第四章 随机变量的数字特征

- §4.1 数学期望
- §4.2 方差
- §4.3 协方差及相关系数
- §4.4 矩、协方差矩阵

§4.4 矩、协方差矩阵

- 本节引入随机变量的另外几个数字特征

定义：X和Y是随机变量

1° 若$E(X^k)$存在，$k=1,2,\ldots$，称它为X的k阶原点矩，简称k阶矩

当$k=1$时即为数学期望$E(X)$,它是X的一阶原点矩

2° 若$E\{[X-E(X)]^k\}$存在，$k=1,2,\ldots$，称它为X的k阶中心矩

当$k=2$时即为方差$D(X)$,它是X的二阶中心矩

3° 若$E(X^kY^l)$存在，$k,l=1,2,\ldots$，称它为X和Y的$k+l$阶混合矩

4° 若$E\{[X-E(X)]^k[Y-E(Y)]^l\}$存在，$k,l=1,2,\ldots$，称它为X和Y的$k+l$阶混合中心矩

当$k=l=1$时即为协方差$Cov(X,Y)$,它是X和Y二阶混合中心矩

§4.4 矩、协方差矩阵

n维随机变量的协方差矩阵

- **首先考虑二维随机变量的协方差矩阵**
- **二维随机变量(X_1,X_2)有四个二阶中心矩，设它们都存在，分别记为**

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

$$c_{12}=E\{[X_1-E(X_1)][X_2-E(X_2)]\} \qquad \mathrm{Cov}(X_1,X_2)$$

$$c_{21}=E\{[X_2-E(X_2)][X_1-E(X_1)]\} \qquad \mathrm{Cov}(X_2,X_1)$$

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

- **将它们排成矩阵的形式，** $\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ **这个矩阵称为随机变量(X_1,X_2)的协方差矩阵**

§ 4.4 矩、协方差矩阵

设 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的二阶混合中心矩

$$c_{ij} = \mathrm{Cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}$$

$i, j = 1,2,\cdots,n.$ 都存在，则称矩阵

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

为 n 维随机变量的协方差矩阵

§4.4 矩、协方差矩阵

- 对于n维随机变量$(X_1,X_2,\ldots,X_n)$的协方差矩阵。由于$c_{ij}=c_{ji}$，$(i\neq j, i, j=1, 2, \ldots, n)$因而是一个**对称矩阵**
- 在实际中，由于n维随机变量的分布是不知道的，或者太复杂，数学上不易处理，因此**协方差矩阵尤为重要**
- **n维正态随机变量的概率密度的协方差矩阵表示**
  - 首先将二维正态随机变量的概率密度改写成另一种形式
  - 二维正态随机变量(X_1,X_2)的概率密度为

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

，$-\infty<x_1<\infty,-\infty<x_2<\infty$。

  - 现在将指数中括号内的部分写成矩阵的行列式形式

§ 4.4 矩、协方差矩阵

- 令$X=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$，$\mu=\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$，记$(X_1,X_2)$的协方差矩阵为

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

其中$\rho\sigma_1\sigma_2=\dfrac{Cov(X_1,X_2)}{\sqrt{D(X_1)}\sqrt{D(X_2)}}\sigma_1\sigma_2$

$$=Cov(X_1,X_2)=Cov(X_2,X_1)$$

- C的行列式$|C|=\sigma_1^2\sigma_2^2(1-\rho^2)$，
- 于是C的逆阵$C^{-1}=\dfrac{1}{|C|}C^*=\dfrac{1}{|C|}\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$

§ 4.4 矩、协方差矩阵

做如下计算

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

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

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

于是(X_1,X_2)的概率密度可写为

- $$f(x_1,x_2)=\frac{1}{(2\pi)^{2/2}|C|^{1/2}}e^{-\frac{1}{2}(X-\mu)'C^{-1}(X-\mu)},$$

$$-\infty<x_1<\infty,-\infty<x_2<\infty$$

§ 4.4 矩、协方差矩阵

- 推广到n维正态随机变量$(X_1,X_2,\ldots,X_n)$的情况
- 引入矩阵

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

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

$-\infty<x_i<\infty$，$i=1$，2，…，n

C是$(X_1,X_2,\ldots,X_n)$的协方差矩阵

§4.4 矩、协方差矩阵

- n维正态随机变量的四条重要性质：
- 1° n维正态随机变量$(X_1,X_2,\ldots,X_n)$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$(X_1,X_2,\ldots,X_n)$是正态变量

证明：对于第一个问题，由归纳法，正态变量的边缘概率密度都是正态变量

反之，若相互独立，则协方差为0，协方差矩阵只有主对角线元素不为0且是各正态变量的方差。

- 2° n维正态随机变量$(X_1,X_2,\ldots,X_n)$服从n维正态分布的充要条件是$X_1,X_2,\ldots,X_n$的任意线性组合：

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

服从一维正态分布(其中$l_1,l_2,\ldots,l_n$不全为0)

§4.4 矩、协方差矩阵

- 3° 正态变量的线性变换不变性：若$(X_1,X_2,\ldots,X_n)$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"与"$X_1,X_2,\ldots,X_n$两两不相关"是等价的。

这是因为，正态变量的性质由协方差矩阵完全确定，而矩阵中的元素为两两分量的协方差，反映了两两相关性

本章小结

数学期望(期望，均值)：级数或广义积分绝对收敛

函数的数学期望 $Y=g(X)$

- 连续型 $E(g(X))=\int_{-\infty}^{\infty} g(x)f(x)dx$
- 离散型 $E(g(X))=\sum_{i=1}^{\infty} g(x_i)p(x_i)$

数学期望的性质

- $1^\circ\ E(C)=C$
- $2^\circ\ E(CX)=CE(X)$
- $3^\circ\ E(X+Y)=E(X)+E(Y)$
- $4^\circ\ E(XY)=E(X)E(Y)$，当 X 和 Y 独立时

变量分解法求数学期望，以上定义可以推广到多维，求解时要用到多维随机变量的联合分布

契比雪夫不等式

设随机变量X具有数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$，则 $\forall\ \varepsilon$，如下不等式成立

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$$

方差(描述分散度) $D(X)=E\{[X-E(X)]^2\}$

方差的计算：(1)根据定义式；(2)利用方差恒等式 $D(X)=E(X^2)-E(X)^2$

标准差 $\sqrt{D(X)}$

六种重要分布的数学期望和方差
参见第二章所述

方差的性质

- $1^\circ\ D(C)=0$
- $2^\circ\ D(CX)=C^2D(X)$
- $3^\circ\ D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X,Y)$
- $4^\circ\ D(X)=0 \Leftrightarrow P\{X=C\}=1$

协方差 $\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$

相关系数 $\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$

相关系数的性质 $1^\circ\ |\rho_{XY}|\leq 1$ $2^\circ\ |\rho_{XY}|=1 \Leftrightarrow \exists\ a,b$ 使 $P\{Y=a+bX\}=1$

ρ_{XY} 仅说明X和Y的线性相关性，因此X，Y不相关，不一定独立　　因此X和Y独立，则一定不相关

对于二维正态随机变量X,Y则存在特例：$\rho=0 \Leftrightarrow$ X和Y相互独立

矩、协方差矩阵

- l 阶中心矩 $E\{[X-E(X)]^l\}$
- l 阶原点矩 $E(X^l)$
- $l+m$ 阶混合原点矩 $E(X^lY^m)$
- $l+m$ 阶混合中心矩 $E\{[X-E(X)]^l[Y-E(Y)]^m\}$

二阶正态随机变量的协方差矩阵

$$\begin{bmatrix} D(X) & Cov(X,Y) \\ Cov(Y,X) & D(Y) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

本章小结（2）

关于正态分布与独立性的若干结论

1°m个标准正态分布的随机变量之和(Z)的分布仍然服从正态分布Z～N(0,m)

$Y = aX + b \sim N(a\mu + b, (a\sigma)^2)$

2°一般，如果X和Y相互独立且$X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，则有

$Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

3°以上结论推广到n个相互独立的正态随机变量之和也成立

$X_i \sim N(\mu_i, \sigma_i^2)$，$i = 1, \ldots, n$，

$Z = X_1 + X_2 +, \ldots, + X_n \sim N(\mu_1 + \mu_2 + \ldots + \mu_n, \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2)$

4°更一般的有限个相互独立的正态随机变量的线性组合仍然服从正态分布

$Z = a_1X_1 + a_2X_2 +, \ldots, + a_nX_n \sim N(a_1\mu_1 + a_2\mu_2 + \ldots + a_n\mu_n, (a_1\sigma_1)^2 + (a_2\sigma_2)^2 + \ldots + (a_n\sigma_n)^2)$

5°n维正态随机变量$(X_1, X_2, \ldots, X_n)$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$(X_1, X_2, \ldots, X_n)$是正态变量

6°n维正态随机变量$(X_1, X_2, \ldots, X_n)$服从n维正态分布的充要条件是$X_1, X_2, \ldots, X_n$的任意线性组合：$l_1X_1 + l_2X_2 + \ldots + l_nX_n$ 服从一维正态分布(其中$l_1, l_2, \ldots, l_n$不全为0)

7°正态变量的线性变换不变性：若$(X_1, X_2, \ldots, X_n)$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

8°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”与“$X_1, X_2, \ldots, X_n$两两不相关”是等价的。

本章作业

- 第一次：
  - P_{113}：1(3)，2，4，6，7
- 第二次：
  - P_{115}：9(2)，10，14，15，18，22(2)，25
- 第三次：
  - P_{116}：26(1)，29，32，33，34，36