Chpt.4 Digital Features of Random Variables

第四章 随机变量的数字特征

上节回顾

- 方差： $D(X)=E(X-E(X))^2=E(X^2)-[E(X)]^2$
- 切比雪夫不等式： $P\{|X-\mu|\geq\varepsilon\}\leq\dfrac{\sigma^2}{\varepsilon^2}$
- 期望与方差的性质： ★常见分布的期望与方差

期望	方差	
$E(C)=C$	$D(C)=0$	
$E(X+C)=E(X)+C$	$D(X+C)=D(X)$	
$E(aX)=aE(X)$	$D(aX)=a^2D(X)$	
$E\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_iE(X_i)$	$D\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_i^2D(X_i)$	←独立
独立→ $E(XY)=E(X)E(Y)$	$D(X)=0\Leftrightarrow P(X=C)=1$	

常见分布及其期望和方差列表

分布名称	数学期望 E(X)	方差 D(X)
0-1分布	p	pq
二项分布	np	npq
泊松分布	λ	λ
均匀分布	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布	μ	σ^2
指数分布	$\frac{1}{\lambda}$ （或θ）	$\frac{1}{\lambda^2}$ （或θ^2）

4.3 协方差与相关系数

单个随机变量的期望与方差分别反映了随机变量取值的平均水平和随机变量取值相对于均值的分散程度.

对于两个随机变量X、Y，他们之间的关系如何描述？

4.3.1 协方差

[Definition] 设（X，Y）是二维随机变量，如果 E(X), E(Y) 都存在, 且 $E\left(X-E(X)\right)\left(Y-E(Y)\right)$ 也存在，则称之为 X 与 Y 的协方差，记为 Cov(X,Y)

易见

(1) $Cov\left(X,X\right)=D\left(X\right)$

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

4.3 协方差与相关系数

4.3.1 协方差

（3）$Cov(X,Y)=E\left[\left(X-E(X)\right)\left(Y-E(Y)\right)\right]$

$$=E\left[XY-XE(Y)-YE(X)+E(X)E(Y)\right]$$

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

如果X,Y独立，则 $E(XY)=E(X)E(Y)$

从而有 $\mathrm{Cov}(X,Y)=0$

当X,Y不独立时, $\mathrm{Cov}(X,Y)$ **也可以为0。**

4.3.1 协方差

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

$$
\begin{aligned}
&D(X+Y)\\
&=E[(X+Y)-E(X+Y)]^2\\
&=E(X-E(X))^2+E(Y-E(Y))^2+2E(X-E(X))(Y-E(Y))\\
&=D(X)+D(Y)+2Cov(X,Y)
\end{aligned}
$$

性质：

(1) $Cov(X,C)=0;$ （C为任意常数）

(2) $Cov(aX,bY)=abCov(X,Y)$，(a，b为常数).

(3) $Cov(a_1X_1+a_2X_2,Y)=a_1Cov(X_1,Y)+a_2Cov(X_2,Y)$

$$Cov(a_1X_1+a_2X_2,Y)=E\big(a_1X_1+a_2X_2-E(a_1X_1+a_2X_2)\big)(Y-EY)$$

$$=E\big(a_1(X_1-EX_1)+a_2(X_2-EX_2)\big)(Y-EY)$$

$$=E\big(a_1(X_1-EX_1)(Y-EY)+a_2(X_2-EX_2)(Y-EY)\big)$$

$$=a_1Cov(X_1,Y)+a_2Cov(X_2,Y)$$

拓展：若X，Y都是组合，会是什么结果

计算D(3X-2Y)=?

4.3.2 相关系数

协方差$Cov(X,Y)$反映了随机变量X与Y的线性相关性：

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

协方差Cov(X,Y)的取值大小与X,Y的量纲有关. 我们考虑消除量纲，同时把这种关系归一化。

4.3.2 相关系数

Definition 称下式为X,Y的相关系数(correlation coefficient).

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

Remark: ρ_{XY} 是无量纲的，它事实上是X,Y的标准化随机变量X*、Y*的协方差

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=E\left[\frac{(X-E(X))}{\sqrt{D(X)}}\ \frac{(Y-E(Y))}{\sqrt{D(Y)}}\right]$$

$$=Cov(X^*,Y^*)$$

相关系数的意义：

假设以X的线性函数 $aX+b$ **逼近**Y，以均方误差

$$e = E[(Y-(aX+b))^2]$$

表示逼近的程度，e越小说明逼近越好，反之越差。

那么最好的逼近是否存在？

如存在是什么样的？

$$e = E[Y-(aX+b)]^2$$

$$= E[Y^2 - 2aXY - 2bY + a^2X^2 + b^2 + 2abX]$$

$$= E(Y^2) - 2aE(XY) - 2bE(Y) + a^2E(X^2) + 2abE(X) + b^2$$

为求得最佳的a,b应该满足偏导为零：

$$0=\frac{\partial e}{\partial a}=-2E(XY)+2aE(X^2)+2bE(X)$$

$$0=\frac{\partial e}{\partial b}=-2E(Y)+2aE(X)+2b$$

$$\begin{cases} aE(X^2)+bE(X)=E(XY) \\ aE(X)+b=E(Y) \end{cases}$$

解出：

$$a^*=\frac{E(XY)-E(X)E(Y)}{E(X^2)-[E(X)]^2}=\frac{Cov(X,Y)}{D(X)}$$

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

相关系数的意义：

将 a^*, b^* 代入得

$$\min e = E(Y - (a^* X + b^*))^2$$

$$= E\left(Y - \frac{Cov(X,Y)}{D(X)} X - E(Y) + E(X) \frac{Cov(X,Y)}{D(X)} \right)^2$$

$$= E\left(Y - E(Y) - \frac{Cov(X,Y)}{D(X)} \left[X - E(X) \right] \right)^2$$

$$\min(e) = E\left[[Y - E(Y)]^2 - 2\frac{Cov(X,Y)}{D(X)}[Y - E(Y)][X - E(X)] \right.$$

$$\left. + \left[\frac{Cov(X,Y)}{D(X)}\right]^2 [X - E(X)]^2 \right]$$

$$= D(Y) - 2\frac{Cov(X,Y)}{D(X)} Cov(X,Y) + \left[\frac{Cov(X,Y)}{D(X)}\right]^2 D(X)$$

$$= D(Y) - \frac{[Cov(X,Y)]^2}{D(X)}$$

$$= D(Y)\left(1 - \frac{[Cov(X,Y)]^2}{D(X)D(Y)}\right)$$

$$= D(Y)[1 - \rho_{XY}^2]$$

以 $aX + b$ 逼近Y，最好的逼近与Y的均方差是 $D(Y)[1 - \rho_{XY}^2]$

$$\min(e) = D(Y)[1-\rho_{XY}^2]$$

[定理]: $(1)\ |\rho_{XY}| \le 1$

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

Proof:

（1） 由e的非负性得到 $|\rho_{XY}| \le 1$

（2）如果$|\rho_{XY}| = 1$ **，则对于上述的** a^*, b^* **得到**

$$E[(Y - a^* X - b^*)^2] = D(Y)\left(1 - \rho_{XY}^2\right) = 0$$

由 D(X)＝E(X²)－(EX)² 得到

$$D(Y-a^*X-b^*)+[E(Y-a^*X-b^*)]^2 \quad =E(Y-a^*X-b^*)^2=0$$

上式中两项全是非负数，因此

$$\begin{cases} D(Y-a^*X-b^*)=0 \\ E(Y-a^*X-b^*)=0 \end{cases}$$

由第一式子得到

$$P\{Y-a^*X-b^*=C\}=1$$

进一步由第二式得到 C=0。因此

$$P\{Y=a^*X+b^*\}=1$$

反之，如果存在 a^*, b^* 使得 $P\{Y = a^* X + b^*\} = 1$，则

$$P\{Y - a^* X - b^* = 0\} = 1$$

$$P\{[Y - a^* X - b^*]^2 = 0\} = 1$$

$$E[(Y - a^* X - b^*)^2] = 0$$

因此

$$0 = E[(Y - a^* X - b^*)^2] = e(a^*, b^*) \geq \min e = [1 - \rho_{XY}^2] D(Y)$$

必然 $|\rho_{XY}| = 1$

$$\min(e) = D(Y)[1-\rho_{XY}^2]$$

Remark: 当$|\rho_{XY}|$较大时，表明X与Y的线性关系程度较好；

当$|\rho_{XY}|$较小时，表明X与Y的线性关系程度较差.

特别地，

当$|\rho_{XY}|=1$ 时，表明X与Y之间以概率1存在线性关系；

当 $\rho_{XY}=0$ 时，表明X与Y之间没有线性关系，称两个变量**不相关**.

Definition 如果 $\rho_{XY}=0$，称 **X、Y不相关**

Remark：

（1）X,Y独立，必然有 $\rho_{XY}=0$，即X，Y不相关；

（2）X,Y不相关未必有X, Y独立。

Example 设随机变量θ服从均匀分布U [0, 2π]，X=cosθ, Y=sinθ，显然$X^2+Y^2=1$, 故X与Y不独立. 但

$$EX = \int_0^{2\pi} \cos\varphi \frac{1}{2\pi} d\varphi = 0$$

$$EY = \int_0^{2\pi} \sin\varphi \frac{1}{2\pi} d\varphi = 0$$

$$E(XY) = E(\cos\theta \sin\theta)$$

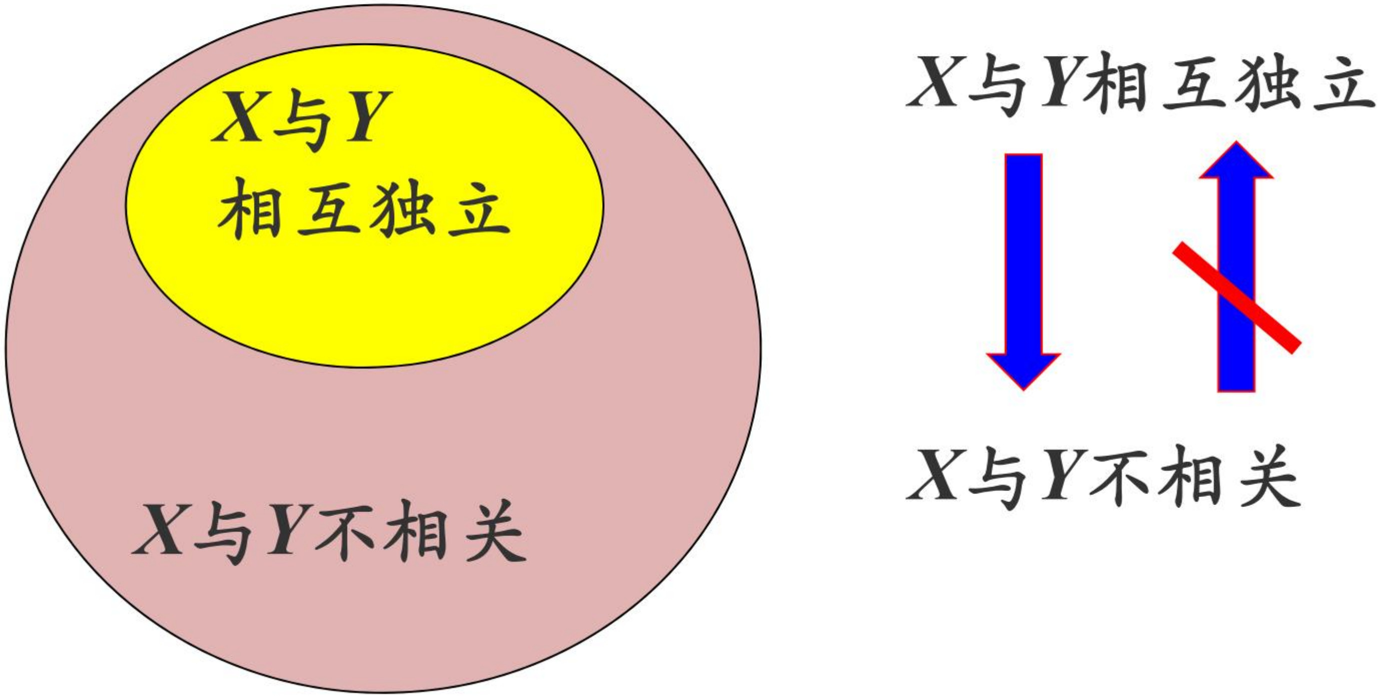
$$= \int_0^{2\pi} \cos\varphi \sin\varphi \frac{1}{2\pi} d\varphi = 0$$

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

因此 $\rho_{XY} = 0$ ， X与Y不相关.

说明: X与Y不相关, 仅针对于线性关系而言;

X与Y相互独立, 是就一般关系而言.



填空题 1分

设置

例3： 抛一枚均匀的硬币十次，若令X，Y分别表示出现正面和反面的次数，求X和Y的相关系数ρ_{XY}． [填空1]

作答

[定理] 对随机变量X和Y, 下列事实等价：

(1) X与Y不相关

(2) $Cov(X,Y)=0$

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

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

期望	方差
$E(C)=C$	$D(C)=0$
$E(X+C)=E(X)+C$	$D(X+C)=D(X)$
$E(aX)=aE(X)$	$D(aX)=a^2D(X)$
$E\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_iE(X_i)$	$D\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_i^2D(X_i)$ $\Longleftrightarrow$ 不相关
不相关 $\Longleftrightarrow$ $E(XY)=E(X)E(Y)$	$D(X)=0\Leftrightarrow P(X=C)=1$

正态分布-- 独立 vs 不相关

Example 设X,Y服从二元正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,

试求 ρ_{XY} **.**

解：前面知X,Y分别服从正态分布 $N(\mu_1,\sigma_1^2),N(\mu_2,\sigma_2^2)$

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

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dxdy$$

$$=\frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)$$

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

$$=\frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)$$

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

$$t=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right) \qquad z=\frac{x-\mu_1}{\sigma_1}$$

$$=\frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sigma_1 z)(\sigma_2\sqrt{1-\rho^2}t+\rho\sigma_2 z)$$

$$\exp\left[-\frac{1}{2}z^2-\frac{1}{2}t^2\right]\left(\sigma_1\sqrt{1-\rho^2}\sigma_2\right)dzdt$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sigma_1\sigma_2\sqrt{1-\rho^2}zt+\rho\sigma_1\sigma_2 z^2)\exp\left[-\frac{1}{2}(z^2+t^2)\right]dzdt$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma_1\sigma_2\sqrt{1-\rho^2}\,zt\exp\left[-\frac{1}{2}(z^2+t^2)\right]dzdt+$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho\sigma_1\sigma_2 z^2\exp\left[-\frac{1}{2}(z^2+t^2)\right]dzdt$$

$$=\frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\int_{-\infty}^{\infty}ze^{-\frac{1}{2}z^2}dz\int_{-\infty}^{\infty}te^{-\frac{1}{2}t^2}dt+$$

$$\frac{\rho\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}z^2e^{-\frac{1}{2}z^2}dz\int_{-\infty}^{\infty}e^{-\frac{1}{2}t^2}dt$$

$$=0+\frac{\rho\sigma_1\sigma_2}{2\pi}\sqrt{2\pi}\sqrt{2\pi}$$

$$=\rho\sigma_1\sigma_2$$

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

Remark 1： 二维正态随机变量的分布函数完全由参数 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 决定；

Remark 2: 如果(X,Y)服从正态分布，则X,Y独立与X,Y不相关是等价的，充要条件都是 $\rho=0$ 。

4.4 矩(moment)与协方差阵

Definition：随机变量X,Y

$E(X^k)\ (k=1,2,\ldots)$ 称为X的k阶原点矩，简称k阶矩。

$E\left[X-E(X)\right]^k\ (k=1,2,\ldots)$ 称为X的k阶中心矩。

$E\left(X^kY^l\right)\ (k,l=1,2,\ldots)$ 称为X, Y的$k+l$阶混合矩。

$E\left[X-E(X)\right]^k\left[Y-E(Y)\right]^l\ (k,l=1,2,\ldots)$

称为X, Y的$k+l$阶混合中心矩。

随机变量的期望是其1阶原点矩，
方差是其2阶中心矩，
协方差就是其1+1阶混合中心矩

Example 设X服从正态分布N (0，σ²)的随机变量，EX=0，且n阶中心距和原点矩

$$m_n = c_n = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^n e^{-x^2/2\sigma^2}\,\mathrm{d}x$$

$$= \begin{cases} 0, & n = 2k+1, \\ 1\times 3\times\cdots\times(n-1)\sigma^n, & n = 2k. \end{cases}$$

Example 如果ξ服从参数为λ的指数分布，那么 对于$k \ge 1$，

$$E\xi^k = \int_0^{+\infty} x^k \lambda e^{-\lambda x}\,\mathrm{d}x = \frac{k}{\lambda}E\xi^{k-1}$$

根据递推关系得 $E\xi^k = \dfrac{k!}{\lambda^k}$

即指数分布的任意阶矩存在.

对于n个随机变量 $X_1, X_2, \cdots, X_n$ ，可以定义：

$$C_{ij} = Cov(X_i, X_j) \quad i, j = 1,2,\cdots,n$$

进一步定义

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

称 *C*为 $X_1, X_2, \cdots, X_n$ 的协方差矩阵。

[定理] 设 $X \sim N(\mu, C)$，*B* 是一个 *n* 维的可逆矩阵,$Y = BX$ 则

$$Y \sim N(B\mu, BCB^T)$$

n维正态分布

n 维正态随机变量具有以下四条重要性质(证略)：

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

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

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

服从一维正态分布(其中 $l_1,l_2,\cdots,l_n$ 不全为零).

3° 若$(X_1,X_2,\cdots,X_n)$服从 n 维正态分布，设 $Y_1,Y_2,\cdots,Y_k$ 是 X_j $(j=1,2,\cdots,n)$的线性函数，则$(Y_1,Y_2,\cdots,Y_k)$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.

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

n 维正态分布在随机过程和数理统计中常会遇到.

思考题

设随机变量(X,Y)服从二元正态分布$N(0,1;1,4;-\frac{1}{2})$，求：

(1)$D(2X-Y)$; (2) $P(2X>Y)$; (3) (Z_1,Z_2)的分布，$Z_1=X+Y, Z_2=X-Y$.