1

概率论

1.1 节主要是对概率论内容的概述；因此关于数学基础知识的阐述以基本概念和有用的结论为主，而命题的严格证明较少涉及，对此感兴趣的读者可以参阅有关参考文献。

1.1 概率论概要

1.1.1 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。例如，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。

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。例如，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 “l" 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 “l" 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，最后对常用的多维随机变量作简单的概述。

定义 1.1 基本事件

观察并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 (结果出现具有随机性)。

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，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，为了方便研究，对试验的有关结果给出如下概念。

定义 1.2 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了解和研究各种事件发生可能性的大小, 并加以度量。我们把刻画事件 A 发生可能性大小的数量指标称为事件 A 的概率，记为 $P(A)$.

下面给出计算 $P(A)$ 的三种主要方法。

定义 1.3 统计概率

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 $P(A) = p$. ♣

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 1.4 古典概率

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

定义 1.5 几何概率

对几何概型中的任一随机事件 A, 以

$$p(A)=\frac{m(A)}{m(\Omega)} \tag{1.1}$$

作为事件 A 的概率，称为几何概率。 ♣

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 1.6 古典概率

设随机试验 E 的样本空间为 Ω, 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 $P(A)\geq 0$;

(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 $P(\sum\limits_{i=1}^{\infty}A)=\sum\limits_{i=1}^{\infty}P(A)$. ♣

定义 1.7 概率空间

在概率论的公理化结构中，称三元组 (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 1.8 古典概率

设 A 和 B 为任意两个随机事件，且 $P(B)>0$, 称

$$P(A|B)=\frac{P(AB)}{P(B)}. \tag{1.2}$$

为事件 B 发生条件下事件 A 发生的条件概率，也称 A 对 B 的条件概率。♣

由条件概率定义 (1.2) 式可知

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

类似地

$$P(AB)=P(A)P(B|A), P(A)>0.$$

以上两式可以称为概率的乘法公式。乘法公式还可以推广到任意有限事件的情况：

$$\begin{aligned}P\left(\prod_{i=1}^{n} A_i\right) &= P\left(A_1A_2\cdots A_{n-1}\right)P\left(A_n|A_1A_2\cdots A_{n-1}\right)\\ &= P\left(A_1\right)P\left(A_2|A_1\right)P\left(A_3|A_1A_2\right)\cdots P\left(A_n|A_1A_2\cdots A_{n-1}\right).\end{aligned} \tag{1.3}$$

利用条件概率定义、乘法公式以及概率的可加性, 可以推导出两个十分有用的公式：全概率公式和贝叶斯公式。

定义 1.9 全概率公式

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 $P\left(A_i\right)>0\,(i=1,2,\cdots)$，则对任一事件 B, 有

$$P(B)=\sum P\left(A_i\right)P\left(B|A_i\right) \tag{1.4}$$

称为全概率公式。♣

定义 1.10 贝叶斯公式

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 $P\left(A_i\right)>0\,(i=1,2,\cdots)$，则对任一事件 $A\subset\Omega, B\subset\Omega, P(B)>0$, 有如下的条件概率：

$$P\left(A_m|B\right)=\frac{P\left(A_m\right)P\left(B|A_m\right)}{\sum P\left(A_i\right)P\left(B|A_i\right)}, \quad m=1,2,\cdots \tag{1.5}$$

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 1.11 事件独立

如果随机事件 A 与 B 满足如下关系

$$P(A|B) = p(B), \tag{1.6}$$

则称事件 A 与 B 是相互独立的，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

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

定义 1.12 试验的独立定义

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。 ♣

1.1.2 随机变量及其分布

在本部分，我们将引入随机变量的概念来表达随机事件。

1) 随机变量

定义 1.13 事件独立

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 $x \in \mathbb{R}$, "$X(\omega) \leq x$" 都是一个随机事件, 并有确定的概率，则称 $X = X(\omega)$ 为随机变量。 ♣

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。

由定义可见，随机变量的每一个取值都对应着随机试验样本空间 Ω 中的一个样本点 e. 这样，引入随机变量之后，就将随机试验的结果数量化，从而把对随机试验 E 以及其中随机事件 A 的研究转化为对随机变量以及其取值的研究。

随机变量可分为两类：离散型随机变量和连续型随机变量。

定义 1.14 分布律

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。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

$$P\left(X=x_i\right)=p_i, i=1,2,\cdots \tag{1.7}$$

我们称 (1.7) 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。

2) 分布函数

定义 1.15 分布函数

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

$$F(x)=P(X \leqslant x), \tag{1.8}$$

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 $[0,1]$. 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 $f(x)\,(-\infty<x<+\infty)$, 有

$$F(x)=\int_{-\infty}^{x} f(t) \mathrm{d} t, \tag{1.9}$$

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 1.16 0-1 分布

假设随机试验 E 只有两个可能结果 A 与 $\bar{A}$ 时，随机变量

$$X=\begin{cases}0 & \bar{A}\text{出现} \\ 1 & A\text{出现}\end{cases} \tag{1.10}$$

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

$$P(X=k)=p^k(1-p)^{1-k}, k=0,1 \tag{1.11}$$

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

定义 1.17 泊松分布

如果随机变量 X 的分布为

$$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \tag{1.12}$$

其中 $\lambda>0$ 为常数，则称 X 服从参数为 λ 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 1.18 均匀分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & \text{其他}\end{cases} \tag{1.13}$$

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 $X\sim U[a,b]$. ♣

定义 1.19 高斯分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right], \quad -\infty<x<+\infty \tag{1.14}$$

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 $X\sim N\left(m_X,\sigma_X^2\right)$. ♣

除此之外，还有如下几种分布：

定义 1.20 瑞利分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), & x\geqslant 0\\ 0, & x<0\end{cases} \tag{1.15}$$

其中 σ 为常数，且 $\sigma>0$，则称 X 服从参数为 σ 的瑞利分布，$X\sim \mathrm{Rayleigh}(\sigma)$。瑞利分布的密度函数如图 1–2. ♣

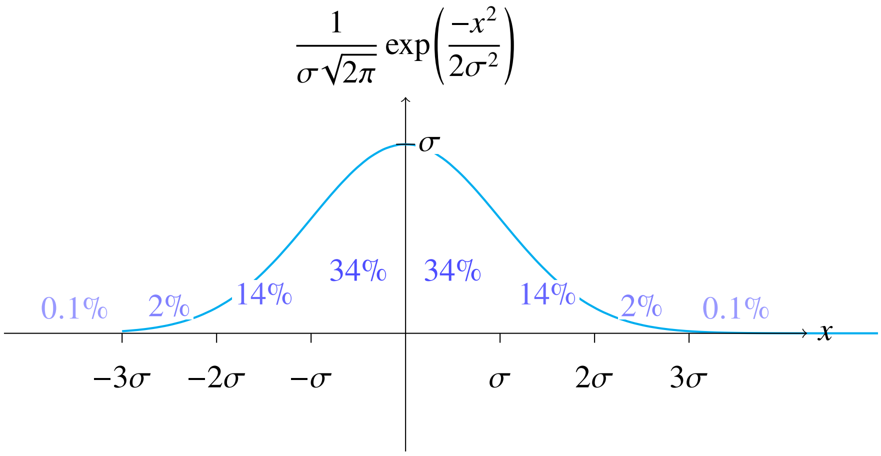

图 1–1 正态分布的分位点

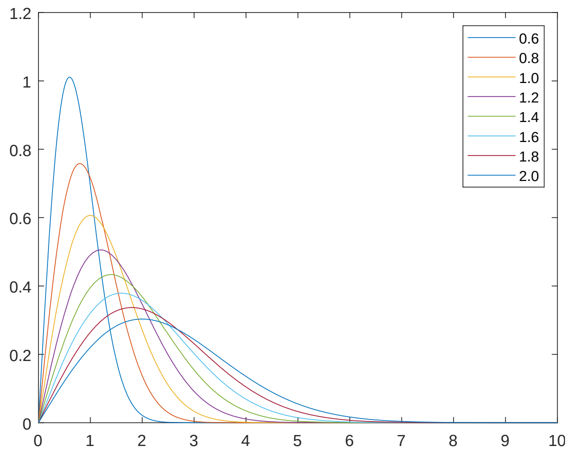

图 1–2 瑞利分布的密度函数

定义 1.21 卡方分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}\exp\left(-\frac{x}{2}\right), & x>0\\ 0, & x\leqslant 0\end{cases}\tag{1.16}$$

其中 n 为正整数，则称 X 服从自由度为 n 的 χ^2 分布，$X\sim\chi^2(n)$。♣

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

定义 1.22 莱斯分布

如果连续型随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{x}{\sigma^2}\exp\left[-\frac{(x^2+v^2)}{2\sigma^2}\right]\mathrm{I}_0\left(\frac{xv}{\sigma^2}\right), & x>0\\ 0, & x\leqslant 0\end{cases} \tag{1.17}$$

其中 σ 为常数且 $\sigma>0$, $\mathrm{I}_0(z)$ 是零阶第一类贝塞尔 (Bessel) 函数，则称 X 服从莱斯分布。♣

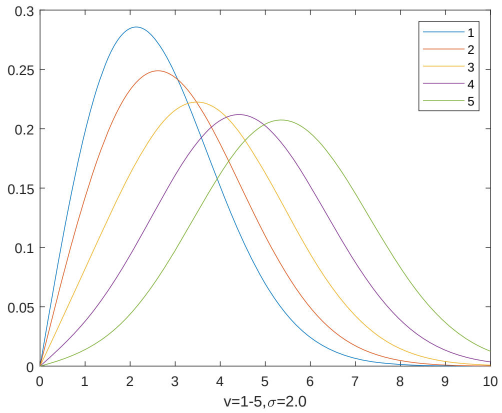

图 1–3 不同 v 下的莱斯分布的密度函数

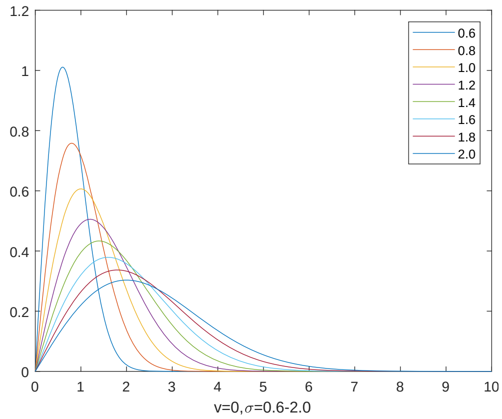

图 1–4 不同 σ 下的莱斯分布的密度函数

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, 均为右连续。

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

$$\begin{aligned}P\{a_1<X\leqslant a_2, b_1<Y\leqslant b_2\}=F_{XY}(a_2,b_2)+F_{MY}(a_1,b_1)-F_{XY}(a_1,b_2)\\ -F_{XY}(a_2,b_1).\end{aligned} \tag{1.18}$$

如图 1–5 所示。

3) 概率密度的定义

定义 1.23 联合概率密度

若 $F_{XY}(x,y)$ 存在二阶偏导数, 则称

$$f_{XY}(x,y)=\frac{\partial F_{XY}(x,y)}{\partial x\partial y} \tag{1.19}$$

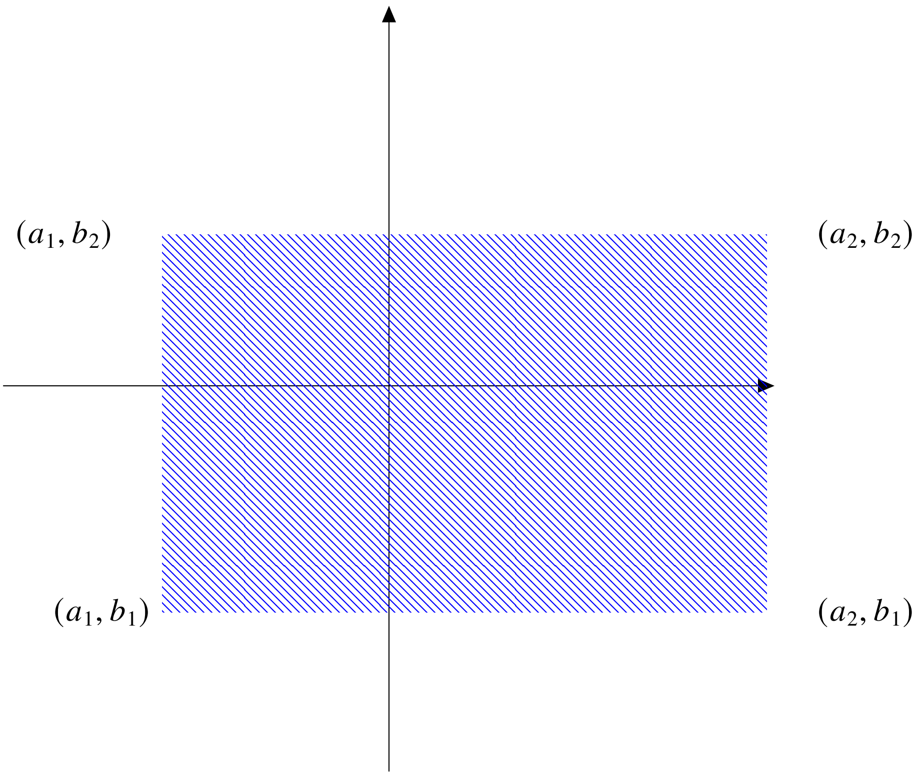

图 1–5 联合分布函数

为二维随机变量 (X, Y) 的联合概率密度。♣

联合概率密度具有以下基本性质:

1° $f_{XY}(x, y) \geqslant 0$.

2° $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x, y)\mathrm{d}x\mathrm{d}y = 1$.

3° $\int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u, v)\mathrm{d}u\mathrm{d}v = F_{XY}(x, y)$.

4° $P\{(x, y) \in D\} = \iint f_{XY}(u, v)\mathrm{d}u\mathrm{d}v$.

在几何上, $P(x, y) \in D$ 表示曲面 $f_x(x, y)$ 与 D 所围的柱体体积, 如图 1–6 所示。

例 1.24 设二维随机变量 (X, Y) 的概率密度

$$f(x, y) = \begin{cases} \mathrm{e}^{-x-y}, & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \tag{1.20}$$

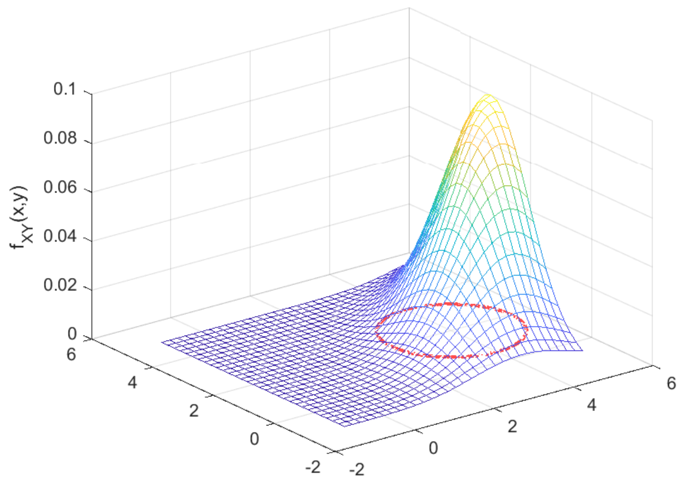

图 1–6 矩形区域上的联合分布函数值

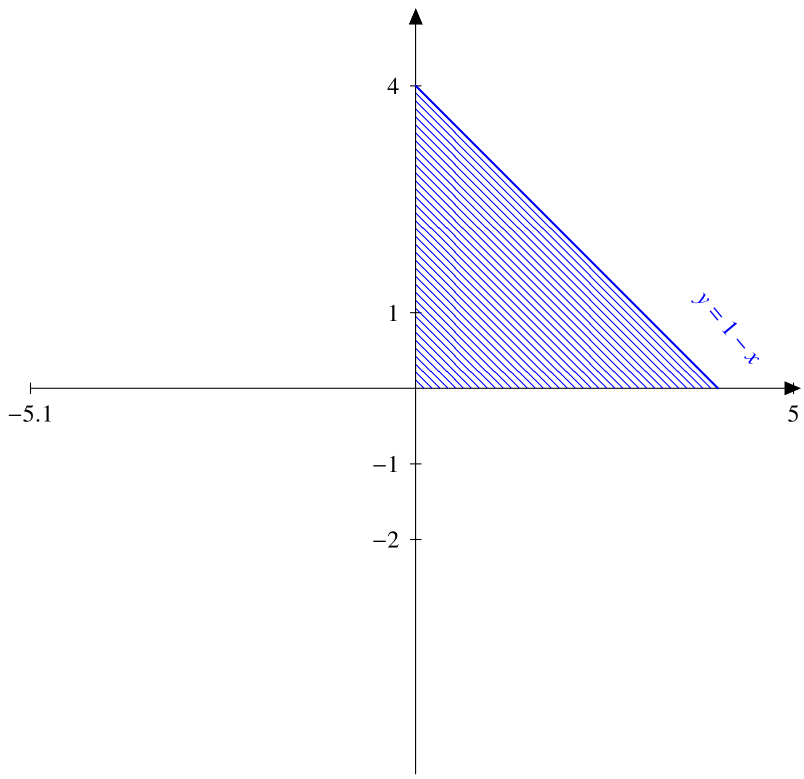

图 1–7 三角积分区域 $G : x > 0, y > 0, x + y < 1$

求: ① 求分布函数 $F_X(x,y)$。② (X,Y) 落在如图 1–7 所示的三角形域 G 内的概率.

解: ① 分布函数

$$
\begin{aligned}
F_{XY}(x,y) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\mathrm{d}u\mathrm{d}v \\
&= \left\{ \int_0^y \int_0^x f(u,v)\mathrm{d}u\mathrm{d}v, \quad 0<x<+\infty, 0<y<+\infty 0, \quad \text{其他} \right. \\
&= \begin{cases} (1-\mathrm{e}^{-\alpha})(1-\mathrm{e}^{-y}), & 0<x<+\infty, 0<y<+\infty \\ 0, & \text{其他} \end{cases}
\end{aligned} \tag{1.21}
$$

② (X,Y) 落在三角形区域 G 内的概率

$$
\begin{aligned}
P\{(x,y)\in G\} &= \iint_{\sigma} f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^1 \int_0^{1-y} \mathrm{e}^{-(x+y)}\mathrm{d}x\mathrm{d}y \\
&= \int_0^1 \mathrm{e}^{-y}\left[\int_0^{1-y}\mathrm{e}^{-x}\mathrm{d}x\right]\mathrm{d}y = \int_0^1 \mathrm{e}^{-y}\cdot\left(1-\mathrm{e}^{-1+y}\right)\mathrm{d}y \\
&= \int_0^1 \left(\mathrm{e}^{-y}-\mathrm{e}^{-1}\right)\mathrm{d}y = 1-2\mathrm{e}^{-1} = 0.2642.
\end{aligned} \tag{1.22}
$$

例 1.1.1

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

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

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

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 $0 \le y < 4$, $x \in [-1,2], 0 \le |x| \le 2$, $F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\}$.
   (a) 当 $0 \le y < 1$ 时, $0 \le \sqrt{y} < 1$, $-1 < -\sqrt{y} \le 0$, 且 $-1 \le x \le 2$.

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

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

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

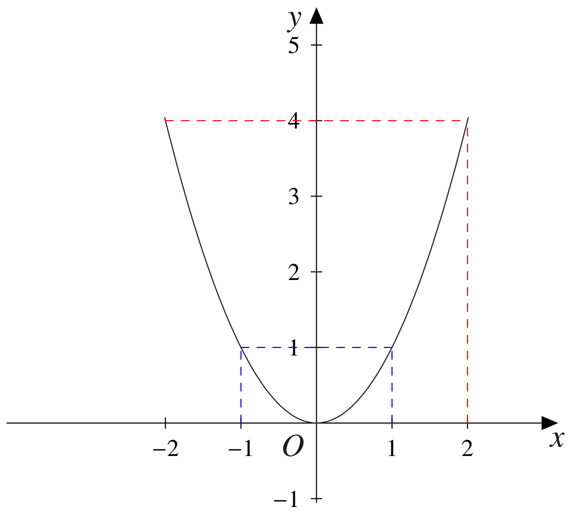

图 1–8 随机变量函数关系 $Y = X^2$

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 Y 的概率密度函数

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

(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) 的所有可能取值为 (x_i, y_j) $(i, j = 1, 2, \cdots)$, 其上的概率记为

$$P\{X = x_i, Y = y,\} = p_{ij}, i, j = 1, 2, \cdots \tag{1.23}$$

根据概率的性质，有

1° $p_{ij} \geqslant 0$,

2° $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = 1$,

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

$$\begin{aligned} F_{XY}(x,y) &= P\{X \leqslant x, Y \leqslant y\} \\ &= \sum_i \sum_j P\left\{X = x_i, Y = y_j\right\} U\left(x - x_i\right) U\left(y - y_j\right) \\ &= \sum_i \sum_j p_{ij} U\left(x - x_i\right) U\left(y - y_j\right). \end{aligned} \tag{1.24}$$

离散型二维随机变量的联合概率密度可表示为

$$\begin{aligned} P_{ij} &= \sum_i \sum_j P\left\{X = x_i, Y = y_j\right\} \delta\left(x - x_i\right) \delta\left(y - y_j\right) \\ &= \sum_i \sum_j p_{ij} \delta\left(x - x_i\right) \delta\left(y - y_j\right). \end{aligned} \tag{1.25}$$

2. 二维随机变量的边缘分布和条件分布

(1) 边缘分布函数和边缘概率密度

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

$$F_X(x) = F_X(x,\infty), \quad F_Y(y) = F_{XY}(\infty, y), \tag{1.26}$$

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。

例 1.1.2

如 1–9 所示，试着解释哪张图可以推出联系分布律？

① 对于连续型随机变量 (X,Y), 有 (X,Y) 对 X 的

$$F_X(x) = F_{XY}(x,\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{XY}(u,y)\mathrm{d}y\mathrm{d}u. \tag{1.27}$$

对 $F_X(x) = F_{XY}(x,\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{XY}(u,y)\mathrm{d}y\mathrm{d}u$ 求导, 得

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\mathrm{d}y, \tag{1.28}$$

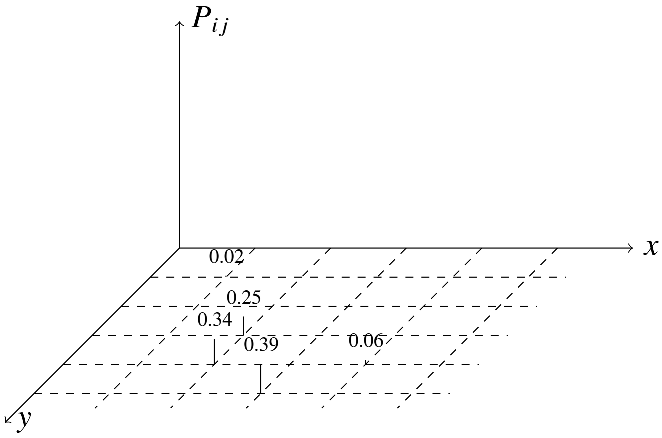

图 1–9 离散型二维随机变量的分布律

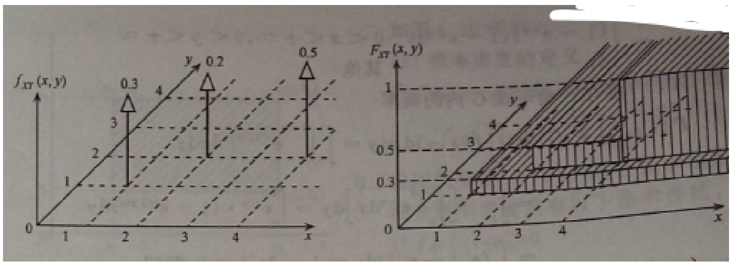

图 1–10 离散型二维随机变量的概率密度 $f_X(x)$ 和分布函数 $F_Y(y)$

称之为 (X,Y) 关于 X 的边缘概率密度。

① 同理, (X,Y) 关于 Y 的边缘概率密度为

$$f_Y(y)=\int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{d}x. \tag{1.29}$$

对于离散型随机变量 (X,Y), 有

$$F_X(x)=F_{XY}(x,\infty)=\sum_i\sum_{j=1}^{\infty}p_{ij}U\left(x-x_i\right)=\sum_{x_i\leqslant x}\sum_{j=1}^{\infty}p_{ij}. \tag{1.30}$$

可得

$$p_{i\cdot}=P\left\{X=x_i\right\}=\sum_{j=1}^{\infty}p_{ij}, i=1,2,\ldots \tag{1.31}$$

称之为 (X,Y) 关于 X 的边缘分布律。同理, (X,Y) 关于 Y 的边缘分布律为

$$p_{\cdot j}=P\left\{Y=y_j\right\}=\sum_{i=1}^{\infty}p_{ij}, j=1,2,\ldots \tag{1.32}$$

边缘分布函数、边缘概率密度和边缘分布律, 反映了二维随机变量中各随机变量本身的统计特征。

例 1.25 (续上例) 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

$$\begin{aligned} F_X(x) = F_X(x,\infty) &= \begin{cases} \left(1-\mathrm{e}^{-x}\right)\left(1-\mathrm{e}^{-\infty}\right), & 0<x<+\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} 1-\mathrm{e}^{-x}, & 0<x<+\infty \\ 0, & \text{其他} \end{cases} \end{aligned} \tag{1.33}$$

同理

$$F_Y(y) = F_{XY}(\infty,y) = \begin{cases} 1-\mathrm{e}^{-y}, & 0<y<+\infty \\ 0, & \text{其他} \end{cases} \tag{1.34}$$

边缘概率密度的计算

$$\begin{aligned} f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\mathrm{d}y &= \begin{cases} \int_0^{+\infty} \mathrm{e}^{-(x+y)}\mathrm{d}y, & 0<x<+\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} \mathrm{e}^{-x}, & 0<x<+\infty \\ 0, & \text{其他} \end{cases} \end{aligned} \tag{1.35}$$

同理

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\mathrm{d}x = \begin{cases} \mathrm{e}^{-y}, & 0<y<+\infty \\ 0, & \text{其他} \end{cases}. \tag{1.36}$$

1) 条件概率的定义

定义 1.26 条件分布函数和条件概率密度

1 对于连续型二维随机变量 (X,Y) 有

$$\begin{aligned} F_Y(y|X=x) &= P\{Y\le y|X\le x\} = P(A|B) = \int_{-\infty}^{y} \frac{f_{XY}(x,y)}{f_X(x)}\mathrm{d}v, \\ f_Y(y|X=x) &= \frac{f_{XY}(x,y)}{f_X(x)}; \end{aligned} \tag{1.37}$$

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。

$$P(A|B)=\frac{P(AB)}{P(B)}. \tag{1.38}$$

把这个概念引入到随机变量的理论中。

1) 先推导 $X\leqslant x$ 条件下的分布函数和分布律:

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 $P(A|B)=P\{Y\leqslant y|B\}=F_Y(y|B)$ 为给定条件 B 下 Y 的分布函数。

若上述讨论，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

$$F_Y(y|X\leqslant x)=P(A|B)=\frac{P(AB)}{P(B)}=\frac{P\{X\leqslant x,Y\leqslant y\}}{P\{X\leqslant x\}}=\frac{F_{XY}(x,y)}{F_X(x)}. \tag{1.39}$$

若上式对 y 的导数存在, 则有

$$f_Y(y|X\leqslant x)=\frac{\partial F_Y(y|X\leqslant x)}{\partial y}=\frac{\partial F_{XY}(x,y)/\partial y}{F_X(x)}=\frac{\int_{-\infty}^{x}f_{XY}(u,y)\mathrm{d}u}{\int_{-\infty}^{x}f_X(u)\mathrm{d}u}. \tag{1.40}$$

2) 先推导 $X=x$ 条件下的分布函数:

若令 $B=\{X=x\}$, 代入式 (1.39)，得

$$\begin{aligned}F_Y(y|X=x)&=\lim_{\Delta x\to 0}F_Y(y|x<X\leq x+\Delta x)=\lim_{\Delta x\to 0}\frac{P\{x<X\leqslant x+\Delta x,Y\leqslant y\}}{P\{x<X\leqslant x+\Delta x\}}\\&=\lim_{\Delta x\to 0}\frac{P\{X\leqslant x+\Delta x,Y\leqslant y\}-P\{X\leqslant x,Y\leqslant y\}}{P\{X\leqslant x+\Delta x\}-P\{X\leqslant x\}}\\&=\lim_{\Delta x\to 0}\frac{[F_{XY}(x+\Delta x,y)-F_{XY}(x,y)]}{[F_X(x+\Delta x)-F_X(x)]}\\&=\lim_{\Delta x\to 0}\frac{[F_{XY}(x+\Delta x,y)-F_{XY}(x,y)]/\Delta x}{[F_X(x+\Delta x)-F_X(x)]/\Delta x}=\frac{\partial F_{XY}(x,y)/dx}{\partial F_X(x)/\partial x}\\&=\frac{\int_{-\infty}^{y}f_{XY}(x,y)\mathrm{d}y}{f_X(x)}.\end{aligned} \tag{1.41}$$

得

$$f_Y(y|X=x)=\frac{\partial}{\partial y}F_Y(y|X=x)=\frac{f_{XY}(x,y)}{f_X(x)}, f_X(x)\neq 0. \tag{1.42}$$

例 1.27 (续上例) 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

$$\begin{aligned}F_X(x|y)&=\int_0^x\frac{\int_{-\infty}^{x}f_X(x,y)\mathrm{d}x}{f_Y(y)}dx=\begin{cases}\int_0^x\frac{\mathrm{e}^{-(x+y)}}{\mathrm{e}^{-y}\mathrm{d}x} & 0<x<+\infty\\ 0, & \text{其他}\end{cases}\\&=\begin{cases}1-\mathrm{e}^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases}.\end{aligned} \tag{1.43}$$

同理

$$F_Y(y|x)=\begin{cases}1-\mathrm{e}^{-y}, & 0<y<+\infty\\ 0, & \text{其他}\end{cases}. \tag{1.44}$$

⑥ 条件概率密度

$$f_X(x|y)=\frac{f_X(x,y)}{f_Y(y)}=\begin{cases}\frac{e^{-(x+y)}}{e^{-y}}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases}=\begin{cases}\mathrm{e}^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases} \tag{1.45}$$

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 $\{Y(\zeta)\leqslant y\}$ 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的性质如下:

条件分布函数性质如下:

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

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

性质 条件密度函数的性质如下：

1° $f_Y(y|B)\geqslant 0$.

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

3° $F_Y(y|B)=\int^{y}f_Y(y|B)\mathrm{d}y$.

4) 离散型随机变量的条件分布律

定义 1.28 条件分布律

1 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

$$P\{Y=y|X=x\}=\frac{P\{X=x,Y=y\}}{P\{X=x\}},\quad P(X=x)>0. \tag{1.46}$$

♣

① 二维随机变量 (X,Y), 对于固定的 j, 若 $P\left\{Y=y_j\right\}>0$, 则

$$P\left\{X=x_i|Y=y_j\right\}=\frac{P\left\{X=x_i,Y=y_i\right\}}{P\left\{Y=y_j\right\}}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\cdots, \tag{1.47}$$

称为在 $Y=y$ 条件下随机变量 X 的条件分布律。

② 对于固定的 i, 若 $P\{X=x\}>0$, 则

$$P\left\{Y=y_j|X=x_i\right\}=\frac{P\left(X=x_i,Y=y_j\right)}{P\left\{X=x_i\right\}}=\frac{p_{ij}}{p_{i\cdot}},\quad j=1,2,\cdots, \tag{1.48}$$

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

现在把事件独立的概念引入到随机变量中来。

定义 1.29 相互独立

X, Y 是两个随机变量, 若对任意实数 x 和 y, 有

$$P\{X < x, Y < y\} = P\{(X < x) \cap (Y < y)\} = P\{X < x\} \cdot P\{Y < y\}, \tag{1.49}$$

则称随机变量 X, Y 相互独立。♣

二维随机变量 (X, Y), X 与 Y 相互独立的条件为

$$F_{XY}(x, y) = F_X(x) \cdot F_Y(y). \tag{1.50}$$

或

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y). \tag{1.51}$$

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

$$\begin{cases} F_Y(y|x) = F_Y(y) \\ f_Y(y|x) = f_Y(y) \end{cases} \tag{1.52}$$

同理可得

$$\begin{cases} F_X(x|y) = F_X(x) \\ f_X(x|y) = f_X(x) \end{cases} \tag{1.53}$$

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

$$P\left\{X = x_i, Y = y_j\right\} = P\left\{X = x_i\right\} \cdot P\left\{Y = y_j\right\}. \tag{1.54}$$

例 1.30 (续上例) ⑦ X 和 Y 是否统计独立?

解: ⑦ 独立。因为存在下面的关系 (只需满足 a, b, c, d 任一条件即可推出统计独立)

a. 由已知条件和求得的结论可知

$$F_{XY}(x, y) = \begin{cases} \left(1 - \mathrm{e}^{-x}\right)\left(1 - \mathrm{e}^{-y}\right), & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases}. \tag{1.55}$$

$$F_X(x)=\begin{cases}1-\mathrm{e}^{-x}, & 0<x<+\infty,\\ 0, & \text{其他}\end{cases},F_Y(y)=\begin{cases}1-\mathrm{e}^{-y}, & 0<y<+\infty\\ 0, & \text{其他}\end{cases} \tag{1.56}$$

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

b. 由已知条件和求得的结论可知

$$F_X(x|y)=\begin{cases}1-\mathrm{e}^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases}, \quad F_X(x)=\begin{cases}1-\mathrm{e}^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases} \tag{1.57}$$

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

c. 由已知条件和求得的结论可知

$$f_{XY}(x,y)=\begin{cases}\mathrm{e}^{-(x+y)}, & 0<x<+\infty,0<y<+\infty\\ 0, & \text{其他}\end{cases} \tag{1.58}$$

$$f_X(x)=\begin{cases}e^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases},f_Y(y)=\begin{cases}\mathrm{e}^{-y}, & 0<y<+\infty\\ 0, & \text{其他}\end{cases} \tag{1.59}$$

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

d. 由已知条件和求得的结论可知

$$f_X(x|y)=\begin{cases}\mathrm{e}^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases},f_X(x)=\begin{cases}\mathrm{e}^{-x}, & 0<x<+\infty\\ 0, & \text{其他}\end{cases} \tag{1.60}$$

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

1.1.3 多维随机变量

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。为此，我们需要引入多维随机量 (也可称为随机矢量) 的概念。

定义 1.31 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 $\omega\in\Omega$, $X_i=X_i(\omega),i=1,2,\cdots$ 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 $[X_1,X_2,\cdots,X_n]^{\mathrm{T}}$ 也称为随机矢量。 ♣

本节我们简要回忆二维随机变量的情况，它的很多结果都可以推广到多维随机变量上。

定义 1.32 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

$$F(x,y)=P(X\leqslant x,Y\leqslant y) \tag{1.61}$$

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。♣

除了分布函数，我们还将简单介绍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。它与二维变量的分布函数具有如下关系：

$$F_X(x)=P(X\leqslant x,Y<+\infty)=F(x,+\infty). \tag{1.62}$$

$$F_Y(y)=P(X<+\infty,Y\leqslant y)=F(+\infty,y). \tag{1.63}$$

同一维随机变量类似，二维随机变量也可以分为离散型和连续型两种形式。

定义 1.33 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 1.34 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

$$P\left(X=x_i,Y=y_j\right)=p_{ij},i,j=1,2,\cdots \tag{1.64}$$

称为 (X,Y) 的联合概率分布。♣

定义 1.35 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

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v, \tag{1.65}$$

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。♣

利用二维随机变量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 1.36 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

$$F(x,y) = F_X(x)F_Y(y), \tag{1.66}$$

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

$$P\left(X = x_i | Y = y_j\right) = \frac{P\left(X = x_i, Y = y_j\right)}{P\left(Y = y_j\right)}. \tag{1.67}$$

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

由于连续型随机变最取任何单点数值的概率都是零，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。因此，考虑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 $P(X \leqslant x | Y = y)$.

定义 1.37 m 维边缘分布函数

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

$$\begin{aligned} &F_X\left(x_1, x_2, \cdots, x_m\right) = F_X\left(x_1, x_2, \cdots, x_m, \infty, \cdots, \infty\right). \\ &F_X\left(x_i\right) = F_X\left(\infty, \cdots, \infty, x_i, \infty, \cdots, \infty\right). \end{aligned} \tag{1.68}$$

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

$$f_X(x_1,x_2,\cdots,x_m)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f_X(x_1,\cdots,x_m,x_{m+1},\cdots,x_n)\,dx_{m+1}dx_{m+2}\cdots dx_n. \tag{1.69}$$

$$f_X(x_i)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f_X(x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_n)\,dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n. \tag{1.70}$$

4. 条件概率密度

定义 1.38 m 维边缘分布函数

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

$$f_X(x_2,\cdots,x_n|x_1)=\frac{f_X(x_1,x_2,\cdots,x_n)}{f_X(x_1)}. \tag{1.71}$$

♣

定义 1.39 m 维边缘分布函数

也可用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

$$f_X(x_3,\cdots,x_n|x_1,x_2)=\frac{f_X(x_1,x_2,\cdots,x_n)}{f_X(x_1,x_2)} \tag{1.72}$$

♣

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

$$f_X(x_n|x_1,x_2,\cdots,x_{n-1})=\frac{f_X(x_1,x_2,\cdots,x_n)}{f_X(x_1,x_2,\cdots,x_{n-1})}. \tag{1.73}$$

利用上述条件概率密度的定义, 可得 n 维随机变量联合概率密度的递推公式:

$$\begin{aligned}f_X(x_1,\cdots,x_n)&=f_X(x_n|x_1,\cdots,x_{n-1})\,f_X(x_{n-1}|x_1,\cdots,x_{n-2})\\&\qquad\cdots f_X(x_2|x_1)\,f_X(x_1)\,.\end{aligned} \tag{1.74}$$

证明. 归纳证明过程如下

① 二维情况 $f_X(x_1,x_2)=f_X(x_2|x_1)\,f_X(x_1)$.

② 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为

$$f_X(x_1,x_2,x_3)=f_X(x_3|x_1,x_2)\,f_X(x_2|x_1)\,f_X(x_1)\,. \tag{1.75}$$

③ 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

$$f_X(x_1,x_2,x_3)=f_X(x_3|x_1,x_2)\,f_X(x_2|x_1)\,f_X(x_1)\,. \tag{1.76}$$

当 $X_1,X_2,\cdots,X_n$ 相互独立时, 可进一步得到

$$f_X(x_1,\cdots,x_n)=f_X(x_n)\,f_X(x_{n-1})\cdots f_X(x_2)\,f_X(x_1)\,. \tag{1.77}$$

□

例 1.40 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

$$f_X(x_i)=\begin{cases}1, & 0<x_i<1\\ 0, & 其他\end{cases} \tag{1.78}$$

且随机变量 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

$$\begin{aligned}f_X(x_1,x_2,x_3,x_4)&=f_X(x_1)\cdot f_X(x_2)\cdot f_X(x_3)\cdot f_X(x_4)\\&=\begin{cases}1, & 0<x_1,x_2,x_3,x_4<1\\ 0, & 其他\end{cases}\end{aligned} \tag{1.79}$$

② 同理可知 X_1,X_2 的联合概率密度为

$$f_X(x_1,x_2)=f_X(x_1)\cdot f_X(x_2)=\begin{cases}1, & 0<x_1,x_2<1\\ 0, & 其他\end{cases} \tag{1.80}$$

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

$$f_X(x_3|x_1,x_2)=f_X(x_3)=\begin{cases}1, & 0<x_3<1\\ 0, & 其他\end{cases} \tag{1.81}$$

$$f_X(x_3,x_4|x_1,x_2)=f_X(x_3,x_4)=f_X(x_3)\cdot f_X(x_4)=\begin{cases}1, & 0<x_3,x_4<1\\ 0, & 其他\end{cases} \tag{1.82}$$

1.1.4 随机变量函数的分布

1. 一维随机变量函数的分布

定理 1.41 一维随机变量

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

$$f(y) = \begin{cases} f(h(y))\,|h'(y)|, & a < y < \beta \\ 0, & 其他 \end{cases} \tag{1.83}$$

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

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

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

3. 当 $\alpha \le y < \beta$, $F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\} = P\{X \le h(y)\} = \int_{-\infty}^{h(y)} f_X(x)dx$.

概率密度函数

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

对于 $g^{'}(x) < 0$ 的情况,

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

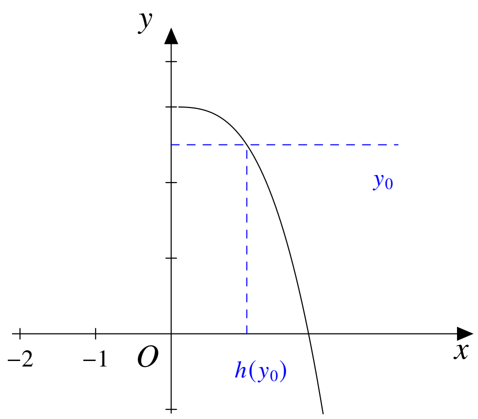

图 1–11 单调递减函数

对于 $g'(x) > 0$ 的情况,

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

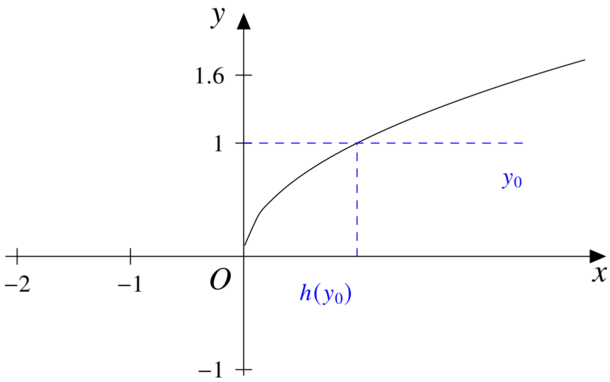

图 1–12 单调递增函数

合并后,得密度函数

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

□

例 1.1.3

设 X 的概率密度为

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

求 $Y = \sin X$ 的概率密度.

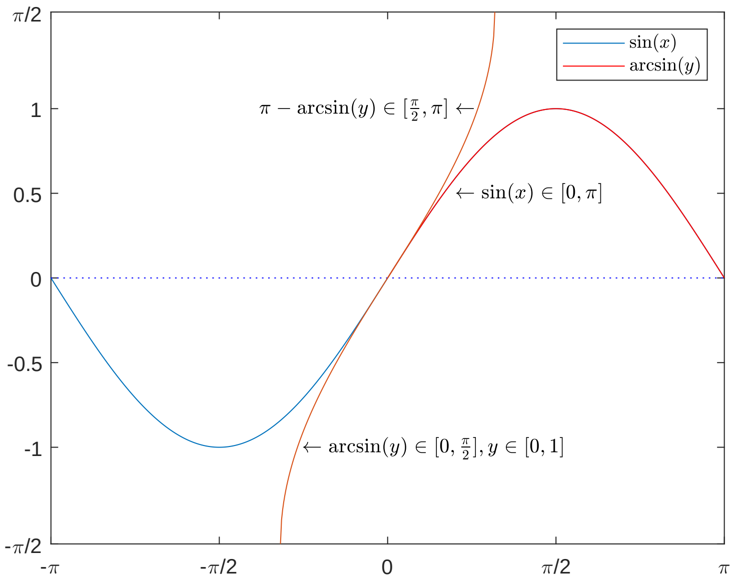

图 1–13 $\sin(x)$ 及其反函数的取值区间.

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1–13 或者 1–14).

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

- 当 $1 < y$ 时: $F_Y(y) = 1$.

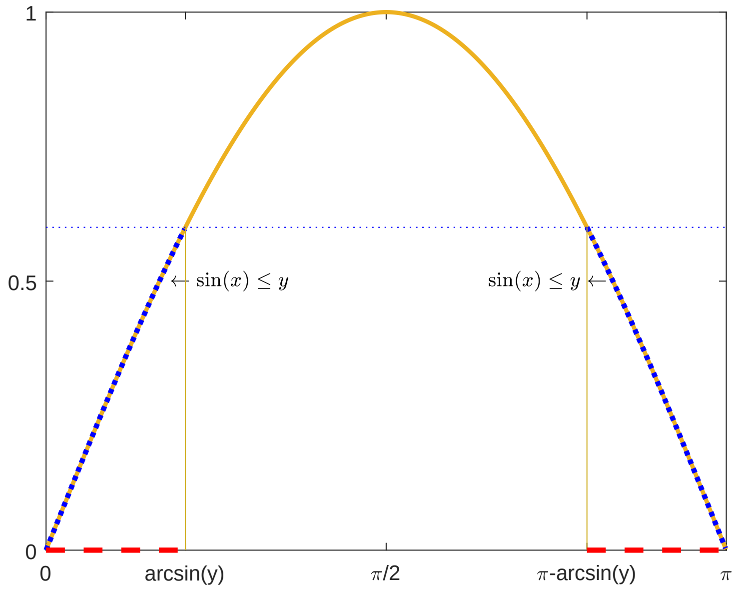

图 1–14 $\sin(x) \leq y$ 的取值范围.

Y 的概率密度 $f(y)$ 为:

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

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

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

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

2. 多维随机变量函数的分布

定理 1.42 多维随机变量函数的分布

$(X_1, X_2, \cdots, X_n)$ 具有概率密度函数 $f(x_1, x_2, \cdots, x_n)$ 的连续型 n 维随机变量,
(1) $y_1 = g_1(x_1, x_2, \cdots, x_n), \cdots, y_n = g_n(x_1, x_2, \cdots, x_n)$ 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:

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

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

$$f(y_1, y_2, \cdots, y_n) = f(h_1(y_1, y_2, \cdots, y_n), \cdots, h_n(y_1, y_2, \cdots, y_n))\,|J|. \tag{1.84}$$

1.1.4.1 随机变量函数的分布

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。例如电子系统中, 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。

$$Y = \begin{cases} X^{1/n}, & X \geqslant 0 \\ -|X|^{1/n}, & X < 0 \end{cases}, n \in \mathbb{Z}^+ \tag{1.85}$$

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。

1.1.4.2 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

$$P\{x_0 < X \leqslant x_0 + \mathrm{d}x\} = P\{y_0 < Y \leqslant y_0 + \mathrm{d}y\}. \tag{1.86}$$

可得

$$f_Y(y)\mathrm{d}y = f_X(x)\mathrm{d}x. \tag{1.87}$$

所以

$$f_Y(y) = f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y} = [h'(y)] \cdot f_X[h(y)]. \tag{1.88}$$

由于概率密度不可能取负值, 所以 dx/dy 应取绝对值, 即

$$f_Y(y) = |h'(y)| \cdot f_X[h(y)]. \tag{1.89}$$

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 1.43 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

$$f_Y(y) = |h'(y)|\, f_X[h(y)] = 1 \cdot f_X(y-5) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-5)^2}{2}}; \tag{1.90}$$

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?

2. 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 $X_2 = h_2(Y)$。所以, 当 X 位于 $(x_1, x_1 + dx_1)$ 内或位于 $(x_2, x_2 + \mathrm{d}x_2)$ 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

$$f_Y(y)dy = f_X(x_1)\,dx_1 + f_X(x_2)\,dx_2. \tag{1.91}$$

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

$$f_Y(y) = \left|h_1'(y)\right| f_X[h_1(y)] + \left|h_2'(y)\right| f_X[h_2(y)]. \tag{1.92}$$

更复杂的是一个 Y 值对应多个值的情况。此时, 将上式作进一步推广, 由概率可加性可得

$$f_Y(y)\mathrm{d}y = f_X(x_1)\,\mathrm{d}x_1 + f_X(x_2)\,\mathrm{d}x_2 + f_X(x_3)\,\mathrm{d}x_3 + \cdots \tag{1.93}$$

则

$$f_Y(y) = \left|h_1'(y)\right| f_X\left[h_1(y)\right] + \left|h_2'(y)\right| .f_X\left[h_2(y)\right] + \left|h_3'(y)\right| f_X\left[h_3(y)\right] + \cdots \tag{1.94}$$

例 1.44 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

$$X = h(Y) = \pm\sqrt{Y}. \tag{1.95}$$

其反函数导数的绝对值为

$$\left|h_1'(y)\right| = \left|h_2'(y)\right| = \left|\frac{dx}{dy}\right| = \frac{1}{2\sqrt{y}}. \tag{1.96}$$

① 当 $y < 0$ 时, $\left\{X^2 \leqslant y\right\}$ 为不可能事件, 所以 $P\left\{X^2 \leqslant y\right\} = 0$, 得 $F_Y(y) = 0$, 因此, 当 $y < 0$ 时, 其概率密度 $f(y) = 0$。

② 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

$$\begin{aligned} f_Y(y) &= \left|h_1'(y)\right| f_X\left[h_1(y)\right] + \left|h_2'(y)\right| f_X\left[h_2(y)\right] \\ &= \frac{1}{2\sqrt{y}}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-(-\sqrt{y})^2/2} + \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-(\sqrt{y})^2/2}\right) \\ &= \frac{1}{2\sqrt{y}} \cdot \frac{2}{\sqrt{2\pi}}\mathrm{e}^{-x/2} = \frac{1}{\sqrt{2\pi}}y^{-1/2}\mathrm{e}^{-y/2}. \end{aligned} \tag{1.97}$$

综合①②可得

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}}y^{-1/2}\mathrm{e}^{-y/2}, & y > 0 \\ 0, & y < 0 \end{cases}. \tag{1.98}$$

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 1–15 所示。

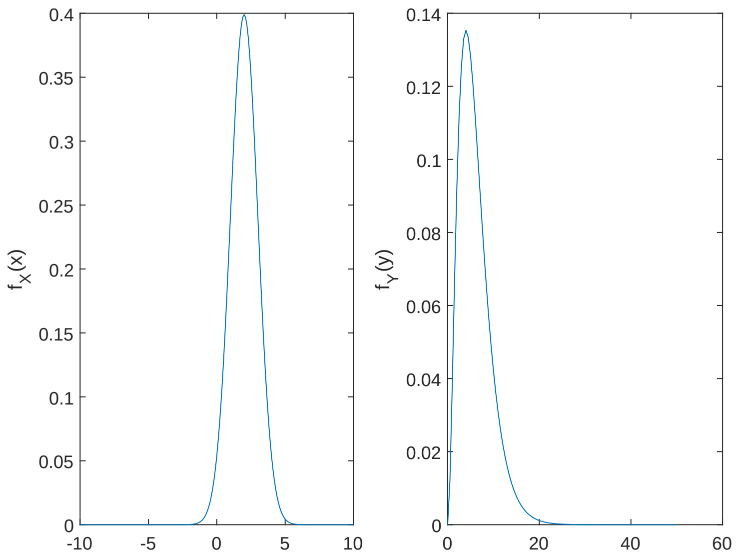

图 1–15 输入、输出随机变量的概率密度: $mu = 2, \sigma = 1, n = 6$

1.1.4.3 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

$$\begin{cases} Y_1 = g_1(X_1, X_2) \\ Y_2 = g_2(X_1, X_2) \end{cases} \tag{1.99}$$

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

$$\begin{cases} X_1 = h_1(Y_1, Y_2) \\ X_2 = h_2(Y_1, Y_2) \end{cases} \tag{1.100}$$

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, 如图 1–16 所示那么 (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 (Y_1, Y_2) 落入 $dS_{y_1y_2}$ 区间的概率 $f_Y(y_1, y_2)\, dS_{y_1y_2}$。

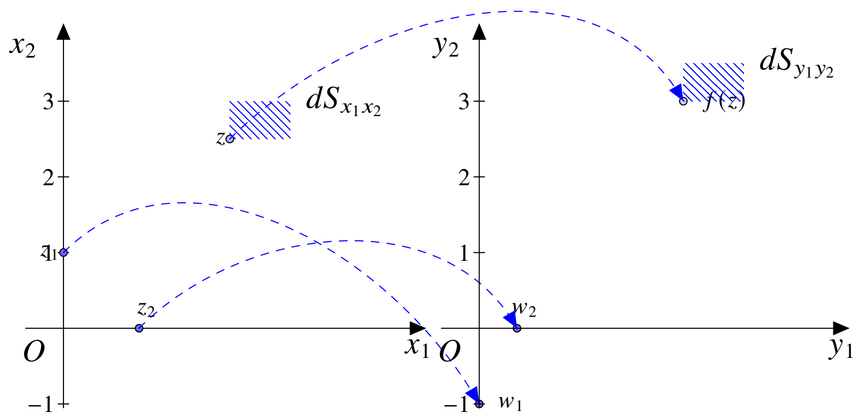

图 1–16 单值映射下，函数变换对应的区间变换

所以, 新的二维随机变量 (Y_1, Y_2) 的概率密度为

$$f_Y(y_1, y_2) = f_X(x_1, x_2) \cdot \frac{\mathrm{d}S_{x_1x_2}}{\mathrm{d}S_{y_1y_2}}. \tag{1.101}$$

坐标转换中 $dS_{x_1x_2}$ 和 $dS_{y_1y_2}$ 之间的变换称为雅可比变换。可得雅可比行列式为

$$J = \frac{\mathrm{d}S_{x_1x_2}}{\mathrm{d}S_{y_1y_z}} = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix}. \tag{1.102}$$

于是

$$\begin{aligned} f_Y(y_1, y_2) = f_X(x_1, x_2) \left|\frac{dS_{x_1x_2}}{dS_{y_1x_2}}\right| &= |J| f_X(x_1, x_2) \\ &= |J| f_X[h_1(y_1, y_2), h_2(y_1, y_2)]. \end{aligned} \tag{1.103}$$

例 1.45 已知二维随机变量 (X_1, X_2) 的联合概率密度

$$f_X(x_1, x_2) = \begin{cases} \mathrm{e}^{-(x_1+x_2)}, & x_1 > 0, x_2 > 0 \\ 0, & \text{其他} \end{cases} \tag{1.104}$$

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

$$Y_1 = \frac{X_1 + X_2}{2}, \quad Y_2 = \frac{X_1 - X_2}{2}. \tag{1.105}$$

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

$$\begin{cases} x_1 = h_1(y_1, y_2) = y_1 + y_2 \\ x_2 = h_2(y_1, y_2) = y_1 - y_2 \end{cases}. \tag{1.106}$$

则其雅可比行列式为

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2. \tag{1.107}$$

可得

$$\begin{aligned} f_Y(y_1, y_2) &= |J| f_X[h_1(y_1, y_2), h_2(y_1, y_2)] \\ &= 2 f_X[y_1 + y_2, y_1 - y_2] = \begin{cases} 2\mathrm{e}^{-2y_1}, & y_1 > |y_2| \geqslant 0 \\ 0, & \text{其他} \end{cases}. \end{aligned} \tag{1.108}$$

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 $y_1 O y_z$ 平面, 找出 (Y_1, Y_2) 的值域, 如图 1–17 所示。(Y_1, Y_2) 的值域满足

$$\begin{cases} x_1 > 0 \\ x_2 > 0 \end{cases} \Rightarrow \begin{cases} y_1 + y_2 > 0 \\ y_1 - y_2 > 0 \end{cases} \Rightarrow y_1 > |y_2| \geqslant 0\ . \tag{1.109}$$

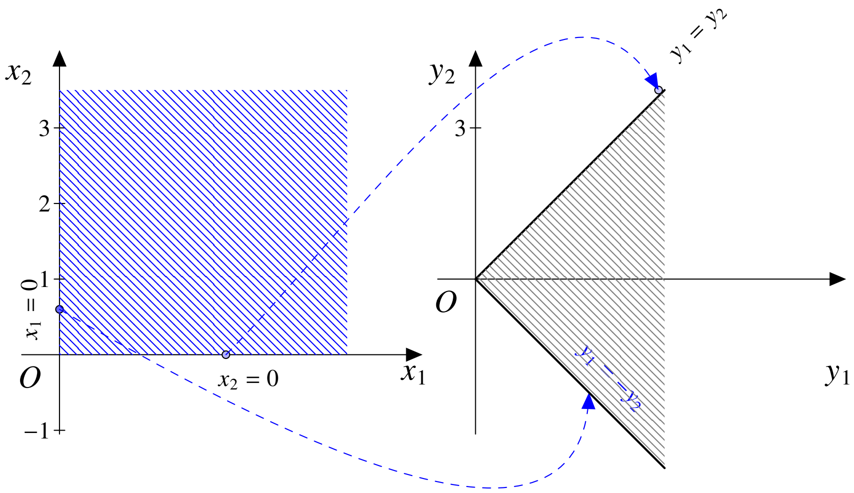

图 1–17 例1.45 的区间变换

② 其边缘分布为

$$
\begin{aligned}
f_Y(y_1) &= \int_{-y_1}^{y_1} f_Y(y_1, y_2)\,\mathrm{d}y_2 = \int_{-y_1}^{y_1} 2\mathrm{e}^{-2y_1}\mathrm{d}y_2 = 4y_1\mathrm{e}^{-2y_1}, y_1 > 0 \\
f_Y(y_2) &= \begin{cases} \int_{y_2}^{\infty} 2\mathrm{e}^{-2y_1}\mathrm{d}y_1 = \mathrm{e}^{-2y_2}, y_2 \geqslant 0 \\ \int_{-y_2}^{\infty} 2\mathrm{e}^{-2y_1}\mathrm{d}y_1 = \mathrm{e}^{2y_2}, y_2 < 0 \end{cases} = \mathrm{e}^{-2|y_2|}, -\infty < y_2 < +\infty
\end{aligned} \tag{1.110}
$$

由于

$$
\begin{aligned}
f_Y(y_1) \cdot f_Y(y_2) &= 4y_1\mathrm{e}^{-2(y_1+|y_2|)} \\
&\neq f_Y(y_1, y_2) = \begin{cases} 2\mathrm{e}^{-2y_1}, & y_1 > |y_2| \geqslant 0 \\ 0, & \text{其他} \end{cases}.
\end{aligned} \tag{1.111}
$$

所以 Y_1 与 Y_2 不是相互独立的。

2. 多值变换

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

$$
\begin{cases} X_{a_1} = h_{a_1}(Y_1, Y_2) \\ X_{a_2} = h_{a_2}(Y_1, Y_2) \end{cases}, \begin{cases} X_{b_1} = h_{b_1}(Y_1, Y_2) \\ X_{b_2} = h_{b_2}(Y_1, Y_2) \end{cases}. \tag{1.112}
$$

这种 (Y_1, Y_2) 有多个 (X_1, X_2) 与其对应的函数变换, 称之为多值变换。与一维的多值变换情况类似, 可用概率的加法定理, 求二维随机变量的多值变换函数的概率密度, 即

$$
f_Y(y_1, y_2) = |J_a|\, f_X\left[h_{a_1}(y_1, y_2), h_{a_2}(y_1, y_2)\right] + |J_b|\, f_X\left[h_{b_1}(y_1, y_2), h_{b_2}(y_1, y_2)\right]. \tag{1.113}
$$

其中, 雅可比行列式

$$
J_a = \frac{\partial(x_{a_1}, x_{a_2})}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_{a_1}}{\partial y_1} & \frac{\partial h_{a_1}}{\partial y_2} \\ \frac{\partial h_{a_2}}{\partial y_1} & \frac{\partial h_{a_2}}{\partial y_2} \end{vmatrix}, \tag{1.114}
$$

$$
J_b = \frac{\partial(x_{b_1}, x_{b_2})}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_{b_1}}{\partial y_1} & \frac{\partial h_{b_1}}{\partial y_2} \\ \frac{\partial h_{b_2}}{\partial y_1} & \frac{\partial h_{b_2}}{\partial y_2} \end{vmatrix}. \tag{1.115}
$$

1.1.4.4 n 维随机变量函数的分布

由上述的一维和二维随机变量函数变换的结论, 可用归纳法将其扩展到 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

$$
\begin{cases} X_1 = h_1(Y_1, Y_2, \cdots, Y_n) \\ X_2 = h_2(Y_1, Y_2, \cdots, Y_n) \\ \qquad \vdots \\ X_n = h_n(Y_1, Y_2, \cdots, Y_n) \end{cases}. \tag{1.116}
$$

则

$$
\begin{aligned}
f_Y\left(y_1,\cdots,y_n\right) &= |J|\cdot f_X\left(x_1,\cdots,x_n\right) \\
&= |J|\cdot f_X\left[h_1\left(y_1,\cdots,y_n\right),\cdots,h_n\left(y_1,\cdots,y_n\right)\right].
\end{aligned}
\tag{1.117}
$$

雅可比行列式为

$$
J=\frac{\partial\left(x_1,\cdots,x_n\right)}{\partial\left(y_1,\cdots,y_n\right)}=\begin{vmatrix}
\frac{\partial h_1(y_1,\cdots,y_n)}{\partial y_1} & \cdots & \frac{\partial h_1(y_1,\cdots,y_n)}{\partial y_n} \\
\vdots & & \vdots \\
\frac{\partial h_n(y_1,\cdots,y_n)}{\partial y_1} & \cdots & \frac{\partial h_n(y_1,\cdots,y_n)}{\partial y_n}
\end{vmatrix}. \tag{1.118}
$$

例 1.46 已知 n 维随机变量 $(X_1,\cdots,X_n)$ 的联合概率密度 $f_X\left(x_1,\cdots,x_n\right)$, 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

$$
Y_1=X_1,Y_2=X_2,\cdots,Y_{n-1}=X_{n-1},Y_n=\sum_{i=1}^{n}X_i=Y. \tag{1.119}
$$

解出反函数

$$
X_1=Y_1,X_2=Y_2,\ldots,X_{n-1}=Y_{n-1},X_n=Y_n-\sum_{i=1}^{n-1}Y_i. \tag{1.120}
$$

则 n 维雅可比行列式为

$$
J=\begin{vmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 \\
-1 & -1 & \cdots & -1 & 1
\end{vmatrix}=1. \tag{1.121}
$$

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

$$
f_Y\left(y_1,\cdots,y_n\right)=f_X\left(y_1,\cdots,y_{n-1},y_n-\sum_{i=1}^{n-1}y_i\right). \tag{1.122}
$$

其关于 Y 的边缘分布，即随机变量 Y 的概率密度 $f_Y(y)$ 为

$$
f_Y(y)=f_Y\left(y_n\right)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f_X\left(y_1,\cdots,y_{n-1},y_n-\sum_{i=1}^{n-1}y_i\right)dy_1\cdots dy_{n-1}. \tag{1.123}
$$

1.1.4.5 随机变量函数分布的广义定积分方法

随机变量的函数分布是概率论与数理统计中的重要内容之一, 关于二维随机变量分布函数与概率密度的求解是一个难点。对于二维随机变量的函数的概率密度的计算，一般教材和文献仅给出随机变量的和、差、积和商的概率密度公式, 其余均须按定义先求其分布函数, 然后对分布函数求导得到概率密度。尽管求解思路不复杂, 但求分布函数时多需要分区域求二重积分, 而正确划分积分区域给出准确表达积分式往往难度较大, 极易出错, 运算量巨大。为了解决这一难题, 文中应用积分变换给出了二维随机变量的函数的概率密度的新计算公式 [**宋明娟 2011**] 。利用这些公式求概率密度, 仅需完成广义定积分计算即可, 较分布函数法降低了积分重数, 简化了计算难度, 提高了计算效率和准确率。

定理 1.1.1

设二维随机变量 (X,Y) 的概率密度为 $f(x,y), Z = g(X,Y)$ 为关于随机变量 X 或 Y 的概率密度，可按如下方式计算

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, y(x,z)) \left|\frac{\partial y}{\partial z}\right| \mathrm{d}x. \tag{1.124}$$

或

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x(y,z), y) \left|\frac{\partial x}{\partial z}\right| \mathrm{d}y. \tag{1.125}$$

证明　若当 $y \in (-\infty, +\infty)$ 时, $\frac{\partial z}{\partial y} > 0$, 则函数 $z = g(x,y)$ 关于变量 y 严格单调增加，它的反函数 $y = y(x,z)$ 及 $\frac{\partial y}{\partial z}$ 必存在 , 且 $\frac{\partial y}{\partial z} > 0$, 于是雅可比行列式

$$J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial y}{\partial z} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} 1 & \frac{\partial y}{\partial x} \\ 0 & \frac{\partial y}{\partial z} \end{vmatrix} = \frac{\partial y}{\partial z}.$$

(X,Z) 的联合分布为

$$\begin{aligned} &F(x,z) = P\{X \leqslant x, Z(X,Y) \leqslant z\} = P\{X \leqslant x, Y \leqslant y(x,z)\} \\ &= \int_{-\infty}^{x} \mathrm{d}x \int_{-\infty}^{y(x,z)} f(x,y)\mathrm{d}y = \frac{\partial y}{\partial z}. \end{aligned}$$

所以

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, y(x,z)) \frac{\partial y}{\partial z} \mathrm{d}x.$$

若当 $y \in (-\infty,+\infty)$ 时，$\frac{\partial z}{\partial y} < 0$, 则

$$
\begin{aligned}
F(x,z) = P\{X \leqslant x, Z(X,Y) \leqslant z\} &= P\{X \leqslant x, Y \geqslant\} y(x,z)\} \\
&= \int_{-\infty}^{x} \mathrm{d}x \int_{y(x,z)}^{+\infty} f(x,y)\mathrm{d}y. \\
&= \int_{-\infty}^{x} \mathrm{d}x \int_{-\infty}^{z} f(x,y(x,z)) \left(-\frac{\partial y}{\partial z}\right) \mathrm{d}z.
\end{aligned}
$$

综上可得, 当 z 为 y 的严格单调函数且 $\frac{\partial y}{\partial z}$ 处处存在, 则有

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x,y(x,z)) \left|\frac{\partial y}{\partial z}\right| \mathrm{d}x.$$

同理, 当 z 为 x 的严格单调函数且 $\frac{\partial x}{\partial z}$ 处处存在, 则有

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x(y,z),y) \left|\frac{\partial x}{\partial z}\right| \mathrm{d}y.$$

推论 1　设二维连续型随机变量 (X,Y) 的概率密度为 $f(x,y)$, a,b 为非零实数, 则
(1) $Z = aX + bY$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax}{b}\right) \left|\frac{1}{b}\right| \mathrm{d}x.$$

或

$$f_z(z) = \int_{-\infty}^{+\infty} f\left(\frac{z-by}{a}, y\right) \left|\frac{1}{a}\right| \mathrm{d}y.$$

(2) $Z = \frac{Y}{aX}$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x,axz)|ax|\mathrm{d}x.$$

(3) $Z = aXY$ 的概率密度为

$$f_z(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z}{ax}\right) \left|\frac{1}{ax}\right| dx.$$

或

$$f_z(z) = \int_{-\infty}^{+\infty} f\left(\frac{z}{ay}, y\right) \left|\frac{1}{ay}\right| \mathrm{d}y.$$

例 1.1.4

设 (X,Y) 服从 $D = \{(x,y) \mid -1 \leqslant x \leqslant 10 \leqslant y < 1\}$ 上的均匀分布。试求 $Z = \frac{Y}{3X}$ 的概率密度 $f_z(z)$.

解 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}\frac{1}{2}, & -1\leqslant x\leqslant 1, 0\leqslant y<1\\ 0, & \text{其他}\end{cases}$$

由 $Z=\frac{Y}{3X}$, 得 $Y=3XZ$, $\frac{\partial Y}{\partial Z}=3X$,

$$f(x,3xz)=\begin{cases}\frac{1}{2}, & -1\leqslant x\leqslant 0, \frac{1}{3x}\leqslant z<0 \text{ 或} 0\leqslant x\leqslant 1, 0\leqslant z<\frac{1}{3x}\\ 0, & \text{其他}\end{cases} \tag{1.126}$$

于是

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,3xz)|3x|\mathrm{d}x=\begin{cases}\int_{\frac{1}{3}}^{0}\frac{-3x}{2}\mathrm{d}x, & -\infty<z<-\frac{1}{3}\\ \int_{-1}^{0}\frac{-3x}{2}\mathrm{d}x, & -\frac{1}{3}\leqslant z<0\\ \int_{0}^{1}\frac{3x}{2}\mathrm{d}x, & 0\leqslant z<\frac{1}{3}\\ \int_{0}^{\frac{1}{3}}\frac{3x}{2}\mathrm{d}x, & \frac{1}{3}\leqslant z<+\infty\end{cases} \tag{1.127}$$

$$=\begin{cases}\frac{1}{12z^2}, & |z|\geqslant\frac{1}{3}\\ \frac{3}{4}, & |z|<\frac{1}{3}\end{cases}$$

定理 1.1.2

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 $^n\ i=1I_i=-\infty,+\infty)$，则随机变量 Z 的概率密度为

$$f_Z(z)=\sum_{i=1}^{n}\int_{-\infty}^{+\infty} f\left(x,y_i(x,z)\right)\left|\frac{\partial y_i}{\partial z}\right|\mathrm{d}x. \tag{1.128}$$

或

$$f_Z(z)=\sum_{i=1}^{n}\int_{-\infty}^{+\infty} f\left(x_i(y,z),y\right)\left|\frac{\partial x_i}{\partial z}\right|\mathrm{d}y. \tag{1.129}$$

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

$$
\begin{aligned}
F(x,z) =& P\{X \leqslant x, Z \leqslant z\} = P\left\{\bigcup_{i=1}(X \leqslant x, g(X,Y) \leqslant z, Y \in I_i)\right\} \\
&= \sum_{i=1}^{n} P\{X \leqslant x, g(X,Y) \leqslant z, Y \in I_i\} \\
&= \sum_{i=1}^{n} \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} f(x, y_i(x,z)) \left|\frac{\partial y_i}{\partial z}\right| \mathrm{d}z \\
&= \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \left[\sum_{i=1}^{n} f(x, y_i(x,z)) \left|\frac{\partial y_i}{\partial z}\right|\right] \mathrm{d}z.
\end{aligned}
\tag{1.130}
$$

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

$$
f_Z(z) = \sum_{i=1}^{n} \int_{-\infty}^{+\infty} f(x, y_i(x,z)) \left|\frac{\partial y_i}{\partial z}\right| \mathrm{d}x. \tag{1.131}
$$

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,

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

例 1.1.5

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

$$
f(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi\sigma^2} \mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}, x, y \in \mathbb{R}. \tag{1.132}
$$

由 $z = \sqrt{x^2 + y^2}$, 解得

$$
\begin{cases}
y_1 = \sqrt{z^2 - x^2}, & \frac{\partial y_1}{\partial z} = \frac{z}{\sqrt{z^2 - x^2}} > 0, \quad |x| \leqslant z \\
y_2 = -\sqrt{z^2 - x^2}, & \frac{\partial y_2}{\partial z} = -\frac{z}{\sqrt{z^2 - x^2}} < 0, \quad |x| \leqslant z
\end{cases}, \tag{1.133}
$$

$$
f\left(x, \pm\sqrt{z^2 - x^2}\right) = \begin{cases}
\frac{z}{2\pi\sigma^2} e^{-\frac{j2}{2\sigma^2}}, & |x| < z \\
0, & \text{其他}
\end{cases}. \tag{1.134}
$$

当 $z > 0$ 时,

$$\begin{aligned}f_z(z) &= \int_{-z}^{z} f\left(x, \sqrt{z^2 - x^2}\right) \frac{z}{\sqrt{z^2 - x^2}} \mathrm{d}x \\ &\quad + \int_{-z}^{2} f\left(x, -\sqrt{z^2 - x^2}\right) \frac{z}{\sqrt{z^2 - x^2}} \mathrm{d}x \\ &= 4 \int_0^z \frac{1}{2\pi\sigma^2} \mathrm{e}^{-\frac{z^2}{2\sigma^2}} \frac{z}{\sqrt{z^2 - x^2}} \mathrm{d}x = \frac{2z}{\pi\sigma^2} \mathrm{e}^{-\frac{z^2}{2\sigma^2}} \left[\arcsin \frac{x}{z}\right]_0^z \\ &= \frac{z}{\sigma^2} \mathrm{e}^{-\frac{z^2}{2\sigma^2}}.\end{aligned} \tag{1.135}$$

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

综上可得

$$f_Z(z) = \begin{cases} \frac{z}{\sigma^2} \mathrm{e}^{-\frac{z^2}{2\sigma^2}}, & z \geqslant 0 \\ 0, & z < 0 \end{cases}, Z \sim \mathrm{Rayleigh}(\sigma). \tag{1.136}$$

1.1.5 随机变量的数字特征

在实用中，概率分布函数 (或概率密度函数) 往往很难获得，有时也仅仅需要随机变量的一些统计特性。下面以连续型随机变量为例，描述随机变量的期望、方差和相关系数等数字特征。

定义 1.47 数学期望

变量 X 的数学期望定义为

$$E\{X\} = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x. \tag{1.137}$$

数学期望是在概率统计意义上的一种平均, 称为集合均值, 简称集平均, 因此 $E\{X\}$ 又称为均值. ♣

1) 数学期望具有如下一些基本的性质：

(1) 常量的数学期望等于常量本身。

(2) 对常数 a_i $(i = 1, 2, \cdots, n)$, 有

$$E\left\{\sum_{i=1}^{n} a_i X_i\right\} = \sum_{i=1}^{n} a_i E\{X_i\}. \tag{1.138}$$

(3) X_i $(i = 1, 2, \cdots, n)$ 相互独立，则

$$E\{X_1 X_2 \cdots X_n\} = \prod_{i=1}^{n} E\{X_i\}. \tag{1.139}$$

例 1.48 随机变量 X 在区间 (a,b) 上服从均匀分布, 求 X 的数学期望。

解: 由于 X 服从均匀分布, 则概率密度为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases} \tag{1.140}$$

则数学期望为

$$E[X]=\int_{-\infty}^{\infty}xf_X(x)\mathrm{d}x=\int_a^b\frac{x}{b-a}\mathrm{d}x=\frac{b+a}{2}. \tag{1.141}$$

2) 一维随机变量函数的数学期望

实际应用中, 不仅要会求随机变量的期望, 还要求随机变量函数的数学期望。例如: 飞机机翼受到的压力 $W=KV\ (K>0)$ 是常数, V 表示风速, 是个随机变量, 若要计算受到的压力 W 的统计平均值, 即求随机变量 V 的函数 W 的数学期望。下面讨论已知随机变量 X 的分布, 求其函数 $Y=g(X)$ 的数学期望。

已知随机变量 X 的概率密度函数 $f_X(x)$, 且随机变量 $Y=g(X)$, 其中 $g(\cdot)$ 是连续型实函数。随机变量 Y 的数学期望为

$$E[Y]=\int_{-\infty}^{\infty}yf_Y(y)dy. \tag{1.142}$$

由于 $f_Y(y)$ 未知, 故下一步利用 $f_X(x)$, 求出 $f_Y(y)$.

① 若 $g(\cdot)$ 是单值变换, 则

$$f_Y(y)dy=f_X(x)\mathrm{d}x\Rightarrow f_Y(y)=f_X(x)\mathrm{d}x/dy. \tag{1.143}$$

代入期望定义式得

$$E[Y]=\int_{-\infty}^{\infty}yf_Y(y)\mathrm{d}y=\int_{-\infty}^{\infty}g(x)f_X(x)\mathrm{d}x=E[g(X)]. \tag{1.144}$$

② 若 $g(\cdot)$ 是多值变换

$$\begin{aligned}f_Y(y)\mathrm{d}y&=f_X(x_1)\,\mathrm{d}x_1+f_X(x_2)\,\mathrm{d}x_2+\ldots\\ f_Y(y)&=[f_X(x_1)\,\mathrm{d}x_1+f_X(x_2)\,\mathrm{d}x_2+\ldots]/\mathrm{d}y.\end{aligned} \tag{1.145}$$

代入期望定义式得

$$\begin{aligned}E[Y]&=\int_{D_{x_1}}g(x_1)\,f_X(x_1)\,\mathrm{d}x_1+\int_{D_{x_2}}g(x_2)\,f_X(x_2)\,\mathrm{d}x_2+\ldots\\&=\int_{-\infty}^{\infty}g(x)f_X(x)\mathrm{d}x=E[g(X)].\end{aligned} \tag{1.146}$$

其中 $D_{x_1},D_{x_2},\cdots$ 为 $f_X(x_1),f_X(x_2),\ldots$ 的定义域。

综合①②可知, 无论函数 $g(X)$ 是单值还是多值变换随机变量函数的期望定义如下:

定义 1.49 随机变量函数的期望

若连续型变量 X 的概率密度为 $f_X(x)$, 且 $\int_{-\infty}^{\infty}|g(x)|f_X(x)\mathrm{d}x<\infty$, 则函数 $g(X)$ 的期望为

$$E[g(X)]=\int_{-\infty}^{\infty}g(x)f_X(x)\mathrm{d}x. \tag{1.147}$$

若 X 为离散型变量, 且 $\sum_{k=1}^{\infty}|g(x_k)|p_k<\infty$, 则函数 $g(X)$ 的期望为

$$E[g(X)]=\sum_{k=1}^{\infty}g(x_k)p_k. \tag{1.148}$$

♣

例 1.50 随机变量 X 在区间 (a,b) 上服从均匀分布, 求 $g(X)=x^2+1$ 的数学期望.

解: 由于 X 服从均匀分布, 则概率密度为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases} \tag{1.149}$$

则函数的数学期望为

$$E[g(X)]=\int_{-\infty}^{\infty}g(x)f_X(x)dx=\int_a^b\frac{x^2+1}{b-a}\mathrm{d}x=\frac{1}{3}\left(a^2+ab+b^2\right)+1. \tag{1.150}$$

3. 二维随机变量及其函数的数学期望

定义 1.51 二维连续型随机变量函数的期望

设 (X,Y) 是定义在概率空间 (Ω,F,P) 上的二维连续型随机变量, 且联合概率密度 $f_{XY}(x,y)$ 已知, 则由联合概率密度与边缘概率密度的关系及期望定义式可得

$$\begin{cases}E[X]=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{XY}(x,y)\mathrm{d}x\mathrm{d}y;\\ E[Y]=\int_{-\infty}^{\infty}yf(y)\mathrm{d}y=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{XY}(x,y)\mathrm{d}x\mathrm{d}y.\end{cases} \tag{1.151}$$

♣

定义 1.52 二维离散型随机变量函数的期望

若 (X,Y) 为离散型随机变量, 且联合概率分布率 $P\{X=x,Y=y\}$ 已知, 则

$$\begin{cases}E[X]=\sum_i x_iP\{X=x_i\}=\sum_i\sum_j x_iP\{X=x_i,Y=y_j\};\\ E[Y]=\sum_j y_jP\{Y=y_j\}=\sum_j\sum_i y_jP\{X=x_i,Y=y_j\}.\end{cases} \tag{1.152}$$

♣

仿照单个随机变量函数求期望的方法, 二维随机变量函数 $g(X,Y)$ 的数学期望定义为

定义 1.53 二维离散型随机变量函数的期望

二维随机变量函数 $g(X,Y)$ 的数学期望为

$$\begin{aligned} E[g(X,Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)\mathrm{d}x\mathrm{d}y. \\ E[g(X,Y)] &= \sum_i \sum_j g\left(x_i,y_j\right) P\left\{X=x_i, Y=y_j\right\}. \end{aligned} \tag{1.153}$$

♣

4. n 维随机变量的数学期望

定义 1.54 n 维连续型随机变量的期望

设 $(X_1, X_2, \cdots, X_n)$ 是定义在概率空间 (Ω, F, P) 上的 n 维连续型随机变量, 若其联合概率密度分布为 $f_X(x_1, x_2, \cdots, x_n)$, 则与二维情况类似, 有

$$E[X_i] = \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}} x_i f_X(x_1, \cdots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n, i=1,2,\cdots,n; \tag{1.154}$$

♣

定义 1.55 n 维连续型随机变量函数的期望

若 n 维随机变量的函数为 $g(X_1, X_2, \cdots, X_n)$, 则 n 维随机变量函数的数学期望为

$$E[g(X_1, X_2, \cdots, X_n)] = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}} g(x_1, \cdots, x_n)\, f_X(x_1, \cdots, x_n)\, dx_1 \cdots dx_n. \tag{1.155}$$

♣

定义 1.56 随机矢量的期望

当 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 用随机矢量 $X = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}^{\mathrm{T}}$ 来表示, 且若随机矢量 $\boldsymbol{X}$ 中的每个分量 $\boldsymbol{X}_i$ 的数学期望均存在 $(E[X_i] = m_i)$, 则随机矢量 $\boldsymbol{X}$ 的数学期望为

$$E(\boldsymbol{X}) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{bmatrix} = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} = \boldsymbol{M}_X. \tag{1.156}$$

♣

可见, 随机矢量 $\boldsymbol{X}$ 的数学期望是一个常数矢量, 常用 $\boldsymbol{M}_X$ 表示。

例 1.57 设 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的函数为 $g(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} a_i X_i$, 其中权重 a_i 是常数。求数学期望 $E[g(X_1, X_2, \cdots, X_n)]$。

解: 由 n 维随机变量函数的期望定义

$$
\begin{aligned}
E\left[g\left(X_1, X_2, \cdots, X_n\right)\right] &= \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} g\left(x_1, \cdots, x_n\right) f_X\left(x_1, x_2, \cdots, x_n\right) dx_1 \cdots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n\right) f_X\left(x_1, \cdots, x_n\right) dx_1 \cdots dx_n \\
&= \sum_{i=1}^{n} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} a_i x_i f_X\left(x_1, \cdots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n. \qquad (1.157)
\end{aligned}
$$

根据边缘概率密度的公式, 和式中每一项都可以化成

$$
\begin{aligned}
\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_i x_i f_X\left(x_1, \cdots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n &= \int_{-\infty}^{\infty} a_i x_i f_{X_i}\left(x_i\right) \mathrm{d}x_i = E\left[a_i X_i\right] \\
&= a_i E\left[X_i\right]. \qquad (1.158)
\end{aligned}
$$

所以

$$
E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E\left[X_i\right]. \qquad (1.159)
$$

由此可见, 随机变量加权求和的均值等于各随机变量均值的加权和。

例 1.58 已知 n 个相互独立的随机变量 $X_1, X_2, \cdots, X_n$ 分别为 $E\left[X_1\right], E\left[X_2\right], \cdots, E\left[X_n\right]$。求其函数 $g\left(X_1, X_2, \cdots, X_n\right) = \prod_{i=1}^{n} X_i$ 的数学期望。

解: 由 n 维随机变量函数的期望定义

$$
\begin{aligned}
E\left[g\left(X_1, X_2, \cdots, X_n\right)\right] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g\left(x_1, \cdots, x_n\right) f_X\left(x_1, x_2, \cdots, x_n\right) dx_1 \cdots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 x_2 \cdots x_n f_X\left(x_1, \cdots, x_n\right) \mathrm{d}x_1 \cdots dx_n \text{ (相互独立条件下)} \\
&= \int_{-\infty}^{\infty} x_1 f_{x_1}\left(x_1\right) \mathrm{d}x_1 \cdot \int_{-\infty}^{\infty} x_2 f_{x_2}\left(x_2\right) \mathrm{d}x_2 \cdots \cdots \int_{-\infty}^{\infty} x_n f_{x_n}\left(x_n\right) \mathrm{d}x_n \\
&= E\left[X_1\right] E\left[X_2\right] \cdots E\left[X_n\right]. \qquad (1.160)
\end{aligned}
$$

即

$$
E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E\left[X_i\right]. \qquad (1.161)
$$

由上式可见, n 个相互独立的随机变量乘积的期望等于 n 个随机变量期望的乘积。

归纳起来，数学期望的基本性质有

1° 若随机变量满足 $a \leqslant X \leqslant b, a, b$ 常数, 则其数学期望 $a \leqslant E[X] \leqslant b$。

2° 常数 C 的期望 $E[C] = C$。

3° 对任意常数 $b, a_i\ (i = 1, 2, \cdots, n)$, 有

$$E\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i E\left[X_i\right] + b. \tag{1.162}$$

4° 若随机变量 X 与随机变量 Y 互不相关, 则

$$E[XY] = E[X]E[Y]. \tag{1.163}$$

5° 若 n 个随机变量 $(X_1, X_2, \cdots, X_n)$ 相互独立, 则

$$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E\left[X_i\right]. \tag{1.164}$$

定义 1.1.1

设 $X \sim N(0,1)$, 若 z_α 满足 $P\{X > z_\alpha\} = \alpha\,(0 < \alpha < 1)$, 则称 z_α 为标准正态分布的**上 α 分位点** (图1–18). $\Phi\left(z_\alpha\right) = P\{X \leq z_\alpha\} = 1 - \alpha$.

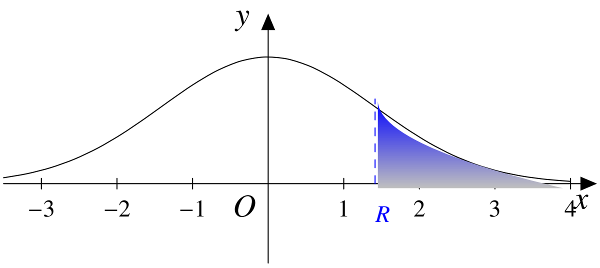

图 1–18 上 α 分位点

注 1.59.

$$z_{0.05} = 1.645, z_{0.005} = 2.57, z_{0.025} = 1.96.$$

在本小节最后，我们先列出一些关于多维随机变量函数经常用到的结论：

性质 (1) 若 $X_i \sim N\left(\mu_i, \sigma_i^2\right)\ (i = 1, 2, \cdots, n)$, 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.

性质 (2) 若 $X_i \sim N\left(0, \sigma^2\right)$ $(i=1,2,\cdots,n)$ 且相互独立，则 $Z=\sum_{i=1}^n X_i^2$ 服从由由度为 n 的中心 χ^2 分布:

$$f(z)=\begin{cases}\frac{1}{\sigma^n 2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} z^{\frac{n}{2}-1} \mathrm{e}^{-\frac{z}{2}}, & z>0 \\ 0, & z \leqslant 0\end{cases} \tag{1.165}$$

性质 (3) 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

$$f(z)=\frac{\sigma^2}{2} \mathrm{e}^{-(x+\lambda)/2}\left(\frac{x}{\lambda}\right)^{n/4-1/2} \mathbf{I}_{n/2-1}(\sqrt{\lambda x}), \tag{1.166}$$

式中，$\lambda=n(\mu/\sigma)^2$ 为非中心参数，$\mathbf{I}_{n/2-1}(\cdot)$ 为 $(n/2-1)$ 阶第一类贝塞尔函数, 具体表达式见 1.168.

n 阶第一类贝塞尔函数, $c=-n$ 时,

$$I_{-n}(x)=\sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(-n+m+1)}\left(\frac{x}{2}\right)^{-n+2m}, n \neq 1,2,\cdots \tag{1.167}$$

$$I_n(x)=\sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(n+m+1)}\left(\frac{x}{2}\right)^{n+2m}, \tag{1.168}$$

性质 (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

$$f(z)=\begin{cases}\frac{z}{\sigma^2} \mathbf{e}^{-\frac{z^2}{2\sigma}}, & z \geqslant 0 \\ 0, & z<0\end{cases} \tag{1.169}$$

1.1.6 条件数学期望

前面,我们已经讨论了条件分布的概念,现在引入条件数学期望的概念。它在随机过程、时间序列分析和统计判决理论中起着重要作用。

1. 随机变量关于某给定值的条件期望

设 (X,Y) 是定义在同一概率空间上的二维连续型随机变量。Y 的数学期望可以通过 Y 本身的概率密度 $f_Y(y)$ 来计算, 也可用 (X,Y) 的联合概率密度 $f_{XY}(x,y)$ 来计算, 也可以借助条件分布 $f_Y(y|x)$ 来计算, 如

$$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y = \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_Y(y|x) f_X(x)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} y f_Y(y|x)\mathrm{d}y\right]\cdot f_X(x)\mathrm{d}x. \end{aligned} \tag{1.170}$$

注意方括号 $[\cdot]$ 内的项, $\int_{-\infty}^{\infty} y f_Y(y|x)dy$ 也是 y 的一种统计平均, 只不过权重不是一般的非条件概率，而是条件概率。常称它为 Y 在 $\{X=x\}$ 下 Y 的条件数学期望, 并用符号 $E[Y|X=x]$ 表示, 即

$$E[Y|X=x] = \int_{-\infty}^{\infty} y f_Y(y|x)dy. \tag{1.171}$$

上述定义方法也适用于 (X,Y) 为离散型随机变量的情况:

$$\begin{aligned} E[Y] &= \sum_j y_j P\left\{Y=y_j\right\} = \sum_i\sum_j y_j P_{ij} = \sum_i\sum_j y_j P\left\{X=x_i, Y=y_j\right\} \\ &= \sum_i\left[\sum_j y_j P\left\{Y=y_j|X=x_i\right\}\right] P\left\{X=x_i\right\}. \end{aligned} \tag{1.172}$$

离散型随机变量 Y 在给定条件 $X=x_i$ 下的条件数学期望表示为

$$E\left[Y|X=x_i\right] = \sum_j y_j P\left\{Y=y_j|X=x_i\right\}. \tag{1.173}$$

此外, 若 X 是离散型随机变量, 而 Y 是连续型随机变量, 而且对所有的 x_i $(i=1,2,\cdots)$ 和 y 取值的条件概率密度 $f_Y\left(y|X=x_i\right)$ 都存在, 则

$$E[Y] = \sum\left[\int_{-\infty}^{\infty} y f_Y\left(y|X=x_i\right)\mathrm{d}y\right] P\left\{X=x_i\right\}. \tag{1.174}$$

连续型随机变量 Y 在给定条件 $X=x_i$ 下的条件数学期望表示为

$$E\left[Y|X=x_i\right] = \int_{-\infty}^{\infty} y f_Y\left(y|X=x_i\right)\mathrm{d}y. \tag{1.175}$$

由于 Y 的条件数学期望是对 Y 的所有取值求统计平均,因此由函数数学期望的求法,可以推得

$$E[g(Y)|X=x]=\int_{-\infty}^{\infty}g(y)f_Y(y|x)\mathrm{d}y. \tag{1.176}$$

$$E[g(X,Y)|X=x]=\int_{-\infty}^{\infty}g(x,y)f_Y(y|x)\mathrm{d}y. \tag{1.177}$$

$$E\left[g_1(X)g_2(Y)|X=x\right]=g_1(x)E\left[g_2(Y)|X=x\right]. \tag{1.178}$$

2. 一随机变量关于另一随机变量的条件期望

由条件期望 $E[Y|X=x]$ 的定义可知, $E[Y|X=x]$ 是与条件 $X=x$ 有关的量。若以随机变量 X 替换给定值 x, 则称 $E[Y|X=x]$ 为随机变量 Y 关于条件 X 的条件数学期望。

对于随机变量 X 的所有取值而言, 条件期望 E[Y|X] 是定义在样本空间 Ω_X 上的函数。$E[Y|X=x]$ 结果取决于 X 的值, 而 $E[Y|X]$ 则是随机变量 X 的函数, 也是个随机变量。所以, 条件期望 $E[Y|X]$ 有如下的性质:

$$E_X\{E[Y|X]\}=E[Y]. \tag{1.179}$$

推导过程:

$$\begin{aligned}E_X\{E[Y|X]\}&=\int_{-\infty}^{\infty}E[Y|X=x]f_X(x)\mathrm{d}x=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}yf_Y(y|x)\mathrm{d}y\right]f_X(x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_Y(y|x)f_X(x)\mathrm{d}y\mathrm{d}x\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{XY}(x,y)\mathrm{d}y\mathrm{d}x=E[Y].\end{aligned} \tag{1.180}$$

即条件期望的期望等于非条件期望。同理可得其他性质:

2° $E\{E[g(X,Y)|X]\}=E[g(X,Y)]$.

3° $E[g(X)Y|X]=g(X)E[Y|X]$.

4° 当随机变量 X 和 Y 相互独立时, $E[Y|X]=E[Y]$, 且 $E[C|X]=E[C]=C$, C 为常数 (常数与一切随机变量独立)。

例 1.60 已知随机变量 X 服从 $(0,1)$ 的均匀分布, 随机变量 Y 服从 $(X,1)$ 的均匀分布。求 ① 条件数学期望 $E[Y|X=x]$。② 条件数学期望 $E[Y|X]$。

解: ① 根据已知条件, 在给定条件 $\{X=x\}$ 下, 随机变量 Y 的概率密度的表达式为

$$f_Y(y|X=x)=\begin{cases}\frac{1}{1-x}, & x<y<1\\ 0, & \text{其他}\end{cases} \tag{1.181}$$

条件数学期望

$$E[Y|X=x]=\int_{-\infty}^{\infty} y f_Y(y|x)\mathrm{d}y=\int_x^1 \frac{y}{1-x}\mathrm{d}y=\frac{1+x}{2}. \tag{1.182}$$

由上可以看出条件数学期望 $E[Y|X=x]$ 是关于给定值 x 的函数。

② 用随机变量 X 替换给定值 x, 则数学期望 $E[Y|X]=\frac{1+X}{2}$ 是随机变量 X 的函数, 也是个随机变量。因为随机变量 X 服从 $(0,1)$ 的均匀分布, 函数 $1+X$ 服从 (1,2) 的均匀分布, 则其函数 $\frac{1+X}{2}$ 也服从 $\left(\frac{1}{2},1\right)$ 的均匀分布, 即 $E[Y|X]$ 服从 $\left(\frac{1}{2},1\right)$ 的均匀分布, $E[Y|X]\sim U\left(\frac{1}{2},1\right)$.

1.1.7 随机变量的方差和矩

定义 1.61 方差

变量 X 的方差定义为

$$\begin{aligned} D(X) &= E\left\{[X-E\{X\}]^2\right\} \\ &= \int_{-\infty}^{+\infty} (x-E\{X\})^2 f(x)\mathrm{d}x = E\left\{X^2\right\}-(E\{X\})^2. \end{aligned} \tag{1.183}$$

方差的大小反映了随机变量相对于其均值的离散程度。随机变量 X 的方差是对“X 的取值与其期望 m_X 差值的平方”求计平均。通常称 $\sqrt{D(X)}$ 为随机变量 X 的均方差或标准差, 习惯上用 σ_X^2 表示 $D(X)$. ♣

方差具有如下一些基本的性质:

(1) 对于任意随机变量 X, 有 $D[X]\geq 0$, 且当 $x=C$ (C 常数) 时, $D[X]=0$.

(2) 对于任意实数 C, 有

$$D[CX]=C^2D[X]. \tag{1.184}$$

(3) 对于常数 a_i $(i=1,2,\cdots,n)$, 随机变量 $X_1,\cdots,X_n$ 的线性组合 $\sum_{i=1}^n a_iX_i$ 的方差为

$$\operatorname{Var}\left\{\sum_{i=1}^n a_iX_i\right\}=\sum_{i=1}^n\sum_{j=1}^n a_ia_jE\left\{(X_i-E\{X_i\})\left(X_j-E\left\{X_j\right\}\right)\right\}. \tag{1.185}$$

(4) X_i $(i=1,2,\cdots,n)$ 独立，则

$$\operatorname{Var}\left\{\sum_{i=1}^n a_iX_i\right\}=\sum_{i=1}^n a_i^2D[X_i]. \tag{1.186}$$

(5) 对于一切实数 μ, 有

$$E\left[(X-\mu)^2\right] \geqslant E\left[(X-m_X)^2\right] = D[X]. \tag{1.187}$$

证:

$$\begin{aligned} E\left[(X-\mu)^2\right] &= E\left\{\left[(X-m_X)+(m_X-\mu)\right]^2\right\} \\ &= E\left[(X-m_X)^2 + 2(X-m_X)(m_X-\mu) + (m_X-\mu)^2\right] \\ &= E\left[(X-m_X)^2\right] + 2E\left[(X-m_X)(m_X-\mu)\right] + E\left[(m_X-\mu)^2\right] \\ &= E\left[(X-m_X)^2\right] + 2(m_X-\mu)E\left[X-m_X\right] + (m_X-\mu)^2 \\ &= E\left[(X-m_X)^2\right] + (m_X-\mu)^2 \geqslant E\left[(X-m_X)^2\right]. \end{aligned} \tag{1.188}$$

这表明, 随机变量的取值相对于任何其他 μ 值的分散程度不小于其相对于期望的分散程度。可知如下两个结论:

① 此统计平均的结果是个大于等于 0 的数。

② 方差是用来表征随机变量取值相对数学期望的分散程度。

例 1.1.6

高斯变量 X 的概率密度为

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-m_X)^2}{2\sigma^2}\right\}. \tag{1.189}$$

可求出其数学期望为 m_X, 方差为 σ^2。显然, 概率密度 $f_X(x)$ 的值与均方差 σ 的大小成反比。又由概率的性质可知, 无论 σ^2 是多少, $f_X(x)$ 曲线下的面积都是相同的, 即

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1. \tag{1.190}$$

在 σ 大小不同的情况下, $f_X(x)$ 的图形如图 1–19 所示。从图中可以看出, 当 σ 小时, X 的分布相对数学期望 m 较集中; 当 σ 大时, X 的分布相对数学期望 m 较分散。由此进一步说明, 方差 σ^2 的大小正比于随机变量的取值相对于其期望分散程度的大小。

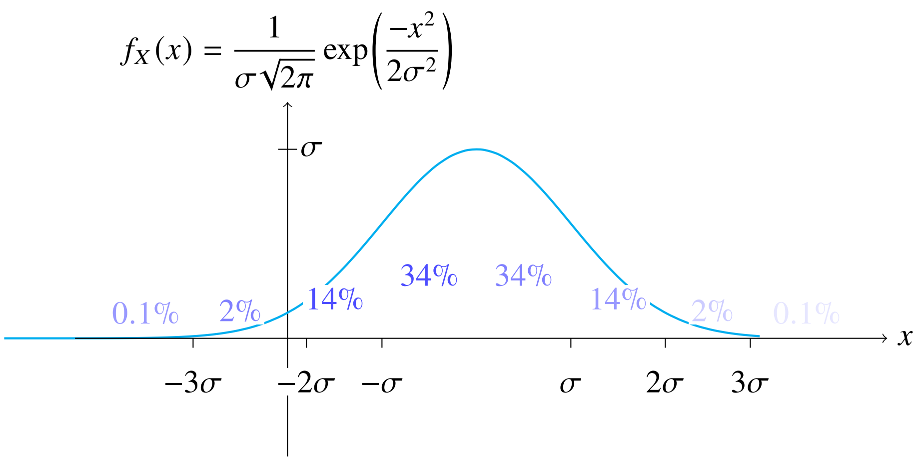

图 1–19 高斯变量的概率密度分布

1. 随机变量的矩

(1) 一维随机变量的矩

随机变量的矩有两类; 原点矩和中心矩

1) k 阶原点矩

定义 1.62 随机变量 X 的 k 阶原点矩

对随机变量 X，求 X 的 k 次幂的统计平均

$$E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f_X(x)\mathrm{d}x. \tag{1.191}$$

称为随机变量 X 的 k 阶原点矩。♣

对随机变量 X 的 k 阶原点矩的讨论如下:

① $k = 0$ 时, $E[1] = 1$.

② $k = 1$ 时, $E[X]$ 为随机变量 X 的数学期望。

③ $k = 2$ 时, $E[X^2]$ 为二阶原点矩, 又称为随机变量 X 的均方值。有时还用到 $E\left[|X|^k\right]$, 被称作 k 阶绝对原点矩，即

$$E\left[|X|^k\right] = \int_{-\infty}^{\infty} |x|^k f_X(x)dx. \tag{1.192}$$

2) k 阶中心矩

随机变量 X 相对于其数学期望 m_X 的差 $(X - m_X)$ 的 k 次幂求统计平均

$$E\left[(X - m_X)^k\right] = \int_{-\infty}^{\infty} (x - m_X)^k f_X(x)\mathrm{d}x. \tag{1.193}$$

称为随机变量 X 的 k 阶中心原点矩。讨论如下情况:

① $k=0$ 时, $E\left[(X-m_X)^0\right]=1$;

② $k=1$ 时, $E[X-m_X]=E[X]-m_X=0$;

③ $k=2$ 时, $E\left[(X-m_X)^2\right]$ 为二阶中心矩, 它是中心矩中最重要也是最常用的一种矩, 通常称为随机变量 X 的方差, 用 $D[X]$ 或 σ_X^2 表示。$D[X]=\sigma_X^2=E\left[(X-m_X)^2\right]$, $D[X]$ 的正平方根 σ_X 称为随机变量 X 的均方差或标准偏差.

有时还常用到 $E\left[|X-m_X|^k\right]$, 被称作 k 阶绝对中心矩, 即

$$E\left[|X-m_X|^k\right]=\int_{-\infty}^{\infty}|x-m_X|^k f_X(x)\mathrm{d}x. \tag{1.194}$$

3) 矩存在的条件

随机变量的各阶矩不是都存在的。

例 1.63 若随机变量的各阶绝对矩都存在, 则它相应的各阶矩都存在。但有的随机变量则不满足“各阶绝对矩都存在”的条件, 如柯西分布 (图1–20), 它的概率密度为

$$f_X(x)=\frac{1}{\pi\left(\sigma^2+\nu_2^2(x-\mu)^2\right)}, \tag{1.195}$$

其一阶绝对矩 $\int_{-\infty}^{\infty}|x|f_X(x)\mathrm{d}x$ 是发散的, 所以它的数学期望就不存在。

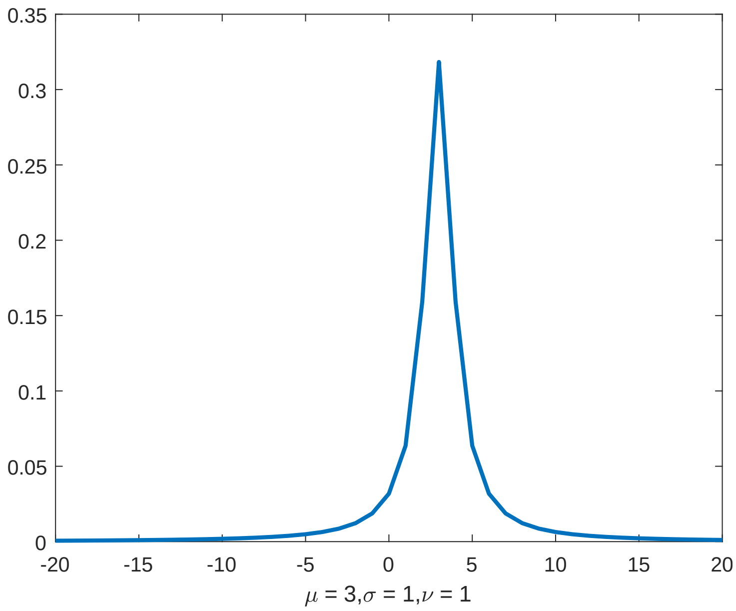

图 1–20 柯西分布

若 $(X_1, X_2, \cdots, X_n)$ 两两互不相关, 则

$$D\left[X_1 \pm X_2 \pm \cdots \pm X_n\right] = D\left[X_1\right] + D\left[X_2\right] + \cdots + D\left[X_n\right]. \tag{1.196}$$

证: 设随机变量 $Y = X_1 \pm X_2 \pm \cdots \pm X_n$, 由方差的定义 $D[Y] = E\left\{[Y - E(Y)]^2\right\}$, 可得

$$\begin{aligned} D\left[X_1 \pm X_2 \pm \cdots \pm X_n\right] &= E\left\{\left[(X_1 \pm X_2 \pm \cdots \pm X_n) - E\left(X_1 \pm X_2 \pm \cdots \pm X_n\right)\right]^2\right\} \\ &= E\left\{\left[(X_1 - m_1) \pm (X_2 - m_2) \pm \cdots \pm (X_n - m_n)\right]^2\right\} \\ &= E\left\{\sum_{i=1}^{n} (X_i - m_i)^2 \pm 2\sum_{i<j} (X_i - m_i)\left(X_j - m_j\right)\right\} \\ &= \sum_{i=1}^{n} E\left[(X_i - m_i)^2\right] \pm 2\sum_{i<j} E\left\{(X_i - m_i)\left(X_j - m_j\right)\right\}. \end{aligned} \tag{1.197}$$

对于所有 $i \neq j$, 随机变量 X_i 和 X_j 互不相关。因此根据数学期望的性质 4°, 有

$$\begin{aligned} E\left[(X_i - m_i)\left(X_j - m_j\right)\right] &= E\left[X_i - m_i\right] E\left[X_j - m_j\right] \\ &= \left\{E\left[X_i\right] - m_i\right\}\left\{E\left[X_j\right] - m_j\right\} = 0. \end{aligned} \tag{1.198}$$

所以

$$D\left[X_1 + X_2 \pm \cdots \pm X_n\right] = \sum_{i=1}^{n} E\left[(X_i - m_i)^2\right] = \sum_{i=1}^{n} D\left[X_i\right]. \tag{1.199}$$

(2) n 维随机变量的矩

1) 联合原点矩

已知二维随机变量 (X, Y) 的联合概率密度函数为 $f_{XY}(x, y)$, 定义二维随机变量 (X, Y) 的 $n + k$ 阶联合原点矩为

$$E\left[X^n Y^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n y^k f_{XY}(x, y) dx dy. \tag{1.200}$$

显然:

① $k = 0$ 时, $E\left[X^n\right]$ 是随机变量 X 的 n 阶原点矩。

② $n = 0$ 时, $E\left[Y^k\right]$ 是随机变量 Y 的 k 阶原点矩。

③ $E[XY]$ 是联合原点矩中最重要的一个, 它反映了 X 与 Y 两个随机变量间的关联程度, 称之为随机变量 X 和 Y 的互相关, 通常用 R_{XY} 表示, 如下所示:

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy = R_{XY}. \tag{1.201}$$

已知 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度函数为 $f_X(x_1, x_2, \cdots, x_n)$，定义 n 维随机变量的 $(k_1 + k_2 + \cdots + k_n)$ 阶联合原点矩为

$$E\left[X_1^{k_1} X_2^{k_2} \cdots X_n^{k}\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^k x_{2^2}^k \cdots x_k^k \cdot f_X(x_1, x_2, \cdots, x_n) \, \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n. \quad (1.202)$$

其中 $k_1, k_2, \cdots, k_n$ 为正整数。

2) 联合中心矩

定义 1.64 二维随机变量 (X, Y) 的 $n+k$ 阶联合中心矩

二维随机变量 (X, Y) 的 $n+k$ 阶联合中心矩为

$$E\left[(X - m_X)^n (Y - m_Y)^k\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)^n (y - m_Y)^k f_{XY}(x, y) \mathrm{d}x \mathrm{d}y. \quad (1.203)$$

♣

显然:

① $E\left[(X - m_X)^2\right] = \sigma_X^2$ 是随机变量 X 的二阶中心矩，即 X 的方差。

② $E\left[(Y - m_Y)^2\right] = \sigma_Y^2$ 是随机变量 Y 的二阶中心矩，即 Y 的方差。

③ $E\left[(X - m_X)(Y - m_Y)\right]$ 是联合中心矩中最重要的一个，它也反映了 X 与 Y 两个随机变量间的关联程度，称之为 X 和 Y 的协方差，通常用 C_{XY} 或 $\mathrm{Cov}(X, Y)$ 表示如下：

对于 n 维随机变量 $(X_1, X_2, \cdots, X_n)$，若其联合概率密度函数为 $f_X(x_1, x_2, \cdots, x_n)$。定义 n 维随机变量的 $(k_1 + k_2 + \cdots + k_n)$ 阶联合中心矩为

$$\begin{aligned} E\left[(X - m_X)(Y - m_Y)\right] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)(y - m_Y) f_{XY}(x, y) \mathrm{d}x \mathrm{d}y \\ &= C_{XY}. \end{aligned} \quad (1.204)$$

其中 $k_1, k_2, \cdots, k_n$ 为正整数。

例 1.65 已知二维随机变量 (X, Y) 的联合概率密度为

$$f_{XY}(x, y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{其他} \end{cases} \quad (1.205)$$

求：① 随机变量 X 的一、二阶原点矩。② 随机变量 X 的二、三阶中心矩。③ 联合原点矩 R_{XY}。④ 联合中心矩 C_{XY}。

解: 由二维随机变量的联合概率密度可得其边缘概率密度

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

$$E[X]=\frac{1}{2}, E\left[X^2\right]=\int_{-\infty}^{\infty}x^2 f_X(x)\mathrm{d}x=\int_0^1 x^2\mathrm{d}x=\frac{1}{3}.$$

$$E\left[(X-m_X)^2\right]=\int_0^1\left(x-\frac{1}{2}\right)^2\mathrm{d}x=\frac{1}{12}.$$

$$E\left[(X-m_X)^3\right]=\int_0^1\left(x-\frac{1}{2}\right)^3\mathrm{d}x=\frac{1}{32}.$$

$$R_{XY}=E[XY]=\int_0^1\int_0^1 xydxdy=\frac{1}{4}.$$

$$\begin{aligned}C_{XY}&=E\left[(X-m_X)(Y-m_y)\right]\\&=\int_0^1\int_0^1\left(x-\frac{1}{2}\right)\left(y-\frac{1}{2}\right)\mathrm{d}x\mathrm{d}y=0.\end{aligned}$$

2. 方差

(1) 二维随机变量的协方差

定义 1.66 二阶联合中心矩

若二维随机变量 (X,Y) 中关于 X 和 Y 的数学期望和方差均存在, 则称

$$E\left[(X-m_X)(Y-m_Y)\right]=C_{XY}. \tag{1.206}$$

为随机变量 X 与 Y 的协方差, 又称为“相关矩”或“二阶联合中心矩”。 ♣

它通常被用来表征两个随机变量间的关联程度。当 $C_{XY}=0$ 时, 称 X 与 Y 互不相关。协方差也可以通过下式进行计算:

$$C_{XY}=E\left[(X-m_X)(Y-m_y)\right]=E[XY]-m_Xm_Y=R_{XY}-m_Xm_Y. \tag{1.207}$$

当 X 与 Y 互不相关时, $C_{XY}=0$, 由上式可得

$$E[XY]=E[X]E[Y]. \tag{1.208}$$

3. 随机矢量的方差

定义 1.67 随机矢量 $\boldsymbol{X}$ 的方差

若将 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 用随机矢量 $\boldsymbol{X} = [X_1, X_2, \cdots, X_n]^{\mathrm{T}} \in \mathbb{R}^{n\times 1}$ 来表示, 则随机矢量 $\boldsymbol{X}$ 的方差定义为

$$D\boldsymbol{X} = E\left[(\boldsymbol{X} - E\boldsymbol{X})(X - E\boldsymbol{X})^{\mathrm{T}}\right]. \tag{1.209}$$

记作 $D(\boldsymbol{X}) = \mathrm{Var}\{\boldsymbol{X}\}$。 ♣

若将上式展开, 可得随机矢量的方差阵

$$\begin{aligned} D(\boldsymbol{X}) &= E\left\{\begin{bmatrix} X_1 - m_1 \\ X_2 - m_2 \\ \vdots \\ X_n - m_n \end{bmatrix}_{n\times 1} \times [X_1 - m_1 \quad X_2 - m_2 \quad \cdots \quad X_n - m_n]_{1\times n}\right\} \\ &= \begin{bmatrix} E\left[(X_1 - m_1)^2\right] & E\left[(X_1 - m_1)(X_2 - m_2)\right] & \cdots & E\left[(X_1 - m_1)(X_n - m_n)\right] \\ E\left[(X_2 - m_2)(X_1 - m_1)\right] & E\left[(X_2 - m_2)^2\right] & \cdots & E\left[(X_2 - m_2)(X_n - m_n)\right] \\ \vdots & \vdots & & \vdots \\ E\left[(X_n - m_n)(X_1 - m_1)\right] & E\left[(X_n - m_n)(X_2 - m_2)\right] & \cdots & E\left[(X_n - m_n)^2\right] \end{bmatrix}. \end{aligned} \tag{1.210}$$

若 n 维变量 $(X_1, X_2, \cdots, X_n)$ 中每个变量之间的协方差 $C_{ij} = E\left[(X_i - m_i)\left(X_j - m_j\right)\right]$ (包括自身的方差) 均存在, 则

$$D(X) = \begin{bmatrix} D\left[X_1\right] & C_{12} & \cdots & C_{1n} \\ C_{21} & D\left[X_2\right] & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & D\left[X_n\right] \end{bmatrix} = C_X, \tag{1.211}$$

由式 (1.211) 可见, 矢量 $\boldsymbol{X}$ 的方差阵中除对角线上均为各分量的方差外, 其余均为分量之间的协方差 $\boldsymbol{C}_{ij}\ (i \neq j)$。若将方差 $D\left[\boldsymbol{X}_i\right]$ 看成协方差的特例 $C_{ij}\ (i = j)$, 将方差阵中的所有元素均用协方差 C_{ij} 表示, 则有

$$D(\boldsymbol{X}) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n\times n} = \boldsymbol{C}_X, \tag{1.212}$$

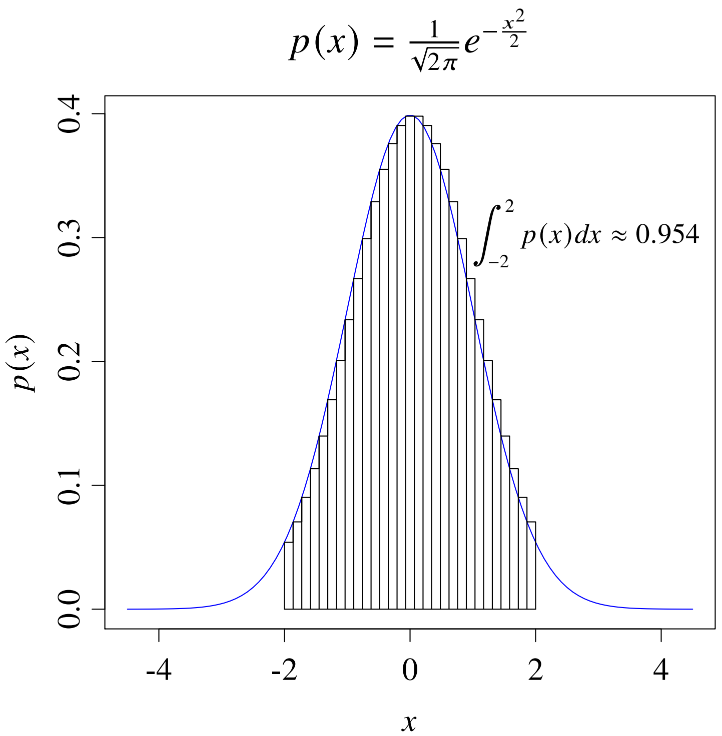

图 1–21 (X, Y) 取值的直方图

又可称为 n 维随机变量的协方差矩阵。

由 $C_{ij} = C_{ji}$, 可得 $C_x = C_X^{\mathrm{T}}$, 随机矢量 $\boldsymbol{X}$ 的方差阵为对称矩阵。

1.1.8 相关、正交和独立

1. 相关系数

(1) 两个随机变量相互关系的描述

在实际中, 为了将两个随机变量 X 和 Y 的相互关系表示出来, 最简单的方法是将两个随机变量的所有取值情况在 xOy 以平面上画出来, 称之为随机变量 X 和 Y 取值的散布图 (图 1–22)。若 X 与 Y 独立, X 的取值与 Y 的取值没有任何关系, 则 (X, Y) 在 xOy 平面上没有样本点存在。若 X 与 Y 相关, 对于 X 取的每一确定值, 都有的取值与其相对应。因此, (X, Y) 在 xOy 平面上有样本点存在。

所以散布图直观反映了 X 与 Y 之间的相关情况。如果要从数学上表示 X 与 Y 的函数关系, 则要设法找到逼近其散布点密集分布的一条回归线。如果这条回归线是直线, 我们就说 X 与 Y 线性相关; 如果这条回归线是曲线, 我们就说 X 与 Y 非线性相关。

最常用的方法是用一根直线去逼近 (X, Y) 取值散布点的密集分布, 即寻找最逼近

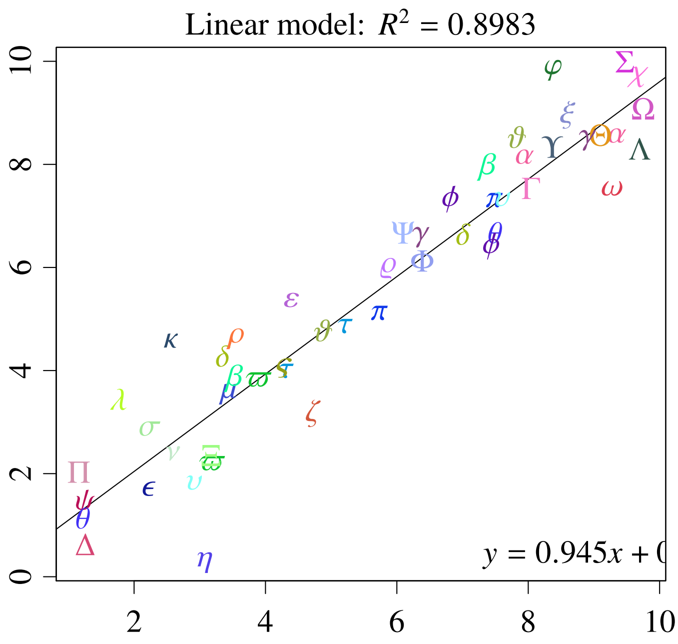

图 1–22 (X,Y) 取值的散布图

(X,Y) 关系的直线 (线性回归)。在这条直线上, 可由 X 的取值预测 Y 的取值, 即

$$Y_p = a + bX, \tag{1.213}$$

当然 Y 并不就是 Y, 而是根据 X 的取值所得到的线性相关预测值。

如果 a,b 构造的是回归线, 那么一定会使得 Y 与 Y_p 的误差最小。

寻找使均方误差 $\varepsilon = E\left[\left(Y - Y_p\right)^2\right]$ 最小的 a,b 的值的方法如下

$$\begin{aligned}
\varepsilon = E\left[\left(Y - Y_p\right)^2\right] &= E\left[(Y - a - bX)^2\right] \\
&= E\left[Y^2 + b^2X^2 + a^2 + 2abX - 2aY - 2bXY\right] \\
&= a^2 - 2aE[Y] + 2abE[X] - 2bE[XY] + b^2E\left[X^2\right] + E\left[Y^2\right].
\end{aligned} \tag{1.214}$$

令

$$\begin{cases} \frac{\partial\varepsilon}{\partial a} = 2a - 2m_Y + 2bm_X = 0 \\ \frac{\partial\varepsilon}{\partial b} = 2am_X - 2R_{XY} + 2bE\left[X^2\right] = 0 \end{cases}. \tag{1.215}$$

a,b 的值为

$$\begin{cases} a &= m_Y - bm_X = m_Y - \frac{C_{XY}}{\sigma_X^2 m_X}. \\ b &= \frac{R_{XY} - m_X m_Y}{E\left[X^2\right] - m_X^2} = \frac{C_{XY}}{\sigma_X^2}. \end{cases} \tag{1.216}$$

将其代入式 (1.214) 中, 可验证给出的 ε 是极小值, 从而得到最好的预测直线方程为

$$Y_p = a + bX = m_Y + \frac{C_{XY}}{\sigma_X^2}(X - m_X). \tag{1.217}$$

由式 (1.217) 可见, 最佳预测线通过 (m_X, m_Y) 点。通过这条直线, 可以根据 X 的取值, 给出 Y 的最佳预测值 Y, 即这条直线最能反映 X 与 Y 之间的线性关系。

为了简便, 引入 X 和 Y 的归一化随机变量

$$X = \frac{X - m_X}{\sigma_X}, Y = \frac{Y_p - m_Y}{\sigma_Y}. \tag{1.218}$$

式中斜率 $\frac{C_{XY}}{\sigma_X \sigma_Y}$ 的大小反映出随机变量 X 与 Y 之间线性相关的程度。

(2) 相关系数的定义

定义 1.68 相关系数

用来表征两随机变量之间线性相关程度的量

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}. \tag{1.219}$$

称之为相关系数. ♣

(3) 相关系数的性质

1° $-1 \leqslant \rho_{XY} \leqslant 1$, 即 $|\rho_{XY}| \leqslant 1$.

证: 为了使证明简化, 这里研究两个归一化随机变量的相关系数。令

$$U = \frac{X - m_X}{\sigma_X}, V = \frac{Y - m_Y}{\sigma_Y}. \tag{1.220}$$

显然有 $E[U] = E[V] = 0, D[U] = D[V] = 1, E\left[V^2\right] = 1$ 成立。由

$$\begin{aligned} &E^2[U \pm V] = \{E[U] \pm E[V]\}^2 = 0. \\ &E\left[(U \pm V)^2\right] = E\left[U^2\right] \pm 2E[UV] + E\left[V^2\right] = 2 \pm 2E[UV]. \\ &\rho_{UV} = \tfrac{C_{UV}}{\sigma_U \sigma_V} = \tfrac{R_{UV} - m_U m_V}{\sigma_U \sigma_V} = R_{UV} = E[UV]. \end{aligned} \tag{1.221}$$

把上述三式代入式 (1.211), 可得

$$D[U \pm V] = E\left[(U \pm V)^2\right] - E^2[U \pm V] \geqslant 0. \tag{1.222}$$

2° 若 X 与 Y 间以概率 1 存在线性关系, 即满足 $P\{Y = aX + b\} = 1$ (a 和 b 为实常数) 时, 则有 $|\rho_{XY}| = 1$。

3° 若 X 与 Y 线性不相关, 则有 $|\rho_{XY}| = 0$。等价于

$$\begin{cases} C_{XY} = 0. \\ R_{XY} = E[X]E[Y]. \\ D[X \pm Y] = D[X] + D[Y]. \end{cases} \tag{1.223}$$

注: 通常说的"互不相关"均指的是线性互不相关。

推广到 n 维的情况。当 n 维随机变量 $X_1, X_2, \cdots, X_n$ 之间互不相关时, 由于 $C_{ij} = 0\ (i \neq j)$, 则 n 维随机变量的协方差矩阵为对角线矩阵, 如下所示

$$C_X = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix}. \tag{1.224}$$

相关系数矩阵

$$\rho_X = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{mn} \end{bmatrix}. \tag{1.225}$$

当 $X_1, X_2, \cdots, X_n$ 互不相关时, 由于 $\rho_{ij} = 0\ (i \neq j)$ 和 $\rho_{ij} = 1\ (i = j)$ 成立, 则相关系数矩阵为单位矩阵。

2. 不相关、独立和正交

(1) 正交的定义

定义 1.69 X 与 Y 正交

若随机变量 X 与 Y 满足 $R_{XY} = E[XY] = 0$, 则称随机变量 X 与 Y 正交。♣

考虑到 $C_{XY} = R_{XY} - m_X m_Y$, 则 X 与 Y 正交满足

$$C_{XY} = -m_X m_Y. \tag{1.226}$$

(2) 不相关、独立和正交的关系

1° 若随机变量 X 与 Y 相互独立, 则 X 与 Y 必定互不相关。

证: 当 X 与 Y 相互独立时, 满足

$$f_{XY}(x,y) = f_X(x) f_Y(y). \tag{1.227}$$

$$\begin{aligned} C_{XY} &= \int_{-\infty}^{\infty} (x-m_X) f_X(x)\mathrm{d}x \cdot \int_{-\infty}^{\infty} (y-m_Y) f_Y(y)\mathrm{d}y \\ &= E\left[(X-m_X)\right] \cdot E\left[(Y-m_Y)\right] = 0. \end{aligned} \tag{1.228}$$

即 X 与 Y 互不相关。

2° 若随机变量 X 与 Y 互不相关 ($\rho_{XY}=0$), 则 X 与 Y 不一定相互独立。

因为独立表示两个随机变量之间既线性不相关, 又非线性不相关。而相关系数 $\rho_{XY}=0$ 仅仅表征两个随机变量线性不相关, 而不能说明它们之间非线性也不相关。

例 1.70 已知随机变量 $X=\cos\varphi$ 和 $Y=\sin\varphi$, 式中 φ 是在 $(0,2\pi)$ 上均匀分布, 讨论 X 和 Y 的相关性及独立性。

解:

① 因为 φ 是在 $(0,2\pi)$ 上均匀分布的随机变量, 其概率密度为

$$f(\varphi) = \begin{cases} \frac{1}{2\pi}, & 0<\varphi<2\pi \\ 0, & \text{其他} \end{cases}. \tag{1.229}$$

则

$$E[X] = \int_{-\infty}^{\infty} \cos\varphi f(\varphi)\mathrm{d}\varphi = \frac{1}{2\pi}\int_0^{2\pi} \cos\varphi \mathrm{d}\varphi = 0.$$

$$E[Y] = \int_{-\infty}^{\infty} \sin\varphi f(\varphi)\mathrm{d}\varphi = \frac{1}{2\pi}\int_0^{2\pi} \sin\varphi \mathrm{d}\varphi = 0.$$

$$E[XY] = E[\cos\varphi\sin\varphi] = \frac{1}{2}E[\sin 2\varphi] = 0.$$

$$C_{XY} = E\left[(X-m_X)(Y-m_Y)\right] = E[XY] = 0.$$

即相关系数 $\rho_{XY}=0$, 表明 X 与 Y 是互不相关的。

② 由于随机变量 X 与 Y 存在 $X^2+Y^2=1$ 的关系 (非线性相关), Y 的取值依赖于 X 的取值, 所以随机变量 X 与 Y 之间相互不独立。

3° 若两个随机变量的联合矩对任意 $n\geq 1$ 和 $k\geq 1$ 均可分解为

$$E\left[X^nY^k\right] = E\left[X^n\right]\cdot E\left[Y^k\right], \tag{1.230}$$

则 X 与 Y 统计独立。

4° 当随机变量 X 与 Y 之间存在线性函数关系 $Y=aX+b$ 时, 则有 $\rho_{XY}=\pm 1$ 当 X 与 Y 间存在非线性函数关系时, 则有 $0<|\rho_{XY}|<1$。

例 1.71 已知随机变量 X 与 Y 有非线性关系 $Y = X^3$, 且

$$E\left(X^n\right) = \begin{cases} 1\cdot 3\cdot 5\cdots\cdots(n-1)\sigma_X', & n\text{为偶数} \\ 0, & n\text{为奇数} \end{cases}, n \geqslant 2\,. \tag{1.231}$$

求相关系数 ρ_{XY}。

解: 由于

$$m_X = E[X] = 0, m_Y = E\left[X^3\right] = 0. \tag{1.232}$$

则协方差为

$$C_{XY} = R_{XY} - m_X m_Y = R_{XY} = E[XY] = E\left[XX^3\right] \ = E\left[X^4\right] = 3\sigma_X^4.$$

相关系数为

$$\rho_{XY} = \frac{3\sigma_X^4}{\sigma_X\sigma_Y} = \frac{3\sigma_X^4}{\sqrt{E\left[X^2\right]\cdot\sqrt{E\left[X^6\right]}}} = \frac{3\sigma_X'}{\sqrt{15\sigma_X^8}} = \frac{3}{\sqrt{15}} = 0.775 < 1. \tag{1.233}$$

5° 针对一般情况而言, 两个随机变量正交不能保证它们不相关; 反之, 两个变量不相关也不能保证正交。由 $C_{XY} = R_{XY} - m_X m_Y$ 可以看出, 若 $m_X = 0$ 或 $m_Y = 0$ 其中一个成立, 则不相关和正交等价。

例 1.72 已知随机变量 X 的均值 $m_X = 3$, 方差 $\sigma_X^2 = 2$, 且另一随机变量 $Y = -6X + 22$。讨论 X 与 Y 的相关性与正交性。

解: 由题可知 X 的均值 $m_X = 3$, 所以 Y 的均值为

$$m_Y = E[-6X + 22] = -6E[X] + 22 = 4. \tag{1.234}$$

① X 与 Y 的互相关为

$$\begin{aligned} R_{XY} &= E[XY] = E\left[-6X^2 + 22X\right] = -6E\left[X^2\right] + 22E[X] \\ &= -6\left(m_X^2 + \sigma_X^2\right) + 22m_X = 0. \end{aligned} \tag{1.235}$$

可见 X 与 Y 是正交的。

由于 $\sigma_X^2 = 2$ 且 $\sigma_Y^2 = E\left[(Y - m_Y)^2\right] = E\left[(-6X + 18)^2\right] = 72$, 所以相关系数为

$$\rho_{XY} = \frac{R_{XY} - m_X m_y}{\sqrt{\sigma_X^2\sigma_Y^2}} = \frac{-12}{\sqrt{144}} = -1. \tag{1.236}$$

这说明 X 与 Y 又是线性相关的。

1.1.9 随机变量的特征函数

随机变量的数学期望、方差等数字特征虽可以反映其概率分布的某些特征, 但一般无法通过它们来确定其概率分布。而随机变量的特征函数既可以确定其概率分布又具有良好的分析性质, 它是概率论中最重要的分析工具之一。它的优越性, 首先在于特征函数与概率密度唯一对应。其次, 由特征函数计算统计特性比由概率密度函数计算更为方便。例如, 由概率密度计算矩需要积分运算, 而由特征函数计算矩则只要微分运算; 由概率密度求独立随机变量和的分布, 需要各个概率密度进行卷积, 而由特征函数求独立随机变量和的特征函数, 只需要各个特征函数相乘即可。这些优点将会在下面的介绍中逐一体现。

1. 一元特征函数及其性质

定义 1.73 随机变量 X 的特征函数

设 X 是定义在概率空间 (Ω, F, P) 上的随机变量, 其概率密度函数为 $f(x)$, 则其函数 e^{juX} 的数学期望为

$$Q_X(u) = E\left[\mathrm{e}^{juX}\right], \mathrm{j} = \sqrt{-1}, -\infty < u < \infty. \tag{1.237}$$

称之为随机变量 X 的特征函数。 ♣

① 当 X 为连续型随机变量时, 其特征函数为

$$Q_X(u) = \int_{-\infty}^{\infty} e^{jux} f(x)\mathrm{d}x. \tag{1.238}$$

② 当 X 为离散型随机变量时, 其特征函数为

$$Q_X(u) = \sum_i \mathrm{e}^{jux_i} P(X = x_i) = \sum_i \mathrm{e}^{jux_i} p_i. \tag{1.239}$$

(2) 性质

1° $|Q_X(u)| \leqslant Q_X(0) = 1$

证: 由特征函数的定义

$$|Q_X(u)| = \left|\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{j}ux} f(x)\mathrm{d}x\right| \leq \int_{-\infty}^{\infty} \left|\mathrm{e}^{\mathrm{j}ux} f(x)\right| \mathrm{d}x. \tag{1.240}$$

由于 $f(x) \geqslant 0$ 且. $\left|\mathrm{e}^{\mathrm{j}ux}\right| = 1$, 可得

$$|Q_X(u)| \leqslant \left|\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{j}ux} f(x)\mathrm{d}x\right| = \int_{-\infty}^{\infty} \left|\mathrm{e}^{\mathrm{j}ux} f(x)\right| \mathrm{d}x = Q_X(0) = 1. \tag{1.241}$$

由于 $|Q_X(u)| \leqslant 1$, 说明特征函数 $Q_X(u)$ 在一切实数 u 上都有定义。

2° 特征函数 $Q_X(u)$ 是实变量 u 在 $(-\infty, +\infty)$ 上的连续函数。

3° 特征函数 $Q_X(u)$ 是实变量 u 的复函数, 有

$$Q_X^*(u) = Q_X(-u). \tag{1.242}$$

4° 随机变量 X 的函数 $y = aX + b$ 的特征函数

$$Q_Y(u) = \mathrm{e}^{\mathrm{j}ub} Q_X(au). \tag{1.243}$$

证:

$$\begin{aligned} Q_Y(u) &= E\left[\mathrm{e}^{\mathrm{j}uY}\right] = E\left[\mathrm{e}^{\mathrm{ju}(aX+b)}\right] = E\left[\mathrm{e}^{\mathrm{j}uaX} \cdot \mathrm{e}^{\mathrm{j}ub}\right] \\ &= \mathrm{e}^{\mathrm{j}ub} E\left[\mathrm{e}^{\mathrm{juaX}}\right] = \mathrm{e}^{\mathrm{j}ub} Q_X(au). \end{aligned}$$

5° 相互独立随机变量和的特征函数等于它们特征函数的积

若随机变量 $X_1, \cdots, X_n$ 相互独立, 且各自相应的特征函数为 $Q_{X_1}(u), Q_{X_2}(u), \cdots, Q_{X_n}(u)$. 随机变量 $Y = \sum_{k=1}^{n} X_i$, 则其特征函数为

$$Q_Y(u) = \prod_{i=1}^{n} Q_{X_i}(u). \tag{1.244}$$

证

$$Q_Y(u) = E\left[\mathrm{e}^{\mathrm{ju}Y}\right] = E\left[\mathrm{e}^{\mathrm{j}u(X_1+X_2+\cdots+X_n)}\right] = E\left[\prod_{i=1}^{n} e^{juX_i}\right]. \tag{1.245}$$

由于 $X_1, \cdots, X_n$ 相互独立, 根据数学期望的性质 5°，可得

$$Q_Y(u) = E\left[\mathrm{e}^{juY}\right] = E\left[\mathrm{e}^{\mathrm{j}u(X_1+X_2+\cdots+X_n)}\right] = \prod_{i=1}^{n} Q_{X_i}(u). \tag{1.246}$$

6° 若随机变量 X 的 n 阶绝对矩存在, 则它的特征函数存在 n 阶导数。并且当 $1 \leqslant k \leqslant n$ 时，有

$$E\left[X^k\right] = (-\mathrm{j})^k \left.\frac{\mathrm{d}^k Q_X(u)}{\mathrm{d}u^k}\right|_{u=0}. \tag{1.247}$$

可见, 利用这一性质由特征函数 $Q_X(a)$, 求随机变量 X 的 k 阶矩, 只要对特征函数 $Q_X(u)$ 求 k 次导数 (微分) 即可, 显然这比由概率密度 $f(x)$ 求 k 阶矩的积分运算要简单得多。

7° 若随机变量 X 的特征函数 $Q_X(u)$ 可以展开成麦克劳林级数, 其特征函数可由该随机变量 X 的各阶矩唯一地确定, 即

$$Q_X(u) = \sum_{n=0}^{\infty} E\left[X^n\right] \frac{(ju)^n}{n!}. \tag{1.248}$$

这个性质常用在理论推导中。实际应用中由随机变量的各阶矩 $E\left(X^n\right)(n=1,2,\cdots)$ 来求 $Q_X(u)$ 不现实，特别是某些随机变量的各阶矩并不一定都存在，如柯西分布。

例 1.74 求下列高斯变量的特征函数：① 标准高斯变量 $Y \sim N(0,1)$。① 高斯变量 $X \sim N(m,\sigma^2)$。

解：先介绍一个从复变函数中得到的积分公式

$$\int_{-\infty}^{\infty} \mathrm{e}^{-Ax^2+2Bx-C}\mathrm{d}x = \sqrt{\pi/A}\,\mathrm{e}^{-\frac{\left(AC-B^2\right)}{A}}. \tag{1.249}$$

① 根据特征函数的定义

$$Q_Y(u) = E\left[\mathrm{e}^{\mathrm{j}\omega Y}\right] = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega y}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{y^2}{2}+2\frac{\mathrm{j}u}{2}}y\mathrm{d}y. \tag{1.250}$$

令 $A=\frac{1}{2}, B=\frac{ju}{2}, C=0$，则利用积分公式可得标准高斯变量的特征函数如下

$$Q_Y(u) = \frac{1}{\sqrt{2\pi}}\left\{\sqrt{\frac{\pi}{1/2}}\cdot\exp\left[\left(\frac{ju}{2}\right)^2\Big/\frac{1}{2}\right]\right\} = \mathrm{e}^{-\frac{u^2}{2}}. \tag{1.251}$$

② 将高斯变量 X 归一化处理得

$$X = \frac{X-m}{\sigma} = Y, X = \sigma_Y^+. \tag{1.252}$$

利用性质 4 及 ① 的结论，得一般高斯变量的特征函数如下

$$Q_X(u) = \mathrm{e}^{\mathrm{j}mu}Q_Y(\sigma u) = \exp\left(jum - \frac{1}{2}\sigma^2u^2\right). \tag{1.253}$$

例 1.75 求二项分布的数学期望、方差和特征函数。

解：方法一：

二项分布的分布律为

$$P\{Y=k\} = C_n^k p^k q^{n-k}, k=0,1,2,\cdots,n. \tag{1.254}$$

由数学期望的定义可得

$$E[Y] = \sum_{k=1}^{n} kP\{Y=k\} = \sum_{k=1}^{n} kC_n^k p^k(1-p)^{n-k} = \cdots = np. \tag{1.255}$$

由方差的定义，可得

$$\begin{aligned} D[Y] &= \sum_{k=1}^{n}(k-m_Y)^2\cdot P\{Y=k\} \\ &= \sum_{k=1}^{n}(k-np)^2\cdot C_n^k p^k(1-p)^{n-k} = \cdots = npq. \end{aligned} \tag{1.256}$$

直接引用特征函数的定义, 可得

$$Q_Y(u)=\sum_{k=1}^{n}\mathrm{e}^{j\omega k}P\{Y=k\}=\sum_{k=1}^{n}e^{juk}C_n^k p^k q^{n-k}=\cdots \tag{1.257}$$

计算较为繁杂。

方法二: 二项分布的随机变量 Y 代表的是 n 重伯努利试验中随机事件发生的次数。而在每次试验中, 随机事件发生的概率为 p, 随机事件不发生的概率为 $1-p=q$, 每次试验中事件发生的次数 x 均服从 $(0,1)$ 分布。$(0,1)$ 分布的特征函数为

$$Q_{X_i}(u)=E\left[\mathrm{e}^{\mathrm{j}uX_i}\right]=\sum_{k=0}^{1}\mathrm{e}^{\mathrm{j}uk}P\left(X_i=k\right)=q+p\mathrm{e}^{\mathrm{j}u}. \tag{1.258}$$

由于 Y 代表 n 重伯努利试验中随机事件发生的次数, 所以有 $Y=\sum_{i=1}^{n}X_i$。且 n 重伯努利试验中各次试验均相互独立, 由性质 5° 可得二项分布的特征函数

$$Q_Y(u)=\prod_{i=1}^{n}Q_{X_i}(u)=\left(q+p\mathrm{e}^{\mathrm{j}u}\right)^n. \tag{1.259}$$

可见, 离散型随机变量的特征函数也是实数的连续函数。利用性质 6° 得

$$\begin{aligned}
&E[Y]=-\mathrm{j}\tfrac{\mathrm{d}}{\mathrm{d}u}\left(p\mathrm{e}^{\mathrm{j}u}+q\right)^n\Big|_{u=0}=np.\\
&E\left[Y^2\right]=(-\mathrm{j})^2\tfrac{\mathrm{d}^2}{\mathrm{d}u^2}\left(p\mathrm{e}^{\mathrm{j}u}+q\right)^n\Big|_{u=0}=npq+n^2p^2.\\
&D[Y]=E\left[Y^2\right]-(E[Y])^2=npq.
\end{aligned} \tag{1.260}$$

2. 特征函数与概率密度的对应关系

由特征函数的定义可知, 已知随机变量 X 的概率密度 $f(x)$ 总可以求得它的特征函数 $Q_X(n)$, 那么由特征函数能否求得概率密度呢? 在讨论之前, 先回顾一下傅立叶变换对

$$\begin{cases}S(\omega)=F[s(t)]=\int_{-\infty}^{\infty}s(t)\mathrm{e}^{-\mathrm{j}\omega t}\mathrm{d}t\\ s(t)=F^{-1}[S(\omega)]=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)\mathrm{e}^{\mathrm{j}\omega t}\mathrm{d}\omega\end{cases}. \tag{1.261}$$

对照特征函数的定义, 可得

$$Q_X(u)=\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{jux}}\mathrm{d}x=2\pi\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{jux}}\mathrm{d}x\right]=2\pi F^{-1}[f(x)]. \tag{1.262}$$

即

$$F^{-1}[f(x)]=\frac{1}{2\pi}Q_X(u). \tag{1.263}$$

则可用傅立叶变换求

$$f(x)=F\left[\frac{1}{2\pi}Q_X(u)\right]=\frac{1}{2\pi}\int_{-\infty}^{+\infty}Q_X(u)\mathrm{e}^{-jux}du. \tag{1.264}$$

称之为**逆转公式**。

由傅立叶变换对之间的唯一性可知, $Q_X(u)$ 与 $f(x)$ 之间也是唯一对应的。有了特征函数与概率密度的变换关系, 通过特征函数间接求独立随机变量和的分布, 比直接用概率密度 $f(x)$ 来求要方便得多。

例 1.76 求 $Y=\sum_{i=1}^{n} X$ 的分布 $f_X(y)$ 时, 若用概率密度直接求, 则必须对 $f_X(x_1,\cdots,x_n)$ 做 $n-1$ 重积分, 即

$$f_Y(y)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_X\left(x_1\cdots x_{n-1}, y-\sum_{i=1}^{n-1} x_i\right) dx_1\cdots dx_{n-1}. \tag{1.265}$$

即使当 $X_1,X_2,\cdots,X_n$ 相互独立时, 也要做 $n-1$ 个积分; 但用特征函数间接求解时, 只要先求 $Q_Y(u)=\prod_{-1}^{n} Q_{X_i}(u)$, 再由逆转公式即可求出 $f_Y(y)$.

例 1.77 证明: 若随机变量 $X_1\sim B(n,p)$ 和 $X_2\sim B(m,p)$ 之间相互独立, 则有 $Y=X_1+X_2\sim B(n+m,p)$.

证: 由特征函数性质 5° 可得

$$\begin{aligned} Q_Y(u)&=Q_{X_1}(u)Q_{X_2}(u)=\left(pe^{i\omega}+q\right)^n\cdot\left(pe^{i*}+q\right)^m \\ &=\left(pe^{i\mu}+q\right)^{n+m}. \end{aligned} \tag{1.266}$$

而右端正是二项分布 $B(m+n,p)$ 的特征函数, 由特征函数和概率密度对应关系的唯一性, 可知 $Y\sim B(n+m,p)$, 即二项分布关于参数 n 具有再生性。

3. 多元特征函数

下面由一维随机变量的特征函数推广出多维随机变量的特征函数。

(1) 定义

若 n 维随机变量 $(X_1,\cdots,X_n)$ 用随机矢量 $\boldsymbol{X}$ 表示, 取值 $(x_1,\cdots,x_n)$ 用矢量 $\boldsymbol{X}$ 表示, n 个参变量 $(u_1,\cdots+u_n)$ 用矢量 U 表示, 即

$$\boldsymbol{X}=\begin{bmatrix} X_1\\ \vdots\\ X_n\end{bmatrix}, \boldsymbol{X}_i=\begin{bmatrix} x_1\\ \vdots\\ x_n\end{bmatrix}, \boldsymbol{U}=\begin{bmatrix} u_1\\ \vdots\\ u_n\end{bmatrix} \tag{1.267}$$

则 n 维随机变量 $(X_1,\cdots,X_n)$ 的联合特征函数 (或随机矢量 $\boldsymbol{X}$ 的特征函数) 为

$$\begin{aligned} Q_X(u_1,\cdots,u_n)&=Q_X\left(\boldsymbol{U}^{\mathrm{T}}\right)=E\left[e^{j\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}}\right] \\ &=E\left[e^{j(u_1X_1+u_2X_2+\cdots+u_nX_n)}\right]=E\left[e^{\sum_{k=1}^{n} j u_k X_k}\right] \\ &=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{j(u_1x_1+\cdots+u_nx_n)} f_X(x_1,\cdots,x_n)\,dx_1\cdots dx_n. \end{aligned} \tag{1.268}$$

逆转公式为

$$
\begin{aligned}
f_X(x_1,\cdots,x_n) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-jU^T x} Q_X\left(U^T\right)\frac{du_1}{2\pi}\cdots\frac{du_n}{2\pi} \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-j\sum\limits_{k=1}^{n} u_k x_k} Q_X(u_1,\cdots,u_n)\frac{du_1}{2\pi}\cdots\frac{du_n}{2\pi}.
\end{aligned}
\tag{1.269}
$$

(2) 性质

1° $|Q_X(u_1,\cdots,u_n)| \leqslant Q_X(0,\cdots,0) = 1.$

2° 特征函数 $Q_X(u_1,\cdots,u_n)$ 在 R^n 中一致连续, 其中 R^n 表示 $u_1,\cdots,u_n$ 的 n 维空间。

3° 特征函数 $Q_X(u_1,\cdots,u_n)$ 是关于实变量 $u_1,\cdots,u_n$ 的复函数, 有

$$
Q_X^*(u_1,\cdots,u_n) = Q_X(-u_1,\cdots,-u_n). \tag{1.270}
$$

4° 若 $Q_X(u_1,\cdots,u_n)$ 是随机矢量 $\boldsymbol{X}$ 的特征函数, 矩阵 $\boldsymbol{A}$ 是 $r\times n$ 常系数矩阵, 矢量 $\boldsymbol{B}$ 是 $r\ (r<n)$ 维常数矢量, 则随机矢量 $\boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{B}$ 的特征函数为

$$
Q_Y(u_1,\cdots,u_r,0,\cdots,0) = \mathrm{e}^{jU^T B'} Q_X\left(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}'\right), \tag{1.271}
$$

其中 $\boldsymbol{B}' = \begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{0} \end{bmatrix}_{n\times 1}$, $\boldsymbol{0}_{(n-r)\times 1}$ 为补充的零向量。$A' = \begin{bmatrix} A_{r\times n} \\ 0_{(n-r)\times n} \end{bmatrix}_{n\times n}$, $\boldsymbol{0}_{(n-r)\times n}$ 为补充的零矩阵。

下面讨论两种特殊情况:

当 $r=n$ 时, 矩阵 $\boldsymbol{A}$ 是 $n\times n$ 对角矩阵, 矢量 $\boldsymbol{B}$ 是 n 维矢量, 即

$$
\boldsymbol{A} = \begin{bmatrix} a_1 & 0 & \\ & \ddots & \\ 0 & & a_n \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \tag{1.272}
$$

此时

$$
\boldsymbol{AX} = \begin{bmatrix} a_1X_1 \\ \vdots \\ a_nX_n \end{bmatrix}, \boldsymbol{Y} = \boldsymbol{AX}+\boldsymbol{B} = \begin{bmatrix} a_1X_1+b_1 \\ \vdots \\ a_nX_n+b_n \end{bmatrix}, \tag{1.273}
$$

则随机矢量 $\boldsymbol{Y}$ 的特征函数为

$$
Q_Y(u_1,u_2,\cdots,u_n) = \mathrm{e}^{\mathrm{j}\sum_{k=1}^{n} u_k b_k} Q_X(a_1u_1,\cdots,a_nu_n). \tag{1.274}
$$

当 $r=1$ 时, $\boldsymbol{A}=(a_{ij})$ 是 $1\times n$ 矩阵, 而 B 是一常数 b, 则

$$
\boldsymbol{Y} = \boldsymbol{AX}+\boldsymbol{B} = a_1X_1+\cdots+a_nX_n+b. \tag{1.275}
$$

是一维随机变量, 一元函数情形的特征函数为

$$Q_Y(u_1) = e^{ju_1 b} Q_X(a_1 u_1, a_2 u_1, \cdots, a_n u_1). \tag{1.276}$$

5° 若 n 维随机变量 $(X_1, \cdots, X_n)$ 的联合特征函数为 $Q_X(u_1, \cdots, u_n)$, 其分量 X 的特征函数为 $Q_{X_k}(u_k), k = 1, 2, \cdots, n$, 则 $X_1, \cdots, X_n$ 相互独立的充要条件为

$$Q_X(u_1, \cdots, u_n) = \prod_{k=1}^{n} Q_{X_k}(u_k). \tag{1.277}$$

由逆转公式可证

$$f_X(x_1, \cdots, x_n) = \prod_{k=1}^{n} f_{x_k}(x_k). \tag{1.278}$$

思考: 一元特征函数性质 5° 和多元特征函数性质 5° 有什么不同?

6° 若随机矢量 $\boldsymbol{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ 的特征函数为 $Q_x(u_1, \cdots, u_n)$, 则其子向量 $\boldsymbol{Y} = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix}$ 的特征函数为

$$Q_Y(u_1, \cdots, u_k) = Q_X(u_1, \cdots, u_k, 0, \cdots, 0). \tag{1.279}$$

称为 n 维随机变量的边缘特征函数，其中 $k < n$。

同理, 任取 k 个分量组成的子向量 $\boldsymbol{Y}_i = \begin{bmatrix} X_{11} \\ \vdots \\ X_{ik} \end{bmatrix}$, 类似也可得到其边缘特征函数。

7° 若矩 $E\left[X_1^n X_2^k\right]$ 存在, 则

$$E\left[X_1^n X_2^k\right] = (-\mathrm{j})^{n+k} \left.\frac{\partial^{n+k} Q_X(u_1, u_2)}{\partial u_1^n \partial u_2^k}\right|_{u_1=0, u_2=0}. \tag{1.280}$$

8° 若对所有 $n = 0, 1, 2, \cdots$ 和所有 $k = 0, 1, 2, \cdots, E\left[X_1^n X_2^k\right]$ 均存在, 则

$$Q_X(u_1, u_2) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} E\left[X_1^n X_2^k\right] \cdot \frac{(ju_1)^n}{n!} \cdot \frac{(ju_2)^k}{k!}. \tag{1.281}$$

1.1.10 高斯随机变量

在实际应用中, 常常遇到大量随机变量和的问题。中心极限定理已证明, 在满足一定条件的情况下, 大量随机变量和的极限分布是高斯分布。因此, 高斯分布占有特殊的地位, 是科学技术领域中最常遇到的分布, 也是无线电技术理论 (包括噪声理论、信号检测

理论、信息理论等) 中最重要的概率分布实际中有许多随机变量，它们是由大量相互独立的随机因素共同作用所形成的，而其中每一个因素的单独作用都很微小。因此高斯随机变量是一类被广泛应用的随机变量。

1. 一维高斯随机变量及其特征

一维高斯随机变量 X 的概率密度函

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right], x\in\mathbb{R}, \tag{1.282}$$

式中，m_X 为 X 的均值，σ_X 为 X 的方差。

如果 $m_X=0,\sigma_X=1$, 则

$$f(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right),\quad -\infty<x<+\infty. \tag{1.283}$$

对于一给定的 x, 高斯随机变量 X 的分布函数为

$$\begin{aligned} F(x)&=P(X\leqslant x)\\ &=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left[-\frac{(u-m_X)^2}{2\sigma_X^2}\right]\mathrm{d}u. \end{aligned} \tag{1.284}$$

令 $\frac{u-m_X}{\sigma_X}=t$, 则

$$F(x)=\int_{-\infty}^{(x-m_X)/\sigma_X}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t=\Phi\left(\frac{x-m_X}{\sigma_X}\right). \tag{1.285}$$

其中 $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t$ 概率积分函数。

高斯分布函数还可以表示为

$$\begin{aligned} F(x)&=\frac{1}{2}+\frac{1}{2}\operatorname{erf}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right).\\ F(x)&=1-\frac{1}{2}\operatorname{erfc}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right). \end{aligned} \tag{1.286}$$

其中 $\operatorname{erf}(x)$ 称为误差函数，$\operatorname{erfc}(x)$ 称为互补误差函数，它们分别为

$$\begin{aligned} &\operatorname{erf}(x)=\tfrac{2}{\sqrt{\pi}}\int_0^x \mathrm{e}^{-t^2}\mathrm{d}t;\\ &\operatorname{erfc}(x)=1-\operatorname{erf}(x)=\tfrac{2}{\sqrt{\pi}}\int_x^{+\infty}\mathrm{e}^{-t^2}\mathrm{d}t. \end{aligned} \tag{1.287}$$

1. 高斯随机变量 X 具有下列性质 (式 (1.282) 所表示的随机变量)：

1) X 的概率密度函数曲线具有单峰对称形式，并且在均值 m_X 处有极大值 $\frac{1}{\sqrt{2\pi}\sigma_X}$, 当 $x\to\pm\infty$ 时，概率密度函数的值趋于零。均值反映了 X 概率密度函数的最大值所在的位

置信息，方差式反映了概率密度函数曲线的陡度，随着 σ_X 的增大曲线变得平坦,随机变量 X 落在 $[m_X-3\sigma_X, m_X+3\sigma_X]$ 区间的概率

$$P\left(m_X-3\sigma_X \leqslant x \leqslant m_X+3\sigma\right)=\frac{1}{\sqrt{2\pi}\sigma_X}\int_{m_X-3\sigma_X}^{m_X+3\sigma_X}.\exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right]\mathrm{d}x$$
$$\approx 99.7\%. \tag{1.288}$$

2) 高斯随机变量 X 的特征函数为

$$\varphi(\omega)=\exp\left(j\omega m_X-\frac{1}{2}\omega^2\sigma_X^2\right). \tag{1.289}$$

(3) 高斯随机变量 X 的 n 阶中心矩

$$E\left\{(X-m_X)^n\right\}=\begin{cases}0, & n=2k-1\\ 1\times3\times5\times\cdots\times(n-1)\sigma_X^n, & n=2k\end{cases}, k=1,2,\cdots \tag{1.290}$$

2. 二维高斯随机变量及特征

设 X_1 和 X_2 分别是均值为 m_{X_1}、方差为 $\sigma_{X_1}^2$ 和均值为 m_{X_2}、方差为 $\sigma_{X_2}^2$ 的高斯随机变量，它们的相关系数为 $\rho_{X_1X_2}$，那么二维高斯随机变量 (X_1,X_2) 的联合概率密度函数为

$$\begin{aligned}f\left(x_1,x_2\right)=&\frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}\sqrt{1-\rho_{X_1}x_2{}^2}}\exp\left\{-\frac{1}{2\left(1-\rho_{x_1}^2x_2\right)}\left[\frac{\left(x_1-m_{X_1}\right)^2}{\sigma_{X_1}^2}\right.\right.\\&\left.\left.-\frac{2\rho_{X_1}x_2\left(x_1-m_{X_1}\right)\left(x_2-m_{X_2}\right)}{\sigma_{X_1}\sigma_{X_2}}+\frac{\left(x_2-m_{X_2}\right)^2}{\sigma_{X_2}^2}\right]\right\}.\end{aligned} \tag{1.291}$$

可见，二维高斯随机变量的概率密度函数由 $m_{X_1}, m_{X_2}, \sigma_{X_1}^2, \sigma_{X_2}^2, \rho_{X_1X_2}$ 数完全确定.

高斯随机变量 X_1 和 X_2 的边缘概率密度函数分别为

$$f_{X_1}\left(x_1\right)=\int_{-\infty}^{+\infty}f\left(x_1,x_2\right)\mathrm{d}x_2=\frac{1}{\sqrt{2\pi}\sigma_{X_1}}\exp\left[-\frac{\left(x_1-m_{X_1}\right)^2}{2\sigma_{X_1}^2}\right]. \tag{1.292}$$

$$f_{X_2}\left(x_2\right)=\int_{-\infty}^{+\infty}f\left(x_1,x_2\right)\mathrm{d}x_1=\frac{1}{\sqrt{2\pi}\sigma_{X_2}}\exp\left[-\frac{\left(x_2-m_{X_2}\right)^2}{2\sigma_{X_2}^2}\right]. \tag{1.293}$$

可见，当相关系数 $\rho_{X_1X_2}=0$ 时，有 $f\left(x_1,x_2\right)=f_{X_1}\left(x_1\right)f_{X_2}\left(x_2\right)$, 即两个高斯随机变量 X_1 和 X_2 互不相关就意味着它们统计独立。如果两个随机变量 X_1 和 X_2 的联合概率密度函数是高斯的，则 X_1 和 X_2 各自的概率密度函数也是高斯的。反之，则不一定。

可以证明，二维高斯随机变量的特征函数

$$\varphi\left(\omega_1,\omega_2\right)=\exp\left[j\left(m_{X_1}\omega_1+m_{X_2}\omega_2\right)-\frac{1}{2}\left(\sigma_{X_1}^2\omega_1^2+2\rho_{X_1X_2}\sigma_{X_1}\sigma_{X_2}\omega_1\omega_2+\sigma_{X_2}^2\omega_2^2\right)\right]. \tag{1.294}$$

3. 多维高斯随机变量的有关性质

设 $X_1, X_2, \cdots, X_n$ 组成 n 维高斯随机变量，令 $\boldsymbol{x} = [X_1, X_2, \cdots, X_n]^{\mathrm{T}}$, 均值矢量为

$$\begin{aligned} m_X &= [E\{X_1\}, E\{X_2\}, \cdots, E\{X_n\}]^{\mathrm{T}} \\ &= \left[m_{X_1}, m_{X_2}, \cdots, m_{X_n}\right]^{\mathrm{T}}. \end{aligned} \tag{1.295}$$

它的协方差矩阵为

$$C_X = \begin{bmatrix} E\left\{\left(X_1 - m_{X_1}\right)^2\right\} & \cdots & E\left\{\left(X_1 - m_{X_1}\right)\left(X_n - m_{x_n}\right)\right\} \\ \vdots & & \vdots \\ E\left\{\left(X_n - m_{x_n}\right)\left(X_1 - m_{X_1}\right)\right\} & \cdots & E\left\{\left(X_n - m_{x_n}\right)^2\right\} \end{bmatrix}. \tag{1.296}$$

则 n 维高斯随机变量的联合概率密度函数可以表示为

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\left|C_X\right|^{\frac{1}{2}}} \exp\left[-\frac{(x - m_X)^{\mathrm{T}} C_X^{-1}(x - m_X)}{2}\right]. \tag{1.297}$$

n 维高斯随机变量的特征函数为

$$\varphi_X(\omega) = \exp\left(j m_X^{\mathrm{T}} \omega - \omega^{\mathrm{T}} C_X \omega / 2\right), \tag{1.298}$$

式中，$\omega = [\omega_1, \omega_2, \cdots, \omega_n]^{\mathrm{T}}$.

n 维高斯随机变量具有以下性质：

(1) n 维高斯随机变量经过线性变换后仍然是高斯随机变量；

(2) n 维互不相关的高斯随机变量一定是彼此统计独立的。

由随机变量的数字特征的有关分析易知：一般的 n 维随机变量，若它们彼此统计独立，则必然互不相关，反之，则不一定成立；对于 n 维高斯随机变量，若它们互不相关，则必然是彼此统计独立的。这一性质对分析与处理高斯随机信号带来很大的方便。

2. 高斯变量的矩

标准高斯变量 Y 的 $n\,(n \geqslant 2)$ 阶矩为

$$E[Y^*] = \begin{cases} 1 \cdot 3 \cdot 5 \cdots (n-1) = (n-1)!!, & n\text{为偶数} \\ 0, & n\text{为奇数} \end{cases}, n \geqslant 2\,. \tag{1.299}$$

一般高斯变量 $X \sim N\left(m, \sigma^2\right)$ 的 $n\,(n \geqslant 2)$ 阶中心距为

$$E[(X - m)^n] = \begin{cases} \sigma^n (n-1)!!, & n\text{为偶数} \\ 0, & n\text{为奇数} \end{cases}.$$

3. 高斯变量的和

① 相互独立的高斯变量之和服从高斯分布

设 n 个相互独立的高斯变量 $X_k \sim N\left(m_k, \sigma_k^2\right), k=1,\cdots,n$, 那么它的和也服从高斯分布即 $Y=\sum_{k=1}^{n} X_k \sim N\left(m_Y, \sigma_Y^2\right)$, 均值为 $m_Y=\sum_{k=1}^{n} m_k$, 方差为 $\sigma_Y^2=\sum_{k=1}^{n} \sigma_k^2$.

② 相关的高斯变量之和服从高斯分布

设 n 个相关的高斯变量 $X_k \sim N\left(m_k, \sigma_k^2\right), k=1,\cdots,n, X_i \leftrightarrows X_j\ (i,j=1,2,\cdots,k)$ 的相关系数为 ρ_{ij}。那么其和 $Y=\sum_{k=1}^{n} X_k$ 也服从高斯分布, 其期望为 $m_Y=\sum_{k=1}^{n} m_k$, 方差为 $\sigma_Y^2=\sum_{k=1}^{n} \sigma_k^2+2\sum_{j<i} \rho_{ij}\sigma_i\sigma_k$.

1. n 维高斯变量的概率密度函数与特征函数

若 n 个随机变量 $X_1,\cdots,X_n$ 均是高斯变量, 则称 $(X_1,\cdots,X_n)$ 为 n 维高斯变量若用随机矢量 X 来表示 n 维随机变量, 用矢量 $\boldsymbol{x}$ 表示 X 的取值, 用矢量 $\boldsymbol{M}_X$ 表示随机矢量 $\boldsymbol{X}$ 的均值

$$X=\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, x=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, M_x=\begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}. \tag{1.300}$$

方差为

$$DX=C_X=E\left[(X-M_X)(X-M_X)^{\mathrm{T}}\right]=\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1\pi} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n\times n}. \tag{1.301}$$

则 n 维高斯变量 $(X_1,\cdots,X_n)$ 的联合概率密度 (或 n 维高斯矢量 $\boldsymbol{X}$ 的概率密度) 可以写成如下矩阵形式:

$$f_X(x_1,\cdots,x_n)=\frac{1}{(2\pi)^{n/2}\left|C_X\right|^{1/2}}\exp\left[-\frac{(\boldsymbol{x}-M_x)^{\mathrm{T}} C_X^{-1}(\boldsymbol{x}-M_x)}{2}\right]. \tag{1.302}$$

式中 $(\boldsymbol{x}-\boldsymbol{M}_X)^{\mathrm{T}}$ 表示 $(\boldsymbol{x}-\boldsymbol{M}_X)$ 的转置。n 维高斯分布可记为 $\boldsymbol{X}\sim N(\boldsymbol{M}_X, \boldsymbol{C}_X)$。

n 维高斯变量 $(X_1,\cdots,X_n)$ 的联合特征函数也可以表示为矩阵形式

$$Q_X(u_1,\cdots,u_n)=E\left[\mathrm{e}^{j\boldsymbol{U}^T\boldsymbol{X}}\right]=\exp\left[j\boldsymbol{M}_X^{\mathrm{T}}\boldsymbol{U}-\frac{\boldsymbol{U}^T\boldsymbol{C}_X\boldsymbol{U}}{2}\right]. \tag{1.303}$$

其中

$$\boldsymbol{U}=\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \boldsymbol{U}^T\boldsymbol{X}=\sum_{i=1}^{n} u_i\boldsymbol{X}_i. \tag{1.304}$$

2. n 维高斯变量的性质

1° n 维高斯变量的互不相关与独立是等价的。

对于一般随机变量而言, 独立必定不相关, 不相关不一定独立。但对于高斯变量而言, 互不相关与独立是等价的。

证: 设 n 维高斯变量 $X_1, \cdots, X_n$ 互不相关, 且 $X_k \sim N\left(m_k, \sigma_k^2\right), k=1, \cdots, n$, 即所有协方差 $C=0\,(i \neq 1)$, 其协方差矩阵为对角阵。

$$C_X=\begin{bmatrix}\sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2\end{bmatrix}_{n \times n} \tag{1.305}$$

则 n 维联合特征函数

$$\begin{aligned} Q_X\left(u_1, \cdots, u_n\right) &= \exp \left\{j\left(m_1, \cdots, m_n\right)\begin{bmatrix}u_1 \\ \vdots \\ u_n\end{bmatrix}-\frac{1}{2}\left(u_1, \cdots, u_n\right)\begin{bmatrix}\sigma_1^2 & 0 \\ & \ddots \\ 0 & \sigma_n^2\end{bmatrix}\begin{bmatrix}u_1 \\ \vdots \\ u_n\end{bmatrix}\right\} \\ &= \exp \left[\sum_{k=1}^n j m_k u_k-\sum_{k=1}^n \frac{u_k^2 \sigma_k^2}{2}\right]=\exp \left[\sum_{k=1}^n\left(j m_k u_k-\frac{u_k^2 \sigma_k^2}{2}\right)\right] \\ &= \prod_{k=1}^n \exp \left[j m_k u_k-\frac{u_k^2 \sigma_k^2}{2}\right]=\prod_{k=1}^n Q_{x_k}\left(u_k\right) . \end{aligned} \tag{1.306}$$

由特征函数性质 5° 可知, $X_1, \cdots, X_n$ 之间相互独立。

2° n 维高斯变量线性变换后仍服从高斯分布。

证: 先用 n 维高斯矢量 $\boldsymbol{X}=\left(X_1, \cdots, X_n\right)^{\top}$ 来表示 n 维高斯变量 $\left(X_1, \cdots, X_n\right)$, 再对 $\boldsymbol{X}$ 作一线性变换 $\boldsymbol{Y}=\boldsymbol{A} \boldsymbol{X}$, 其中 $A=\left(a_{i j}\right)_{m \times n}$ 为 $m \times n$ 阶常系数矩阵。线性变换后

$$\boldsymbol{Y}=\boldsymbol{A}_{m \times n} \boldsymbol{X}_{n \times 1}=\begin{bmatrix}Y_1=a_{11} X_1+\cdots+a_{1 n} X_n \\ Y_2=a_{21} X_1+\cdots+a_{2 n} X_n \\ \vdots \\ Y_m=a_{m 1} X_1+\cdots+a_{m n} X_n\end{bmatrix} . \tag{1.307}$$

Y 表示新的 m 维随机变量 $\{Y_1,\cdots,Y_m\}$. 若 $X\sim N(M_X,C_X)$, 则线性变换后的数字特征为

$$\begin{aligned}
\boldsymbol{M}_Y &= E[Y]=E[\boldsymbol{AX}]=\boldsymbol{A}E[\boldsymbol{X}]=\boldsymbol{AM}_X\\
\boldsymbol{C}_Y &= D[Y]=E\left[(\boldsymbol{Y}-\boldsymbol{M}_Y)(\boldsymbol{Y}-\boldsymbol{M}_Y)^{\mathrm{T}}\right]\\
&= E\left[(\boldsymbol{AX}-\boldsymbol{AM}_X)(\boldsymbol{AX}-\boldsymbol{AM}_X)^{\mathrm{T}}\right]\\
&= E\left[\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{M}_X)(\boldsymbol{X}-\boldsymbol{M}_X)^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\right]\\
&= \boldsymbol{A}E\left[(\boldsymbol{X}-\boldsymbol{M}_X)(\boldsymbol{X}-\boldsymbol{M}_X)^{\mathrm{T}}\right]\boldsymbol{A}^{\mathrm{T}}\\
&= \boldsymbol{AC}_X\boldsymbol{A}^{\mathrm{T}}.
\end{aligned}\tag{1.308}$$

$$\begin{aligned}
Q_Y(u_1,\cdots,u_n) &= E\left[\exp\left(j\boldsymbol{U}^{\top}\boldsymbol{Y}\right)\right]=E\left[\exp\left(j\boldsymbol{U}^{T}\boldsymbol{AX}\right)\right]\\
&= E\left(\exp\left[j\left(\boldsymbol{A}^{\top}\boldsymbol{U}\right)^{\top}\right.\right.\\
&= Q_x\left(\boldsymbol{U}^{T}\boldsymbol{A}\right)=\exp\left[j\boldsymbol{M}_X^{T}\left(\boldsymbol{A}^{\top}\boldsymbol{U}\right)-\frac{\left(\boldsymbol{A}^{\top}\boldsymbol{U}\right)^{\top}\boldsymbol{C}_X\left(\boldsymbol{A}^{\top}U\right)}{2}\right]\\
&= \exp\left[j\left(\boldsymbol{AM}_x\right)^{\top}\boldsymbol{U}-\frac{\boldsymbol{U}^{T}\left(\boldsymbol{AC}_X\boldsymbol{A}^{T}\right)U}{2}\right]\\
&= \exp\left[j\boldsymbol{M}_Y^{\top}\boldsymbol{U}-\frac{\boldsymbol{U}^{\top}\boldsymbol{C}_X\boldsymbol{U}}{2}\right].
\end{aligned}\tag{1.309}$$

由特征函数的形式可知, 线性变换后的 m 维随机变量 $(Y_1,\cdots,Y_n)$ 仍服从高斯分布。

由于高斯变量线性变换后仍服从高斯分布, 因此, 只要常系数矩阵 A 选择适当, 使常系数矩阵 A 成为对角阵, 则可以使不独立的 $X_1,\cdots,X_n$ 通过线性变换构成相互独立的 $Y_1,\cdots,Y_m$

$$\boldsymbol{C}_Y=AC_XA^{\top}=\begin{bmatrix}\sigma_{X_1}^2 & & 0\\ & \ddots & \\ 0 & & \sigma_{X_n}^2\end{bmatrix}.\tag{1.310}$$

3° n 维高斯变量的边缘分布仍服从高斯分布

证: 设 n 维高斯矢量 $\boldsymbol{X}=\begin{bmatrix}X_1\\ \vdots\\ X_n\end{bmatrix}$, 其中任一子矢量为 $\overline{\boldsymbol{X}}=\begin{bmatrix}X_i\\ \vdots\\ X_{ik}\end{bmatrix}$, $k<n$。数学期望 $\boldsymbol{M}_X=\begin{bmatrix}m_1\\ \vdots\\ m_n\end{bmatrix}$, $\boldsymbol{M}_X=\begin{bmatrix}m_n\\ \vdots\\ m_{ik}\end{bmatrix}$ 为 M 中相应的子矢量。$C_{\tilde{X}}$ 为协方差矩阵 C_X 中的子矩阵, $C_X=\begin{bmatrix}C_X & C_1\\ C_2 & C_3\end{bmatrix}_{n\times n}$ 其中 $C_{\tilde{X}}$ 为 k 阶矩阵。根据 n 维随机变量特征函数的性质可知, n 维

随机矢量 X 的 k 维边缘特征函数为

$$Q_{\tilde{X}}\left(u_{i1},\cdots,u_{ik}\right)=Q_X\left(u_n,\cdots,u_{ik},0,\cdots,0\right), \tag{1.311}$$

其中 $\tilde{U}=(u_{i1},\cdots,u_{ik})^{\mathrm{T}}$ 为 $U=(u_1,\cdots,u_n)^T$ 的子矢量。

若 $\boldsymbol{X}$ 为 n 维高斯矢量, 则由特征函数的矩阵形式与边缘特征函数的求法得

$$\begin{aligned}Q_{\tilde{X}}(u_{i1},\cdots,u_{ik})&=\exp\left[j(m_1,\cdots,m_n)\begin{pmatrix}u_{i1}\\ \vdots\\ u_{ik}\\ 0\\ \vdots\\ 0\end{pmatrix}-\frac{1}{2}(u_{i1},\cdots,u_{ik},0,\cdots,0)\,\boldsymbol{C}_X\begin{pmatrix}u_{i1}\\ \vdots\\ u_{ik}\\ 0\\ \vdots\\ 0\end{pmatrix}\right]\\ &=\exp\left[\mathrm{j}\boldsymbol{M}_{\tilde{X}}^{\mathrm{T}}\bar{\tilde{\boldsymbol{U}}}-\frac{1}{2}\tilde{\mathrm{U}}^{\mathrm{T}}\mathrm{C}_{\tilde{X}}\bar{\tilde{\boldsymbol{U}}}\right].\end{aligned} \tag{1.312}$$

由逆转公式可知, n 维高斯变量的 k 维边缘分布也服从高斯分布。

例 1.78 已知三维随机矢量 $X=\begin{pmatrix}X_1\\X_2\\X_3\end{pmatrix}\sim N(0,B)$, 其中 $B^{-1}=\begin{pmatrix}3&2&0\\2&4&-2\\0&-2&5\end{pmatrix}$. 令二维随机矢量 $\boldsymbol{Y}=\begin{pmatrix}Y_1\\Y_2\end{pmatrix}$, 其中 $\begin{cases}Y_1=X_1\\Y_2=X_1+X_2\end{cases}$。求: ① 随机矢量 X 的方差 $D[X]$。② 随机矢量 Y 的概率密度和特征函数。③ 随机变量 Y_1 和 Y_2 服从什么分布? 它们是否独立? 给出理由。

解:

① 随机矢量 $\boldsymbol{X}$ 的方差

$$D[\boldsymbol{X}]=\boldsymbol{B}=\left[\boldsymbol{B}^{-1}\right]^{-1}=\begin{bmatrix}3&2&0\\2&4&-2\\0&-2&5\end{bmatrix}^{-1}=\begin{bmatrix}\frac{4}{7}&-\frac{5}{14}&-\frac{1}{7}\\-\frac{5}{14}&\frac{15}{28}&\frac{3}{14}\\-\frac{1}{7}&\frac{3}{14}&\frac{2}{7}\end{bmatrix}. \tag{1.313}$$

① 随机矢量 $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$, 其中 $\boldsymbol{A}$ 由性质 2° 可知

$$\boldsymbol{M}_Y=\boldsymbol{A}\boldsymbol{M}_X=\begin{pmatrix}1&0&0\\1&1&0\end{pmatrix}\begin{bmatrix}0\\0\\0\end{bmatrix}=\begin{pmatrix}0\\0\end{pmatrix}. \tag{1.314}$$

差的逆矩阵为

$$
\begin{aligned}
\boldsymbol{C}_Y &= \boldsymbol{A}\boldsymbol{C}_X\boldsymbol{A}^T \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ -\frac{5}{14} & \frac{15}{28} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\
&= \begin{pmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ \frac{3}{14} & \frac{5}{28} & \frac{1}{14} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
&= \begin{bmatrix} \frac{4}{7} & \frac{3}{14} \\ \frac{3}{14} & \frac{11}{28} \end{bmatrix}.
\end{aligned} \tag{1.315}
$$

可得 $|\boldsymbol{C}_Y| = \begin{vmatrix} \frac{4}{7} & \frac{3}{14} \\ \frac{3}{14} & \frac{11}{28} \end{vmatrix} = \frac{5}{28}$, 方差的逆矩阵为 $\boldsymbol{C}_Y^{-1} = \begin{pmatrix} \frac{11}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{16}{5} \end{pmatrix}$.

则其概率密度

$$
\begin{aligned}
f_Y(y_1, j_z) &= \frac{1}{2\pi\,|\boldsymbol{C}_Y|\,\frac{1}{2}} \exp\left[-\frac{(y - \boldsymbol{M}_Y)^T\,\boldsymbol{C}_{\bar{Y}}^{-1}\,(y - \boldsymbol{M}_Y)}{2}\right] \\
&= \frac{1}{2\pi\,|\boldsymbol{C}_Y|^{\frac{1}{2}}} \exp[\underbrace{y^{\mathrm{T}}\boldsymbol{C}_\gamma^{-1}\boldsymbol{y}}] \\
&= \frac{1}{\pi}\sqrt{\frac{7}{5}} \exp\left[-\frac{11y_1^2 - 12y_1y_2 + 16y_7^1}{10}\right].
\end{aligned} \tag{1.316}
$$

其特征函数为

$$
Q_Y(u_1, u_2) = \exp\left[\mathrm{j}\boldsymbol{M}_Y^{\mathrm{T}}\boldsymbol{U} - \frac{\boldsymbol{U}^T\boldsymbol{C}_Y\boldsymbol{U}}{2}\right] = \exp\left[-\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56}\right]. \tag{1.317}
$$

② 由性质 2° 可知, 随机矢量 Y 也服从高斯分布; 由性质 3° 可知, 高斯分布的

$$
Q_Y(u_1, u_2) = \exp\left[\mathrm{j}M_Y^{\mathrm{T}}\boldsymbol{U} - \frac{U^{\mathrm{T}}C_\gamma\boldsymbol{U}}{2}\right] = \exp\left[-\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56}\right]. \tag{1.318}
$$

边缘分布也服从高斯分布, 即 Y_1 和 Y_2 服从高斯分布。由于随机矢量 Y 的方差 C 不是对角阵, 所以 Y_1 和 Y_2 是相关的, 也表明 Y_1 和 Y_2 不是相互独立的。

1.1.11 复随机变量

上述的讨论是针对实随机变量的，在实际中我们还经常用到复随机变量中，这里先定义复随机变量为

$$Z = X_1 + jX_2, \tag{1.319}$$

其中，X_1 和 X_2 为两个实随机变最这样一个复随机变量的概率密度函数, 可以定义为如下的联合概率密度函数

$$f(z) = f(x_1, x_2). \tag{1.320}$$

一个复随机变量的均值为

$$\begin{aligned} m_Z &= E\{Z\} = E\{X_1 + jX_2\} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x_1 + jx_2) f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2. \end{aligned} \tag{1.321}$$

或者，定义 $\mathrm{d}z = \mathrm{d}x_1 \mathrm{d}x_2$ 作为复 z 平面上的积分微元，是一块无穷小区域，有

$$m_Z = E\{Z\} = \int_{(z)} z f(z) \mathrm{d}z. \tag{1.322}$$

且

$$\begin{aligned} E\left\{|Z|^2\right\} &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |x_1 + jx_2|^2 f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 \\ &= E\left\{X_1^2\right\} + E\left\{X_2^2\right\}. \end{aligned} \tag{1.323}$$

则复随机变量的方差为

$$\sigma_Z^2 = E\left\{|Z - m_Z|^2\right\} = \sigma_{X_1}^2 + \sigma_{X_2}^2 = E\left\{|Z|^2\right\} - m_Z^2. \tag{1.324}$$

对于两个复随机变量，$W = U_1 + \mathrm{j}U_2, Z = X_1 + jX_2$，概率密度函数为

$$f(w, z) = f(x_1, x_2, u_1, u_2). \tag{1.325}$$

其协方差定义为

$$C_{WZ} = E\{(W - m_W)(Z - m_Z)\}. \tag{1.326}$$

相关系数为

$$\rho_{WZ} = \frac{C_{WZ}}{\sigma_W \sigma_Z}. \tag{1.327}$$

一般情况下相关系数为复值。

对于复随机变量 $\boldsymbol{w}$ 和 $\boldsymbol{z}$, 我们定义协方差矩阵和互协方差矩阵为

$$\begin{aligned} \boldsymbol{C}_Z &= E\left\{(z-m_Z)(z-m_Z)^H\right\} \\ \boldsymbol{C}_{WZ} &= E\left\{(w-m_W)(z-m_Z)^H\right\} \end{aligned} . \tag{1.328}$$

特别地，对 n 维复高斯随机矢量 $\boldsymbol{z}$, 若其均值和协方差矩阵用 $\boldsymbol{m}_Z$ 和 $\boldsymbol{C}_z$ 表示，则复高斯随机矢量的概率密度函数为

$$f(z) = \frac{1}{\pi^n |\boldsymbol{C}_Z|} \exp\left[-(z-m_Z)^H \boldsymbol{C}_Z^{-1}(z-m_Z)\right]. \tag{1.329}$$

高斯随机变量的特征函数由下式给出

$$\varphi(\omega) = \exp\left(j\,\mathrm{Re}\left\{\omega^{\mathrm{H}} m_Z\right\} - \frac{\omega^{\mathrm{H}} \boldsymbol{C}_Z \omega}{4}\right). \tag{1.330}$$

其中 $\mathrm{Re}\{\cdot\}$ 表示复数的取实部操作。

1.2 习题

✍ **练习 1.1** 写出下列随机试验的样本空间: ① 10 只产品中有 3 只次品, 每次从中取只 (不放回), 直到将 3 只次品都取出, 记录抽取的次数。② 甲、乙两人下棋一局, 观察棋赛的结果。③ 口袋中有许多红色、白色、蓝色的乒乓球, 任取 4 只, 观察它们有哪几种颜色。

✍ **练习 1.2** 已知 A, B, C 为三个事件, 用 A, B, C 的运算关系表示下列事件: ① A 发生, B 与 C 不发生。② A, B, C 都发生。③ A, B, C 至少有一个发生。④ A, B, C 不多于两个发生。

✍ **练习 1.3** 已知样本空间 $\Omega = \{1, 2, \cdots, 10\}$, 事件 $A = \{2, 3, 4\}$, $B = \{3, 4, 5\}$, $C = \{5, 6, 7\}$, 写出下列事件的表达式: ① $\overline{AB}$。② $\bar{A}B$ 。③ $A\overline{BC}$。④ $\overline{A(B \cup C)}$。

✍ **练习 1.4** 我方对敌方雷达设备同时采取 A, B, C 三种干扰措施, 它们之间相互独立。根据以往的作战经验可估算出 A, B, C 三种干扰措施的成功率分别为 0.2、0.3 和 0.4。求: ① 恰好只有一种干扰措施成功的概率。② 敌方雷达被干扰后失效的总概率。

✍ **练习 1.5** 考察甲、乙两个城市十月份下雨的情况, A, B 分别表示甲、乙两市出现雨天的事件。根据以往气象记录得知 $P(A) = P(B) = 0.4, P(AB) = 0.28$。求 $P(A|B), P(B|A)$ 和 $P(A \cup B)$。

✍ **练习 1.6** 信息 A 和信息 B 通过发射机发送。接收机接收时, A 被误收作 B 的概率为 0.02; 而 B 被误收作 A 的概率为 0.01。已知发送信息 A 和信息 B 的次数比为 3:1, 若接收机收到的信息是 A 时, 原发送信息是 A 的概率为多少。

练习 1.7 已知事件 A, B 相互独立, 分别证明: A 和 B 相互独立; A 和 B 相互独立; $\bar{A}$ 和 $\bar{B}$ 相互独立。

练习 1.8 有朋自远方来, 她乘火车、轮船、汽车或飞机来的概率分别是 0.3、0.20.1 和 0.4。如果她乘火车、轮船或汽车来, 迟到的概率分别是 0.25、0.4 和 0.1, 乘飞机来则不会迟到。结果她迟到了, 问她乘火车来的概率是多少?

练习 1.9 已知随机变量 X 的分布函数为

$$F_X(x) = \begin{cases} 0, & x < 0 \\ kx^2, & 0 \leqslant x \leqslant 1 \\ 1, & x > 1 \end{cases} \tag{1.331}$$

练习 1.10 已知随机变量 X 的概率密度为 $f_X(x) = ke^{-|x|}, -\infty < x < +\infty$ 为拉普拉斯分布, 求:① 系数 k。② X 落在区间 $(0, 1)$ 内的概率。③ 随机变量 X 的分布函数。

练习 1.11 某繁忙的汽车站, 每天有大量的汽车进出。设每辆汽车在一天内出事故的概率为 1E-4, 若每天有 1000 辆汽车进出汽车站, 问汽车站出事故的次数不小于 2 的概率是多少?

练习 1.12 已知随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} k\mathrm{e}^{-(3x+4y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases} \tag{1.332}$$

求: ① 系数 k。② (X, Y) 的分布函数。③ $P\{0 < X \leq 1, 0 < y \leq 2\}$。

练习 1.13 已知随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 1, & 0 < x < 1, |y| < x \\ 0, & \text{其他} \end{cases} \tag{1.333}$$

① 求条件概率密度 $f_X(x|y)$ 和 $f_Y(y|x)$。② 判断 X 和 Y 是否独立, 给出理由。

练习 1.14 已知离散型随机变量 X 的分布律为

$$\begin{array}{c|ccc} X & 3 & 6 & 7 \\ \hline P & 0.2 & 0.1 & 0.7 \end{array} \tag{1.334}$$

求: ① X 的分布函数。② 随机变量 $Y = 3X + 1$ 的分布律。

练习 1.15 已知随机变量 X 服从标准高斯分布。求: ① 随机变量 $Y = e^X$ 的概率密度。② 随机变量 $Z = |X|$ 的概率密度。

✍ **练习 1.16** 已知随机变量 X_1 和 X_2 相互独立, 概率密度分别为

$$f_{X_1}(x_1) = \begin{cases} \frac{1}{2}e^{-\frac{1}{2}x_1}, & x_1 \geqslant 0 \\ 0, & x_1 < 0 \end{cases}, \quad f_{X_2}(x_2) = \begin{cases} \frac{1}{3}e^{-\frac{1}{3}x_2}, & x_2 \geqslant 0 \\ 0, & x_2 < 0 \end{cases} \tag{1.335}$$

求随机变量 $Y = X_1 + X_2$ 的概率密度。

✍ **练习 1.17** 已知随机变量 X, Y 的联合分布律为

$$P\{X = m, Y = n\} = \frac{3^m 2^n e^{-5}}{m!n!}, m, n = 0, 1, 2, \ldots \tag{1.336}$$

求: ① 边缘分布律 $P\{X = m\}, m = 0, 1, 2, \cdots$ 和 $P\{Y = n\}, n = 0, 1, 2, \cdots$。② 条件分布律 $P\{X = m, Y = n\}$ 和 $P\{Y = n, X = m\}$。

✍ **练习 1.18** 已知随机变量 $X_1, X_2, \cdots, X_n$ 相互独立, 概率密度分别为 $f_2(x_2), \cdots, f_n(x_n)$。又随机变量

$$\begin{cases} Y_1 = X_1 \\ Y_2 = X_1 + X_2 \\ \cdots \\ Y_n = X_1 + X_2 + \cdots + X_n \end{cases} \tag{1.337}$$

证明: 随机变量 $Y_1, Y_2, \cdots, Y_n$ 的联合概率密度为 $f_Y(y_1, y_2, \cdots, y_n) = f_1(y_1)f_2(y_2 - y_1)\cdots f_n(y_n - y_{n-1})$.

✍ **练习 1.19** 已知随机变量 X 服从拉普拉斯分布, 其概率密度为

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < +\infty \tag{1.338}$$

求其数学期望与方差。

✍ **练习 1.20** 已知随机变量 X 的可能取值为-4,-1,2,3,4, 且每个值出现的概率均为 1/5。求: ① 随机变量 X 的数学期望和方差。② 随机变量 $Y = 3X^2$ 的概率密度。③ Y 的数学期望和方差。

✍ **练习 1.21** 已知随机变量 X 服从 $(0, 1)$ 的均匀分布, 随机变量 Y 服从高斯分布 $Y \sim N(3X, 1)$。求: ① 条件数学期望 $E[Y|X = x]$。② 条件数学期望 $E[Y|X]$。

✍ **练习 1.22** 已知两个随机变量 X, Y 的数学期望为 $m_X = 1, m_Y = 2$, 方差为 $\sigma_X = 4, \sigma_Y = 1$, 相关系数 $\rho_{XY} = 0.4$。现定义新随机变量 V, W 为

$$\begin{cases} V = -X + 2Y \\ W = X + 3Y \end{cases} \tag{1.339}$$

求 V, W 的期望, 方差以及它们的相关系数。

✍ **练习 1.23** 已知随机变量 X, Y 满足 $Y = aX + b, a, b$ 皆为常数。证明: ① $C_{XY} = a\sigma_X^2$。② $\rho_{XY} = \begin{cases} 1, & a > 0 \\ -1, & a < 0 \end{cases}$ ③ 当 $m_X \neq 0$ 且 $b = -\frac{aE[X^2]}{E[X]}$ 时, 随机变量 X, Y 正交。

✍ **练习 1.24** 已知二维随机变量 (X, Y) 的联合概率密度为

$$f_{XY}(x, y) = \begin{cases} \frac{xy}{9}, & 0 < x < 2, 0 < y < 3 \\ 0, & \text{其他} \end{cases} \tag{1.340}$$

判断随机变量 X 和 Y 是否正交、不相关和独立, 给出理由。

✍ **练习 1.25** 已知随机变量 X, Y 相互独立, 分别服从参数为 λ_1 和 λ_2 的泊松分布。① 求随机变量 X 的期望和方差。② 证明 $Z = X + Y$ 服从参数为 $\lambda_1 + \lambda_2$ 的泊松分布。

✍ **练习 1.26** 已知随机变量 X, Y 的联合特征函数为

$$Q_{XY}(u, v) = \frac{6}{6 - 2ju - 3jv - uv}. \tag{1.341}$$

求: ① 随机变量 X 的特征函数。② 随机变量 Y 的期望和方差。

✍ **练习 1.27** 已知随机变量 X, Y 对于任意 $n \geq 1, m \geq 1$ 都有 $E[XY] = E[X^m]E[Y]$, 证明 X, Y 独立

✍ **练习 1.28** 已知两个独立的随机变量 X, Y 的特征函数分别是 $Q_X(n)$ 和 $Q_Y(u)$, 求随机变量 $Z = 3(X + 1) + 2(Y - 4)$ 的特征函数 $Q_Z(u)$.

✍ **练习 1.29** 已知二维高斯变量 (X_1, X_2) 中, 高斯变量 X_1, X_2 的期望分别为方差分别为 σ_1^2, σ_2^2, 相关系数为 ρ。令

$$Y_1 = \frac{X_1 - m_1}{\sigma_1}, Y_2 = \frac{1}{\sqrt{1 - \rho^2}} \left(\frac{X_2 - m_2}{\sigma_2} - \rho \frac{X_1 - m_1}{\sigma_1} \right). \tag{1.342}$$

① 写出二维高斯变量 (X_1, X_2) 的概率密度和特征函数的矩阵形式, 并展开。② 证明 (Y_1, Y_2) 相互独立, 皆服从标准高斯分布。

✍ **练习 1.30** 已知二维高斯变量 (X_1, Y_1) 的两个分量相互独立, 期望皆为 0, 方差皆为 σ^2。令

$$\begin{cases} Y_1 = \alpha X_1 + \beta X_2 \\ Y_2 = \alpha X_1 - \beta X_2 \end{cases} \tag{1.343}$$

其中 $a \neq 0, B \neq 0$ 为常数。① 证明: (Y_1, Y_2) 服从二维高斯分布。② 求 (Y_1, Y_2) 的均值和协方差矩阵。③ 证明: Y_1, Y_2 相互独立的条件为 $\alpha = \pm\beta$。

✍ **练习 1.31** 已知三维高斯随机矢量 $\boldsymbol{X}=\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix}$ 的均值为常矢量 σ, 其方差阵为

$$\boldsymbol{B}=\begin{bmatrix}2&2&-2\\2&5&-4\\-2&-4&4\end{bmatrix}, \tag{1.344}$$

证明: $X_1, X_2-X_1, X_1/3+2X_2/3+X_3$ 相互独立。

✍ **练习 1.32** 已知三维高斯随机变量 (X_1,X_2,X_3) 各分量相互独立, 皆服从标准高斯分布。求 $Y_1=X_1+X_2$ 和 $Y_2=X_1+X_3$ 的联合特征函数。