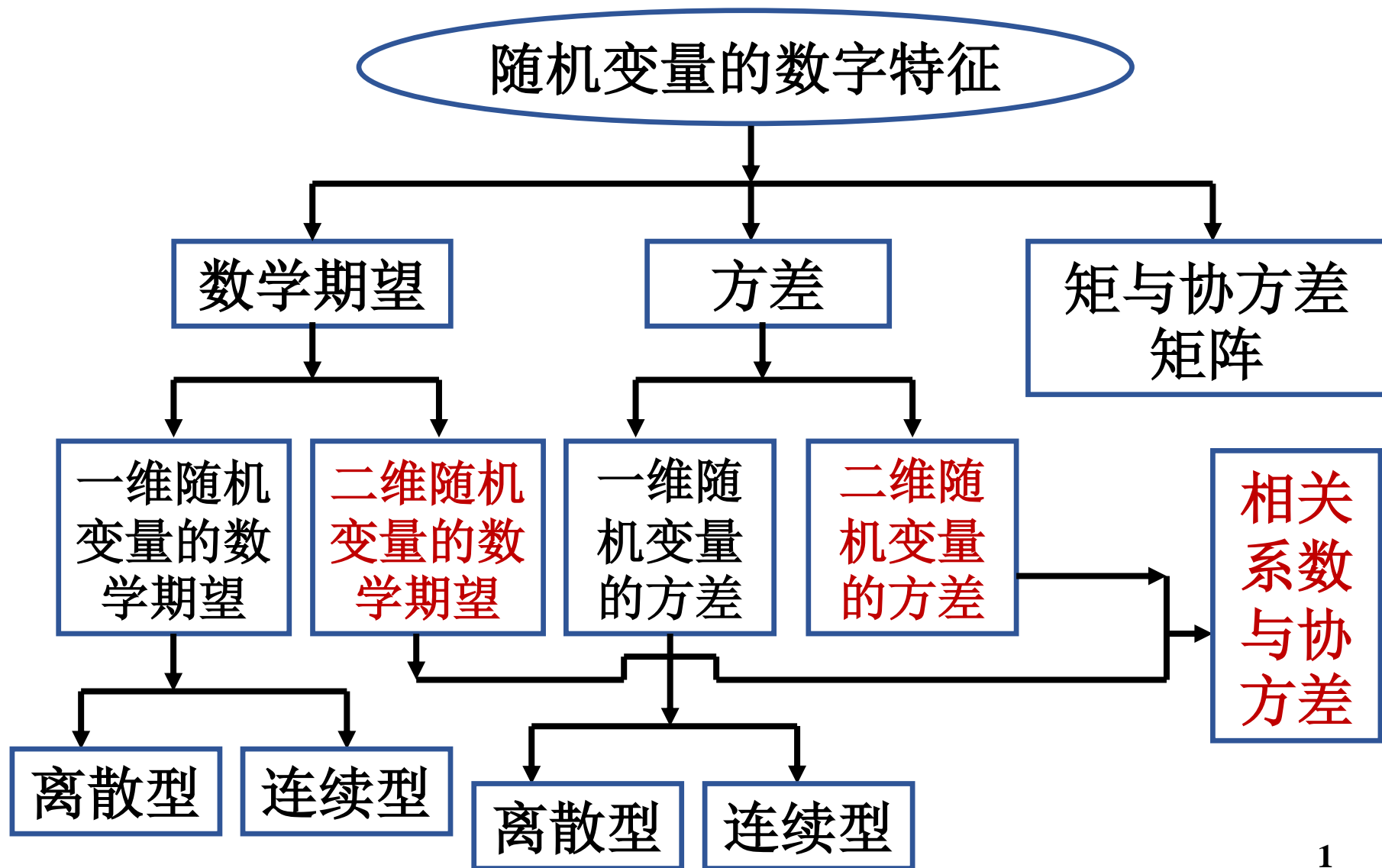




## 第三节 协方差及相关系数



## ▶ 第三节 协方差及相关系数

### 问题的引出

- 随机变量的数学期望及方差都只刻画了一个随机变量的某一方面的特征，而协方差与相关系数是刻画两个随机变量之间关系的数字特征。
- 对于多维随机变量，反映分量之间关系的数字特征中最重要的，就是本节要讨论的内容。
- 如果两个随机变量  $X, Y$  相互独立，则：

$$E(XY) = E(X)E(Y)$$

$$E\{[X - E(X)][Y - E(Y)]\} = 0$$

## 一. 协方差

**1.定义1.** 量  $E\{[X - E(X)][Y - E(Y)]\}$  称为随机变量  $X$  与  $Y$  的**协方差**, 记为:  $Cov(X, Y)$

即:  $Cov(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$

**注:** 协方差中当  $X = Y$  时即为方差的定义, 即:

$Cov(X, Y) = D(X)$ , 故方差是协方差的**特例**。

## 2. 协方差的简单性质

(1).  $Cov(X, Y) = Cov(Y, X)$

(2).  $Cov(aX, bY) = ab Cov(X, Y)$   $a, b$  是常数

(3).  $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$

### 3. 计算协方差的一个简单公式

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

证明：由协方差的定义及期望的性质，可得：

$$\begin{aligned} Cov(X, Y) &= E\{[X - E(X)][Y - E(Y)]\} \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

注：显然，若  $X$  与  $Y$  相互独立则：  $Cov(X, Y) = 0$

### 4. 随机变量和的方差与协方差的关系（教材P104）

$$(1). D(X + Y) = D(X) + D(Y) + 2Cov(X, Y)$$

$$(2). D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j)$$

若  $X_1, X_2, \dots, X_n$  **两两独立**，上式化为：

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i)$$

问题：

- 协方差的大小在一定程度上反映了  $X$  和  $Y$  相互间的关系，但它还受  $X$  与  $Y$  本身**度量单位**的影响。

**例如：**  $\text{Cov}(kX, kY) = k^2 \text{Cov}(X, Y)$

- 为了克服这一缺点，对协方差进行**标准化**，这就引入了**相关系数的概念**。

## 二. 相关系数

1. 定义2. 量(无量纲)  $\frac{Cov(X,Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$  称为随机变量

$X, Y$  的**相关系数**，记为： $\rho_{XY}$

$$\text{即: } \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$$

### 2. 相关系数的简单性质

(1).  $|\rho_{XY}| \leq 1$

(2).  $|\rho_{XY}| = 1 \Leftrightarrow$  存在常数  $a, b$ , 使得:  
 $P(Y = aX + b) = 1$

$X$  和  $Y$  以  
概率 1 **线性**  
**相关**

说明:

相关系数刻画了  $X$  和  $Y$  间“线性相关”的程度。

若考虑以  $X$  的线性函数  $a + bX$  来近似表示  $Y$ ,

以均方误差  $e = E\{[Y - (a + bX)]^2\}$  来衡量以

$a + bX$  近似表示  $Y$  的好坏程度。

则： $e$  值越小表示  $a + bX$  与  $Y$  的近似程度越好。现用**求极值**的方法，求出使  $e$  达到最小时的  $a, b$ ：

$$\begin{aligned} e &= E \{ [Y - (a + bX)]^2 \} \\ &= E(Y^2) + b^2 E(X^2) + a^2 - 2b E(XY) \\ &\quad + 2ab E(X) - 2a E(Y) \end{aligned}$$

令：

$$\begin{cases} \frac{\partial e}{\partial a} = 2a + 2b E(X) - 2E(Y) = 0 \\ \frac{\partial e}{\partial b} = 2b E(X^2) - 2E(XY) + 2a E(X) = 0 \end{cases}$$



解得：

$$\begin{cases} b_0 = \frac{\text{Cov}(X, Y)}{D(X)} \\ a_0 = E(Y) - b_0 E(X) \end{cases}$$

这样求出的  
最佳逼近为：  
 $L(X) = a_0 + b_0 X$

这一逼近的**剩余**  $E\{[Y - L(X)]^2\} = D(Y)(1 - \rho_{XY}^2)$   
(P109)：

可见：若  $\rho_{XY} = \pm 1$ ，则  $Y$  与  $X$  **以概率1具有**线性关系；

若  $\rho_{XY} = 0$ ，则  $Y$  与  $X$  **无线性**关系；

若  $0 \leq |\rho_{XY}| \leq 1$ ，则当：

$|\rho_{XY}|$  的值越接近于**1**， $Y$  与  $X$  的线性相关程度**越高**；

$|\rho_{XY}|$  的值越接近于**0**， $Y$  与  $X$  的线性相关程度**越弱**。

性质证明: (1).  $|\rho_{XY}| \leq 1$  (1).  $D(X+Y) = D(X) + D(Y) + 2Cov(X, Y)$

由方差的性质和协方差的定义知, 对任意实数 **$b$** , 有

$$0 \leq D(Y - bX) = b^2 D(X) + D(Y) - 2b Cov(X, Y)$$

令  $b = \frac{Cov(X, Y)}{D(X)}$ , 则上式为:

$$\begin{aligned} D(Y - bX) &= D(Y) - \frac{[Cov(X, Y)]^2}{D(X)} \\ &= D(Y) \left[ 1 - \frac{[Cov(X, Y)]^2}{D(X)D(Y)} \right] = D(Y) [1 - \rho_{XY}^2] \end{aligned}$$

由于方差 **$D(Y)$** 是正的, 故必有:  $1 - \rho_{XY}^2 \geq 0$

所以证得:  $|\rho_{XY}| \leq 1$

证明: (2).  $|\rho_{XY}| = 1 \Leftrightarrow$  存在常数  $a, b$  使得:

$$P(Y = aX + b) = 1$$

→  $\because |\rho_{XY}| = 1$

由方差与协方差协关系有:

$D(X \pm Y) = D(X) + D(Y) \pm 2Cov(X, Y)$  因此有:

$$D\left(\frac{X - E(X)}{\sqrt{D(X)}} \pm \frac{Y - E(Y)}{\sqrt{D(Y)}}\right) = D\left(\frac{X - E(X)}{\sqrt{D(X)}}\right) + D\left(\frac{Y - E(Y)}{\sqrt{D(Y)}}\right)$$

$$\pm 2Cov\left(\frac{X - E(X)}{\sqrt{D(X)}}, \frac{Y - E(Y)}{\sqrt{D(Y)}}\right)$$

$$= 1 + 1 \pm 2\rho_{XY} = 2(1 \pm \rho_{XY})$$

$\frac{X - E(X)}{\sqrt{D(X)}}$  与  $\frac{Y - E(Y)}{\sqrt{D(Y)}}$

是**标准化随机变量**, 故  
其均值为 0, 方差为 1

当  $\rho_{XY} = 1$  时有:

$$D\left(\frac{X - E(X)}{\sqrt{D(X)}} - \frac{Y - E(Y)}{\sqrt{D(Y)}}\right) = 0$$

由方差的性质, 可知:

$$P\left(\frac{X - E(X)}{\sqrt{D(X)}} - \frac{Y - E(Y)}{\sqrt{D(Y)}} = c\right) = 1 \quad c \text{ 为常数}$$

整理得:

$$P(Y = aX + b) = 1$$

$$\text{其中: } a = \frac{\sqrt{D(Y)}}{\sqrt{D(X)}}, \quad b = E(Y) - \sqrt{D(Y)}\left(\frac{E(X)}{\sqrt{D(X)}} + c\right)$$

同理, 当  $\rho_{XY} = -1$  时也可推出此结论。因此得证。

← ∴  $P(Y = aX + b) = 1$ , 即  $P(Y - aX = b) = 1$

∴  $D(Y - aX) = 0$

又 ∴  $D(Y - aX) = D(Y) + a^2 D(X) - 2Cov(Y, aX)$   
 $= D(aX + b) + a^2 D(X) - 2E[Y - E(Y)][aX - aE(X)]$   
 $= a^2 D(X) + a^2 D(X) - 2aE[X - E(X)][Y - E(Y)]$   
 $= 2a^2 D(X) - 2a Cov(X, Y) = 0$

所以:  $Cov(X, Y) = a D(X)$

$$\begin{aligned} &= a \sqrt{D(X)} \sqrt{D\left(\frac{Y-b}{a}\right)} = a \sqrt{D(X)} \sqrt{\frac{D(Y)}{a^2}} \\ &= \frac{a}{|a|} \sqrt{D(X)} \sqrt{D(Y)} \end{aligned}$$

于是得:  $|Cov(X, Y)| = \left| \sqrt{D(X)} \sqrt{D(Y)} \right|$

即:  $|\rho_{XY}| = 1$

注:  $X$  和  $Y$  独立时,  $\rho_{XY} = 0$  但其逆不真。

由于当  $X$  和  $Y$  独立时,  $Cov(X, Y) = 0$ , 故  $\rho_{XY} = 0$ ;

但  $\rho_{XY} = 0$  ( $X$ 与 $Y$ 不相关) 并不一定能推出 $X$ 和 $Y$ 独立。

例1. 设  $X, Y$  在  $x^2 + y^2 \leq 1$  上服从均匀分布, 即:

$$f(x, y) = \begin{cases} \frac{1}{\pi} & x^2 + y^2 \leq 1 \\ 0 & x^2 + y^2 > 1 \end{cases}$$

验证:  $X$  与  $Y$  是不相关的, 但不是相互独立的。

证明: 由已知,  $X, Y$  的边缘概率密度为:

$$f_X(x) = \begin{cases} \frac{2}{\pi} \sqrt{1-x^2}, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases} \quad \text{与} \quad f_Y(y) = \begin{cases} \frac{2}{\pi} \sqrt{1-y^2}, & |y| \leq 1 \\ 0, & |y| > 1 \end{cases}$$

显然,  $f(x, y) \neq f_X(x)f_Y(y)$

所以:  $X$  与  $Y$  是不独立的

奇函数在对称区间上的  
积分为 0

又因为:  $E(X) = \int_{-1}^1 x \frac{2}{\pi} \sqrt{1-x^2} dx = 0$

$$E(Y) = \int_{-1}^1 y \frac{2}{\pi} \sqrt{1-y^2} dy = 0$$

$$E(XY) = \iint_{x^2+y^2 \leq 1} xy \frac{1}{\pi} dx dy = \frac{1}{\pi} \int_{-1}^1 dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy dy = 0$$

从而有:  $Cov(X, Y) = E(XY) - E(X)E(Y) = 0$

于是得:  $\rho_{xy} = 0$

故得:  $X, Y$  是不相关的 (不存在线性关系)。



一般：

1. 当  $\rho_{XY} = 0$  时，称  $X$  与  $Y$  不相关。

故有：若  $X$  与  $Y$  相互独立，则  $X$  与  $Y$  不相关；但反之不真。

2. 但对下述情形，独立与不相关等价

若  $(X, Y)$  服从二维正态分布，则  
 $X$  与  $Y$  相互独立  $\Leftrightarrow X$  与  $Y$  不相关（后证）

例2. 设  $(X, Y)$  服从**二维正态分布**，它的概率密度为：

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

求：  $X$  与  $Y$  的相关系数

解：由已知，  $X, Y$  的边缘概率密度为：

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \quad -\infty < x < +\infty$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}, \quad -\infty < y < +\infty$$

其数学期望与方差分别为：

$$E(X) = \mu_1, E(Y) = \mu_2, D(X) = \sigma_1^2, D(Y) = \sigma_2^2$$

$$\text{而: } Cov(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_1)(y - \mu_2) f(x, y) dx dy$$

$$= \rho \sigma_1 \sigma_2$$

$$\text{于是: } \rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}} = \rho$$

其积分过程  
见教材（第  
五版）P111

结论： 若  $(X, Y)$  服从二维正态分布，则  
 $X$  与  $Y$  相互独立  $\Leftrightarrow X$  与  $Y$  不相关

$$E(X) = \mu_1, E(Y) = \mu_2, D(X) = \sigma_1^2, D(Y) = \sigma_2^2$$

$$\text{而: } Cov(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_1)(y - \mu_2) f(x, y) dx dy$$

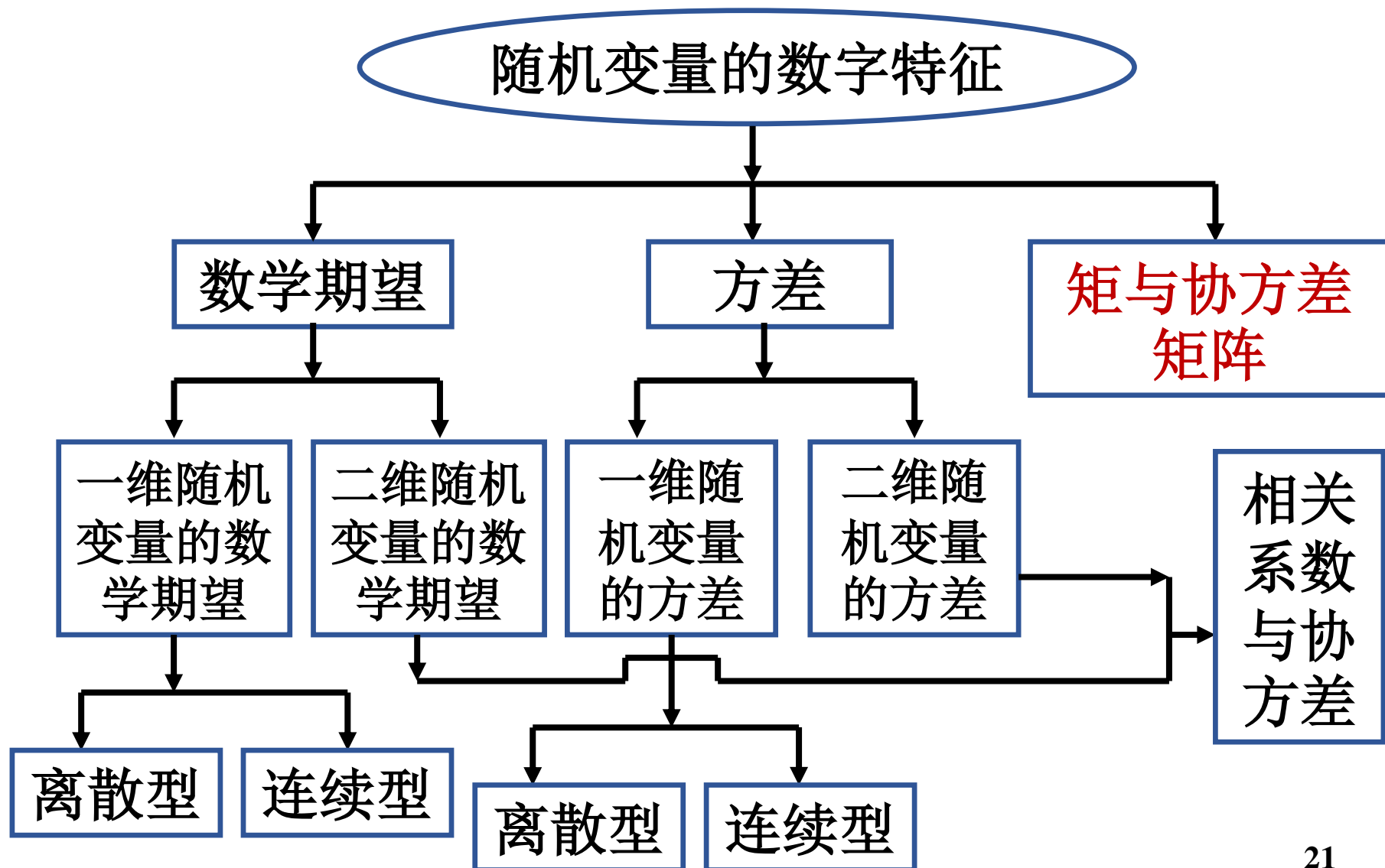
$$= \rho \sigma_1 \sigma_2$$

$$\text{于是: } \rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}} = \rho$$

其积分过程  
见教材（第  
五版）P111

结论：二维正态分布的概率密度函数中的五个参数的意义分别为：X 的数学期望与方差；Y 的数学期望与方差；X 与 Y 的相关系数。二维正态分布完全由这五个参数所确定。

## 第四节 矩、协方差矩阵





## 第四节 矩、协方差矩阵

### 一. 矩

矩是随机变量的**更为广泛**的一种数字特征，前面介绍的数学期望及方差都是某种矩。

**定义：** 设  $X$  和  $Y$  是随机变量

(1). 若  $E(X^k)$  存在，则称它为  $X$  的 **$k$  阶原点矩**， $k = 1, 2, \dots$  简称  $k$  阶矩。

(2). 若  $E\{[X - E(X)]^k\}$  存在， $k = 2, 3, \dots$   
则称它为  $X$  的 **$k$  阶中心矩**。

(3). 若  $E(X^k Y^l)$  存在,  $k, l = 1, 2, \dots$  则称它为  $X$  和  $Y$  的  $k + l$  阶混合矩。

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

注:

数学期望  $E(X)$  是随机变量  $X$  的一阶原点矩;  
方差  $D(X)$  是随机变量  $X$  的二阶中心矩;  
协方差  $Cov(X, Y)$  是随机变量  $X$  和  $Y$  的二阶混合中心矩。

## 二. 协方差矩阵

**定义：**若二维随机变量  $(X_1, X_2)$  的四个二阶中心矩都存在，分别记为：

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

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

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

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

将它们排成矩阵的形式：

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

这是一个  
对称  
矩阵

称此矩阵为  $(X_1, X_2)$  的**协方差矩阵**。



**注：** 类似可定义  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  的协方差矩阵。

$$\begin{aligned} \text{若 } c_{ij} &= \text{Cov}(X_i, X_j) \quad i, j = 1, 2, \dots, n \\ &= E\{[X_i - E(X_i)][X_j - E(X_j)]\} \end{aligned}$$

都存在，则**称**矩阵：

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

为  $(X_1, X_2, \dots, X_n)$   
的**协方差矩阵**  
(**对称矩阵**)

▲  $n$  维正态分布的概率密度的定义。

设  $X = (X_1, X_2, \dots, X_n)^T$  是一个  $n$  维随机向量，若它的概率密度为：

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

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

推导过程见教材P113

则称  $X$  服从  $n$  维正态分布。

其中：  $C$  是  $(X_1, X_2, \dots, X_n)$  的协方差矩阵。

$|C|$  是它的行列式， $C^{-1}$  表示  $C$  的逆矩阵。

$X$  和  $\mu$  是  $n$  维列向量， $X^T$  表示  $X$  的转置。

## ▲ $n$ 维正态分布的四条重要性质

(1).  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  服从正态分布的充分必要条件是：对一切不全为零的实数：

$a_1, a_2, \dots, a_n$ ， $(X_1, X_2, \dots, X_n)$  的任意线性组合： $a_1 X_1 + a_2 X_2 + \dots + a_n X_n$  均服从一维正态分布。

(2). 若  $X = (X_1, X_2, \dots, X_n)$  服从  $n$  维正态分布， $Y_1, Y_2, \dots, Y_k$  是  $X_j$  ( $j = 1, 2, \dots, n$ ) 的线性函数，则  $(Y_1, Y_2, \dots, Y_k)$  也服从多维正态分布。

这一性质称为正态变量的线性变换不变性。

(3). 若  $X = (X_1, X_2, \dots, X_n)$  服从  $n$  维正态分布, 则它的每一个分量  $X_j$  ( $j = 1, 2, \dots, n$ ) 都服从正态分布; 反之, 若  $X_1, X_2, \dots, X_n$  都服从正态分布, 且相互独立, 则  $(X_1, X_2, \dots, X_n)$  服从  $n$  维正态分布。

(4). 设  $(X_1, X_2, \dots, X_n)$  服从  $n$  维正态分布, 则:  
“ $X_1, X_2, \dots, X_n$  相互独立” 与  
“ $X_1, X_2, \dots, X_n$  两两不相关” 是等价的。

上述的四条性质在后续的“随机过程”与“数理统计”课程中会经常用到。

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

**故：**  $X$  和  $Y$  的联合分布为正态分布（**性质3**），  
 $X$  和  $Y$  的任意线性组合是正态分布（**性质1**）。

**即：**  $Z \sim N(E(Z), D(Z))$

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

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

所以:  $Z \sim N(5, 3^2)$

故:  $Z$  的概率密度为:

$$f_Z(z) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(z-5)^2}{18}}$$

$$-\infty < z < \infty$$

## 第四章作业（教材第五版）：

**P115: 3、4、5**

**P116: 8、9、10**

**P117: 14、18、19**

**P118: 22、26、27**

**P119: 30、32、33、34**

注：作业不得抄袭；写上姓名、班级、学号和页码（如1/5），11月10日前提交至教学云平台。