Lec-15. 协方差与相关系数、矩

主讲教师： 吴利苏 **(wulisu@sdust.edu.cn)**
主　　页： wulisu.cn

本次课内容

1. 协方差及相关系数

2. 矩、协方差矩阵

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

其中 $Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$.

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

其中 $Cov(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$.

- 若 X, Y 相互独立, 则
  $Cov(X, Y) = E\{[X - E(X)][Y - E(Y)]\} = 0.$

$$D(X+Y) = D(X) + D(Y).$$

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

其中 $Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$.

- 若 X,Y 相互独立, 则
  $Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=0.$

$$D(X+Y)=D(X)+D(Y).$$

- $Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$定义为协方差.

协方差和相关系数

定义

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

称为随机变量 X 与 Y 的协方差(covariance).

$$\boldsymbol{\rho}_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

称为 X 与 Y 的相关系数(correlation coefficient).

本节公式

- 协方差的计算公式:

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

- 相关系数

$$\boldsymbol{\rho}_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)D(Y)}}.$$

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

练习

设随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \mathrm{e}^{-y}, & 0 < x < y, \\ 0, & \text{else} \end{cases}$$

求 $\boldsymbol{\rho}_{XY}$.

解:

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\,dxdy \\
&= \int_0^{\infty} dy \int_0^{y} xe^{-y}dx = \int_0^{\infty} \frac{y^2}{2}e^{-y}dy = \frac{\Gamma(3)}{2} = \frac{2!}{2} = 1
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)\,dxdy \\
&= \int_0^{\infty} dy \int_0^{y} x^2 e^{-y}dx = \int_0^{\infty} \frac{y^3}{3}e^{-y}dy = \frac{\Gamma(4)}{3} = \frac{3!}{3} = 2
\end{aligned}
$$

所以 $D(X) = E(X^2) - (E(X))^2 = 1$.

$$
\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dxdy \\
&= \int_{0}^{\infty} dy \int_{0}^{y} y e^{-y} dx = \int_{0}^{\infty} y^2 e^{-y} dy = \Gamma(3) = 2! = 2
\end{aligned}
$$

$$
\begin{aligned}
E(Y^2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2 f(x,y)\,dxdy \\
&= \int_{0}^{\infty} dy \int_{0}^{y} y^2 e^{-y} dx = \int_{0}^{\infty} y^3 e^{-y} dy = \Gamma(4) = 3! = 6
\end{aligned}
$$

所以 $D(Y) = E(Y^2) - (E(Y))^2 = 2.$

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dxdy \\
&= \int_0^{\infty} dy \int_0^{y} xye^{-y}dx = \int_0^{\infty} \frac{y^3}{2}e^{-y}dy = \frac{\Gamma(4)}{2} = \frac{3!}{2} = 3
\end{aligned}
$$

所以 $Cov(X,Y) = E(XY) - E(X)E(Y) = 1$,

$$
\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)\cdot D(Y)}} = \frac{\sqrt{2}}{2}. \qquad \square
$$

$Cov(X, Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X, Y) > 0$ 时, X 与 Y 正相关.

$Cov(X, Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X, Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X, Y) < 0$ 时, X 与 Y 负相关.

$Cov(X, Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X, Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X, Y) < 0$ 时, X 与 Y 负相关.
- 当 $Cov(X, Y) = 0$ 时, X 与 Y (线性) 不相关.

$Cov(X, Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X, Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X, Y) < 0$ 时, X 与 Y 负相关.
- 当 $Cov(X, Y) = 0$ 时, X 与 Y (线性) 不相关.
- 若 X 与 Y 相互独立, 则 $Cov(X, Y) = 0$.

$Cov(X,Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X,Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X,Y) < 0$ 时, X 与 Y 负相关.
- 当 $Cov(X,Y) = 0$ 时, X 与 Y (线性) 不相关.
- 若 X 与 Y 相互独立, 则 $Cov(X,Y) = 0$.
    - 相互独立 $\Rightarrow$ 不相关;

$Cov(X,Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X,Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X,Y) < 0$ 时, X 与 Y 负相关.
- 当 $Cov(X,Y) = 0$ 时, X 与 Y (线性) 不相关.
- 若 X 与 Y 相互独立, 则 $Cov(X,Y) = 0$.
    - 相互独立 $\Rightarrow$ 不相关;
    - 相互独立 $\not\Leftarrow$ 不相关;

$Cov(X,Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X,Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X,Y) < 0$ 时, X 与 Y 负相关.
- 当 $Cov(X,Y) = 0$ 时, X 与 Y (线性) 不相关.
- 若 X 与 Y 相互独立, 则 $Cov(X,Y) = 0$.
    - 相互独立 $\Rightarrow$ 不相关;
    - 相互独立 $\not\Leftarrow$ 不相关;

注: 这里 X, Y 不相关是指 X 和 Y 没有线性关系, 而不是指完全没有关系. X, Y 相互独立是指 X 和 Y 完全没有关系 (包括线性关系).

$Cov(X, Y)$ 反应了 X 与 Y 的线性相关性.

- 当 $Cov(X, Y) > 0$ 时, X 与 Y 正相关.
- 当 $Cov(X, Y) < 0$ 时, X 与 Y 负相关.
- 当 $Cov(X, Y) = 0$ 时, X 与 Y (线性) 不相关.
- 若 X 与 Y 相互独立, 则 $Cov(X, Y) = 0$.
  - 相互独立 $\Rightarrow$ 不相关;
  - 相互独立 $\not\Leftarrow$ 不相关;

注: 这里 X, Y 不相关是指 X 和 Y 没有线性关系, 而不是指完全没有关系. X, Y 相互独立是指 X 和 Y 完全没有关系 (包括线性关系).

注: 对于二维正态分布 (X, Y), X 和 Y 不相关当且仅当它们相互独立 ($\rho_{XY} = 0$).

例 (反例)

(X, Y) 的分布律为

$Y \backslash X$	-2	-1	1	2	$P\{Y=j\}$
1	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
4	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{2}$
$P\{X=i\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

则 X 与 Y 不相关，且不相互独立.

例 (反例)

(X, Y) 的分布律为

$Y \backslash X$	-2	-1	1	2	$P\{Y=j\}$
1	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
4	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{2}$
$P\{X=i\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

则 X 与 Y 不相关，且不相互独立.

实际上，X 与 Y 非线性相关，但 $Y = X^2$.

协方差的性质

性质（对称性）

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

协方差的性质

性质 (对称性)

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

性质 (双线性)

3. $Cov(aX, bY) = ab \cdot Cov(X, Y)$, $a, b \in \mathbb{R}$.
4. $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$.

协方差的性质

性质（对称性）

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

2. $Cov(X, X) = D(X)$.

性质（双线性）

3. $Cov(aX, bY) = ab \cdot Cov(X, Y)$, $a, b \in \mathbb{R}$.

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

5. $Cov(X + C, Y) = Cov(X, Y)$.

例

计算 $Cov(3X + 2Y, 2X)$.

性质

下列说法等价（*TFAE*）:

1. X, Y 不相关.
2. $\boldsymbol{\rho}_{XY} = 0.$
3. $Cov(X, Y) = 0.$
4. $E(XY) = E(X)E(Y).$

相关系数的意义

- 相关系数刻画了两个随机变量的线性近似程度, 即 $a+bX$ 与 Y 的最大近似程度.

相关系数的意义

- 相关系数刻画了两个随机变量的线性近似程度, 即 $a + bX$ 与 Y 的最大近似程度.
  - $-1 \leq \boldsymbol{\rho}_{XY} \leq 1$.

相关系数的意义

- 相关系数刻画了两个随机变量的线性近似程度, 即 $a+bX$ 与 Y 的最大近似程度.
  - $-1 \leq \rho_{XY} \leq 1$.
  - 当 $\rho_{XY} > 0$ 时, X 和 Y 是正相关的.
    特别地, 若 $\rho_{XY} = 1$, 则存在 $a > 0$, 使得 $Y = aX + b$ 几乎完全成立.

相关系数的意义

- 相关系数刻画了两个随机变量的线性近似程度, 即 $a+bX$ 与 Y 的最大近似程度.
  - $-1 \le \rho_{XY} \le 1$.
  - 当 $\rho_{XY} > 0$ 时, X 和 Y 是正相关的.
    特别地, 若 $\rho_{XY} = 1$, 则存在 $a > 0$, 使得
    $Y = aX + b$ 几乎完全成立.
  - 当 $\rho_{XY} < 0$ 时, X 和 Y 是负相关的.
    特别地, 若 $\rho_{XY} = 1$, 则存在 $a < 0$, 使得
    $Y = aX + b$ 是几乎完全成立.

相关系数的意义

- 相关系数刻画了两个随机变量的线性近似程度, 即 $a + bX$ 与 Y 的最大近似程度.
  - $-1 \leq \rho_{XY} \leq 1$.
  - 当 $\rho_{XY} > 0$ 时, X 和 Y 是正相关的.
    特别地, 若 $\rho_{XY} = 1$, 则存在 $a > 0$, 使得 $Y = aX + b$ 几乎完全成立.
  - 当 $\rho_{XY} < 0$ 时, X 和 Y 是负相关的.
    特别地, 若 $\rho_{XY} = 1$, 则存在 $a < 0$, 使得 $Y = aX + b$ 是几乎完全成立.
  - 当 $\rho_{XY} = 0$ 时, X 和 Y 是不相关, 即它们不存在线性关系.

相关系数的意义

- 定义**均方误差**

$$e = E\{[Y - (a + bX)]^2\}$$

e 越小, $a + bX$ 与 Y 近似程度越好. 确定 a 与 b 的值, 使得 e 达到最小.

相关系数的意义

- 定义均方误差

$$e = E\{[Y - (a + bX)]^2\}$$

  e 越小, $a + bX$ 与 Y 近似程度越好. 确定 a 与 b 的值, 使得 e 达到最小.

- 求 $e = e(a, b)$ 的最小值.

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

相关系数的意义

- 对 e 分别关于 a, b 求偏导数, 并令它们等于零.

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

得

$$\begin{cases} b_0 = \frac{Cov(X,Y)}{D(X)}, \\ a_0 = E(Y) - b_0E(X) = E(Y) - E(X)\frac{Cov(X,Y)}{D(X)} \end{cases}$$

则

$$e_{\min} = e(a_0, b_0) = E([Y-(a_0+b_0X)]^2) = (1-\boldsymbol{\rho}_{XY}^2)D(Y).$$

$\boldsymbol{\rho}_{XY}$ 是表示 X, Y 之间线性相关程度的数字特征.

- 当 $|\boldsymbol{\rho}_{XY}|$ 较大时, e 较小, 表明 X, Y 的线性关系较紧密.
- 当 $|\boldsymbol{\rho}_{XY}|$ 较小时, e 较大, 表明 X, Y 的线性关系较差.
- 当 $\boldsymbol{\rho}_{XY} = 0$ 时, X, Y 不相关.

相关系数的性质

定理

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

(2.) $|\rho_{XY}| = 1 \Leftrightarrow \exists a, b$ s.t

$$P\{Y = a + bX\} = 1.$$

相关系数的性质

定理

(1.) $|\boldsymbol{\rho}_{XY}| \leq 1$;

(2.) $|\boldsymbol{\rho}_{XY}| = 1 \Leftrightarrow \exists a, b$ **s.t**

$$P\{Y = a + bX\} = 1.$$

证明: (1) $E[Y-(a_0+b_0X)]^2 = (1-\boldsymbol{\rho}_{XY}^2)D(Y)$, $E[Y-(a_0+b_0X)]^2 \geq 0$, $D(Y) \geq 0$, 所以

$$1-\boldsymbol{\rho}_{XY}^2 \geq 0 \Rightarrow |\boldsymbol{\rho}_{XY}| \leq 1.$$

(2) "⇒" 若 $|\boldsymbol{\rho}_{XY}| = 1$, 则

$$D(Y - (a_0 + b_0 X)) = E\{[Y - (a_0 + b_0 X)]^2\} \\ - \{E(Y - (a_0 + b_0)X)\}^2 \\ = 0 - \{E(Y - (a_0 + b_0)X)\}^2$$

因此

$$D(Y - (a_0 + b_0 X)) = 0$$

$$E(Y - (a_0 + b_0 X)) = 0.$$

由方差的性质 4, 知

$$P\{Y = a_0 + b_0 X\} = 1.$$

"$\Leftarrow$" 若 $\exists a^*, b^*$, s.t $P\{Y = a^* + b^*X\} = 1$, 即

$$P\{Y - (a^* + b^*X) = 0\} = 1.$$

则 $E\{[Y - (a^* + b^*X)]^2\} = 0$.

$$\begin{aligned} 0 = E\{[Y - (a^* + b^*X)]^2\} &\geq \min E\{[Y - (a + bX)]^2\} \\ &= E\{Y - (a_0 + b_0X)^2\} \\ &= (1 - \boldsymbol{\rho}_{XY}^2)D(Y) \end{aligned}$$

故 $|\boldsymbol{\rho}_{XY}| = 1$. $\square$

性质

取标准化变量

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}}, \; Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}.$$

则$\boldsymbol{\rho}_{XY} = Cov(X^*, Y^*)$.

性质

取标准化变量

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}}, \ Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}.$$

则$\boldsymbol{\rho}_{XY} = Cov(X^*, Y^*)$.

证：

$$\begin{aligned} Cov(X^*, Y^*) &= Cov\left(\frac{X - E(X)}{\sqrt{D(X)}}, \frac{Y - E(Y)}{\sqrt{D(Y)}}\right) \\ &= \frac{Cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \\ &= \boldsymbol{\rho}_{XY}. \end{aligned}$$

□

例

设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 求 X 与 Y 的相关系数.

例

设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 求 X 与 Y 的相关系数.

解: $\forall x, y \in \mathbb{R}$,

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

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

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

令 $X^* = \frac{X-\mu_1}{\sigma_1}$, $Y^* = \frac{Y-\mu_2}{\sigma_2}$,

则 $E(X^*) = E(Y^*) = 0$, $D(X^*) = D(Y^*) = 1$.

$\rho_{XY} = Cov(X^*, Y^*) = E(X^*Y^*) - E(X^*)E(Y^*) = E(X^*Y^*)$.

$$
\begin{aligned}
E(X^* Y^*) &= E\left(\frac{X-\mu_1}{\sigma_1}\cdot\frac{Y-\mu_2}{\sigma_2}\right) \\
&= \iint_{\mathbb{R}^2} \frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2}\cdot f(x,y)\,dxdy \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\iint_{\mathbb{R}^2} \frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2} \\
&\quad\times\exp\left(\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right)dxdy
\end{aligned}
$$

令 $u=\frac{x-\mu_1}{\sigma_1}, v=\frac{y-\mu_2}{\sigma_2}, |J|=\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\sigma_1\sigma_2,$

$$
\begin{aligned}
&= \frac{1}{2\pi\sqrt{1-\rho^2}}\iint_{\mathbb{R}^2} uv\times\exp\left(-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)\right)dudv \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}}\iint_{\mathbb{R}^2} uv\times\exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)}-\frac{v^2}{2}\right)dudv
\end{aligned}
$$

令 $t=\frac{u-\boldsymbol{\rho} v}{\sqrt{1-\boldsymbol{\rho}^2}}, w=v, |J|=\left|\frac{\partial(u,v)}{\partial(w,t)}\right|=\sqrt{1-\boldsymbol{\rho}^2}$,

$$\begin{aligned}
&=\frac{1}{2\boldsymbol{\pi}}\iint\limits_{\mathbb{R}^2}(\sqrt{1-\boldsymbol{\rho}^2}t+\boldsymbol{\rho} w)w\times\exp\left(-\frac{t^2}{2}-\frac{w^2}{2}\right)dtdw\\
&=\frac{1}{2\boldsymbol{\pi}}\iint\limits_{\mathbb{R}^2}\sqrt{1-\boldsymbol{\rho}^2}te^{-\frac{t^2}{2}}\cdot e^{-\frac{w^2}{2}}+e^{-\frac{t^2}{2}}\cdot\boldsymbol{\rho} w^2e^{-\frac{w^2}{2}}dtdw\\
&=\frac{\sqrt{1-\boldsymbol{\rho}^2}}{2\boldsymbol{\pi}}\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt\cdot\int_{-\infty}^{\infty}e^{-\frac{w^2}{2}}dw+\frac{\boldsymbol{\rho}}{2\boldsymbol{\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt\cdot\int_{-\infty}^{\infty}w^2e^{-\frac{w^2}{2}}dw\\
&=0+\frac{\boldsymbol{\rho}}{2\boldsymbol{\pi}}\sqrt{2\boldsymbol{\pi}}\cdot\sqrt{2\boldsymbol{\pi}}=\boldsymbol{\rho}
\end{aligned}$$

所以 $\boldsymbol{\rho}_{XY}=Cov(X^*,Y^*)=E(X^*Y^*)-E(X^*)E(Y^*)=E(X^*Y^*)=\boldsymbol{\rho}$. □

令 $t = \frac{u-\rho v}{\sqrt{1-\rho^2}}, w = v, |J| = \left|\frac{\partial(u,v)}{\partial(w,t)}\right| = \sqrt{1-\rho^2}$,

$$= \frac{1}{2\pi}\iint\limits_{\mathbb{R}^2}(\sqrt{1-\rho^2}t + \rho w)w \times \exp\left(-\frac{t^2}{2} - \frac{w^2}{2}\right) dtdw$$

$$= \frac{1}{2\pi}\iint\limits_{\mathbb{R}^2} \sqrt{1-\rho^2}te^{-\frac{t^2}{2}} \cdot e^{-\frac{w^2}{2}} + e^{-\frac{t^2}{2}} \cdot \rho w^2 e^{-\frac{w^2}{2}} dtdw$$

$$= \frac{\sqrt{1-\rho^2}}{2\pi}\int_{-\infty}^{\infty} te^{-\frac{t^2}{2}} dt \cdot \int_{-\infty}^{\infty} e^{-\frac{w^2}{2}} dw + \frac{\rho}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \cdot \int_{-\infty}^{\infty} w^2 e^{-\frac{w^2}{2}} dw$$

$$= 0 + \frac{\rho}{2\pi}\sqrt{2\pi} \cdot \sqrt{2\pi} = \rho$$

所以 $\rho_{XY} = Cov(X^*, Y^*) = E(X^* Y^*) - E(X^*)E(Y^*) = E(X^* Y^*) = \rho$. □

- 二维正态分布参数 ρ 就是相关系数.
- 二维正态分布中,

X与Y相互独立 $\Leftrightarrow \rho = 0 \Leftrightarrow$ X与Y不相关.

例

设 $X \sim N(1, 3^2)$, $Y \sim N(0, 4^2)$, $\boldsymbol{\rho}_{XY} = -\frac{1}{2}$. 令 $Z = \frac{X}{3} + \frac{Y}{2}$.

(1) 求 Z 的数学期望和方差.

(2) 求 X 与 Z 的相关系数.

(3) 问 X 与 Z 是否相互独立, 为什么?

解: (1) $E(X)=1$, $E(Y)=0$,

$D(X)=9$, $D(Y)=16$.

$E(Z)=\frac{1}{3}E(X)+\frac{1}{2}E(Y)=\frac{1}{3}$.

$D(Z)=D(\frac{X}{3})+D(\frac{Y}{2})+2Cov(\frac{X}{3},\frac{Y}{2})=3$.

(2) $Cov(X,Z)= Cov(X,(\frac{X}{3}+\frac{Y}{2}))$

$=\frac{1}{3}Cov(X,X)+\frac{1}{2}Cov(X,Y)=0$.

$\boldsymbol{\rho}_{XZ}=0$.

(3) X,Z 为正态分布，$\boldsymbol{\rho}_{XZ}=0\Rightarrow X$ 与 Z 相互独立. $\square$

矩

- $D(X) = E(X^2) - [E(X)]^2$.
- $Cov(X, Y) = E(XY) - E(X)E(Y)$.
- 本质上都是计算期望！

矩

- $D(X) = E(X^2) - [E(X)]^2$.
- $Cov(X, Y) = E(XY) - E(X)E(Y)$.
- 本质上都是计算期望！

矩

设 X, Y 为随机变量,

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

设 X, Y 为随机变量,

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

- 若
$$E\{[X - E(X)]^k\}$$
存在, 称它为 X 的k 阶中心距.

设 X, Y 为随机变量,

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

- 若
$$E\{[X-E(X)]^k\}$$
存在, 称它为 X 的k 阶中心距.

- 若
$$E(X^k Y^l)$$
存在, 称它为 X 和 Y 的$k+l$ 阶混合矩.

设 X, Y 为随机变量,

- 若

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

存在, 则称它为 X 的 k 阶原点矩, 简称 k 阶矩.

- 若

$$E\{[X - E(X)]^k\}$$

存在, 称它为 X 的 k 阶中心距.

- 若

$$E(X^k Y^l)$$

存在, 称它为 X 和 Y 的 $k + l$ 阶混合矩.

- 若

$$E\{[X - E(X)]^k[Y - E(Y)]^l\}$$

存在, 则称它为 X 和 Y 的 $k + l$ 阶混合中心距.

矩

1. $E(X)$ 是 X 的一阶原点矩,
   $D(X)$ 为二阶中心距,
   $Cov(X, Y)$ 是 X 和 Y 的二阶混合中心距.

矩

1. $E(X)$ 是 X 的一阶原点矩,
   $D(X)$ 为二阶中心距,
   $Cov(X,Y)$ 是 X 和 Y 的二阶混合中心距.
2. 以上数字特征都是随机变量函数的数学期望.

矩

1. $E(X)$ 是 X 的一阶原点矩,
   $D(X)$ 为二阶中心距,
   $Cov(X, Y)$ 是 X 和 Y 的二阶混合中心距.
2. 以上数字特征都是随机变量函数的数学期望.
3. 在实际应用中, 高于 4 阶的矩很少用.
   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.

协方差矩阵

二维随机变量 (X_1, X_2) 的协方差矩阵为

$$C = \begin{pmatrix} D(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & D(X_2) \end{pmatrix}.$$

协方差矩阵

n 维随机变量 $(X_1, \cdots, X_n)$ 的协方差矩阵为

$$C = \begin{pmatrix} D(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & D(X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & D(X_n) \end{pmatrix}.$$

- 协方差矩阵 $C = C^T$ 为对称非负定矩阵.
- 作用：刻画高维随机变量.
  比如定义 n 维正态分布.

二维正态分布

二维正态分布 (X, Y) 的概率密度：

$$
\begin{aligned}
f(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right] \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x-\mu_1, y-\mu_2)\begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}\right] \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}(x-\mu_1, y-\mu_2)\frac{\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}}{(1-\rho^2)\sigma_1^2\sigma_2^2}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}\right] \\
&= \frac{1}{2\pi\sqrt{\det C}} \exp\left[-\frac{1}{2}(x-\mu_1, y-\mu_2)\, C^{-1}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}\right]
\end{aligned}
$$

其中 $C = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ 为协方差矩阵, $\det C = (1-\rho^2)\sigma_1^2\sigma_2^2$.

n 维正态分布

n 维正态分布 $\mathbf{X} = (X_1, \cdots, X_n)^T$ 的概率密度：

$$f(X) = \frac{1}{(2\boldsymbol{\pi})^{n/2}\sqrt{\det C}} \exp\left[-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T C^{-1}(\mathbf{X}-\boldsymbol{\mu})\right]$$

其中 $\boldsymbol{\mu} = (E(X_1), \cdots, E(X_n))^T$, C 为协方差矩阵.

n 维正态分布的性质

1. n 维正态随机变量 $\mathbf{X} = (X_1, ..., X_n)^T, n \geq 1$, 其任意子向量 $(X_{i_1}, \cdots, X_{i_k})^T (1 \leq k \leq n)$ 均服从 k 维正态分布.

> 例如: $\mathbf{X} = (X_1, X_2, X_3)^T$ 为 3 维正态随机变量, 则 $(X_1, X_2)^T, (X_1, X_3)^T, (X_2, X_3)^T$ 均为二维正态随机变量.

特别地, 每一个分量 X_i 都是一维正态变量.
反之, 若 X_i 均为一维正态变量, 且相互独立, 则 $\mathbf{X} = (X_1, \cdots, X_n)^T$ 为 n 维正态随机变量.

n 维正态分布的性质

2. n 维随机变量 $\mathbf{X}=(X_1,X_2,\cdots,X_n)^T, n\geq 1$, 服从 n 维正态分布 $\Leftrightarrow X_1,X_2,\cdots,X_n$ 的任意线性组合

$$l_0+l_1X_1+l_2X_2+\cdots+l_nX_n$$

均服从一维正态分布，其中 $l_1,l_2,\ldots,l_n$ 不全为 0.

例如: $\mathbf{X}=(X_1,X_2,X_3)^T$ 为 3 维正态随机变量，则 $3X_1-X_2, 2X_1+4X_3+1, X_2-3X_1-X_3-2$ 均为一维正态随机变量.

3. n 维正态随机变量 $\mathbf{X} = (X_1, \cdots, X_n)^T, n \geq 1,$ 若 $Y_1, \cdots, Y_k$ 均为 X_i 的线性函数，则 $(Y_1, Y_2, \cdots, Y_k)^T$ 也服从 k 维正态分布.

这一性质称为正态变量的线性变换不变性.

例如: $\mathbf{X} = (X_1, X_2, X_3)^T$ 为 3 维正态随机变量，则 $(3X_1 - X_2,\ 2X_1 + 4X_3 + 1,\ X_2 - 3X_1 - X_3 - 2, X_2)^T$ 服从 4 维正态分布.

4. 设 $\mathbf{X} = (X_1, \cdots, X_n)^T, n \geq 1$ 服从 n 维正态分布，TFAE:

1. $X_1, X_2, \cdots, X_n$ 相互独立;
2. $X_1, X_2, ..., X_n$ 两两不相关;
3. $\mathbf{X}$ 的协方差矩阵为对角矩阵.

例

设 $(X, Y) \sim N(0, 1, 1, 4, -\frac{1}{2})$, 求:

(1) $D(2X - Y)$,

(2) $P\{2X > Y\}$,

(3) (Z_1, Z_2) 的分布, $Z_1 = X + Y, Z_2 = X - Y$.