西安電子科技大學
XIDIAN UNIVERSITY

第四章 随机变量的数字特征



西安电子科技大学 XIDIAN UNIVERSITY

3、方差的性质

性质 1 设 C 是常数, 则有 $D(C)=0$。

性质 2 $D(aX+b)=a^2D(X)$

$$E(aX+b)=aE(X)+b$$

$$D(aX+b)=E\left[aX+b-E(aX+b)\right]^2=E\left[aX-aE(X)\right]^2$$

$$=a^2E\left[X-E(X)\right]^2 \quad =a^2D(X)$$

性质 3 设 X, Y 相互独立, $D(X)$, $D(Y)$ 存在, 则 $D(X\pm Y)=DX+DY$ 。

证明 $D(X\pm Y)=E\{[(X\pm Y)-E(X\pm Y)]^2\}=E\{[X-E(X)]\pm[Y-E(Y)]\}^2$

$$=E[X-E(X)]^2\pm 2E\{[X-E(X)][Y-E(Y)]\}+E[Y-E(Y)]^2$$

$$=DX\pm 2E\{[X-E(X)][Y-E(Y)]\}+DY$$

中间项 $2E\{[X-E(X)][Y-E(Y)]\}\quad=2E[XY-XE(Y)-YE(X)+E(X)E(Y)]$

$$=2[E(XY)-E(X)E(Y)-E(Y)E(X)+E(X)E(Y)]\quad=2[E(XY)-E(X)E(Y)]$$

X, Y 相互独立，该项为0。

3、方差的性质

推广： 若 $X_1, X_2, \cdots, X_n$ 相互独立且各自的方差都存在，则：

$$D(X_1 \pm X_2 \pm \cdots \pm X_n) = DX_1 + DX_2 + \cdots + DX_n$$

利用性质求：$X \sim B(n, p)$ 时的 $D(X)$。

设 $X_i = \begin{cases} 1 & \text{如第}i\text{次试验成功} \quad i=1,2, \ldots, n \\ 0 & \text{如第}i\text{次试验失败} \end{cases}$

$X = \sum_{i=1}^{n} X_i$ ，$X_1, X_2, \cdots, X_n$ 独立同分布

$$D(X) = D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D(X_i) = nD(X_1) = np(1-p)$$

X_1	0	1
P	$1-p$	p

$E(X_1) = p$ $\quad E(X_1^2) = 1^2 \cdot p = p$

$D(X_1) = p - p^2 = p(1-p)$

3、方差的性质

例1： $X_i \sim N\left(\mu_i, \sigma_i^2\right)(i=1,2,\cdots,n)$ 且它们相互独立。

$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ ，求 $E(\overline{X}), D(\overline{X})$

解

$$E(\overline{X})=E(\sum_{i=1}^{n}\frac{1}{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\cdot n\mu=\mu$$

$$D(\overline{X})=D(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\sigma^2=\frac{\sigma^2}{n}$$

$\overline{X}$ 仍服从正态分布，即：$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$

对称中心未变，方差变小。顶点变高。

	$E(X)$	$D(X)$
0–1分布	p	$p(1-p)$
二项分布	np	$np(1-p)$
泊松分布	λ	λ
均匀分布	$(a+b)/2$	$(b-a)^2/12$
指数分布	$1/\lambda$	$1/\lambda^2$
正态分布	μ	σ^2

4、标准差

$$\sigma=\sqrt{D(X)}$$

取 $Y=\dfrac{X-E(X)}{\sigma}$，$E(Y)=0$，$D(Y)=1$ 。Y 称为随机变量X 的标准化随机变量。

目的： 为了消除由于计量单位的不同而给随机变量带来的影响。

正态分布标准化后成为标准正态分布。

5、矩

矩与期望、方差有密切的联系。

$g_1(X)=X^k, g_2(X)=[X-E(X)]^k$ ，求这两个函数的期望。

$E(X^k)$ ——k阶原点矩

$E[X-E(X)]^k$ ——k阶中心矩

数

4.3 协方差和相关系数

1、协方差

定义1 设X，Y是随机变量，若 $E[(X-EX)(Y-EY)]$ 存在，则称它为X与Y的协方差，记作 $Cov(X,Y)$ ，即：

$$Cov(X,Y)=E[(X-EX)(Y-EY)]=E(XY)-E(X)E(Y)$$

由方差的性质3：$D(X+Y)=E\left[X+Y-E(X+Y)\right]^2$

$$=E\left[X-E(X)\right]^2+2E\left[X-E(X)\right]\left[Y-E(Y)\right]+E\left[Y-E(Y)\right]^2$$

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

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

若Y=X，$Cov(X,X)=D(X)$ 。协方差是方差的一种推广，一半与X有关，一半与Y有关，反映X与Y之间的某种联系。

若从矩的关系看，就是混合矩

$E(X^mY^n)$ 混合原点矩 $m+n$阶

$E\left\{\left[X-E(X)\right]^m\left[Y-E(Y)\right]^n\right\}$ 混合中心矩 $m+n$阶

2、协方差的性质

性质1 $Cov(aX+b,cY+d)=acCov(X,Y)$，其中$a,b,c,d$为常数。

证明 $$Cov(aX+b,cY+d)=E\{[aX+b-E(aX+b)][cY+d-E(cY+d)]\}$$
$$=E\{a[X-E(X)]c[Y-E(Y)]\}$$
$$=acCov(X,Y)$$

性质2 $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

请自行证明

3、相关系数

定义 若 $D(X)>0, D(Y)>0$，则称 $\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}$ 为 X 和 Y 的相关系数。

若令：$X^*=\dfrac{X-E(X)}{\sqrt{D(X)}}, Y^*=\dfrac{Y-E(Y)}{\sqrt{D(Y)}}$ 分别为 X，Y 相应的标准化随机变量，

则 $\rho_{XY}=Cov(X^*,Y^*)$ 。

$\rho_{XY}=0$ 时，称 X 与 Y 不相关。

例 计算二维随机变量$(X,Y)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$的相关系数。

解 $Cov(X,Y)=E[(X-EX)(Y-EY)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)f(x,y)\mathrm{d}x\mathrm{d}y$

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

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

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

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

令 $u=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right),v=\frac{y-\mu_2}{\sigma_2}$

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

令 $u=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right),v=\frac{y-\mu_2}{\sigma_2}$

原式 $$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\sigma_2 ve^{-\frac{v^2}{2}}\int_{-\infty}^{+\infty}\sigma_1\left(\sqrt{1-\rho^2}u+\rho v\right)e^{-\frac{u^2}{2}}du$$

$$=\frac{\sigma_1\sigma_2}{2\pi}\left\{\int_{-\infty}^{\infty}ve^{-\frac{v^2}{2}}dv\int_{-\infty}^{\infty}\sqrt{1-\rho^2}ue^{-\frac{u^2}{2}}du+\int_{-\infty}^{\infty}\rho v^2e^{-\frac{v^2}{2}}dv\int_{-\infty}^{\infty}e^{-\frac{u^2}{2}}du\right\}$$

$$=\frac{\rho\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}v^2e^{-\frac{v^2}{2}}dv\int_{-\infty}^{\infty}e^{-\frac{u^2}{2}}du$$

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

$$=\rho\sigma_1\sigma_2$$

故 $\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}=\frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2}=\rho$

4、独立和不相关

相关系数 ρ_{XY} 是刻画随机变量 X、Y 之间线性相关程度的一个数字特征。

回顾： 二维正态分布证明过：随机变量X、Y相互独立 $\Leftrightarrow \rho_{XY}=0$

即对于二维正态分布，独立和不相关等价。那么对于更一般情况呢？

定理 1 若随机变量 X、Y 相互独立，则$Cov(X,Y)=0$，即 X、Y 不相关，但反之不成立。

证明 若随机变量 X、Y 相互独立，则：

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

$$=[E(X)-E(X)][E(Y)-E(Y)]=0 \qquad \text{不相关}$$

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

不独立

例1 设 X, Y 在单位圆盘上服从均匀分布，即 $f(x,y)=\begin{cases}\frac{1}{\pi}, & 0\le x\le 1, 0\le y\le 1\\ 0, & \text{其他}\end{cases}$

在第三章中已知该例 X, Y 的边缘概率密度求得分别为：

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

显然 $f(x,y)\ne f_X(x)f_Y(y)$，故 X、Y 不相互独立。

现讨论 X 与 Y的相关性。

解

$$\because\ EY=EX=\int_{-1}^{1}x\frac{2}{\pi}\sqrt{1-x^2}\mathrm{d}x=0$$

$$E(XY)=\iint\limits_{x^2+y^2\le 1}xy\frac{1}{\pi}\mathrm{d}x\mathrm{d}y=\frac{1}{\pi}\int_{-1}^{1}x\mathrm{d}x\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}y\mathrm{d}y=0$$

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

即 $\rho_{XY}=0$

故 X、Y 是不相关的。

设 X，Y 是方差均大于零的随机变量，则

(1) X，Y 不相关 $\Leftrightarrow$ $\mathrm{cov}(X, Y) = 0$；

(2) X，Y 不相关 $\Leftrightarrow D(X \pm Y) = DX + DY$；

(3) X，Y 不相关 $\Leftrightarrow E(XY) = EX \cdot EY$；

什么叫X，Y不相关？

X、Y 线性不相关，随机变量X、Y不是没有关系，而是没有严格的线性关系。

例 设 $X \sim U(-\pi,\pi), Y = \cos X$

解

$$E(X) = \int_{-\infty}^{+\infty} x \frac{1}{2\pi} I_{(-\pi,\pi)}(x)dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} xdx = 0$$

$$E(Y) = \int_{-\infty}^{+\infty} \cos x \frac{1}{2\pi} I_{(-\pi,\pi)}(x)dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos xdx = 0$$

$$Cov(X,Y) = E(XY) = E(X\cos X) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\cos xdx = 0$$

所以X，Y不相关，但明显X与Y之间存在着非线性的关系。

定理 2 $Cov^2(X,Y) \le D(X)D(Y)$ 即 $|\rho| \le 1$

若等号成立 $\Leftrightarrow$ X与Y之间存在严格的线性关系

如 $Y = a + bX$

$\rho_{XY} = 1$ 存在常数 $b>0$，使 $Y = a + bX$

$\rho_{XY} = -1$ 存在常数 $b<0$，使 $Y = a + bX$

$\rho_{XY} = 0$ 无线性关系

$\rho_{XY} \ne 0$ ， $|\rho_{XY}| \ne 1$ X与Y有相关关系，但不能严格确定。

$0 < \rho_{XY} < 1$ X与Y正相关

$-1 < \rho_{XY} < 0$ X与Y负相关

$|\rho_{XY}|$ 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。