### 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

1) 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

$$\begin{aligned}E[Y] &= \int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y = \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_Y(y|x) f_X(x)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} y f_Y(y|x)\mathrm{d}y\right]\cdot f_X(x)\mathrm{d}x.\end{aligned} \tag{1.169}$$

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

$$E[Y|X=x] = \int_{-\infty}^{\infty} y f_Y(y|x)dy. \tag{1.170}$$

2) 上述定义方法也适用于 $(X,Y)$ 为离散型随机变量的情况:

$$\begin{aligned}E[Y] &= \sum_j y_j P\left\{Y=y_j\right\} = \sum_i\sum_j y_j P_{ij} = \sum_i\sum_j y_j P\left\{X=x_i, Y=y_j\right\} \\ &= \sum_i\left[\sum_j y_j P\left\{Y=y_j|X=x_i\right\}\right] P\left\{X=x_i\right\}.\end{aligned} \tag{1.171}$$

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

$$E\left[Y|X=x_i\right] = \sum_j y_j P\left\{Y=y_j|X=x_i\right\}. \tag{1.172}$$

3) 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

$$E[Y] = \sum_i\left[\int_{-\infty}^{\infty} y f_Y\left(y|X=x_i\right)\mathrm{d}y\right] P\left\{X=x_i\right\}. \tag{1.173}$$

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

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

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

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

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

$$E\left[g_1(X)g_2(Y)|X=x\right]=g_1(x)E\left[g_2(Y)|X=x\right]. \tag{1.177}$$

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 $\Omega_X$ 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:

1° $E_X\{E[Y|X]\}=E[Y]$.

推导过程:

$$\begin{aligned}E_X\{E[Y|X]\}&=\int_{-\infty}^{\infty} E[Y|X=x]f_X(x)\mathrm{d}x=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} y f_Y(y|x)\mathrm{d}y\right]f_X(x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_Y(y|x)f_X(x)\mathrm{d}y\mathrm{d}x\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x,y)\mathrm{d}y\mathrm{d}x=E[Y].\end{aligned} \tag{1.178}$$

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 1.61** 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

$$f_Y(y|X=x)=\begin{cases}\frac{1}{1-x}, & x<y<1\\ 0, & \text{其他}\end{cases} \tag{1.179}$$

条件数学期望

$$E[Y|X=x]=\int_{-\infty}^{\infty} y f_Y(y|x)\mathrm{d}y=\int_x^1 \frac{y}{1-x}\mathrm{d}y=\frac{1+x}{2}. \tag{1.180}$$

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 1.62 方差**

变量 $X$ 的方差定义为

$$\begin{aligned}D(X)&=E\left\{[X-E\{X\}]^2\right\}\\&=\int_{-\infty}^{+\infty}(x-E\{X\})^2 f(x)\mathrm{d}x=E\left\{X^2\right\}-(E\{X\})^2.\end{aligned} \tag{1.181}$$

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”计算平均。通常称 $\sqrt{D(X)}$ 为随机变量 $X$ 的均方差或标准差，习惯上用 $\sigma_X^2$ 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

$$D[CX]=C^2D[X]. \tag{1.182}$$

(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

$$\operatorname{Var}\left\{\sum_{i=1}^n a_iX_i\right\}=\sum_{i=1}^n\sum_{j=1}^n a_ia_jE\left\{(X_i-E\{X_i\})\left(X_j-E\left\{X_j\right\}\right)\right\}. \tag{1.183}$$

(4) $X_i$ $(i=1,2,\cdots,n)$ 独立，则

$$\operatorname{Var}\left\{\sum_{i=1}^n a_iX_i\right\}=\sum_{i=1}^n a_i^2D[X_i]. \tag{1.184}$$

(5) 对于一切实数 $\mu$, 有

$$E\left[(X-\mu)^2\right] \geqslant E\left[(X-m_X)^2\right] = D[X]. \tag{1.185}$$

证:

$$\begin{aligned}
E\left[(X-\mu)^2\right] &= E\left\{\left[(X-m_X)+(m_X-\mu)\right]^2\right\} \\
&= E\left[(X-m_X)^2+2(X-m_X)(m_X-\mu)+(m_X-\mu)^2\right] \\
&= E\left[(X-m_X)^2\right]+2E\left[(X-m_X)(m_X-\mu)\right]+E\left[(m_X-\mu)^2\right] \\
&= E\left[(X-m_X)^2\right]+2(m_X-\mu)E\left[X-m_X\right]+(m_X-\mu)^2 \\
&= E\left[(X-m_X)^2\right]+(m_X-\mu)^2 \geqslant E\left[(X-m_X)^2\right].
\end{aligned} \tag{1.186}$$

这表明, 随机变量的取值相对于任何其他 $\mu$ 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 1.1.8**

高斯变量 $X$ 的概率密度为

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\frac{-(x-m_X)^2}{2\sigma^2}\right\}. \tag{1.187}$$

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

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1. \tag{1.188}$$

在 $\sigma$ 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 $\sigma$ 小时, $X$ 的分布相对数学期望 $m$ 较集中; 当 $\sigma$ 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 $\sigma^2$ 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。

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

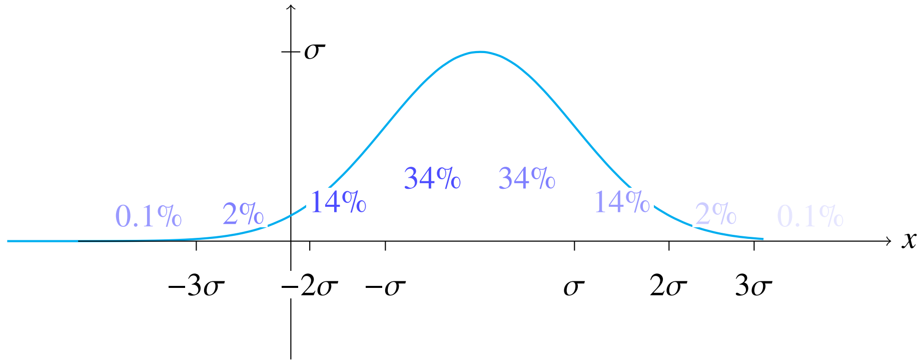


图 1–19 高斯变量的概率密度分布

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 1.63 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

$$E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f_X(x)\mathrm{d}x. \tag{1.189}$$

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

$$E\left[|X|^k\right] = \int_{-\infty}^{\infty} |x|^k f_X(x)dx. \tag{1.190}$$

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

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

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 $\sigma_X^2$ 表示。$D[X]=\sigma_X^2=E\left[(X-m_X)^2\right]$, $D[X]$ 的正平方根 $\sigma_X$ 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

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

3) 矩存在的条件

随机变量的各阶矩不是都存在的。

**例 1.64** 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

$$f_X(x)=\frac{1}{\pi\left(\sigma^2+\nu_2^2(x-\mu)^2\right)}, \tag{1.193}$$

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。

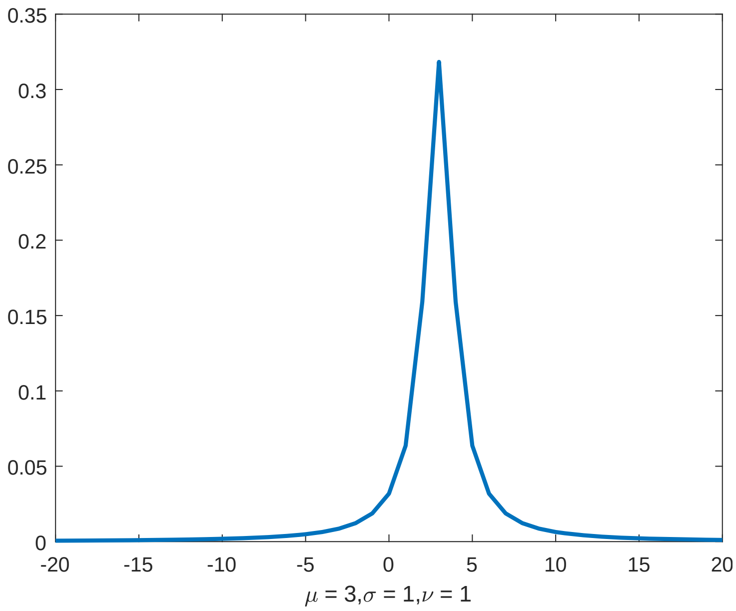


图 1–20 柯西分布

若 $(X_1, X_2, \cdots, X_n)$ 两两互不相关, 则

$$D\left[X_1 \pm X_2 \pm \cdots \pm X_n\right] = D\left[X_1\right] + D\left[X_2\right] + \cdots + D\left[X_n\right]. \tag{1.194}$$

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

$$\begin{aligned} D\left[X_1 \pm X_2 \pm \cdots \pm X_n\right] =& E\left\{\left[\left(X_1 \pm X_2 \pm \cdots \pm X_n\right) - E\left(X_1 \pm X_2 \pm \cdots \pm X_n\right)\right]^2\right\} \\ =& E\left\{\left[\left(X_1 - m_1\right) \pm \left(X_2 - m_2\right) \pm \cdots \pm \left(X_n - m_n\right)\right]^2\right\} \\ =& E\left\{\sum_{i=1}^{n}\left(X_i - m_i\right)^2 \pm 2\sum_{i<j}\left(X_i - m_i\right)\left(X_j - m_j\right)\right\} \\ =& \sum_{i=1}^{n} E\left[\left(X_i - m_i\right)^2\right] \pm 2\sum_{i<j} E\left\{\left(X_i - m_i\right)\left(X_j - m_j\right)\right\}. \end{aligned} \tag{1.195}$$

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

$$\begin{aligned} E\left[\left(X_i - m_i\right)\left(X_j - m_j\right)\right] &= E\left[X_i - m_i\right] E\left[X_j - m_j\right] \\ &= \left\{E\left[X_i\right] - m_i\right\}\left\{E\left[X_j\right] - m_j\right\} = 0. \end{aligned} \tag{1.196}$$

所以

$$D\left[X_1 \pm X_2 \pm \cdots \pm X_n\right] = \sum_{i=1}^{n} E\left[\left(X_i - m_i\right)^2\right] = \sum_{i=1}^{n} D\left[X_i\right]. \tag{1.197}$$

(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

$$E\left[X^n Y^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n y^k f_{XY}(x, y) dx dy. \tag{1.198}$$

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:

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy = R_{XY}. \tag{1.199}$$

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 $(k_1 + k_2 + \cdots + k_n) \triangleq k$ 阶联合原点矩为

$$E\left[X_1^{k_1} X_2^{k_2} \cdots X_n^{k}\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^k x_{2^2}^k \cdots x_k^k \cdot f_X(x_1, x_2, \cdots, x_n)\, \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n. \quad (1.200)$$

其中 $k_1, k_2, \cdots, k_n$ 为正整数。

2) 联合中心矩

**定义 1.65 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

$$E\left[(X - m_X)^n (Y - m_Y)^k\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)^n (y - m_Y)^k f_{XY}(x, y) \mathrm{d}x \mathrm{d}y. \quad (1.201)$$

♣

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:

**定义 1.66 联合中心矩**

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

$$\begin{aligned} E\left[(X - m_X)(Y - m_Y)\right] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)(y - m_Y) f_{XY}(x, y) \mathrm{d}x \mathrm{d}y \\ &= C_{XY}. \end{aligned} \quad (1.202)$$

其中 $k_1, k_2, \cdots, k_n$ 为正整数。

♣

**例 1.67** 已知二维随机变量 $(X, Y)$ 的联合概率密度为

$$f_{XY}(x, y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{其他} \end{cases} \quad (1.203)$$

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

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

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

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

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

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

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

2. 方差

(1) 二维随机变量的协方差

**定义 1.68 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

$$E\left[(X - m_X)(Y - m_Y)\right] = C_{XY}. \tag{1.204}$$

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"。 ♣

二阶联合中心矩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:

$$C_{XY} = E\left[(X - m_X)(Y - m_Y)\right] = E[XY] - m_X m_Y = R_{XY} - m_X m_Y. \tag{1.205}$$

当 $X$ 与 $Y$ 互不相关时, 可知 $C_{XY} = 0$, 由上式可得

$$E[XY] = E[X]E[Y]. \tag{1.206}$$

3. 随机矢量的方差

**定义 1.69 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 $\boldsymbol{X} = [X_1, X_2, \cdots, X_n]^{\mathrm{T}} \in \mathbb{R}^{n\times 1}$ 表示，则随机矢量 $\boldsymbol{X}$ 的方差定义为

$$D(\boldsymbol{X}) = E\left[(\boldsymbol{X} - E\boldsymbol{X})(X - E\boldsymbol{X})^{\mathrm{T}}\right]. \tag{1.207}$$

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

$$\begin{aligned}
D(\boldsymbol{X}) &= E\left\{\begin{bmatrix} X_1 - m_1 \\ X_2 - m_2 \\ \vdots \\ X_n - m_n \end{bmatrix}_{n\times 1} \times [X_1 - m_1 \quad X_2 - m_2 \quad \cdots \quad X_n - m_n]_{1\times n}\right\} \\
&= \begin{bmatrix} E\left[(X_1 - m_1)^2\right] & E\left[(X_1 - m_1)(X_2 - m_2)\right] & \cdots & E\left[(X_1 - m_1)(X_n - m_n)\right] \\ E\left[(X_2 - m_2)(X_1 - m_1)\right] & E\left[(X_2 - m_2)^2\right] & \cdots & E\left[(X_2 - m_2)(X_n - m_n)\right] \\ \vdots & \vdots & & \vdots \\ E\left[(X_n - m_n)(X_1 - m_1)\right] & E\left[(X_n - m_n)(X_2 - m_2)\right] & \cdots & E\left[(X_n - m_n)^2\right] \end{bmatrix}.
\end{aligned} \tag{1.208}$$

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

$$D(X) = \begin{bmatrix} D[X_1] & C_{12} & \cdots & C_{1n} \\ C_{21} & D[X_2] & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & D[X_n] \end{bmatrix} = C_X, \tag{1.209}$$

由式 (1.209) 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 $D(\boldsymbol{X}_i)$ 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

$$D(\boldsymbol{X}) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n\times n} = \boldsymbol{C}_X, \tag{1.210}$$

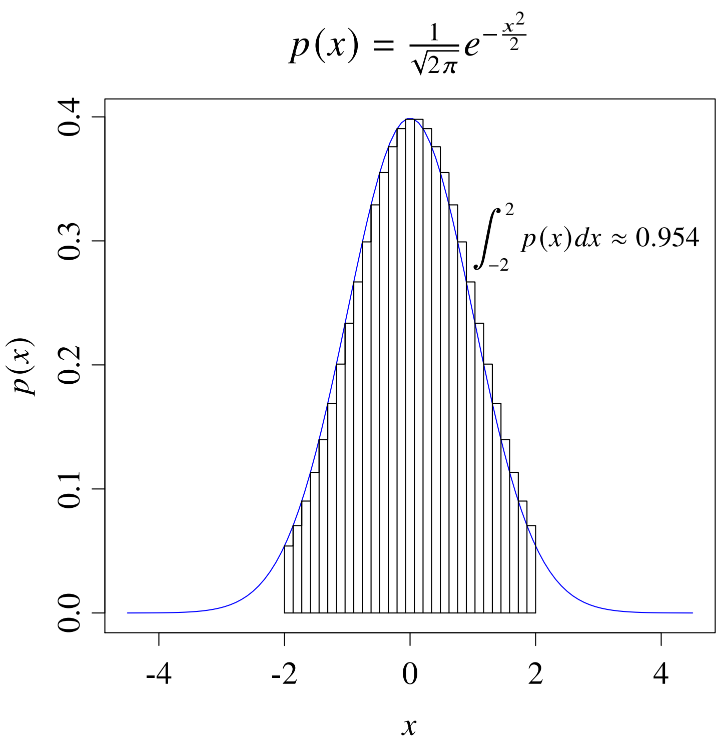


图 1–21 $(X,Y)$ 取值的直方图

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 $\boldsymbol{X}$ 的方差阵是对称阵。

### 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 的平面上画出来, 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 $Y$ 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

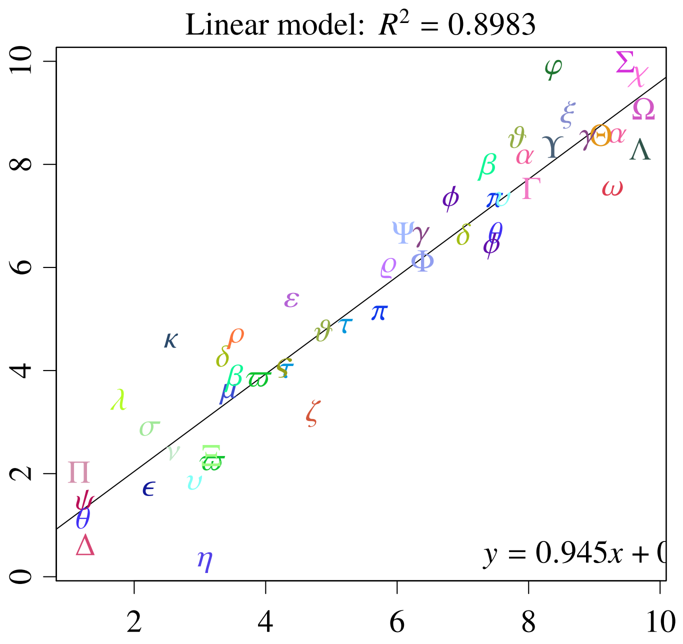


图 1–22 $(X, Y)$ 取值的散布图

$(X, Y)$ 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

$$Y_p = a + bX, \tag{1.211}$$

当然 $Y_p$ 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 $a, b$ 构造的是回归直线, 那么一定会使得 $Y$ 与 $Y_p$ 的误差最小。

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

$$\begin{aligned} \varepsilon = E\left[(Y - Y_p)^2\right] &= E\left[(Y - a - bX)^2\right] \\ &= E\left[Y^2 + b^2X^2 + a^2 + 2abX - 2aY - 2bXY\right] \\ &= a^2 - 2aE[Y] + 2abE[X] - 2bE[XY] + b^2E\left[X^2\right] + E\left[Y^2\right]. \end{aligned} \tag{1.212}$$

令

$$\begin{cases} \frac{\partial \varepsilon}{\partial a} = 2a - 2m_Y + 2bm_X = 0 \\ \frac{\partial \varepsilon}{\partial b} = 2am_X - 2R_{XY} + 2bE\left[X^2\right] = 0 \end{cases}. \tag{1.213}$$

$a, b$ 的值为

$$\begin{cases} a &= m_Y - bm_X = m_Y - \frac{C_{XY}}{\sigma_X^2 m_X}. \\ b &= \frac{R_{XY} - m_X m_Y}{E\left[X^2\right] - m_X^2} = \frac{C_{XY}}{\sigma_X^2}. \end{cases} \tag{1.214}$$

将其代入式 (1.212) 中, 可验证给出的 $\varepsilon$ 是极小值, 从而得到最优预测结果，直线方程为

$$Y_p = a + bX = m_Y + \frac{C_{XY}}{\sigma_X^2}(X - m_X). \tag{1.215}$$

由式 (1.215) 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

$$X = \frac{X - m_X}{\sigma_X}, Y = \frac{Y_p - m_Y}{\sigma_Y}. \tag{1.216}$$

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 1.70 相关系数**

用来表征两随机变量之间线性相关程度的量

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X\sigma_Y}. \tag{1.217}$$

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

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

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

$$\begin{aligned}
&E^2[U \pm V] = \{E[U] \pm E[V]\}^2 = 0. \\
&E\left[(U \pm V)^2\right] = E\left[U^2\right] \pm 2E[UV] + E\left[V^2\right] = 2 \pm 2E[UV]. \\
&\rho_{UV} = \tfrac{C_{UV}}{\sigma_U\sigma_V} = \tfrac{R_{UV} - m_U m_V}{\sigma_U\sigma_V} = R_{UV} = E[UV].
\end{aligned} \tag{1.219}$$

把上述三式代入式 (1.209), 可得

$$D[U \pm V] = E\left[(U \pm V)^2\right] - E^2[U \pm V] \geqslant 0. \tag{1.220}$$

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

$$\begin{cases} C_{XY} = 0. \\ R_{XY} = E[X]E[Y]. \\ D[X \pm Y] = D[X] + D[Y]. \end{cases} \quad . \tag{1.221}$$

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

$$C_X = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix}. \tag{1.222}$$

相关系数矩阵

$$\rho_X = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix}. \tag{1.223}$$

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 1.71 $X$ 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 $Y$ 正交满足下式

$$C_{XY} = -m_X m_Y. \tag{1.224}$$

(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

$$f_{XY}(x,y)=f_X(x)f_Y(y). \tag{1.225}$$

$$\begin{aligned} C_{XY} &= \int_{-\infty}^{\infty}(x-m_X)\,f_X(x)\mathrm{d}x\cdot\int_{-\infty}^{\infty}(y-m_Y)\,f_Y(y)\mathrm{d}y \\ &= E\left[(X-m_X)\right]\cdot E\left[(Y-m_Y)\right]=0. \end{aligned} \tag{1.226}$$

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 1.72** 已知随机变量 $X=\cos\varphi$ 和 $Y=\sin\varphi$, 式中 $\varphi$ 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 $\varphi$ 是在 $(0,2\pi)$ 上均匀分布的随机变量, 其概率密度为

$$f(\varphi)=\begin{cases}\frac{1}{2\pi}, & 0<\varphi<2\pi\\ 0, & \text{其他}\end{cases}. \tag{1.227}$$

则

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty}\cos\varphi f(\varphi)\mathrm{d}\varphi=\frac{1}{2\pi}\int_0^{2\pi}\cos\varphi\mathrm{d}\varphi=0.\\ E[Y] &= \int_{-\infty}^{\infty}\sin\varphi f(\varphi)\mathrm{d}\varphi=\frac{1}{2\pi}\int_0^{2\pi}\sin\varphi\mathrm{d}\varphi=0.\\ E[XY] &= E[\cos\varphi\sin\varphi]=\frac{1}{2}E[\sin 2\varphi]=0.\\ C_{XY} &= E\left[(X-m_X)(Y-m_Y)\right]=E[XY]=0. \end{aligned}$$

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

$$E\left[X^nY^k\right]=E\left[X^n\right]\cdot E\left[Y^k\right], \tag{1.228}$$

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

$$E\left(X^n\right)=\begin{cases}1\cdot 3\cdot 5\cdots\cdots(n-1)\sigma_X^n, & n\text{为偶数}\\ 0, & n\text{为奇数}\end{cases}, n \geqslant 2\,. \tag{1.229}$$

求相关系数 $\rho_{XY}$。

解: 由于

$$m_X = E[X] = 0, m_Y = E\left[X^3\right] = 0. \tag{1.230}$$

则协方差为

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

相关系数为

$$\rho_{XY} = \frac{3\sigma_X^4}{\sigma_X\sigma_Y} = \frac{3\sigma_X^4}{\sqrt{E\left[X^2\right]\cdot\sqrt{E\left[X^6\right]}}} = \frac{3\sigma_X'}{\sqrt{15\sigma_X^8}} = \frac{3}{\sqrt{15}} = 0.775 < 1. \tag{1.231}$$

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 1.74** 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

$$m_Y = E[-6X + 22] = -6E[X] + 22 = 4. \tag{1.232}$$

① $X$ 与 $Y$ 的互相关为

$$\begin{aligned} R_{XY} &= E[XY] = E\left[-6X^2 + 22X\right] = -6E\left[X^2\right] + 22E[X] \\ &= -6\left(m_X^2 + \sigma_X^2\right) + 22m_X = 0. \end{aligned} \tag{1.233}$$

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

$$\rho_{XY} = \frac{R_{XY} - m_X m_y}{\sqrt{\sigma_X^2\sigma_Y^2}} = \frac{-12}{\sqrt{144}} = -1. \tag{1.234}$$

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和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 1.75** 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 1.76 随机变量 $X$ 的特征函数**

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

$$Q_X(u) = E\left[\mathrm{e}^{juX}\right], \mathrm{j} = \sqrt{-1}, -\infty < u < \infty. \tag{1.235}$$

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

$$Q_X(u) = \int_{-\infty}^{\infty} e^{jux} f(x)\mathrm{d}x. \tag{1.236}$$

② 当 $X$ 为离散型随机变量时,其特征函数为

$$Q_X(u) = \sum_i \mathrm{e}^{jux_i} P(X = x_i) = \sum_i \mathrm{e}^{jux_i} p_i. \tag{1.237}$$

(2) 性质

1° $|Q_X(u)| \leqslant Q_X(0) = 1$。

证: 由特征函数的定义

$$|Q_X(u)| = \left|\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{j}ux} f(x)\mathrm{d}x\right| \leq \int_{-\infty}^{\infty} \left|\mathrm{e}^{\mathrm{j}ux} f(x)\right| \mathrm{d}x. \tag{1.238}$$

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

$$|Q_X(u)| \leqslant \left|\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{j}ux} f(x)\mathrm{d}x\right| = \int_{-\infty}^{\infty} \left|\mathrm{e}^{\mathrm{j}ux} f(x)\right| \mathrm{d}x = Q_X(0) = 1. \tag{1.239}$$

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

$$Q_X^*(u) = Q_X(-u). \tag{1.240}$$

4° 随机变量 $X$ 的函数 $y = aX + b$ 的特征函数

$$Q_Y(u) = \mathrm{e}^{\mathrm{j}ub} Q_X(au). \tag{1.241}$$

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

$$Q_Y(u) = \prod_{i=1}^n Q_{X_i}(u). \tag{1.242}$$

证

$$Q_Y(u) = E\left[\mathrm{e}^{\mathrm{ju}Y}\right] = E\left[\mathrm{e}^{\mathrm{j}u(X_1+X_2+\cdots+X_n)}\right] = E\left[\prod_{i=1}^n e^{juX_i}\right]. \tag{1.243}$$

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

$$Q_Y(u) = E\left[\mathrm{e}^{juY}\right] = E\left[\mathrm{e}^{\mathrm{j}u(X_1+X_2+\cdots+X_n)}\right] = \prod_{i=1}^n Q_{X_i}(u). \tag{1.244}$$

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

$$E\left[X^k\right] = (-\mathrm{j})^k \left|\frac{\mathrm{d}^k Q_X(u)}{\mathrm{d}u^k}\right|_{u=0}. \tag{1.245}$$

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

$$Q_X(u) = \sum_{n=0}^{\infty} E\left[X^n\right] \frac{(ju)^n}{n!}. \tag{1.246}$$

这个性质常用在理论推导中。再实际应用中，可由随机变量的各阶矩 $E(X^n)$ $(n=1,2,\cdots)$ 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 1.77** 随机变量 $X \sim U[a,b]$, $f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0 & \text{其他}\end{cases}$.

**解:**

$$\begin{aligned}Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x)dx = \int_a^b e^{jux}\frac{1}{b-a}dx\\ &= \frac{1}{(b-a)ju} e^{jux}\Big|_a^b\\ &= \frac{1}{(b-a)ju}\left[e^{jub}-e^{jua}\right]\\ &= \frac{e^{jua}}{(b-a)ju}\left[e^{ju(b-a)}-1\right].\end{aligned} \tag{1.247}$$

**例 1.78** 随机变量 $X \sim B(1,p)$, 求特征函数 $Q_X(u)$.

**解:**

$$Q_X(u) = e^{ju\times 0}q + e^{ju\times 1}p = q + p(\cos(u) + j\sin(u)). \tag{1.248}$$

**例 1.79** 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

$$P\{Y=k\} = C_n^k p^k q^{n-k}, k=0,1,2,\cdots,n. \tag{1.249}$$

由数学期望的定义，可得

$$E[Y] = \sum_{k=1}^{n} kP\{Y=k\} = \sum_{k=1}^{n} kC_n^k p^k (1-p)^{n-k} = \cdots = np. \tag{1.250}$$

由方差的定义, 可得

$$\begin{aligned}D[Y] &= \sum_{k=1}^{n} (k-m_Y)^2 \cdot P\{Y=k\}\\ &= \sum_{k=1}^{n}(k-np)^2 \cdot C_n^k p^k (1-p)^{n-k} = \cdots = npq.\end{aligned} \tag{1.251}$$

直接引用特征函数的定义, 可得

$$Q_Y(u)=\sum_{k=1}^{n}\mathrm{e}^{j\omega k}P\{Y=k\}=\sum_{k=1}^{n}e^{juk}C_n^k p^k q^{n-k}=\cdots \tag{1.252}$$

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

$$Q_{X_i}(u)=E\left[\mathrm{e}^{\mathrm{j}uX_i}\right]=\sum_{k=0}^{1}\mathrm{e}^{\mathrm{j}uk}P\left(X_i=k\right)=q+p\mathrm{e}^{\mathrm{j}u}. \tag{1.253}$$

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

$$Q_Y(u)=\prod_{i=1}^{n}Q_{X_i}(u)=\left(q+p\mathrm{e}^{\mathrm{j}u}\right)^n. \tag{1.254}$$

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

$$\begin{aligned}
&E[Y]=-\mathrm{j}\tfrac{\mathrm{d}}{\mathrm{d}u}\left(p\mathrm{e}^{\mathrm{j}u}+q\right)^n\Big|_{u=0}=np.\\
&E\left[Y^2\right]=(-\mathrm{j})^2\tfrac{\mathrm{d}^2}{\mathrm{d}u^2}\left(p\mathrm{e}^{\mathrm{j}u}+q\right)^n\Big|_{u=0}=npq+n^2p^2.\\
&D[Y]=E\left[Y^2\right]-(E[Y])^2=npq.
\end{aligned} \tag{1.255}$$

**例 1.80** 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

$$\int_{-\infty}^{\infty}\mathrm{e}^{-Ax^2+2Bx-C}\mathrm{d}x=\sqrt{\frac{\pi}{A}}\,\mathrm{e}^{-\frac{\left(AC-B^2\right)}{A}}. \tag{1.256}$$

① 根据特征函数的定义

$$Q_Y(u)=E\left[\mathrm{e}^{\mathrm{j}uY}\right]=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{j}uy}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{y^2}{2}+2\frac{\mathrm{j}u}{2}y}\mathrm{d}y. \tag{1.257}$$

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

$$Q_Y(u)=\frac{1}{\sqrt{2\pi}}\left\{\sqrt{\frac{\pi}{1/2}}\cdot\exp\left[\left(\frac{ju}{2}\right)^2\Big/\frac{1}{2}\right]\right\}=\mathrm{e}^{-\frac{u^2}{2}}. \tag{1.258}$$

② 将高斯变量 $X$ 归一化处理得

$$X^{\cdot}=\frac{X-m_X}{\sigma_X}=Y, X=\sigma_X Y+m_X. \tag{1.259}$$

利用性质 4 及 ① 的结论 $(X=\sigma_X Y+m_X)$, 得一般高斯变量的特征函数如下

$$Q_X(u)=\mathrm{e}^{\mathrm{j}mu}Q_Y(\sigma u)=\exp\left(\mathrm{j}um-\frac{1}{2}\sigma^2u^2\right). \tag{1.260}$$

**例 1.81** 随机变量 $X\sim\mathrm{N}(\theta,1)$, 求 $X$ 的特征函数 $Q_X(u)$.

因为 $X\sim N(0,1)$, 所以 $X$ 的概率密度函数为

$$\begin{aligned}p(x,\theta)\mathrm{d}x&=(2\pi)^{-1/2}\exp\left\{-\frac{x^2}{2}+x\theta-\frac{\theta^2}{2}\right\}\mathrm{d}x\\&=(2\pi)^{-1/2}\mathrm{e}^{-\theta^2/2}\mathrm{e}^{\theta x}\mathrm{e}^{-x^2/2}\mathrm{d}x\doteq C(\theta)\mathrm{e}^{\theta x}\mathrm{d}\mu(x),\end{aligned} \tag{1.261}$$

其中 $C(\theta)=(2\pi)^{-1/2}\mathrm{e}^{-\theta^2/2}, \mathrm{d}\mu(x)=\mathrm{e}^{-x^2/2}\mathrm{d}x$.

易见，正态分布族 $N(\theta,1)$ 满足指数型分布族 $p_\theta(x)\mathrm{d}\mu(x)=C(\theta)\mathrm{e}^{\theta T(x)}\mathrm{d}\mu(x)$, 其中 $\theta\in\Theta=\{\theta:\theta=a+bi, a,b\in\mathbb{R}\}, T(x)$ 为有限的 $B_S$ 可测的实函数 [陈艳 **2013 指数型分布族的特征函数**]。由定理知，$T(X)=X$ 的特征函数为

$$Q_X(u)=\frac{C(\theta)}{C(\theta+iu)}=\exp\left\{iu\theta-\frac{u^2}{2}\right\}. \tag{1.262}$$

**例 1.82** 随机变量 $X$ 服从伽马分布，即 $X\sim G(\alpha,\beta)$, 则 $X$ 的概率密度函数

$$p(x,\alpha,\beta)\mathrm{d}x=\frac{\alpha^\beta}{\Gamma(\beta)}x^{\beta-1}\mathrm{e}^{-\alpha x}\mathrm{d}x, x>0, \tag{1.263}$$

$$\Gamma(\beta)=\int_0^\infty x^{\beta-1}\mathrm{e}^{-x}\mathrm{d}x, \tag{1.264}$$

其中 $\alpha>0,\beta>0$ 为参数, 则 $X$ 的特征函数为 $Q_X(u)=(1-\mathrm{i}u/\alpha)^{-\beta}$.

① 当 $\beta=1$ 时, $G(\alpha,1)$ 为指数分布 $E(\alpha)$, 其特征函数为 $Q_X(u)=(1-\mathrm{i}u/\alpha)^{-1}$;

② 当 $\alpha=1/2,\beta=n/2$ 时, $G(1/2,n/2)$ 即为自由度 $n$ 的中心化 $\chi^2$ 分布 $\chi^2(n)$, 其特征函数为 $Q_X(u)=(1-2\mathrm{i}u)^{-n/2}$. 这里, 可以顺带得到了指数分布和中心化 $\chi^2$ 分布的特征函数.

**例 1.83** 随机变量 $X\sim E(\lambda)$, 求特征函数 $Q_X(u)$.

**解:**

$$\begin{aligned}Q_X(u)&=\int_{-\infty}^{\infty}e^{jux}f(x)dx=\int_0^\infty e^{jux}\lambda e^{-\lambda x}dx\\&=\frac{\lambda}{ju-\lambda}e^{ju-\lambda}\Big|_0^\infty\\&=\frac{\lambda}{ju-\lambda}[e^{(ju-\lambda)\infty}-1]=\frac{-\lambda}{-\lambda+ju}=\frac{1}{1-j\frac{u}{\lambda}}.\end{aligned} \tag{1.265}$$

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 $Q_X(u)$, 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的定义:

$$\begin{cases} S(\omega) = F[s(t)] = \int_{-\infty}^{\infty} s(t)\mathrm{e}^{-\mathrm{j}\omega t}\mathrm{d}t \\ s(t) = F^{-1}[S(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\mathrm{e}^{\mathrm{j}\omega t}\mathrm{d}\omega \end{cases}. \tag{1.266}$$

对照特征函数的定义, 可得

$$Q_X(u) = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{j}ux}\mathrm{d}x = 2\pi\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{j}ux}\mathrm{d}x\right] = 2\pi F^{-1}[f(x)]. \tag{1.267}$$

即

$$F^{-1}[f(x)] = \frac{1}{2\pi}Q_X(u). \tag{1.268}$$

则可用傅立叶变换，求出 $X$ 的密度函数

$$f(x) = F\left[\frac{1}{2\pi}Q_X(u)\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} Q_X(u)\mathrm{e}^{-jux}du. \tag{1.269}$$

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, 就可以通过特征函数来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 1.84** 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

$$f_Y(y) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_X\left(x_1\cdots x_{n-1}, y - \sum_{i=1}^{n-1} x_i\right) dx_1\cdots dx_{n-1}. \tag{1.270}$$

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 1.85** 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: 由特征函数的性质 5° 可得

$$\begin{aligned} Q_Y(u) &= Q_{X_1}(u)Q_{X_2}(u) = \left(p\mathrm{e}^{ju} + q\right)^n \cdot \left(p\mathrm{e}^{ju} + q\right)^m \\ &= \left(p\mathrm{e}^{ju} + q\right)^{n+m}. \end{aligned} \tag{1.271}$$

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

下面将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) 随机矢量的定义

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

$$\boldsymbol{X}=\begin{bmatrix} X_1\\ \vdots\\ X_n \end{bmatrix}, \boldsymbol{X}_i=\begin{bmatrix} x_1\\ \vdots\\ x_n \end{bmatrix}, \boldsymbol{U}=\begin{bmatrix} u_1\\ \vdots\\ u_n \end{bmatrix} \tag{1.272}$$

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

$$\begin{aligned} Q_X(u_1,\cdots,u_n) &= Q_X\left(\boldsymbol{U}^{\mathrm{T}}\right)=E\left[\mathrm{e}^{j\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}}\right]\\ &=E\left[\mathrm{e}^{\mathrm{j}(u_1X_1+u_2X_2+\cdots+u_nX_n)}\right]=E\left[\mathrm{e}^{\mathrm{j}\sum_{k=1}^{n}u_kX_k}\right]\\ &=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{j(u_1x_1+\cdots+u_nx_n)}f_X(x_1,\cdots,x_n)\,dx_1\cdots dx_n. \end{aligned} \tag{1.273}$$

逆转公式为

$$\begin{aligned} f_X(x_1,\cdots,x_n) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{-jU^Tx}Q_X\left(\boldsymbol{U}^T\right)\frac{du_1}{2\pi}\cdots\frac{du_n}{2\pi}\\ &=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{-j\sum\limits_{k=1}^{n}u_kx_k}Q_X(u_1,\cdots,u_n)\frac{du_1}{2\pi}\cdots\frac{du_n}{2\pi}. \end{aligned} \tag{1.274}$$

(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

$$Q_X^*(u_1,\cdots,u_n)=Q_X(-u_1,\cdots,-u_n). \tag{1.275}$$

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 $\boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{B}\in\mathbb{R}^{r\times 1}$ 的特征函数为

$$Q_Y(u_1,\cdots,u_r,0,\cdots,0)=\mathrm{e}^{j\boldsymbol{U}^T\boldsymbol{B}'}Q_X\left(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}'\right), \tag{1.276}$$

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

下面讨论两种特殊情况:

① 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

$$\boldsymbol{A}=\begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \tag{1.277}$$

此时

$$\boldsymbol{AX}=\begin{bmatrix} a_1X_1 \\ \vdots \\ a_nX_n \end{bmatrix}, \boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{B}=\begin{bmatrix} a_1X_1+b_1 \\ \vdots \\ a_nX_n+b_n \end{bmatrix}, \tag{1.278}$$

则随机矢量 $\boldsymbol{Y}$ 的特征函数为

$$Q_Y\left(u_1,u_2,\cdots,u_n\right)=\mathrm{e}^{\mathrm{j}\sum_{k=1}^{n}u_kb_k}Q_X\left(a_1u_1,\cdots,a_nu_n\right). \tag{1.279}$$

② 当 $r=1$ 时, 矩阵 $\boldsymbol{A}=(a_{ij})_{1\times n}$, 而 $B$ 是一常数 $b$, 则

$$\boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{B}=a_1X_1+\cdots+a_nX_n+b. \tag{1.280}$$

是一维随机变量, 一元函数情形的特征函数为

$$Q_Y\left(u_1\right)=e^{ju_1b}Q_X\left(a_1u_1,a_2u_1,\cdots,a_nu_1\right). \tag{1.281}$$

5° $X_1,\cdots,X_n$ 相互独立的充要条件 (联合特征函数刻画)

**定义 1.86 $n$ 维随机变量相互独立的充要条件**

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

$$Q_X\left(u_1,\cdots,u_n\right)=\prod_{k=1}^{n}Q_{X_k}\left(u_k\right). \tag{1.282}$$

由逆转公式, 可证

$$f_X\left(x_1,\cdots,x_n\right)=\prod_{k=1}^{n}f_{x_k}\left(x_k\right). \tag{1.283}$$

♣

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

$$Q_Y(u_1,\cdots,u_k)=Q_X(u_1,\cdots,u_k,0,\cdots,0)\,. \tag{1.284}$$

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 $E\left[X_1^nX_2^k\right]$ 存在, 则

$$E\left[X_1^nX_2^k\right]=(-\mathrm{j})^{n+k}\left.\frac{\partial^{n+k}Q_X(u_1,u_2)}{\partial u_1^n\partial u_2^k}\right|_{u_1=0,u_2=0}. \tag{1.285}$$

8° 若对所有 $n=0,1,2,\cdots$ 和所有 $k=0,1,2,\cdots$, $E\left[X_1^nX_2^k\right]$ 均存在, 则

$$Q_X(u_1,u_2)=\sum_{n=0}^{\infty}\sum_{k=0}^{\infty}E\left[X_1^nX_2^k\right]\cdot\frac{(ju_1)^n}{n!}\cdot\frac{(ju_2)^k}{k!}. \tag{1.286}$$

### 1.1.10 高斯随机变量

在实际应用中, 常常遇到的随机变量事多个随机变量的和，需要对其进行刻画，给出统计特征。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, 是科技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

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right],x\in\mathbb{R}, \tag{1.287}$$

式中，$m_X$ 为 $X$ 的均值，$\sigma_X$ 为 $X$ 的方差。

如果 $m_X=0,\sigma_X=1$, 则

$$f(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right),\quad -\infty<x<+\infty. \tag{1.288}$$

对于一给定的 $x$, 高斯随机变量 $X$ 的分布函数为

$$\begin{aligned} F(x) &= P(X \leqslant x) \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(u-m_X)^2}{2\sigma_X^2}\right] \mathrm{d}u. \end{aligned} \tag{1.289}$$

令 $\frac{u-m_X}{\sigma_X} = t$, 则

$$F(x) = \int_{-\infty}^{(x-m_X)/\sigma_X} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t = \Phi\left(\frac{x-m_X}{\sigma_X}\right). \tag{1.290}$$

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

高斯分布函数还可以表示为

$$\begin{aligned} F(x) &= \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right). \\ F(x) &= 1 - \frac{1}{2}\operatorname{erfc}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right). \end{aligned} \tag{1.291}$$

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

$$\begin{aligned} &\operatorname{erf}(x) = \tfrac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \mathrm{d}t; \\ &\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \tfrac{2}{\sqrt{\pi}} \int_x^{+\infty} \mathrm{e}^{-t^2} \mathrm{d}t. \end{aligned} \tag{1.292}$$

1. 高斯随机变量 $X$ 具有下列性质 (式 (1.287) 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，方差反映了概率密度函数曲线的陡度，随着 $\sigma_X$ 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

$$\begin{aligned} P\left(m_X - 3\sigma_X \leqslant x \leqslant m_X + 3\sigma\right) &= \frac{1}{\sqrt{2\pi}\sigma_X} \int_{m_X-3_X^{\sigma}}^{m_X+3_X^{\sigma}} \exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right] \mathrm{d}x \\ &\approx 99.7\%. \end{aligned} \tag{1.293}$$

2) 高斯随机变量 $X$ 的特征函数为

$$\varphi(\omega) = \exp\left(j\omega m_X - \frac{1}{2}\omega^2\sigma_X^2\right). \tag{1.294}$$

(3) 高斯随机变量 $X$ 的 $n$ 阶中心矩

$$E\left\{(X-m_X)^n\right\} = \begin{cases} 0, & n = 2k-1 \\ 1\times 3\times 5\times\cdots\times(n-1)\sigma_X^n, & n = 2k \end{cases}, k = 1, 2, \cdots \tag{1.295}$$

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

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}\sqrt{1-\rho_{X_1}x_2}}\exp\left\{-\frac{1}{2\left(1-\rho_{x_1}^2x_2\right)}\left[\frac{\left(x_1-m_{X_1}\right)^2}{\sigma_{X_1}^2}-\frac{2\rho_{X_1}x_2\left(x_1-m_{X_1}\right)\left(x_2-m_{X_2}\right)}{\sigma_{X_1}\sigma_{X_2}}+\frac{\left(x_2-m_{X_2}\right)^2}{\sigma_{X_2}^2}\right]\right\}. \tag{1.296}$$

可见，二维高斯随机变量的概率密度函数由 $m_{X_1}, m_{X_2}, \sigma_{X_1}^2, \sigma_{X_2}^2, \rho_{X_1X_2}$ 五个参数完全确定.

高斯随机变量 $X_1$ 和 $X_2$ 的边缘概率密度函数分别为

$$f_{X_1}(x_1)=\int_{-\infty}^{+\infty}f(x_1,x_2)\,\mathrm{d}x_2=\frac{1}{\sqrt{2\pi}\sigma_{X_1}}\exp\left[-\frac{\left(x_1-m_{X_1}\right)^2}{2\sigma_{X_1}^2}\right]. \tag{1.297}$$

$$f_{X_2}(x_2)=\int_{-\infty}^{+\infty}f(x_1,x_2)\,\mathrm{d}x_1=\frac{1}{\sqrt{2\pi}\sigma_{X_2}}\exp\left[-\frac{\left(x_2-m_{X_2}\right)^2}{2\sigma_{X_2}^2}\right]. \tag{1.298}$$

可见，当相关系数 $\rho_{X_1X_2}=0$ 时，有 $f(x_1,x_2)=f_{X_1}(x_1)\,f_{X_2}(x_2)$, 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

$$\varphi(\omega_1,\omega_2)=\exp\left[j\left(m_{X_1}\omega_1+m_{X_2}\omega_2\right)-\frac{1}{2}\left(\sigma_{X_1}^2\omega_1^2+2\rho_{X_1X_2}\sigma_{X_1}\sigma_{X_2}\omega_1\omega_2+\sigma_{X_2}^2\omega_2^2\right)\right]. \tag{1.299}$$

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

$$\begin{aligned} m_X &= \left[E\{X_1\}, E\{X_2\}, \cdots, E\{X_n\}\right]^{\mathrm{T}} \\ &= \left[m_{X_1}, m_{X_2}, \cdots, m_{X_n}\right]^{\mathrm{T}}. \end{aligned} \tag{1.300}$$

它的协方差矩阵为

$$C_X=\begin{bmatrix} E\left\{\left(X_1-m_{X_1}\right)^2\right\} & \cdots & E\left\{\left(X_1-m_{X_1}\right)\left(X_n-m_{x_n}\right)\right\} \\ \vdots & & \vdots \\ E\left\{\left(X_n-m_{x_n}\right)\left(X_1-m_{X_1}\right)\right\} & \cdots & E\left\{\left(X_n-m_{x_n}\right)^2\right\} \end{bmatrix}. \tag{1.301}$$

则 $n$ 维高斯随机变量的联合概率密度函数 $f(x)$ 可以表示为

$$f(x)=\frac{1}{(2\pi)^{\frac{n}{2}}\left|C_X\right|^{\frac{1}{2}}}\exp\left[-\frac{(x-m_X)^{\mathrm{T}}C_X^{-1}(x-m_X)}{2}\right]. \tag{1.302}$$

$n$ 维高斯随机变量的特征函数为

$$\varphi_X(\omega)=\exp\left(jm_X^{\mathrm{T}}\omega-\omega^{\mathrm{T}}C_X\omega/2\right), \tag{1.303}$$

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

$$E\left[Y^*\right]=\begin{cases}1\cdot 3\cdot 5\cdots(n-1)=(n-1)!!, & n\text{为偶数}\\ 0, & n\text{为奇数}\end{cases}, n\geqslant 2\,. \tag{1.304}$$

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

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

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 $X_k\sim N\left(m_k,\sigma_k^2\right), k=1,\cdots,n$, $X_i$ 和 $X_j\,(i,j=1,2,\cdots,k)$ 的相关系数为 $\rho_{ij}$。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

$$X=\begin{bmatrix} X_1\\ \vdots\\ X_n \end{bmatrix}, x=\begin{bmatrix} x_1\\ \vdots\\ x_n \end{bmatrix}, M_x=\begin{bmatrix} m_1\\ \vdots\\ m_n \end{bmatrix}. \tag{1.305}$$

方差为

$$D[X]=C_X=E\left[(X-M_X)(X-M_X)^{\mathrm{T}}\right]=\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1\pi}\\ C_{21} & C_{22} & \cdots & C_{2n}\\ \vdots & \vdots & & \vdots\\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n\times n}. \tag{1.306}$$

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:

$$f_X(x_1,\cdots,x_n)=\frac{1}{(2\pi)^{n/2}\left|C_X\right|^{1/2}}\exp\left[-\frac{(\boldsymbol{x}-M_x)^{\mathrm{T}}C_X^{-1}(\boldsymbol{x}-M_x)}{2}\right]. \tag{1.307}$$

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

$$Q_X(u_1,\cdots,u_n)=E\left[\mathrm{e}^{j\boldsymbol{U}^T\boldsymbol{X}}\right]=\exp\left[j\boldsymbol{M}_X^{\mathrm{T}}\boldsymbol{U}-\frac{\boldsymbol{U}^T\boldsymbol{C}_X\boldsymbol{U}}{2}\right]. \tag{1.308}$$

其中

$$\boldsymbol{U}=\begin{bmatrix} u_1\\ \vdots\\ u_n \end{bmatrix}, \boldsymbol{U}^T\boldsymbol{X}=\sum_{i=1}^{n}u_i\boldsymbol{X}_i. \tag{1.309}$$

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。

$$C_X=\begin{bmatrix} \sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}_{n\times n}. \tag{1.310}$$

则 $n$ 维联合特征函数

$$
\begin{aligned}
Q_X(u_1,\cdots,u_n) &= \exp\left\{ j(m_1,\cdots,m_n)\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} - \frac{1}{2}(u_1,\cdots,u_n)\begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}\right\} \\
&= \exp\left[\sum_{k=1}^{n} jm_k u_k - \sum_{k=1}^{n}\frac{u_k^2\sigma_k^2}{2}\right] = \exp\left[\sum_{k=1}^{n}\left(jm_k u_k - \frac{u_k^2\sigma_k^2}{2}\right)\right] \\
&= \prod_{k=1}^{n}\exp\left[jm_k u_k - \frac{u_k^2\sigma_k^2}{2}\right] = \prod_{k=1}^{n} Q_{x_k}(u_k).
\end{aligned}
\tag{1.311}
$$

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 $\boldsymbol{X}=(X_1,\cdots,X_n)^{\mathrm{T}}$ 来表示 $n$ 维高斯变量 $(X_1,\cdots,X_n)$, 再对 $\boldsymbol{X}$ 作一线性变换 $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$, 其中 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 为 $m\times n$ 阶常系数矩阵。线性变换后

$$
\boldsymbol{Y}=\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\begin{bmatrix} Y_1 = a_{11}X_1+\cdots+a_{1n}X_n \\ Y_2 = a_{21}X_1+\cdots+a_{2n}X_n \\ \vdots \\ Y_m = a_{m1}X_1+\cdots+a_{mn}X_n \end{bmatrix}. \tag{1.312}
$$

$Y$ 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

$$
\begin{aligned}
\boldsymbol{M}_Y &= E[Y] = E[\boldsymbol{A}\boldsymbol{X}] = \boldsymbol{A}E[\boldsymbol{X}] = \boldsymbol{A}\boldsymbol{M}_X. \\
\boldsymbol{C}_Y &= D[Y] = E\left[(\boldsymbol{Y}-\boldsymbol{M}_Y)(\boldsymbol{Y}-\boldsymbol{M}_Y)^{\mathrm{T}}\right]. \\
&= E\left[(\boldsymbol{A}\boldsymbol{X}-\boldsymbol{A}\boldsymbol{M}_X)(\boldsymbol{A}\boldsymbol{X}-\boldsymbol{A}\boldsymbol{M}_X)^{\mathrm{T}}\right] \\
&= E\left[\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{M}_X)(\boldsymbol{X}-\boldsymbol{M}_X)^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\right] \\
&= \boldsymbol{A}E\left[(\boldsymbol{X}-\boldsymbol{M}_X)(\boldsymbol{X}-\boldsymbol{M}_X)^{\mathrm{T}}\right]\boldsymbol{A}^{\mathrm{T}} \\
&= \boldsymbol{A}\boldsymbol{C}_X\boldsymbol{A}^{\mathrm{T}}.
\end{aligned}
\tag{1.313}
$$

$$
\begin{aligned}
Q_Y\left(u_1, \cdots, u_n\right) &= E\left[\exp\left(j \boldsymbol{U}^{\top} \boldsymbol{Y}\right)\right] = E\left[\exp\left(j \boldsymbol{U}^T \boldsymbol{A} \boldsymbol{X}\right)\right] \\
&= E\left(\exp\left[j\left(\boldsymbol{A}^{\top} \boldsymbol{U}\right)^{\top}\right.\right. \\
&= Q_x\left(\boldsymbol{U}^T \boldsymbol{A}\right) = \exp\left[j \boldsymbol{M}_X^T\left(\boldsymbol{A}^{\top} \boldsymbol{U}\right) - \frac{\left(\boldsymbol{A}^{\top} \boldsymbol{U}\right)^{\top} \boldsymbol{C}_X\left(\boldsymbol{A}^{\top} U\right)}{2}\right] \\
&= \exp\left[j\left(\boldsymbol{A} \boldsymbol{M}_x\right)^{\top} \boldsymbol{U} - \frac{\boldsymbol{U}^T\left(\boldsymbol{A} \boldsymbol{C}_X \boldsymbol{A}^T\right) U}{2}\right] \\
&= \exp\left[j \boldsymbol{M}_Y^{\top} \boldsymbol{U} - \frac{\boldsymbol{U}^{\top} \boldsymbol{C}_X \boldsymbol{U}}{2}\right].
\end{aligned} \tag{1.314}
$$

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。

由于高斯变量经过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变量 $Y_1, \cdots, Y_m$。

$$
\boldsymbol{C}_Y = A C_X A^{\top} = \begin{bmatrix} \sigma_{X_1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{X_n}^2 \end{bmatrix}. \tag{1.315}
$$

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 $M_X = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}$, $M_X = \begin{bmatrix} m_n \\ \vdots \\ m_{ik} \end{bmatrix}$ 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 $\boldsymbol{X}$ 的 $k$ 维边缘特征函数为

$$
Q_{\tilde{X}}\left(u_{i1}, \cdots, u_{ik}\right) = Q_X\left(u_n, \cdots, u_{ik}, 0, \cdots, 0\right), \tag{1.316}
$$

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

$$\begin{aligned} Q_{\tilde{X}}(u_{i1},\cdots,u_{ik}) &= \exp\left[ j(m_1,\cdots,m_n)\begin{pmatrix} u_{i1}\\ \vdots \\ u_{ik}\\ 0\\ \vdots\\ 0\end{pmatrix} - \frac{1}{2}(u_{i1},\cdots,u_{ik},0,\cdots,0)\,\boldsymbol{C}_X \begin{pmatrix} u_{i1}\\ \vdots \\ u_{ik}\\ 0\\ \vdots\\ 0\end{pmatrix}\right] \\ &= \exp\left[ \mathrm{j}\boldsymbol{M}_{\tilde{\mathrm{X}}}^{\mathrm{T}}\bar{\tilde{\boldsymbol{U}}} - \frac{1}{2}\tilde{\mathrm{U}}^{\mathrm{T}}\mathrm{C}_{\tilde{\mathrm{X}}}\bar{\tilde{\boldsymbol{U}}}\right]. \end{aligned} \tag{1.317}$$

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 1.87** 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

$$D[\boldsymbol{X}] = \boldsymbol{B} = \left[\boldsymbol{B}^{-1}\right]^{-1} = \begin{bmatrix} 3 & 2 & 0\\ 2 & 4 & -2\\ 0 & -2 & 5\end{bmatrix}^{-1} = \begin{bmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7}\\ -\frac{5}{14} & \frac{15}{28} & \frac{3}{14}\\ -\frac{1}{7} & \frac{3}{14} & \frac{2}{7}\end{bmatrix}. \tag{1.318}$$

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

$$\boldsymbol{M}_Y = \boldsymbol{A}\boldsymbol{M}_X = \begin{pmatrix} 1 & 0 & 0\\ 1 & 1 & 0\end{pmatrix}\begin{bmatrix} 0\\ 0\\ 0\end{bmatrix} = \begin{pmatrix} 0\\ 0\end{pmatrix}. \tag{1.319}$$

$$
\begin{aligned}
\boldsymbol{C}_Y &= \boldsymbol{A}\boldsymbol{C}_X\boldsymbol{A}^T \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ -\frac{5}{14} & \frac{15}{28} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\
&= \begin{pmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ \frac{3}{14} & \frac{5}{28} & \frac{1}{14} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
&= \begin{bmatrix} \frac{4}{7} & \frac{3}{14} \\ \frac{3}{14} & \frac{11}{28} \end{bmatrix}.
\end{aligned} \tag{1.320}
$$

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

$$
\begin{aligned}
f_Y(y_1, j_z) &= \frac{1}{2\pi\,|\boldsymbol{C}_Y|^{\frac{1}{2}}} \exp\left[-\frac{(y-\boldsymbol{M}_Y)^T\,\boldsymbol{C}_{\bar{Y}}^{-1}\,(y-\boldsymbol{M}_Y)}{2}\right] \\
&= \frac{1}{2\pi\,|\boldsymbol{C}_Y|^{\frac{1}{2}}} \exp[\underbrace{y^{\mathrm{T}}\boldsymbol{C}_\gamma^{-1}\boldsymbol{y}}] \\
&= \frac{1}{\pi}\sqrt{\frac{7}{5}} \exp\left[-\frac{11y_1^2 - 12y_1y_2 + 16y_7^1}{10}\right].
\end{aligned} \tag{1.321}
$$

其特征函数为

$$
Q_Y(u_1, u_2) = \exp\left[\mathrm{j}\boldsymbol{M}_Y^{\mathrm{T}}\boldsymbol{U} - \frac{\boldsymbol{U}^T\boldsymbol{C}_Y\boldsymbol{U}}{2}\right] = \exp\left[-\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56}\right]. \tag{1.322}
$$

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

$$
Q_Y(u_1, u_2) = \exp\left[\mathrm{j}M_Y^{\mathrm{T}}\boldsymbol{U} - \frac{U^{\mathrm{T}}C_\gamma\boldsymbol{U}}{2}\right] = \exp\left[-\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56}\right]. \tag{1.323}
$$

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

$$
Z = X_1 + jX_2, \tag{1.324}
$$

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

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

一个复随机变量的均值为

$$\begin{aligned} m_Z &= E\{Z\} = E\{X_1 + jX_2\} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x_1 + jx_2) f(x_1, x_2)\,\mathrm{d}x_1\mathrm{d}x_2. \end{aligned} \tag{1.326}$$

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

$$m_Z = E\{Z\} = \int_{(z)} z f(z)\mathrm{d}z. \tag{1.327}$$

且

$$\begin{aligned} E\left\{|Z|^2\right\} &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |x_1 + jx_2|^2 f(x_1, x_2)\,\mathrm{d}x_1\mathrm{d}x_2 \\ &= E\left\{X_1^2\right\} + E\left\{X_2^2\right\}. \end{aligned} \tag{1.328}$$

则复随机变量的方差为

$$\sigma_Z^2 = E\left\{|Z - m_Z|^2\right\} = \sigma_{X_1}^2 + \sigma_{X_2}^2 = E\left\{|Z|^2\right\} - m_Z^2. \tag{1.329}$$

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

$$f(w, z) = f(x_1, x_2, u_1, u_2). \tag{1.330}$$

其协方差定义为

$$C_{WZ} = E\{(W - m_W)(Z - m_Z)\}. \tag{1.331}$$

相关系数为

$$\rho_{WZ} = \frac{\boldsymbol{C}_{WZ}}{\sigma_W \sigma_Z}. \tag{1.332}$$

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

$$\begin{aligned} \boldsymbol{C}_Z &= E\left\{(\boldsymbol{Z} - \boldsymbol{m}_Z)(\boldsymbol{Z} - \boldsymbol{m}_Z)^H\right\}. \\ \boldsymbol{C}_{WZ} &= E\left\{(\boldsymbol{W} - \boldsymbol{m}_W)(\boldsymbol{Z} - \boldsymbol{m}_Z)^H\right\}. \end{aligned} \tag{1.333}$$

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

$$f(z) = \frac{1}{\pi^n |\boldsymbol{C}_Z|} \exp\left[-(\boldsymbol{z}-\boldsymbol{m}_Z)^H \boldsymbol{C}_Z^{-1} (\boldsymbol{z}-\boldsymbol{m}_Z)\right]. \tag{1.334}$$

高斯随机变量的特征函数由下式给出

$$\varphi(\boldsymbol{\omega}) = \exp\left(j\,\mathrm{Re}\left\{\boldsymbol{\omega}^{\mathrm{H}}\boldsymbol{m}_Z\right\} - \frac{\boldsymbol{\omega}^{\mathrm{H}}\boldsymbol{C}_Z\boldsymbol{\omega}}{4}\right). \tag{1.335}$$

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 1 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 $\overline{A}B$ 。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; $\overline{A}$ 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

$$F_X(x)=\begin{cases}0, & x<0\\ kx^2, & 0\leqslant x\leqslant 1\\ 1, & x>1\end{cases}\tag{1.336}$$

① 求参数 $k$。② 求区间概率 $P\{0.3\leq X\leq 0.7\}$。③ 求概率密度 $f(X)$。

**练习 1.10** 已知随机变量 $X$ 的概率密度为 $f_X(x)=ke^{-|x|},-\infty<x<+\infty$，$X$ 服从拉普拉斯分布, 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

$$f(x,y)=\begin{cases}k\mathrm{e}^{-(3x+4y)}, & x>0,y>0\\ 0, & \text{其他}\end{cases}\tag{1.337}$$

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

$$f(x,y)=\begin{cases}1, & 0<x<1,|y|<x\\ 0, & \text{其他}\end{cases}\tag{1.338}$$

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

$$\begin{array}{c|ccc} X & 3 & 6 & 7\\ \hline P & 0.2 & 0.1 & 0.7\end{array}\tag{1.339}$$

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

$$f_{X_1}(x_1)=\begin{cases}\frac{1}{2}\mathrm{e}^{-\frac{1}{2}x_1}, & x_1\geqslant 0\\ 0, & x_1<0\end{cases},\quad f_{X_2}(x_2)=\begin{cases}\frac{1}{3}\mathrm{e}^{-\frac{1}{3}x_2}, & x_2\geqslant 0\\ 0, & x_2<0\end{cases}\tag{1.340}$$

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

$$P\{X=m,Y=n\}=\frac{3^m2^n\mathrm{e}^{-5}}{m!n!},m,n=0,1,2,\ldots \tag{1.341}$$

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

$$\begin{cases} Y_1=X_1\\ Y_2=X_1+X_2\\ \cdots\\ Y_n=X_1+X_2+\cdots+X_n \end{cases} \tag{1.342}$$

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

$$f_X(x)=\frac{1}{2}\mathrm{e}^{-|x|},\quad -\infty<x<+\infty \tag{1.343}$$

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

$$\begin{cases} V=-X+2Y\\ W=X+3Y \end{cases} \tag{1.344}$$

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

$$f_{XY}(x,y)=\begin{cases}\frac{xy}{9}, & 0<x<2, 0<y<3\\ 0, & \text{其他}\end{cases} \tag{1.345}$$

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 $\lambda_1$ 和 $\lambda_2$ 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

$$Q_{XY}(u,v)=\frac{6}{6-2ju-3jv-uv}. \tag{1.346}$$

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 $Q_X(u)$ 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 $\sigma_1^2,\sigma_2^2$，相关系数为 $\rho$。令

$$Y_1=\frac{X_1-m_1}{\sigma_1}, Y_2=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{X_2-m_2}{\sigma_2}-\rho\frac{X_1-m_1}{\sigma_1}\right). \tag{1.347}$$

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 $\sigma^2$。令

$$\begin{cases}Y_1=\alpha X_1+\beta X_2\\ Y_2=\alpha X_1-\beta X_2\end{cases} \tag{1.348}$$

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 $\sigma$，其方差阵为

$$\boldsymbol{B}=\begin{bmatrix}2 & 2 & -2\\ 2 & 5 & -4\\ -2 & -4 & 4\end{bmatrix}, \tag{1.349}$$

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。

✍ **练习 1.33** 已知随机变量 $X$ 服从拉普拉斯分布, 其概率密度为

$$f_X(x) = \frac{1}{2}\mathrm{e}^{-|x|}, \quad -\infty < x < +\infty \tag{1.350}$$

求其特征函数。