概率论与数理统计

第四章 随机变量的数字特征

§4 矩、协方差矩阵

- 原点矩、中心矩
- 协方差矩阵
- n维正态随机变量的概率密度

一、 原点矩　中心矩

定义　设X和Y是随机变量，若

$$E(X^k), k = 1,2,\cdots$$

存在，称它为X的k阶原点矩，简称 k阶矩.

若　$$E\{[X - E(X)]^k\}, k = 2,3,\cdots$$

存在，称它为X的k阶中心矩.

可见，均值 E（X）是X一阶原点矩，方差D（X）是X的二阶中心矩.

设 X 和 Y 是随机变量，若

$$E(X^k Y^l) \quad k,l=1,2,\ldots$$ 存在，

称它为 X 和 Y 的 $k+l$ 阶混合（原点）矩.

若 $E\{[X-E(X)]^k[Y-E(Y)]^l\}$ 存在，

称它为 X 和 Y 的 $k+l$ 阶混合中心矩.

可见，

协方差 $Cov(X,Y)$ 是 X 和 Y 的二阶混合中心矩.

说明

(1) 以上数字特征都是随机变量函数的数学期望;

(2) 随机变量 X 的数学期望 $E(X)$ 是 X 的一阶原点矩, 方差为二阶中心矩, 协方差 $\mathrm{Cov}(X,Y)$是 X 与 Y 的二阶混合中心矩;

(3) 在实际应用中, 高于 4 阶的矩很少使用.

三阶中心矩$E\{[X-E(X)]^3\}$主要用来衡量随机变量的分布是否有偏.

四阶中心矩 $E\{[X-E(X)]^4\}$主要用来衡量随机变量的分布在均值附近的陡峭程度如何.

二、协方差矩阵

将二维随机变量（X_1,X_2）的四个二阶中心矩

$$c_{11} = E\{[X_1 - E(X_1)]^2\}$$

$$c_{12} = E\{[X_1 - E(X_1)][X_2 - E(X_2)]\}$$

$$c_{21} = E\{[X_2 - E(X_2)][X_1 - E(X_1)]\}$$

$$c_{22} = E\{[X_2 - E(X_2)]^2\}$$

排成矩阵的形式: $\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$

这是一个

对称矩阵

称此矩阵为（X_1,X_2）的协方差矩阵.

类似定义n维随机变量$(X_1,X_2,\ldots,X_n)$的协方差矩阵.

若 $c_{ij} = Cov(X_i, X_j)$

$$= E\{[X_i - E(X_i)][X_j - E(X_j)]\}$$

$$(\, i, j=1,2,\ldots,n \,)$$

都存在，称

矩阵 $$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为$(X_1,X_2,\ldots,X_n)$的协方差矩阵

协方差矩阵的应用

协方差矩阵可用来表示多维随机变量的概率密度，从而可通过协方差矩阵达到对多维随机变量的研究

以二维正态随机变量 (X_1, X_2) 为例.概率密度

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}\right.\right.$$

$$\left.\left.-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

现在将上式中花括号内的式子写成矩阵形式,为此引入下面的列矩阵

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

(X_1, X_2) 的协方差矩阵为

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

它的行列式 $\det C = \sigma_1^2\sigma_2^2(1-\rho^2)$, C 的逆矩阵为

$$C^{-1} = \frac{1}{\det C}\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$$

经过计算可知 (这里矩阵$(X-\mu)^{\mathrm{T}}$是$(X-\mu)$的转置矩阵).

$$(X-\mu)^{\mathrm{T}}C^{-1}(X-\mu)$$

$$=\frac{1}{\det C}(x_1-\mu_1,x_2-\mu_2)\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}\cdot\begin{pmatrix}x_1-\mu_1\\ x_2-\mu_2\end{pmatrix}$$

$$=\frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right].$$

于是 (X_1,X_2) 的概率密度可写成

$$f(x_1,x_2)$$

$$=\frac{1}{(2\pi)^{2/2}(\det C)^{1/2}}\exp\left\{-\frac{1}{2}(X-\mu)^{\mathrm{T}}C^{-1}(X-\mu)\right\}.$$

引入列矩阵

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ 和 } \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{pmatrix},$$

n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的概率密度定义为

$$f(x_1, x_2, \cdots, x_n)$$

$$= \frac{1}{(2\pi)^{n/2}(\det C)^{1/2}} \exp\left\{-\frac{1}{2}(X-\mu)^{\mathrm{T}} C^{-1}(X-\mu)\right\}.$$

其中 C 是 $(X_1, X_2, \cdots, X_n)$ 的协方差矩阵.

三、n 维正态随机变量的概率密度

引入矩阵记号

$$X=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},\quad \mu=\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}=\begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{bmatrix}$$

n维正态随机向量$(X_1, X_2, \ldots, X_n)$的概率密度定义为

$$f(x_1,x_2,\ldots,x_n)=\frac{1}{(2\pi)^{n/2}(\det C)^{1/2}}\exp\{-\frac{1}{2}(X-\mu)'C^{-1}(X-\mu)\}$$

其中C是$(X_1, X_2, \ldots, X_n)$）的协方差矩阵，$\det C$是矩阵C的行列式，C^{-1}表示C的逆矩阵.

n维正态变量具有以下重要性质

1、n维正态变量$(X_1,X_2,\ldots,X_n)$ 的每一个分量X_i, $i=1,2,\ldots,n$都是正态变量；反之，若$X_1,X_2,\ldots,X_n$都是正态变量，且相互独立，则$(X_1,X_2,\ldots,X_n)$是n维正态变量.

2、n维随机变量$(X_1,X_2,\ldots,X_n)$ 服从n维正态分布的充要条件是$X_1,X_2,\ldots,X_n$的任意的线性组合：

$$l_1X_1+l_2X_2+\ldots+l_nX_n$$

服从一维正态分布（其中实数 $l_1,l_2,\ldots,l_n$不全为0 ）.

3. 正态变量的线性变换不变性.

若 $X=(X_1, X_2, \ldots, X_n)$ 服从 n 维正态分布，

设 $Y_1, Y_2, \ldots, Y_k$ 是 X_j（$j=1,2,\ldots,n$）的线性函数，

则 $(Y_1, Y_2, \ldots, Y_k)$ 也服从多维正态分布.

4. 设 $(X_1, X_2, \ldots, X_n)$ 服从 n 维正态分布，则

"$X_1, X_2, \ldots, X_n$ 相互独立"

等价于

"$X_1, X_2, \ldots, X_n$ 两两不相关"

n 维正态分布在随机过程和数理统计中常会遇到.

例1　设随机变量X和Y相互独立且$X\sim N(1,2),Y\sim N(0,1)$.试求$Z=2X-Y+3$的概率密度.

解: $X\sim N(1,2),Y\sim N(0,1)$，且 X 与 Y 独立,

故X 和Y 的联合分布为正态分布，X 和Y 的任意线性组合是正态分布.

即　$Z\sim N(E(Z), D(Z))$

$E(Z)=2E(X)-E(Y)+3=2+3=5$

$D(Z)=4D(X)+D(Y)=8+1=9$

$\mathbf{Z\sim N(5, 3^2)}$

故 Z 的概率密度是 $f_Z(z)=\dfrac{1}{3\sqrt{2\pi}}e^{-\frac{(z-5)^2}{18}}$，$-\infty < z < \infty$