随机信号分析

概率论概要

主讲: 赵国亮

内蒙古大学电子信息工程学院

October 18, 2020

目录

1 随机变量的数字特征
- 空气弹性变形模型
- 二维连续型随机变量函数的数学期望
- 条件数学期望

2 随机变量的矩和方差
- n 维随机变量的联合中心矩

3 相关、正交和独立
- 2. 不相关、独立和正交
- 随机变量的特征函数
- 2. 特征函数与概率密度的对应关系

4 高斯随机变量
- 多维高斯随机变量的有关性质
- n 维高斯变量的性质

5 复随机变量

目录

1 随机变量的数字特征
- 空气弹性变形模型
- 二维连续型随机变量函数的数学期望
- 条件数学期望

2 随机变量的矩和方差
- n 维随机变量的联合中心矩

3 相关、正交和独立
- 2. 不相关、独立和正交
- 随机变量的特征函数
- 2. 特征函数与概率密度的对应关系

4 高斯随机变量
- 多维高斯随机变量的有关性质
- n 维高斯变量的性质

5 复随机变量

第二次教案下载二维码

Github 下载

第 1 章　概率论基础: 概率部分和统计部分

智慧树课堂二维码和项目地址

图 1: 《随机信号分析》课程号:k213654

Github 项目地址

下载地址：
https://github.com/zggl/random-signal-processing2020-autumn

在使用中，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。

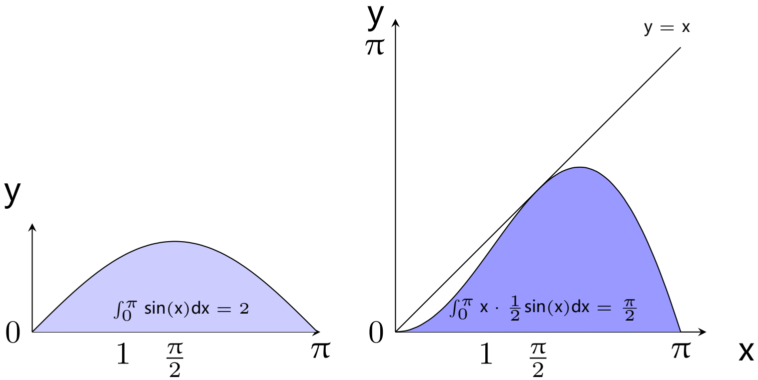
定义 .1 数学期望

变量 X 的数学期望定义为

$$E\{X\} = \int_{-\infty}^{+\infty} xf(x)dx. \tag{1}$$

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.

概率统计意义上的平均


y
0
$\int_0^\pi \sin(x)dx = 2$
1
$\frac{\pi}{2}$
π
y
π
y = x
0
$\int_0^\pi x \cdot \frac{1}{2}\sin(x)dx = \frac{\pi}{2}$
1
$\frac{\pi}{2}$
π
x

数学期望一些基本的性质

(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

$$E\left\{\sum_{i=1}^{n} a_i X_i\right\} = \sum_{i=1}^{n} a_i E\{X_i\}. \tag{2}$$

(3) $X_i\ (i=1,2,\cdots,n)$ 独立，则

$$E\{X_1X_2\cdots X_n\} = \prod_{i=1}^{n} E\{X_i\}. \tag{3}$$

定义 .2 方差和标准差

变量 X 的方差定义为

$$\begin{aligned}\mathrm{Var}\{X\} &= \mathrm{E}\left\{[X-\mathrm{E}\{X\}]^2\right\}\\ &= \int_{-\infty}^{+\infty}(x-\mathrm{E}\{X\})^2 f(x)dx = \mathrm{E}\left\{X^2\right\}-(\mathrm{E}\{X\})^2.\end{aligned} \tag{4}$$

方差的大小反映了随机变量相对于其均值的离散程度。通常称 $\sqrt{\mathrm{Var}\{X\}}$ 为随机变量 X 的均方差或标准差，习惯上用 σ_X^2 表示 $\mathrm{Var}\{X\}$.

方差的一些基本的性质

(1) $\mathrm{Var}\{C\}=0$, C 常数
(2) 对于常数 $a_i\ (i=1,2,\cdots,n)$ 有

$$\mathrm{Var}\left\{\sum_{i=1}^{n} a_i X_i\right\}=\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j E\left\{\left(X_i-E\left\{X_i\right\}\right)\left(X_j-E\left\{X_j\right\}\right)\right\}. \tag{5}$$

(3) $X_i\ (i=1,2,\cdots,n)$ 独立，则

$$\mathrm{Var}\left\{\sum_{i=1}^{n} a_i X_i\right\}=\sum_{i=1}^{n} a_i^2\,\mathrm{Var}\left\{X_i\right\}. \tag{6}$$

(4) $E\left\{(X-C)^2\right\}\geqslant \mathrm{Var}\{X\}$, C 为任意常数。

例 .1

随机变量 X 在区间 (a,b) 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

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}. \tag{7}$$

则数学期望为

$$E[X]=\int_{-\infty}^{\infty}xf_X(x)dx=\int_a^b\frac{x}{b-a}dx=\frac{b+a}{2}. \tag{8}$$

随机变量函数的期望

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 .2

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

空气弹性变形模型

空气弹性变形模型:

$$\begin{pmatrix} m & mx_\alpha b \\ mx_\alpha b & I_\alpha \end{pmatrix}\begin{pmatrix} \ddot{h} \\ \ddot{\alpha} \end{pmatrix} + \begin{pmatrix} c_h & 0 \\ 0 & c_\alpha \end{pmatrix}\begin{pmatrix} \dot{h} \\ \dot{\alpha} \end{pmatrix} + \begin{pmatrix} k_h & 0 \\ 0 & k_\alpha(\alpha) \end{pmatrix}\begin{pmatrix} h \\ \alpha \end{pmatrix} = \begin{pmatrix} -L \\ M \end{pmatrix} \tag{9}$$

系统参数设置

$$b = 0.135\mathrm{m}, \mathrm{span} = 0.6\mathrm{m}, k_h = 2844.4\mathrm{N/m}$$
$$c_h = 27.43\mathrm{Ns/m}, c_\alpha = 0.036\mathrm{Ns}, \rho = 1.225\mathrm{kg/m}^3$$
$$c_{l_\alpha} = 6.28, c_{l_\beta} = 3.358, c_{m_\alpha} = (0.5 + a)c_{l_\alpha}$$
$$c_{m_\beta} = -0.635, m = 12.387\mathrm{kg}, x_\alpha = -0.3533 - a$$
$$I_\alpha = 0.065\,\mathrm{kg\,m}^2, c_\alpha = 0.036$$

系统变量的解释

$$d = m(I_\alpha - mx_\alpha^2 b^2), k_1 = I_\alpha k_h/d, k_3 = -mx_\alpha b k_h/d$$

$$k_2 = (I_\alpha \rho b c_{l_\alpha} + mx_\alpha b^3 \rho c_{m_\alpha})/d$$

$$k_4 = (-mx_\alpha b^2 \rho c_{l_\alpha} - m\rho b^2 c_{m_\alpha})/d, p(\alpha) = (-mx_\alpha b/d) k_\alpha(\alpha)$$

$$q(\alpha) = (m/d) k_\alpha(\alpha), c_1(U) = [I_\alpha(c_h + \rho U b c_{l_\alpha}) + mx_\alpha \rho U^3 c_{m_\alpha}]/d$$

$$c_2(U) = (I_\alpha \rho U b^2 c_{l_\alpha}\left(\frac{1}{2} - a\right) - mx_\alpha b c_\alpha + mx_\alpha \rho U b^4 c_{m\alpha}\left(\frac{1}{2} - a\right))/d$$

$$c_3(U) = (-mx_\alpha b c_h - mx_\alpha \rho U b^2 c_{l_\alpha} - m\rho U b^2 c_{m_\alpha})/d$$

$$c_4(U) = (mc_\alpha - mx_\alpha \rho U b^3 c_{l_\alpha}\left(\frac{1}{2} - a\right) - m\rho U b^3 c_{m_\alpha}\left(\frac{1}{2} - a\right))/d$$

$$g_3 = (-I_\alpha \rho b c_{l_\beta} - mx_\alpha b^3 \rho c_{m_\beta})/d, g_4 = (mx_\alpha b^2 \rho c_{l_\beta} + m\rho b^2 c_{m_\beta})/d.$$

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

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy. \tag{10}$$

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

$$dF_Y(y) = f_Y(y)dy = f_X(x)dx \Rightarrow f_Y(y) = f_X(x)dx/dy, \tag{11}$$

代入期望定义 (10), 得

$$
\begin{aligned}
E[Y] &= \int_{-\infty}^{\infty} y f_Y(y) dy \\
&= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\
&= E[g(X)].
\end{aligned} \tag{12}
$$

② 若 $g(\cdot)$ 是多值变换

$$\begin{aligned} dF_Y(y) &= dF_Y(x_1, x_2, \cdots, x_n) \\ &= f_Y(y)dy = f_X(x_1)\, dx_1 + f_X(x_2)\, dx_2 + \ldots \\ f_Y(y) &= \left[f_X(x_1)\, dx_1 + f_X(x_2)\, dx_2 + \ldots\right] / dy. \end{aligned} \tag{13}$$

代入期望定义 (10), 得

$$\begin{aligned} E[Y] &= \int_{D_{x_1}} g(x_1)\, f_X(x_1)\, dx_1 + \int_{D_{x_2}} g(x_2)\, f_X(x_2)\, dx_2 + \ldots \\ &= \int_{-\infty}^{\infty} g(x) f_X(x) dx = E[g(X)]. \end{aligned} \tag{14}$$

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:

① 若连续型变量 X 的概率密度为 $f_X(x)$, 且 $\int_{-\infty}^{\infty}|g(x)|f_X(x)dx < \infty$, 则函数 $g(X)$ 的期望为)

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx. \tag{15}$$

② 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

$$E[g(X)] = \sum_{k=1}^{\infty} g(x_k)p_k. \tag{16}$$

例 .3

随机变量 X 在区间 (a, b) 呈均匀分布, 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

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases} \tag{17}$$

则函数的数学期望为

$$\begin{aligned} E[g(X)] &= \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_a^b \frac{x^2+1}{b-a} dx \\ &= \frac{1}{3}(a^2 + ab + b^2) + 1. \end{aligned} \tag{18}$$

3. 二维随机变量及其函数的数学期望

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

$$\begin{cases} E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf_{XY}(x, y)dxdy. \\ E[Y] = \int_{-\infty}^{\infty} yf(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf_{XY}(x, y)dxdy. \end{cases} \tag{19}$$

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

$$\begin{cases} E[X] = \sum_i x_i P\{X = x_i\} & = \sum_i \sum_j x_i P\{X = x_i, Y = y_j\}. \\ & = \sum_j x_i P_{ij}. \\ E[Y] = \sum_i y_j P\{Y = y_j\} & = \sum_j \sum_i y_j P\{X = x_i, Y = y_j\}. \\ & = \sum_i y_j P_{ij}. \end{cases} \tag{20}$$

二维连续型随机变量函数的数学期望

仿照单个随机变量函数求期望的方法, 二维连续型随机变量函数 $g(X,Y)$ 的数学期望为:

$$\begin{aligned} E[g(X,Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{XY}(x,y)dxdy. \\ E[g(X,Y)] &= \sum_i\sum_j g\left(x_i,y_j\right)P\left\{X=x_i,Y=y_j\right\}. \end{aligned} \tag{21}$$

4. n 维随机变量的数学期望

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

$$E[X_i] = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}} x_i f_X(x_1, \cdots, x_n)\, dx_1 \cdots dx_n, i = 1, 2, \cdots, n \tag{22}$$

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

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

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

$$E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{bmatrix} = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} = \mathbf{M}_x. \tag{23}$$

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

例 .4

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

$$E[g(X_1, X_2, \cdots, X_n)] = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} g(x_1, \cdots, x_n) f_X(x_1, x_2, \cdots, x_n)$$

$$dx_1 \cdots dx_n$$

$$
\begin{aligned}
E\left[g\left(X_1, X_2, \cdots, X_n\right)\right] &= \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} g\left(x_1, \cdots, x_n\right) f_X\left(x_1, x_2, \cdots, x_n\right) \\
&\qquad dx_1 \cdots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n\right) \\
&\qquad f_X\left(x_1, \cdots, x_n\right) dx_1 \cdots dx_n \\
&= \sum_{i=1}^{n} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} a_i x_i f_X\left(x_1, \cdots, x_n\right) dx_1 \cdots dx_n .
\end{aligned} \tag{24}
$$

根据边缘概率密度的公式, 和式中每一项都可以化成

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} a_i x_i f_X(x_1,\cdots,x_n)\,dx_1\cdots dx_n = \int_{-\infty}^{\infty} a_i x_i f_{X_i}(x_i)\,dx_i = E[a_i X_i] = a_i E[X_i]. \tag{25}$$

所以

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E[X_i]. \tag{26}$$

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 .5

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

解: 由 n 维随机变量函数的期望定义 (相互独立条件下)

$$
\begin{aligned}
E[g(X_1, X_2, \cdots, X_n)] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n) \\
&\qquad f_X(x_1, x_2, \cdots, x_n)\, dx_1 \cdots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 x_2 \cdots x_n \\
&\qquad f_X(x_1, \cdots, x_n)\, dx_1 \cdots dx_n.
\end{aligned} \tag{27}
$$

$$E\left[g\left(X_1, X_2, \cdots, X_n\right)\right] = \int_{-\infty}^{\infty} x_1 f_{X_1}\left(x_1\right) dx_1 \cdot \int_{-\infty}^{\infty} x_2 f_{X_2}\left(x_2\right) dx_2 \cdot \int_{-\infty}^{\infty} x_n f_{X_n}\left(x_n\right) dx_n = E\left[X_1\right] E\left[X_2\right] \cdots E\left[X_n\right]. \quad (28)$$

即

$$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E\left[X_i\right]. \quad (29)$$

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。

5. 数学期望的基本性质

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

$$E\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i E[X_i] + b. \tag{30}$$

4° 若随机变量 X 与随机变量 Y 互不相关, 则

$$E[XY] = E[X]E[Y]. \tag{31}$$

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

$$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E[X_i]. \tag{32}$$

现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 $f_X(x,y)$ 来计算，也可以借助条件分布 $f_Y(y|x)$ 来计算，如

$$
\begin{aligned}
E[Y] &= \int_{-\infty}^{\infty} y f_Y(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_Y(y|x) f_X(x)dxdy \qquad (33)\\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} y f_Y(y|x)dy\right]\cdot f_X(x)dx.
\end{aligned}
$$

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

$$E[Y|X = x] = \int_{-\infty}^{\infty} y f_Y(y|x) dy. \tag{34}$$

上述定义方法也适用于 (X, Y) 为离散型随机变量的情况:

$$\begin{aligned} E[Y] &= \sum_j y_j P\{Y = y_j\} = \sum_i \sum_j y_j P\{X = x_i, Y = y_j\} \\ &= \sum_i \left[\sum_j y_j P\{Y = y_j | X = x_i\} \right] P\{X = x_i\}. \end{aligned} \tag{35}$$

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

$$E\left[Y|X = x_i\right] = \sum_j y_j P\left\{Y = y_j|X = x_i\right\}. \tag{36}$$

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

$$E[Y] = \sum \left[\int_{-\infty}^{\infty} y f_Y\left(y|X = x_i\right) dy\right] P\left\{X = x_i\right\}. \tag{37}$$

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

$$E\left[Y|X=x_i\right]=\int_{-\infty}^{\infty} y f_Y\left(y|X=x_i\right)dy. \tag{38}$$

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

$$E[g(Y)|X=x]=\int_{-\infty}^{\infty} g(y)f_Y(y|x)dy. \tag{39}$$

$$E[g(X,Y)|X=x]=\int_{-\infty}^{\infty} g(x,y)f_Y(y|x)dy. \tag{40}$$

$$E\left[g_1(X)g_2(Y)|X=x\right]=g_1(x)E\left[g_2(Y)|X=x\right]. \tag{41}$$

2. 一随机变量关于另一随机变量的条件期望

(1) 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。

(2) 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 X 的样本空间 2x 上的函数。$E[Y|X=x]$ 是个取决于 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:

$$E_x\{E[Y|X]\} = E[Y] \tag{42}$$

推导过程:

$$
\begin{aligned}
E_X\langle E[Y|X]\rangle &= \int_{-\infty}^{\infty} E[Y|X=x]f_X(x)dx \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} yf_Y(y|x)dy\right]f_X(x)dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf_Y(y|x)f_X(x)dydx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf_X(x,y)dydx = E[Y].
\end{aligned}
\tag{43}
$$

即条件期望的期望等于非条件期望。

条件期望的期望的其他性质

2° $\mathrm{E}\{\mathrm{E}[g(X,Y)|X]\} = \mathrm{E}[g(X,Y)]$.

3° $\mathrm{E}[g(X)Y|X] = g(X)\mathrm{E}[Y|X]$.

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

例 .6

已知随机变量 X 服从 $(0,1)$ 的均匀分布, 随机变量 Y 服从 $(X,1)$ 上的均匀分布。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

$$f_Y(y|X=x) = \begin{cases} \frac{1}{1-x}, & x<y<1 \\ 0, & \text{其他} \end{cases} . \tag{44}$$

条件数学期望

$$E[Y|X=x]=\int_{-\infty}^{\infty} y f_Y(y|x)dy=\int_x^1 \frac{y}{1-x}dy \\ =\frac{1+x}{2}. \tag{45}$$

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 $(\frac{1}{2},1)$ 的均匀分布, 即 $E[Y|X]$ 服从 $(\frac{1}{2},1)$ 的均匀分布.

目录

1 随机变量的数字特征
- 空气弹性变形模型
- 二维连续型随机变量函数的数学期望
- 条件数学期望

2 随机变量的矩和方差
- n 维随机变量的联合中心矩

3 相关、正交和独立
- 2. 不相关、独立和正交
- 随机变量的特征函数
- 2. 特征函数与概率密度的对应关系

4 高斯随机变量
- 多维高斯随机变量的有关性质
- n 维高斯变量的性质

5 复随机变量

一维随机变量的矩（原点矩和中心矩）

定义 .3 k 阶原点矩

随机变量 X 的 k 次幂求统计平均

$$E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f_X(x)dx, \tag{46}$$

称为随机变量 X 的 k 阶原点矩。

讨论如下情况

不同 k

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

$$E\left[|X|^k\right]=\int_{-\infty}^{\infty}|x|^k f_X(x)dx. \tag{47}$$

定义 .4 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

$$E\left[(X-m_X)^k\right]=\int_{-\infty}^{\infty}(x-m_X)^k f_X(x)dx, \tag{48}$$

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种矩, 通常称为随机变量 X 的方差, 用 $D[x]$ 或 σ_X^2 表示。

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

$$E\left[|X-m_X|^k\right]=\int_{-\infty}^{\infty}|x-m_X|^k f_X(x)dx. \tag{49}$$

矩存在的条件

随机变量的各阶矩不是都存在的。若一随机变量的各阶绝对矩都存在，则它相应的各阶矩都存在。

例 .1

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，它的概率密度为

$$f_X(x) = \frac{\sigma_1\sigma_2}{\pi\left(\sigma_1^2 + \sigma_2^2 x^2\right)}, \tag{50}$$

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

定义 .5 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

$$E\left[X^nY^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^ny^kf_{XY}(x, y)dxdy. \tag{51}$$

① $k=0$ 时, $E[X^n]$ 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_{xx} 表示, 如下所示:

$$E[XY]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{x\gamma}(x,y)dxdy=R_{xy}. \tag{52}$$

定义 .6

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

$$E\left[X_1^{k_1} X_2^{k_2} \cdots X_n^{k}\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{k} x_{2^2}^{k} \cdots x_k^{k} \cdot f_x(x_1, x_2, \cdots, x_n)\, dx_1 dx_2 \cdots dx_n. \tag{53}$$

其中 $k_1, k_2, \cdots, k_n$ 为正整数。

定义 .7 联合中心矩

二维随机变量 (X, Y) 的 $n+k$ 阶联合中心矩为

$$E\left[(X-m_X)^n (Y-m_Y)^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-m_X)^n (y-m_\gamma)^k f_{XY}(x,y)dxdy. \tag{54}$$

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。

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

$$
\begin{aligned}
E\left[(X - m_X)(Y - m_Y)\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - m_X)(y - m_Y) \\
&\qquad \cdot f_{XY}(x, y)dxdy \\
&= C_{XY}.
\end{aligned}
\tag{55}
$$

其中 $k_1, k_2, \cdots, k_n$ 为正整数。

例 .2

已知二维随机变量 (X, Y) 的联合概率密度为

$$f_{XY}(x, y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{其他} \end{cases} \tag{56}$$

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

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

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

$$E\left[(X - m_X)^2\right] = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12}, E\left[(X - m_X)^3\right]$$

$$= \int_0^1 \left(x - \frac{1}{2}\right)^3 dx$$

$$= \frac{1}{32}.$$

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

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

$$= \int_0^1 \int_0^1 \left(x - \frac{1}{2}\right)\left(y - \frac{1}{2}\right) dxdy = 0.$$

2. 方差

(1) 维随机变量的方差

二阶中心矩通常被称为随机变量的方差, 记作 D[x] 或 σ_X^2 $D[X] = \sigma_X^2 = E\left[(X - m_X)^2\right]$. 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.

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。可知

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。

举例

高斯变量 X 的概率密度为

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-m)^2}{2\sigma^2}\right\}. \tag{57}$$

可求出其数学期望为 m, 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

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{58}$$

在 σ 大小不同的情况下, $f_X(x)$ 的图形如图 2 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。

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

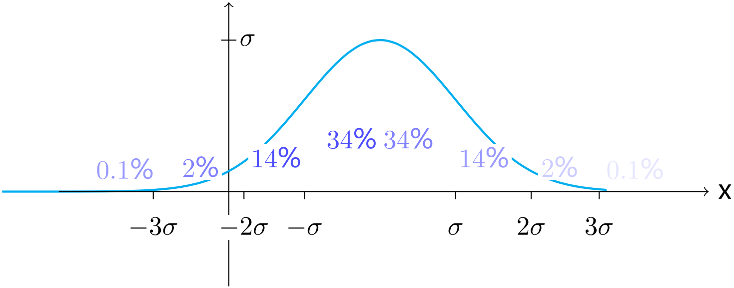


图 2: 高斯变量的概率密度分布

随机变量的方差有下列基本性质:

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

$$D[CX] = C^2 D[X]. \tag{59}$$

③ 若 $(X_1, X_2, \cdots, X_n)$ 两两互不相关, 则

$$D[X_1 \pm X_2 \pm \cdots \pm X_n] = D[X_1] + D[X_2] + \cdots + D[X_n]. \tag{60}$$

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

$$
\begin{aligned}
& D\left[X_1 \pm X_2 \pm \cdots \pm X_n\right] \\
= & E\left\{\left[\left(X_1 \pm X_2 \pm \cdots \pm X_n\right) - E\left(X_1 \pm X_2 \pm \cdots \pm X_n\right)\right]^2\right\} \\
= & E\left\{\left[\left(X_1 - m_1\right) \pm \left(X_2 - m_2\right) \pm \cdots \pm \left(X_n - m_n\right)\right]^2\right\} \\
= & E\left\{\sum_{i=1}^{n}\left(X_i - m_i\right)^2 \pm 2\sum_{i<j}\left(X_i - m_i\right)\left(X_j - m_j\right)\right\} \\
= & \sum_{i=1}^{n} E\left[\left(X_i - m_i\right)^2\right] \pm 2\sum_{i<j} E\left\{\left(X_i - m_i\right)\left(X_j - m_j\right)\right\}.
\end{aligned}
\tag{61}
$$

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

$$\begin{aligned} E\left[(X_i - m_i)(X_j - m_j)\right] &= E\left[X_i - m_i\right] E\left[X_j - m_j\right] \\ &= \langle E\left[X_i\right] - m_i\} \{E\left[X_j\right] - m_j\} \\ &= 0. \end{aligned} \tag{62}$$

所以

$$D\left[X_1 + X_2 \pm \cdots \pm X_n\right] = \sum_{i=1}^{n} E\left[(X_i - m_i)^2\right] = \sum_{i=1}^{n} D\left[X_i\right]. \tag{63}$$

④ 对于一切实数 μ, 有

$$E\left[(X-\mu)^2\right] \geqslant E\left[\left(X-m_X\right)^2\right] = D[X] \tag{64}$$

证:

$$\begin{aligned} E\left[(X-\mu)^2\right] &= E\left\{\left[\left(X-m_X\right)+\left(m_X-\mu\right)\right]^2\right\} \\ &= E\left[\left(X-m_X\right)^2+2\left(X-m_X\right)\left(m_X-\mu\right)+\left(m_X-\mu\right)^2\right] \\ &= E\left[\left(X-m_X\right)^2\right]+2E\left[\left(X-m_X\right)\left(m_X-\mu\right)\right]+E\left[\left(m_X-\mu\right)^2\right] \\ &= E\left[\left(X-m_X\right)^2\right]+2\left(m_X-\mu\right)E\left[X-m_X\right]+\left(m_X-\mu\right)^2 \\ &= E\left[\left(X-m_X\right)^2\right]+\left(m_X-\mu\right)^2 \geqslant E\left[\left(X-m_X\right)^2\right]. \end{aligned} \tag{65}$$

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。

(2) 二维随机变量的协方差

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

$$E\left[(X-m_X)(Y-m_Y)\right]=C_{XY} \tag{66}$$

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:

$$C_{xy}=E\left[(X-m_X)(Y-m_Y)\right]=E[XY]-m_Xm_Y=R_{XY}-m_Xm_Y. \tag{67}$$

当 X 与 Y 互不相关时, $C_{XY}=0$, 由上式可得

$$E[XY]=E[X]E[Y] \tag{68}$$

3. 随机矢量的方差

若 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 用随机矢量 $\mathbf{X} = [X_1, X_2, \cdots, X_n]^T$ 来表示, 则随机矢量 $\mathbf{X}$ 的方差定义为

$$\mathrm{D}\mathbf{X} = \mathrm{E}\left[(\mathbf{X} - \mathrm{E}\mathbf{X})(X - \mathrm{E}\mathbf{X})^T\right], \tag{69}$$

记作 $\mathrm{D}[\mathbf{X}] = \mathrm{Var}\{\mathbf{X}\}$。

若将上式展开可得随机矢量的方差阵

$$D(\mathbf{X}) = E\begin{bmatrix} X_1 - m_1 \\ X_2 - m_2 \\ \vdots \\ X_n - m_n \end{bmatrix}_{n\times 1} \begin{pmatrix} X_1 - m_1 & X_2 - m_2 & \cdots & X_n - m_n \end{pmatrix}_{1\times n} \tag{70}$$

$$= \begin{bmatrix} E\left[(X_1 - m_1)^2\right] & E\left[(X_1 - m_1)(X_2 - m_2)\right] & \cdots & E\left[(X_1 - \right. \\ E\left[(X_2 - m_2)(X_1 - m_1)\right] & E\left[(X_2 - m_2)^2\right] & \cdots & E\left[(X_2 - \right. \\ \vdots & \vdots & & \\ E\left[(X_n - m_n)(X_1 - m_1)\right] & E\left[(X_n - m_n)(X_2 - m_2)\right] & \cdots & E\left[(X\right. \end{bmatrix} \tag{71}$$

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

$$D[X] = \begin{bmatrix} D[X_1] & C_{12} & \cdots & C_{1n} \\ C_{21} & D[X_2] & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & D[X_n] \end{bmatrix} = C_X. \tag{72}$$

由式 (1-186) 可见，矢量 $\mathbf{X}$ 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 $\mathbf{C}_{ij}(i \neq j)$。若将方差 $D[\mathbf{X}_i]$ 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

$$D[\mathbf{X}] = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n \times n} = \mathbf{C}_x. \tag{73}$$

又可称为 n 维随机变量的协方差矩阵。
由 $C_{ij} = C_k$ 可得 $C_x = C_X^T$ 随机矢量 $\mathbf{X}$ 的方差阵为对称矩阵。

目录

1 随机变量的数字特征

- 空气弹性变形模型
- 二维连续型随机变量函数的数学期望
- 条件数学期望

2 随机变量的矩和方差

- n 维随机变量的联合中心矩

3 相关、正交和独立

- 2. 不相关、独立和正交
- 随机变量的特征函数
- 2. 特征函数与概率密度的对应关系

4 高斯随机变量

- 多维高斯随机变量的有关性质
- n 维高斯变量的性质

5 复随机变量

1. 相关系数

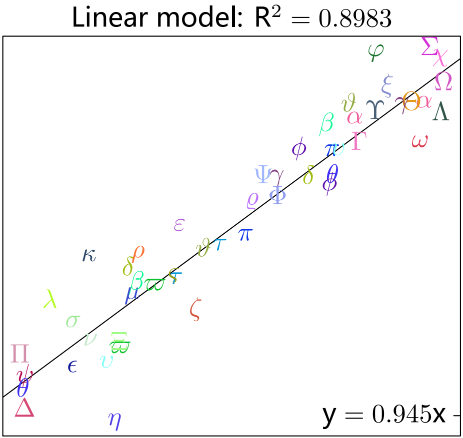
(1) 两个随机变量相互关系的描述

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3)。

1) 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。

1) 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。



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 (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 $\hat{Y}=aX+b$. 当然 $\hat{Y}$ 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 a, b 值的方法如下

$$
\begin{aligned}
\varepsilon &= E\left[\left(Y - Y_p\right)^2\right] = E\left[(Y - a - bX)^2\right] \\
&= E\left[Y^2 + b^2X^2 + a^2 + 2abX - 2aY - 2bXY\right] \\
&= a^2 - 2aE[Y] + 2abE[X] - 2bE[XY] + b^2E\left[X^2\right] + E\left[Y^2\right].
\end{aligned} \tag{74}
$$

令

$$\left\{ \begin{array}{l} \frac{\partial \varepsilon}{\partial a} = 2a - 2m_Y + 2bm_X = 0 \\ \frac{\partial \varepsilon}{\partial b} = 2am_X - 2R_{XY} + 2bE\left[X^2\right] = 0 \end{array} \right. \tag{75}$$

a, b 的值为

$$\left\{ \begin{array}{l} a = m_Y - bm_X = m_Y - \frac{C_{XY}}{\sigma_X^2 m_X} \\ b = \frac{R_{XY} - m_X m_Y}{E[X^2] - m_X^2} = \frac{C_{XY}}{\sigma_X^2} \end{array} \right. \tag{76}$$

将其代入式 (74)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

$$Y_p = a + bX = m_Y + \frac{C_{XY}}{\sigma_X^2}(X - m_X)\,. \tag{77}$$

由式 (77) 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

$$\dot{X} = \frac{X - m_X}{\sigma_X}, Y = \frac{Y_p - m_Y}{\sigma_Y}, \tag{78}$$

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。

定义 .8 相关系数

用来表征两个随机变量之间线性相关程度的量

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} \tag{79}$$

称之为相关系数.

(3) 相关系数的性质

1° $-1 \leqslant \rho_{xy} \leqslant 1$, 即 $|\rho_{x\gamma}| \leqslant 1$.
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

$$U = \frac{X - m_X}{\sigma_X}, V = \frac{Y - m_Y}{\sigma_Y}. \tag{80}$$

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

$$E^2[U \pm V] = \{E[U] \pm E[V]\}^2 = 0.$$
$$E\left[(U \pm V)^2\right] = E\left[U^2\right] \pm 2E[UV] + E\left[V^2\right] = 2 \pm 2E[UV].$$
$$\rho_{UV} = \frac{C_{UV}}{\sigma_U \sigma_V} = \frac{R_{UV} - m_U m_V}{\sigma_U \sigma_V} = R_{UV} = E[UV]. \tag{81}$$

线性互不相关

把上述三式代入方差定义：可得

$$D[U \pm V] = E\left[(U \pm V)^2\right] - E^2[U \pm V] \geqslant 0. \tag{82}$$

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

$$\begin{cases} C_{xy} = 0 \\ R_{XY} = E[X]E[Y] \\ D[X \pm Y] = D[X] + D[Y] \end{cases}. \tag{83}$$

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

$$C_X = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix}. \tag{84}$$

相关系数矩阵

$$\rho_{x} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix}. \tag{85}$$

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

正交的定义

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

$$C_{XY} = -m_X m_Y. \tag{86}$$

(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

$$f_{XY}(x, y) = f_X(x) f_Y(y). \tag{87}$$

互不相关

$$C_{XY} = \int_{-\infty}^{\infty} (x - m_X) f_X(x) dx \cdot \int_{-\infty}^{\infty} (y - m_Y) f_Y(y) dy \tag{88}$$
$$= E[(X - m_X)] \cdot E[(Y - m_Y)] = 0.$$

即 X 与 Y 互不相关。

解释

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, 即 X 与 Y 互不相关。

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

例 .1

已知随机变量 X = $\cos\varphi$ 和 Y = $\sin\varphi$, 式中 φ 是在 $(0, 2\pi)$ 上的均匀分布, 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

$$f(\varphi) = \begin{cases} \frac{1}{2\pi}, & 0 < \varphi < 2\pi \\ 0, & \text{其他} \end{cases} \quad . \tag{89}$$

则

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

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

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

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

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, 随机变量 X 与 Y 之间相互不独立。

统计独立

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
则 X 与 Y 统计独立

$$E\left[X^n Y^k\right] = E\left[X^n\right] \cdot E\left[Y^k\right]. \tag{90}$$

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 .2

已知随机变量 X 与 Y 有非线性关系 $Y = X^3$, 且

$$E(X^n) = \begin{cases} 1\cdot 3\cdot 5\cdots\cdots(n-1)\sigma_X^n, & n\text{为偶数} \\ 0, & n\text{为奇数} \end{cases}, n \geqslant 2\,. \tag{91}$$

求相关系数 ρ_{XY}。

解: 由于

$$\mathsf{m_X} = \mathsf{E}[\mathsf{X}] = 0, \quad \mathsf{m_Y} = \mathsf{E}\left[\mathsf{X}^3\right] = 0. \tag{92}$$

则协方差为

$$\begin{aligned} \mathsf{C_{xY}} &= \mathsf{R_{xy}} - \mathsf{m_X m_Y} = \mathsf{R_{xy}} = \mathsf{E}[\mathsf{XY}] = \mathsf{E}\left[\mathsf{XX}^3\right] \\ &= \mathsf{E}\left[\mathsf{X}^4\right] = 3\sigma_{\mathsf{X}}^4. \end{aligned} \tag{93}$$

相关系数为

$$\begin{aligned} \rho_{\mathsf{xy}} &= \frac{3\sigma_{\mathsf{x}}^4}{\sigma_{\mathsf{x}}\sigma_{\mathsf{y}}} = \frac{3\sigma_{\mathsf{x}}^4}{\sqrt{\mathsf{E}\left[\mathsf{X}^2\right] \cdot \sqrt{\mathsf{E}\left[\mathsf{X}^6\right]}}} \\ &= \frac{3\sigma_{\mathsf{x}}'}{\sqrt{15\sigma_{\mathsf{x}}^8}} = \frac{3}{\sqrt{15}} = 0.775 < 1. \end{aligned} \tag{94}$$

性质 5

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 .3

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

$$\begin{aligned} m_Y &= E[-6X + 22] \\ &= -6E[X] + 22 = 4. \end{aligned} \tag{95}$$

① X 与 Y 的互相关为

$$\begin{aligned} R_{xy} &= E[XY] = E\left[-6X^2 + 22X\right] = -6E\left[X^2\right] + 22E[X] \\ &= -6\left(m_X^2 + \sigma_X^2\right) + 22m_X = 0. \end{aligned} \tag{96}$$

可见 X 与 Y 是正交的。

② 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

$$\rho_{xy} = \frac{R_{xy} - m_X m_Y}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{-12}{\sqrt{144}} = -1. \tag{97}$$

这说明 X 与 Y 又是线性相关的。

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。

特征函数的优越性

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。

特征函数举例

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 .9 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

$$Q_X(u) = E\left[e^{juX}\right], j = \sqrt{-1}, -\infty < u < \infty \tag{98}$$

称之为随机变量 X 的特征函数。

当 X 为连续型随机变量时, 其特征函数为

$$Q_X(u) = \int_{-\infty}^{\infty} e^{jux} f(x) dx. \tag{99}$$

5 当 X 为离散型随机变量时, 其特征函数为

$$Q_X(u) = \sum_i e^{jux_i} P(X = x_i) = \sum_i e^{jux_i} p_i.. \tag{100}$$

(2) 性质 1

1° $|Q_X(u)| \leqslant Q_X(0) = 1.$
证: 由特征函数的定义

$$|Q_X(u)| \leqslant \int_{-\infty}^{\infty} \left|e^{jux}\right| f(x)dx = \int_{-\infty}^{\infty} f(x)dx = Q_X(0) = 1. \quad (101)$$

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

$$|Q_X(u)| \leqslant \left|\int_{-\infty}^{\infty} e^{jux} f(x)dx\right| = \int_{-\infty}^{\infty} \left|e^{jux} f(x)\right| dx = Q_X(0) = 1. \quad (102)$$

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) 性质 2

2° 特征函数 $Q_x(u)$ 是实变量 v 在 $(-\infty, +\infty)$ 上的连续函数。
3° 特征函数 $Q_x(u)$ 是实变量 u 的复函数时, 有

$$Q_x^*(u) = Q_x(-u). \tag{103}$$

4° 随机变量 X 的函数 $y = aX + b$ 的特征函数

$$Q_Y(u) = e^{juy} Q_X(au). \tag{104}$$

证:

$$\begin{aligned} Q_Y(u) = E\left[e^{juY}\right] &= E\left[e^{ju(aX+b)}\right] \\ &= E\left[e^{juaX} \cdot e^{jub}\right] \\ &= e^{jub} E\left[e^{juaX}\right] \\ &= e^{jub} Q_X(au). \end{aligned} \tag{105}$$

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

$$Q_Y(u) = \prod_{i=1}^{n} Q_{X_i}(u). \tag{106}$$

证

$$Q_Y(u) = E\left[e^{juY}\right] = E\left[e^{ju(X_1+X_2+\cdots+X_n)}\right] = E\left[\prod_{i=1}^{n} e^{juX_i}\right]. \tag{107}$$

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

$$Q_Y(u) = E\left[e^{juY}\right] = E\left[e^{ju(X_1+X_2+\cdots+X_n)}\right] = \prod_{i=1}^{n} Q_{X_i}(u). \tag{108}$$

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

$$E\left[X^k\right] = (-j)^k \left.\frac{d^k Q_X(u)}{du^k}\right|_{u=0}. \tag{109}$$

可见, 利用这一性质由特征函数 $Q_X(u)$ 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, 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

$$Q_X(u) = \sum_{k=0}^{\infty} E\left[X^k\right] \frac{(ju)^k}{k!}. \tag{110}$$

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 $E(X^n)$ $(n = 1, 2, \cdots)$ 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 .4

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

$$\int_{-\infty}^{\infty} e^{-Ax^2+2Bx-C} dx = \sqrt{\frac{\pi}{A}}\, e^{-\frac{\left(AC-B^2\right)}{A}}. \tag{111}$$

例 .5

给定区域 $0 \leq \rho \leq$ a, $0 \leq \theta \leq \frac{\pi}{2}$, 积分区域如图4.

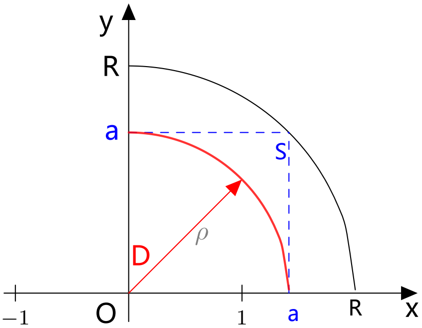


图 4: 重积分积分区域

$$
\begin{aligned}
\iint_D e^{-x^2-y^2}dxdy &= \iint_D e^{-\rho^2}\rho d\rho d\theta = \int_0^{\frac{\pi}{2}}\int_0^a e^{-\rho^2}\rho d\rho d\theta \\
&= \int_0^{2\pi} -\frac{1}{2}e^{-\rho^2}\Big|_0^a d\theta = \frac{1}{2}\left(1-e^{-a^2}\right)\int_0^{\frac{\pi}{2}} d\theta \\
&= \frac{\pi}{4}\left(1-e^{-a^2}\right).
\end{aligned}
$$

对 a 取极限, 有

$$
\begin{aligned}
\iint_S e^{-x^2-y^2}dxdy &= \int_0^R e^{-x^2}dx\int_0^R e^{-y^2}dy = \left(\int_0^R e^{-x^2}dx\right)^2 \\
&\Rightarrow \int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2},\ (R\to\infty).
\end{aligned}
$$

换元后有

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 2\int_0^{\infty} e^{-\frac{x^2}{2}} dx \underline{\underline{t^2 = \frac{x^2}{2}}} \, 2\int_0^{\infty} e^{-t^2} d\sqrt{2}\,t$$

$$= 2\sqrt{2}\int_0^{\infty} e^{-t^2} dt = 2\sqrt{2} \times \frac{\sqrt{\pi}}{2} = \sqrt{2\pi}.$$

例 .6

(菲涅耳 **(fresnel)** 积分) 已知 $\int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$, 证明: $\int_0^{\infty} \sin x^2 dx = \int_0^{\infty} \cos x^2 dx = \frac{1}{2}\sqrt{\frac{\pi}{2}}$.

① 根据特征函数的定义

$$\begin{aligned} Q_Y(u) &= E\left[e^{juY}\right] = \int_{-\infty}^{\infty} e^{juy} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}+2\frac{juy}{2}} dy. \end{aligned} \tag{112}$$

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

$$Q_Y(u) = \frac{1}{\sqrt{2\pi}} \left\{ \sqrt{\frac{\pi}{1/2}} \cdot \exp\left[\left(\frac{ju}{2}\right)^2 \Big/ \frac{1}{2}\right] \right\} = e^{-\frac{u^2}{2}}. \tag{113}$$

② 将高斯变量 X 归一化处理得

$$Y = \frac{X - m}{\sigma}, \quad X = \sigma Y + m. \tag{114}$$

利用性质 4 及 ①的结论, 得一般高斯变量的特征函数, 形式如下

$$Q_X(u) = e^{jmu} Q_Y(\sigma u) = \exp\left(jum - \frac{1}{2}\sigma^2 u^2\right). \tag{115}$$

例 .7

求二项分布的数学期望、方差和特征函数。

方法一

二项分布的分布律为

$$P(Y=k)=C_n^k p^k q^{n-k}, \quad k=0,1,2,\cdots,n \tag{116}$$

由数学期望的定义可得

$$E[Y]=\sum_{k=1}^{n} kP(Y=k)=\sum_{k=1}^{n} kC_n^k p^k(1-p)^{n-k}=\cdots=np. \tag{117}$$

由方差的定义可得

$$\begin{aligned} D[Y] &= \sum_{k=1}^{n} (k - m_Y)^2 \cdot P(Y = k) \\ &= \sum_{k=1}^{n} (k - np)^2 \cdot C_n^k p^k (1 - p)^{n-k} = \cdots = npq. \end{aligned} \tag{118}$$

直接引用特征函数的定义, 可得

$$Q_Y(u) = \sum_{k=1}^{n} e^{juk} P(Y = k) = \sum_{k=1}^{n} e^{juk} C_n^k p^k q^{n-k} = \cdots \tag{119}$$

计算较烦。

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 $1-\mathrm{p}=\mathrm{q}$，第 i 次实验的随机变量记为 X_i，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

$$Q_{X_i}(u)=E\left[e^{juX_i}\right]=\sum_{k=0}^{1}e^{juk}P\left(X_i=k\right)=q+pe^{ju}. \tag{120}$$

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

$$Q_Y(u) = \prod_{i=1}^{n} Q_{X_i}(u) = \left(q + pe^{ju}\right)^n. \tag{121}$$

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

$$\begin{aligned} E[Y] &= -j\frac{d}{du}\left(pe^{ju} + q\right)^n\Big|_{u=0} = np. \\ E\left[Y^2\right] &= (-j)^2\frac{d^2}{du^2}\left(pe^{ju} + q\right)^n\Big|_{u=0} = npq + n^2p^2 \\ D[Y] &= E\left[Y^2\right] - (E[Y])^2 = npq. \end{aligned} \tag{122}$$

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?

傅里叶变换对

$$\begin{cases} S(\omega) = F[s(t)] = \int_{-\infty}^{\infty} s(t)e^{-j\omega t}dt. \\ s(t) = F^{-1}[S(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{j\omega t}d\omega. \end{cases} \quad (123)$$

对照特征函数的定义, 可得

$$\begin{aligned} Q_X(u) &= \int_{-\infty}^{\infty} f(x)e^{jux}dx \\ &= 2\pi\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{jux}dx\right] \\ &= 2\pi F^{-1}[f(x)]. \end{aligned} \tag{124}$$

即

$$F^{-1}[f(x)] = \frac{1}{2\pi}Q_x(u). \tag{125}$$

定义 .10 逆转公式

可用傅里叶变换求

$$f(x)=F\left[\frac{1}{2\pi}Q_X(u)\right]=\frac{1}{2\pi}\int_{-\infty}^{+\infty}Q_X(u)e^{-jux}du. \qquad (126)$$

称之为逆转公式。

由傅里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。

经典方法举例

例如: 求 $Y=\sum_{i=1}^{n} X$ 的分布 $f_Y(y)$ 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

$$f_Y(y)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_X\left(x_1,\cdots,x_{n-1},y-\sum_{i=1}^{n-1}x_i\right)dx_1\cdots dx_{n-1}. \tag{127}$$

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 .8

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

$$Q_Y(u) = Q_{X_1}(u)Q_{X_2}(u) = \left(pe^{ju}+q\right)^n \cdot \left(pe^{ju}+q\right)^m = \left(pe^{ju}+q\right)^{n+m}. \tag{128}$$

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。

多维随机变量的特征函数

3. 多元特征函数

(1) 定义

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

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad U = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}. \tag{129}$$

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

$$Q_X(u_1, \cdots, u_n) = Q_X\left(U^T\right) = E\left[e^{jU^T x}\right]$$
$$= E\left[e^{j\left(u_1 X_1 + u_2 X_2 + \cdots + u_n X_n'\right)}\right] = E\left[e^{\sum_{k=1}^{n} j u_k X_k}\right]$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{j(u_1 x_1 + \cdots + u_n x_n)} f_X(x_1, \cdots, x_n) \, dx_1 \cdots dx_n. \quad (130)$$

逆转公式

逆转公式为

$$
\begin{aligned}
f_X(x_1,\cdots,x_n) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-jU^T x} Q_X\left(U^T\right)\frac{du_1}{2\pi}\cdots\frac{du_n}{2\pi} \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-j\sum\limits_{k=1}^{n} u_k x_k} Q_X(u_1,\cdots,u_n) \\
&\qquad \cdot\frac{du_1}{2\pi}\cdots\frac{du_n}{2\pi}.
\end{aligned}
\tag{131}
$$

(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

$$Q_X^*(u_1,\cdots,u_n)=Q_X(-u_1,\cdots,-u_n). \tag{132}$$

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

$$Q_Y(u_1,\cdots,u_r,0,\cdots,0)=e^{jU^T\mathbf{B}'}Q_X\left(U^T\mathbf{A}'\right), \tag{133}$$

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

特殊情况 1

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

$$A = \begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \tag{134}$$

此时

$$AX = \begin{bmatrix} a_1X_1 \\ \vdots \\ a_nX_n \end{bmatrix}, Y = AX + B = \begin{bmatrix} a_1X_1 + b_1 \\ \vdots \\ a_nX_n + b_n \end{bmatrix}, \tag{135}$$

则随机矢量 Y 的特征函数为

$$Q_Y(u_1, u_2, \cdots, u_n) = e^{j\sum_k^n u_k b_k} Q_X(a_1u_1, \cdots, a_nu_n). \tag{136}$$

特殊情况 2

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

$$Y=AX+B=a_1X_1+\cdots+a_nX_n+b \tag{137}$$

是一维随机变量, 其特征函数为

$$Q_Y(u_1)=e^{ju_1b}Q_X(a_1u_1,a_2u_1,\cdots,a_nu_1)\ (\text{一元}). \tag{138}$$

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

$$Q_X(u_1, \cdots, u_n) = \prod_{k=1}^{n} Q_{X_k}(u_k). \tag{139}$$

由逆转公式可证

$$f_X(x_1, \cdots, x_n) = \prod_{k=1}^{n} f_{X_k}(x_k). \tag{140}$$

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 $\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ 的特征函数为 $Q_X(u_1, \cdots, u_n)$, 则

其子向量 $Y = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix}$ 的特征函数为

$$Q_Y(u_1, \cdots, u_k) = Q_X(u_1, \cdots, u_k, 0, \cdots, 0) \tag{141}$$

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 $Y = \begin{bmatrix} X_{11} \\ \vdots \\ X_{ik} \end{bmatrix}$, 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

$$E\left[X_1^n X_2^k\right] = (-j)^{n+k} \left.\frac{\partial^{n+k} Q_X(u_1, u_2)}{\partial u_1^n \partial u_2^k}\right|_{u_1=0, u_2=0}. \tag{142}$$

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

$$Q_X(u_1, u_2) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} E\left[X_1^n X_2^k\right] \cdot \frac{(ju_1)^n}{n!} \cdot \frac{(ju_2)^k}{k!}. \tag{143}$$

习题

例 .9

随机变量 $X \sim U[a,b]$, $f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0 & \text{其他} \end{cases}$.

例 .10

随机变量 $X \sim B(1,p)$, 求特征函数 $Q_X(u)$.

解:

$$
\begin{aligned}
Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x)dx = \int_a^b e^{jux} \frac{1}{b-a} dx \\
&= \frac{1}{(b-a)ju} e^{jux}\Big|_a^b \\
&= \frac{1}{(b-a)ju} \left[e^{jub} - e^{jua} \right] \\
&= \frac{e^{jua}}{(b-a)ju} [e^{ju(b-a)} - 1].
\end{aligned} \tag{144}
$$

解:

$$
Q_X(u) = e^{ju\times 0} q + e^{ju \times 1} p = q + p(\cos(u) + j\sin(u)). \tag{145}
$$

例 .11

随机变量 $X \sim N(\theta, 1)$, 求 X 的特征函数 $Q_X(u)$.

因为 $X(\theta, 1)$, 所以 X 的概率密度函数为

$$\begin{aligned} p(x,\theta)dx &= (2\pi)^{-1/2}\exp\left\{-\frac{x^2}{2} + x\theta - \frac{\theta^2}{2}\right\}dx \\ &= (2\pi)^{-1/2}e^{-\theta^2/2}e^{\theta x}e^{-x^2/2}dx \doteq C(\theta)e^{\theta x}d\mu(x) \qquad (146) \end{aligned}$$

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

正态分布族和指数型分布族的关系

易见，正态分布族 $N(\theta,1)$ 满足指数型分布族 $p_\theta(x)d\mu(x) = C(\theta)e^{\theta T(x)}d\mu(x)$, 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 [?]。

由定理知，$T(X) = X$ 的特征函数为

$$Q_X(u) = \frac{C(\theta)}{C(\theta + iu)} = \exp\left\{iu\theta - \frac{u^2}{2}\right\}. \tag{147}$$

伽马分布

例 .12

随机变量 $X \sim G(\alpha, \beta)$, 即 X 有概率密度函数

$$p(x, \alpha, \beta)dx = \frac{\alpha^{\beta}}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} dx, x > 0, \tag{148}$$

$$\Gamma(\beta) = \int_0^{\infty} x^{\beta-1} e^{-x} dx, \tag{149}$$

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

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

② 当 $\alpha = 1/2, \beta = \mathsf{n}/2$ 时, $\mathsf{G}(1/2, \mathsf{n}/2)$ 即为自由度 n 的中心化 χ^2 分布 $\chi^2(\mathsf{n})$, 其特征函数为 $\mathsf{Q_X(u)} = (1 - 2\mathsf{iu})^{-\mathsf{n}/2}$.

上式顺带得到了指数分布和中心化 χ^2 分布的特征函数.

指数分布

例 .13

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

解:

$$
\begin{aligned}
Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x) dx = \int_{0}^{\infty} e^{jux} \lambda e^{-\lambda x} dx \\
&= \frac{\lambda}{ju-\lambda} e^{ju-\lambda} \Big|_0^{\infty} \\
&= \frac{\lambda}{ju-\lambda} \left[e^{(ju-\lambda)\infty} - 1\right] = \frac{-\lambda}{-\lambda+ju} = \frac{1}{1-j\frac{u}{\lambda}}.
\end{aligned}
\tag{150}
$$

例 .14

已知随机变量 X 服从拉普拉斯分布, 其概率密度为

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < +\infty, \tag{151}$$

求其特征函数。

目录

1 随机变量的数字特征
- 空气弹性变形模型
- 二维连续型随机变量函数的数学期望
- 条件数学期望

2 随机变量的矩和方差
- n 维随机变量的联合中心矩

3 相关、正交和独立
- 2. 不相关、独立和正交
- 随机变量的特征函数
- 2. 特征函数与概率密度的对应关系

4 高斯随机变量
- 多维高斯随机变量的有关性质
- n 维高斯变量的性质

5 复随机变量

1. 一维高斯随机变量及其特征

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。由中心极限定理可知这类随机变量都 (近似) 服从高斯分布，因此高斯随机变量是一类被广泛应用的随机变量。

一维高斯随机变量 X 的概率密度函数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right], \quad x \in \mathbb{R}. \qquad (152)$$

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

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in \mathbb{R} \tag{153}$$

对于一给定的 x, 高斯随机变量 X 的分布函数为

$$\begin{aligned} F(x) &= P(X \leqslant x) \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(u-m_X)^2}{2\sigma_X^2}\right] du. \end{aligned} \tag{154}$$

令 $\frac{u-m_X}{\sigma_X}=t$，则

$$F(x)=\int_{-\infty}^{(x-m_X)/\sigma_X}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt=\Phi\left(\frac{x-m_X}{\sigma_X}\right), \tag{155}$$

其中 $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt$ 的关于概率密度的变上限积分.

高斯分布函数还可以表示为

$$\begin{aligned} F(x)&=\frac{1}{2}+\frac{1}{2}\operatorname{erf}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right)\\ F(x)&=1-\frac{1}{2}\operatorname{erfc}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right), \end{aligned} \tag{156}$$

其中 erf(x) 称为误差函数和互补误差函数。

定义 .11 误差函数和互补误差函数

$$\mathrm{erf}(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt.$$
$$\mathrm{erfc}(x)=1-\mathrm{erf}(x)=\frac{2}{\sqrt{\pi}}\int_x^{+\infty}e^{-t^2}dt. \tag{157}$$

(152) 式所表示的高斯随机变量 X 具有下列性质：

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 $\frac{1}{\sqrt{2\pi}\sigma_x}$；

2) 当 $x\to\pm\infty$ 时，概率密度函数的值趋于零；

3) 均值反映了 X 概率密度函数的最大值所在的位置信息；

4) 方差反映了概率密度函数曲线的陡度, 随着方差的增大曲线变得平坦；

5) 随机变量 X 落在 $[m_X - 3\sigma_X, m_X + 3\sigma_X]$ 区间的概率

$$P(m_X - 3\sigma_X \leqslant x \leqslant m_X + 3\sigma) = \frac{1}{\sqrt{2\pi}\sigma_X}\int_{m_X-3_X^\sigma}^{m_X+3_X^\sigma} \exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right] dx \approx 99.7\%. \quad (158)$$

高斯随机变量 X 的特征函数为

$$\varphi(\omega) = \exp\left(j\omega m_X - \frac{1}{2}\omega^2\sigma_X^2\right). \tag{159}$$

高斯随机变量 X 的 n 阶中心矩

$$E\left\{(X - m_X)^n\right\} = \begin{cases} 0, & n = 2k - 1 \\ 1 \times 3 \times 5 \times \cdots \times (n-1)\sigma_X^n, & n = 2k \end{cases} \tag{160}$$

其中 $k = 1, 2, \cdots$.

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 $\sigma_{X_2}^2$ 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

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}\sqrt{1-\rho_{X_1X_2}^2}} \exp\left\{-\frac{1}{2\left(1-\rho_{X_1X_2}^2\right)}\left[\frac{\left(x_1-m_{X_1}\right)^2}{\sigma_{X_1}^2} - \frac{2\rho_{X_1X_2}\left(x_1-m_{X_1}\right)\left(x_2-m_{X_2}\right)}{\sigma_{X_1}\sigma_{X_2}} + \frac{\left(x_2-m_{X_2}\right)^2}{\sigma_{X_2}^2}\right]\right\}. \tag{161}$$

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

$$
\begin{aligned}
f_{X_1}(x_1) &= \int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_2 = \frac{1}{\sqrt{2\pi}\sigma_{X_1}} \exp\left[-\frac{(x_1 - m_{X_1})^2}{2\sigma_{X_1}^2}\right]. \\
f_{X_2}(x_2) &= \int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_1 = \frac{1}{\sqrt{2\pi}\sigma_{X_2}} \exp\left[-\frac{(x_2 - m_{X_2})^2}{2\sigma_{X_2}^2}\right].
\end{aligned}
\tag{162}
$$

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

$$\varphi(\omega_1,\omega_2)=\exp\Big[\mathrm{j}\left(m_{X_1}\omega_1+m_{X_2}\omega_2\right)\\
-\frac{1}{2}\left(\sigma_{X_1}^2\omega_1^2+2\rho_{X_1X_2}\sigma_{X_1}\sigma_{X_2}\omega_1\omega_2+\sigma_{X_2}^2\omega_2^2\right)\Big]. \tag{163}$$

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 $\mathbf{x} = [X_1, X_2, \cdots, X_n]^T$，均值矢量为

$$\begin{aligned} m_X &= [E\{X_1\}, E\{X_2\}, \cdots, E\{X_n\}]^T \\ &= [m_{X_1}, m_{X_2}, \cdots, m_{X_n}]^T \end{aligned} \tag{164}$$

它的协方差矩阵为

$$C_X = \begin{bmatrix} E\left\{(X_1 - m_{x_1})^2\right\} & \cdots & E\{(X_1 - m_{x_1})(X_n - m_{x_n})\} \\ \vdots & & \vdots \\ E\{(X_n - m_{x_n})(X_1 - m_{x_1})\} & \cdots & E\left\{(X_n - m_{x_n})^2\right\} \end{bmatrix}. \tag{165}$$

则 n 维高斯随机变量的联合概率密度函数可以表示为

$$f(\mathrm{x})=\frac{1}{(2\pi)^{\frac{n}{2}}\left|\mathrm{C_x}\right|^{\frac{1}{2}}}\exp\left[-\frac{(\mathrm{x}-\mathrm{m_X})^{\mathrm{T}}\mathrm{C_x^{-1}}(\mathrm{x}-\mathrm{m_X})}{2}\right]. \quad (166)$$

n 维高斯随机变量的特征函数为

$$\varphi_{\mathrm{x}}(\omega)=\exp\left(\mathrm{j}\mathrm{m_X^T}\omega-\omega^{\mathrm{T}}\mathrm{C_x}\omega/2\right), \quad (167)$$

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

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 ⩾ 2)) 阶矩为

$$E\left[Y^*\right]=\begin{cases}1\cdot 3\cdot 5\cdots(n-1)=(n-1)!, & n\text{为偶数}\\ 0, & n\text{为奇数}\end{cases}, n\geqslant 2 \tag{168}$$

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

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

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 $X_k \sim N\left(m_k, \sigma_k^2\right), k = 1, \cdots, n$, 那么其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 $\sigma_Y^2 = \sum_{k=1}^{n} \sigma_k^2$.

3. 高斯变量的和的分布

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 **x** 表示 X 的取值, 用矢量 $\mathbf{M}_X$ 表示随机矢量 **X** 的均值

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad M_x = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}. \tag{169}$$

方差

$$
\begin{aligned}
D(\mathbf{X}) &= C_X \\
&= E\left[(\mathbf{X}-\mathbf{M}_X)(\mathbf{X}-\mathbf{M}_X)^T\right] \\
&= \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1\pi} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n\times n}
\end{aligned} \tag{170}
$$

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 **X** 的概率密度) 可以写成如下矩阵形式:

$$f_X(x_1, \cdots, x_n) = \frac{1}{(2\pi)^{n/2} |C_X|^{1/2}} \exp\left[-\frac{(x - M_x)^T C_X^{-1} (x - M_x)}{2}\right] \tag{171}$$

式中 $(x - M_X)^T$ 表示 $(x - M_X)$ 的转置。n 维高斯分布可记为

$$\mathbf{X} \sim N(\mathbf{M}_X, \mathbf{C}_X).$$

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

$$\begin{aligned} Q_X(u_1, \cdots, u_n) &= E\left[e^{jU^T X}\right] \\ &= \exp\left[jM_X^T U - \frac{U^T C_X U}{2}\right], \end{aligned} \tag{172}$$

其中

$$U = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \quad U^T X = \sum_{i=1}^{n} u_i X_i. \tag{173}$$

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。

$$C_X = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}_{n\times n} . \tag{174}$$

n 维联合特征函数

$$
\begin{aligned}
Q_X(u_1,\cdots,u_n) &= \exp\left\{ j\begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}^T \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} - \frac{1}{2}\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}^T \Sigma \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \right\} \\
&= \exp\left[\sum_{k=1}^{n} j m_k u_k - \sum_{k=1}^{n} \frac{u_k^2\sigma_k^2}{2} \right] \\
&= \exp\left[\sum_{k=1}^{n} \left(j m_k u_k - \frac{u_k^2\sigma_k^2}{2} \right) \right].
\end{aligned}
\tag{175}
$$

$$
\begin{aligned}
Q_X(u_1, \cdots, u_n) &= \prod_{k=1}^{n} \exp\left[j m_k u_k - \frac{u_k^2 \sigma_k^2}{2}\right] \\
&= \prod_{k=1}^{n} Q_{X_k}(u_k),
\end{aligned} \tag{176}
$$

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

$$
\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}.
$$

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

$$\mathbf{Y} = \mathbf{A}_{m\times n}\mathbf{X}_{n\times 1} = \begin{bmatrix} Y_1 = a_{11}X_1 + \cdots + a_{1n}X_n \\ Y_2 = a_{21}X_1 + \cdots + a_{2n}X_n \\ \vdots \\ Y_m = a_{m1}X_1 + \cdots + a_{mn}X_n \end{bmatrix}. \quad (177)$$

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

$$
\begin{aligned}
\mathbf{M}_Y &= E[Y] = E[\mathbf{AX}] = \mathbf{A}E[\mathbf{X}] = \mathbf{AM}_X \\
\mathbf{C}_Y &= D[Y] = E\left[(\mathbf{Y} - \mathbf{M}_Y)(\mathbf{Y} - \mathbf{M}_Y)^T\right] \\
&= E\left[(\mathbf{AX} - \mathbf{AM}_X)(\mathbf{AX} - \mathbf{AM}_X)^T\right] \\
&= E\left[\mathbf{A}(\mathbf{X} - \mathbf{M}_X)(\mathbf{X} - \mathbf{M}_X)^T \mathbf{A}^T\right] \\
&= \mathbf{A}E\left[(\mathbf{X} - \mathbf{M}_X)(\mathbf{X} - \mathbf{M}_X)^T\right] \mathbf{A}^T \\
&= \mathbf{AC}_X\mathbf{A}^T.
\end{aligned} \tag{178}
$$

$$
\begin{aligned}
Q_Y(u_1,\cdots,u_n) &= E\left[\exp\left(j\mathbf{U}^\top\mathbf{Y}\right)\right] = E\left[\exp\left(j\mathbf{U}^T\mathbf{A}\mathbf{X}\right)\right] \\
&= E\left(\exp\left[j\left(\mathbf{A}^\top\mathbf{U}\right)^\top\mathbf{X}\right]\right. \\
&= Q_x\left(\mathbf{U}^T\mathbf{A}\right) = \exp\left[j\mathbf{M}_X^T\left(\mathbf{A}^\top\mathbf{U}\right) - \frac{\left(\mathbf{A}^\top\mathbf{U}\right)^\top \mathbf{C}_X\left(\mathbf{A}^\top U\right)}{2}\right] \\
&= \exp\left[j\left(\mathbf{A}\mathbf{M}_x\right)^\top\mathbf{U} - \frac{\mathbf{U}^T\left(\mathbf{A}\mathbf{C}_X\mathbf{A}^T\right)U}{2}\right] \\
&= \exp\left[j\mathbf{M}_Y^\top\mathbf{U} - \frac{\mathbf{U}^\top\mathbf{C}_X\mathbf{U}}{2}\right]. \qquad (179)
\end{aligned}
$$

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 **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 $Y_1, \cdots, Y_m$

$$\mathbf{C}_Y = \mathbf{A}\mathrm{C}_X\mathrm{A}^\top = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{X_2}^2 \end{bmatrix}. \tag{180}$$

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 $X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$，其中任一子矢量为 $\tilde{\mathbf{X}} = \begin{bmatrix} X_i \\ \vdots \\ X_{ik} \end{bmatrix}$，$k < n$。数学期望 $M_X = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}$，$M_{\tilde{X}} = \begin{bmatrix} m_n \\ \vdots \\ m_{ik} \end{bmatrix}$ 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，$C_X = \begin{bmatrix} C_{\tilde{X}} & C_1 \\ C_2 & C_3 \end{bmatrix}_{n\times n}$ 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

$$Q_{\tilde{X}}\left(u_{i1},\cdots,u_{ik}\right)=Q_{X}\left(u_{n},\cdots,u_{ik},0,\cdots,0\right), \tag{181}$$

其中 $\tilde{U}=\left(u_{i1},\cdots,u_{ik}\right)^{T}$ 为 $U=\left(u_{1},\cdots,u_{n}\right)^{T}$ 的子矢量。

若 **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

$$
\begin{aligned}
Q_X(u_{i1},\cdots,u_{ik}) &= \exp\left[j(m_1,\cdots,m_n)\begin{pmatrix} u_{i1} \\ \vdots \\ u_{ik} \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} u_{i1} \\ \cdots \\ u_{ik} \\ 0 \\ \cdots \\ 0 \end{pmatrix}^T C_X \begin{pmatrix} u_{i1} \\ \vdots \\ u_{ik} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right] \\
&= \exp\left[j\mathbf{M}_X^T\bar{\mathbf{U}} - \frac{1}{2}\mathbf{U}^T\mathbf{C}_S\bar{\mathbf{U}} \right]. \qquad (182)
\end{aligned}
$$

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

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 $\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, 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。

解法

① 随机矢量 **X** 的方差

$$\mathrm{D}[\mathbf{X}] = \left[\mathbf{B}^{-1}\right]^{-1} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ -\frac{5}{14} & \frac{15}{28} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{2}{7} \end{bmatrix}. \tag{183}$$

① 随机矢量 $\mathbf{Y} = \mathbf{AX}$, 其中 $\mathbf{A}$。由性质 2°, 可知

$$\mathbf{M_Y} = \mathbf{AM_X} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{184}$$

差的逆矩阵为

$$\mathbf{C}_Y=\mathbf{A}\mathbf{C}_X\mathbf{A}^T,=\begin{pmatrix}1&0&0\\1&1&0\end{pmatrix}\begin{bmatrix}\frac{4}{7}&-\frac{5}{14}&-\frac{1}{7}\\-\frac{5}{14}&\frac{15}{28}&\frac{3}{14}\\-\frac{1}{7}&\frac{3}{14}&\frac{2}{7}\end{bmatrix}\begin{bmatrix}1&1\\0&1\\0&0\end{bmatrix}$$

$$=\begin{pmatrix}\frac{4}{7}&-\frac{5}{14}&-\frac{1}{7}\\\frac{3}{14}&\frac{5}{28}&\frac{1}{14}\end{pmatrix}\begin{pmatrix}1&1\\0&1\\0&0\end{pmatrix}=\begin{bmatrix}\frac{4}{7}&\frac{3}{14}\\\frac{3}{14}&\frac{11}{28}\end{bmatrix}. \tag{185}$$

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

则其概率密度

$$
\begin{aligned}
f_Y(y_1, j_z) &= \frac{1}{2\pi \left|\mathbf{C}_Y\right|^{\frac{1}{2}}} \exp\left[-\frac{(y - \mathbf{M}_Y)^T \mathbf{C}_{\bar{Y}}^{-1} (y - \mathbf{M}_Y)}{2}\right] \\
&= \frac{1}{2\pi \left|\mathbf{C}_Y\right|^{\frac{1}{2}}} \exp[\underbrace{y^T \mathbf{C}_\gamma^{-1} \mathbf{y}}] \\
&= \frac{1}{\pi}\sqrt{\frac{7}{5}} \exp\left[-\frac{11y_1^2 - 12y_1y_2 + 16y_7^1}{10}\right].
\end{aligned}
\tag{186}
$$

其特征函数为

$$Q_Y(u_1, u_2) = \exp\left[jM_Y^{\top}U - \frac{U^{T}C_{Y}U}{2}\right] = \exp\left[-\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56}\right]. \tag{187}$$

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

$$Q_Y(u_1, u_2) = \exp\left[jM_Y^{\top}U - \frac{U^{T}C_{\gamma}U}{2}\right] = \exp\left[-\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56}\right]. \tag{188}$$

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。

目录

1 随机变量的数字特征
- 空气弹性变形模型
- 二维连续型随机变量函数的数学期望
- 条件数学期望

2 随机变量的矩和方差
- n 维随机变量的联合中心矩

3 相关、正交和独立
- 2. 不相关、独立和正交
- 随机变量的特征函数
- 2. 特征函数与概率密度的对应关系

4 高斯随机变量
- 多维高斯随机变量的有关性质
- n 维高斯变量的性质

5 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。

定义 .12 复随机变量

$$Z = X_1 + jX_2, \tag{189}$$

其中，X_1 和 X_2 为两个实随机变。

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

$$f(z) = f(x_1, x_2). \tag{190}$$

一个复随机变显的均值为

$$\begin{aligned} m_Z &= E\{Z\} = E\{X_1 + jX_2\} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x_1 + jx_2) f(x_1, x_2)\, dx_1 dx_2. \end{aligned} \tag{191}$$

或者，定义 $dz = dx_1 dx_2$ 作为复 z 平面上的无穷小区域，有

$$m_Z = E\{Z\} = \int_{(z)} z f(z) dz, \tag{192}$$

且

$$\begin{aligned} E\left\{|Z|^2\right\} &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |x_1 + jx_2|^2 f(x_1, x_2)\, dx_1 dx_2 \\ &= E\left\{X_1^2\right\} + E\left\{X_2^2\right\}. \end{aligned} \tag{193}$$

则复随机变量的方差为

$$\sigma_Z^2 = E\left\{|Z - m_Z|^2\right\} = \sigma_{X_1}^2 + \sigma_{X_2}^2 = E\left\{|Z|^2\right\} - m_Z^2. \tag{194}$$

对于两个复随机变量，$W = U_1 + jU_2, Z = X_1 + jX_2$，概率密度函数为

$$f(w, z) = f\left(x_1, x_2, u_1, u_2\right). \tag{195}$$

其协方差定义为

$$C_{WZ} = E\left\{\left(W - m_W\right)\left(Z - m_Z\right)\right\}. \tag{196}$$

相关系数

$$\rho_{WZ} = \frac{C_{WZ}}{\sigma_W \sigma_Z}. \tag{197}$$

一般情况下相关系数为复值。

对于复随机变量 **w** 和 **z**, 我们定义协方差矩阵和互协方差矩阵为

$$\begin{aligned} C_Z &= E\left\{(\mathbf{Z}-\mathbf{m}_Z)(\mathbf{Z}-\mathbf{m}_Z)^H\right\}. \\ C_{WZ} &= E\left\{(\mathbf{W}-\mathbf{m}_W)(\mathbf{Z}-\mathbf{m}_Z)^H\right\}. \end{aligned} \tag{198}$$

复高斯随机矢量的概率密度函数

特别地，对 n 维复高斯随机矢量 **z**, 若其均值和协方差矩阵用 $\mathbf{m}_Z$ 和 $\mathbf{C}_Z$ 表示，

$$f(z) = \frac{1}{\pi^n |C_Z|} \exp\left[-(z - m_Z)^H C_Z^{-1} (z - m_Z)\right]. \qquad (199)$$

高斯随机变量的特征函数由下式给出

$$\varphi(\omega) = \exp\left(j\,\mathrm{Re}\left\{\omega^H m_Z\right\} - \frac{\omega^H C_Z \omega}{4}\right), \qquad (200)$$

其中 Re$\{\cdot\}$ 表示复数的取实部操作。

例 5

例 .1

求积分 $\int_C (2z^2+8z+1)dz$, 其中 C 是连接 0 到 $2\pi a$ 的摆线, 方程为 $x=a(\theta-\sin\theta)$, $y=a(1-\cos\theta)$.

解: 函数 $f(z)=2z^2+8z+1$ 在复平面内解析, 所以积分与路线无关, 根据牛—莱公式:

$$\int_C (2z^2+8z+1)dz = \int_0^{2\pi a}(2z^2+8z+1)dz$$
$$= \frac{2}{3}z^3+4z^2+z\Big|_0^{2\pi a} = \frac{16}{3}\pi^3a^3+16\pi^2a^2+2\pi a.$$

```
clc
clear
syms z x y; pz=2*z^2+8*z+1
pzxy= expand(subs(pz,z,x+i*y))
syms a theta
pzxy1=subs(pzxy,x,a*theta-a*sin(theta));
pzxy2=subs(pzxy1,y,a-a*cos(theta))
pzxy3=pzxy2*(a*(1-cos(theta))+i*a*sin(theta))
collect(pzxy3,i)
imfun1=- a*sin(theta)*(8*a - 8*a*cos(theta) + 4*(a*theta - a*sin(theta))*(a - a*cos(theta))) + (a*sin(theta)*(8*a*theta + 2*(a*theta - a*sin(theta))^2 - 8*a*sin(theta) - 2*(a - a*cos(theta))^2 + 1) - a*(cos(theta) - 1)*(8*a - 8*a*cos(theta) + 4*(a*theta - a*sin(theta))*(a - a*cos(theta))))
realfun1= - a*(cos(theta) - 1)*(8*a*theta + 2*(a*theta - a*sin(theta))^2 - 8*a*sin(theta) - 2*(a - a*cos(theta))^2 + 1)
intim=int(imfun1,theta,0,2*pi)
intreal=int(realfun1,theta,0,2*pi)
resl=simplify(intim+intreal)
pretty(collect(resl,a))
```

代 码 1: 复变量 z 参数方程的线积分计算.

运行结果见图

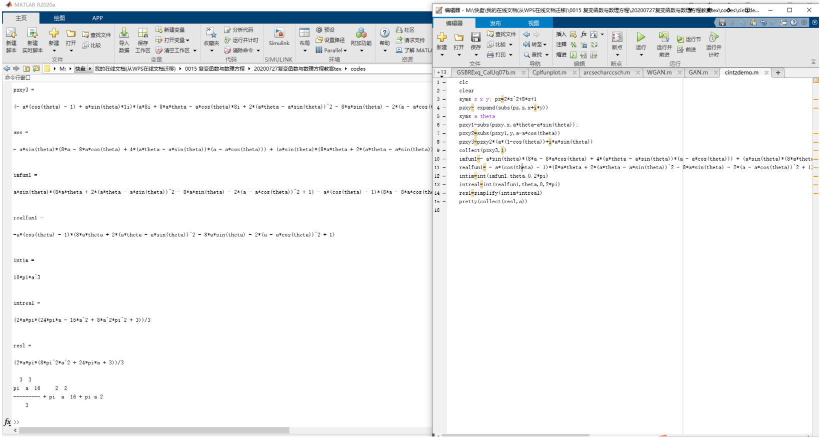


图 5: matlab 符号计算复积分 $\int_C (2z^2 + 8z + 1)dz$.