严格单调函数时,

$$f_Z(z) = \sum_{i=1}^{n} \int_{-\infty}^{+\infty} f(x_i(y, z), y) \left| \frac{\partial x_i}{\partial z} \right| dy.$$

例 .8

设随机变量 X 和 Y 相互独立，并且都服从正态分布 $N(0,\sigma^2)$，求随机变量 $Z=\sqrt{X^2+Y^2}$ 的概率密度。

解由于 X 和 Y 相互独立，因此 (X,Y) 的概率密度为

$$f(x,y)=f_X(x)f_Y(y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}},x,y\in\mathbb{R}. \tag{140}$$

由 $z=\sqrt{x^2+y^2}$，解得

$$\begin{cases} y_1=\sqrt{z^2-x^2}, & \frac{\partial y_1}{\partial z}=\frac{z}{\sqrt{z^2-x^2}}>0, & |x|\leqslant z \\ y_2=-\sqrt{z^2-x^2}, & \frac{\partial y_2}{\partial z}=-\frac{z}{\sqrt{z^2-x^2}}<0, & |x|\leqslant z \end{cases}, \tag{141}$$

$$f\left(x,\pm\sqrt{z^2-x^2}\right)=\begin{cases} \frac{z}{2\pi\sigma^2}e^{-\frac{j2}{2\sigma^2}}, & |x|<z \\ 0, & 其他 \end{cases}. \tag{142}$$

当 $z>0$ 时,

$$
\begin{aligned}
f_z(z) &= \int_{-z}^{z} f\left(x, \sqrt{z^2-x^2}\right) \frac{z}{\sqrt{z^2-x^2}} dx \\
&\quad + \int_{-z}^{2} f\left(x, -\sqrt{z^2-x^2}\right) \frac{z}{\sqrt{z^2-x^2}} dx \\
&= 4\int_0^z \frac{1}{2\pi\sigma^2} e^{-\frac{z2}{2\sigma^2}} \frac{z}{\sqrt{z^2-x^2}} dx = \frac{2z}{\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} \left[\arcsin\frac{x}{z}\right]_0^z \\
&= \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}.
\end{aligned}
\tag{143}
$$

显然, 当 $z \leqslant 0$ 时 , $f_Z(z) = 0$.
综上可得

$$
f_Z(z) = \begin{cases} \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}, & z \geqslant 0 \\ 0, & z < 0 \end{cases}, Z \sim \text{Rayleigh}(\sigma). \tag{144}
$$

多维随机变量函数经常用到的结论 1

(1) 若 $X_i \sim N(\mu_i, \sigma_i^2)$ $(i = 1, 2, \cdots, n)$, 则随机变量 $Z = \sum_{i=1}^{n} c_i X_i \sim N\left(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right)$, 其中 c_i $(i = 1, 2, \cdots, n)$ 是任意常数.

多维随机变量函数经常用到的结论 2

(2) 若 $X_i \sim N(0, \sigma^2)$ $(i = 1, 2, \cdots, n)$ 且相互独立，则 $Z = \sum_{i=1}^{n} X_i^2$ 服从自由度为 n 的中心 χ^2 分布, 密度函数为

$$f(z) = \begin{cases} \dfrac{1}{\sigma^n 2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}}, & z > 0 \\ 0, & z \leqslant 0 \end{cases} \tag{145}$$

注：$E\left(\frac{X_i^2}{\sigma^2}\right) = 1$, $E\left(\frac{\sum_{i=1}^{n} X_i^2}{\sigma^2}\right) = n$, $D\left(\frac{\sum_{i=1}^{n} X_i^2}{\sigma^2}\right) = 2n$.

卡方分布的说明

若 $X_i \sim e(\lambda), i = 1, 2, \cdots, n$, 且 X_i 独立, 则 $\sum_{i=1}^{n} e(\lambda) \sim \Gamma(n, \lambda)$. n 个独立同分布的指数变量之和为伽玛变量。用 $\chi^2(n) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ 来定义卡方分布。

非中心 χ^2 分布

若 $X_i \sim N\left(\mu, \sigma^2\right)$ $(i = 1, 2, \cdots, n)$ 且相互独立，则 $Z = \sum_{i=1}^{n} X_i^2$ 服从自由度为 n 的非中心 χ^2 分布。

$$f(z) = \frac{\sigma^2}{2} e^{-\frac{(x+\lambda)}{2}} \left(\frac{x}{\lambda}\right)^{\frac{n}{4}-1/2} I_{\frac{n}{2}-1}(\sqrt{\lambda x}), \tag{146}$$

式中，$\lambda = n(\mu/\sigma)^2$ 为非中心参数，$I_{n/2-1}(\cdot)$ 为 $(n/2-1)$ 阶第一类贝塞尔函数.

当 p 为正整数时, $\Gamma(p+1)=p!$, 当 p 为负整数或零时, $\Gamma(p)\to\infty$.

$$a_{2m}=\frac{(-1)^m}{2^{n+2m}m!\Gamma(n+m+1)}, n\geq 0. \tag{147}$$

n 阶第一类贝塞尔函数

n 阶第一类贝塞尔函数为

$$I_n(x)=\sum_{m=0}^{\infty}\frac{(-1)^m}{m!\Gamma(n+m+1)}\left(\frac{x}{2}\right)^{n+2m}, n\geq 0, \tag{148}$$

瑞利分布

瑞利分布

(3) 若随机变量 X 和 Y 相互独立，且都服从 $N(0,\sigma^2)$，则 $Z=\sqrt{X^2+Y^2}$ 服从参数为 σ 的瑞利分布

$$f(z)=\begin{cases}\frac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma}}, & z\geqslant 0\\ 0, & z<0\end{cases} \tag{149}$$