5.5 多维随机变量的数学特征

5.5.1 多维随机变量的期望

定理5.12. 对离散型二维随机变量(X,Y), 设随机变量$Z=g(X,Y)$, 则

$$E[Z]=E[g(X,Y)]=\sum_{i,j}g(x_i,y_j)p_{ij};$$

对连续型二维随机变量(X,Y), 设联合概率密度为$f(x,y)$, 以及随机变量$Z=g(X,Y)$, 则

$$E[Z]=E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy.$$

例5.13. 设$X\sim\mathcal{N}(0,1)$, $Y\sim\mathcal{N}(0,1)$且X与Y独立, 求$E[\max(X,Y)]$.

解. 根据独立性定义得到随机变量X和Y的联合概率密度为$f(x,y)=f_X(x)f_Y(y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$.

$$\begin{aligned}E[\max(X,Y)] &= \int\int_{D_1}xf(x,y)dxdy+\int\int_{D_2}yf(x,y)dxdy\\ &= \int_{-\infty}^{+\infty}dy\int_{y}^{+\infty}xf(x,y)dx+\int_{-\infty}^{+\infty}dx\int_{x}^{+\infty}yf(x,y)dy\\ &= 2\int_{-\infty}^{+\infty}dy\int_{y}^{+\infty}xf(x,y)dx=\frac{1}{\pi}\int_{-\infty}^{+\infty}dy\int_{y}^{+\infty}xe^{-\frac{x^2+y^2}{2}}dx\\ &= \frac{1}{\pi}\int_{-\infty}^{+\infty}e^{-y^2}dy=\frac{1}{\sqrt{\pi}}\end{aligned}$$

最后一个等式成立是因为$\int_{-\infty}^{+\infty}e^{-y^2}dy=\sqrt{\pi}$. □

习题. 在长为1的线段上任取两点X,Y, 求$E[\min(X,Y)],E[|X-Y|]$.

定理5.13. *1)* 对任意随机变量X,Y和常数a,b, 有$E[aX+bY]=aE[X]+bE[Y]$.
2) 对独立随机变量X和Y, 有$E[XY]=E[X]E[Y]$, 对任意函数h,g有$E[h(X)g(Y)]=E[h(X)]E[g(Y)]$;
3) 对非独立随机变量X和Y, 有*Cauchy-Schwartz*不等式$E[XY]\leq\sqrt{E[X^2]E[Y^2]}$.

Proof. 设X,Y的联合概率密度为$f(x,y)$, 则

$$\begin{aligned}E[aX+bY] &= \int\int(ax+by)f(x,y)dxdy\\ &= a\int\int xf(x,y)dxdy+b\int\int yf(x,y)dxdy=aE(X)+bE(Y).\end{aligned}$$

若X与Y独立, 则有

$$\begin{aligned}E[XY] &= \int\int xyf(x,y)dxdy=\int\int xf_X(x)yf_Y(y)dxdy\\ &= \int_{-\infty}^{+\infty}xf_X(x)dx\int_{-\infty}^{+\infty}yf_Y(y)dy=E(X)E(Y).\end{aligned}$$

若X与Y不独立, 对任意$t\in\mathbb{R}$有$E[(X+tY)^2]\geq0$成立, 即任意$t\in\mathbb{R}$,

$$t^2E[Y^2]+E[X^2]+2tE[XY]\geq0.$$

因此有$\Delta=4[E(XY)]^2-4E[X^2]E[Y^2]\leq0$, 即$E(XY)\leq\sqrt{E(X^2)E(Y^2)}$. □

5.5.2 协方差

定理5.14. 对随机变量X与Y, 有

$$Var(X+Y) = Var(X) + Var(Y) + 2E[(X-E(X))(Y-E(Y))].$$

特别地, 当X与Y独立时, 有

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

Proof. 令$Z = X + Y$, 则

$$\begin{aligned} Var(Z) &= E[(Z-EZ)^2] = E[(X-EX+Y-EY)^2] \\ &= E(X-EX)^2 + E(Y-EY)^2 + 2E[(X-EX)(Y-EY)] \\ &= Var(X) + Var(Y) + 2E[(X-EX)(Y-EY)]. \end{aligned}$$

若X与Y独立, 则$2E[(X-EX)(Y-EY)] = 0$, 所以$Var(X+Y) = Var(X) + Var(Y)$. □

定义5.9. 定义随机变量X和Y的协方差为

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

根据协方差定义和定理5.14有

$$Cov(X,X) = Var(X) \quad \text{和} \quad Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

下面研究协方差的性质.

性质5.1. 对任意常数c, 有$Cov(X,c) = 0$; $Cov(X,Y) = Cov(Y,X)$. (交换律)

性质5.2. 对任意常数a和b, 随机变量X和Y, 有

$$Cov(aX, bY) = abCov(X,Y), \quad Cov(X+a, Y+b) = Cov(X,Y).$$

Proof. 根据协方差定义有

$$\begin{aligned} Cov(aX,bY) &= E[(aX-E(aX))(bY-E(bY))] = abE[(X-E(X))(Y-E(Y))]; \\ Cov(X+a,Y+b) &= E[(X+a-E(X+a))(Y+b-E(Y+b))] = E[(X-E(X))(Y-E(Y))]. \end{aligned}$$

□

性质5.3. 对任意随机变量X_1, X_2, Y, 有

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

Proof. 我们有

$$
\begin{aligned}
&Cov(X_1+X_2,Y)=E[(X_1+X_2-E(X_1)-E(X_2))(Y-E(Y))]\\
=\;& E[(X_1-E(X_1))(Y-E(Y))]+E[(X_2-E(X_2))(Y-EY)]=Cov(X_1,Y)+Cov(X_2,Y).
\end{aligned}
$$

□

由此性质可进一步得到: 对随机变量$X_1,X_2,\ldots,X_n$和$Y_1,Y_2,\ldots,Y_m$, 有

$$Cov\left(\sum_i^n X_i,\sum_j^m Y_j\right)=\sum_i^n\sum_j^m Cov(X_i,Y_j),$$

以及进一步有

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

性质5.4. *随机变量X与Y独立$\Rightarrow Cov(X,Y)=0$; 但反之不成立.*

Proof. 若X与Y独立, 则

$$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E[X-E(X)]E[Y-E(Y)]=0.$$

反之不成立, 例如随机变量X的分布列为

X	-1	0	1
P_i	$1/3$	$1/3$	$1/3$

当$X\neq 0$时随机变量$Y=0$, 否则$Y=0$; 则X与Y不独立. 同时有

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

□

性质5.5. *对任意随机变量X与Y, 有*

$$(Cov(X,Y))^2\le Var(X)Var(Y)$$

等号成立的充要条件是$Y=aX+b$ (即X与Y之间有线性关系).

Proof. 由Cauchy-Schwartz不等式有

$$
\begin{aligned}
Cov(X,Y) &= E[(X-E(X))(Y-E(Y))]\\
&\le \sqrt{E[(X-E(X))^2]E[(Y-E(Y))^2]}=\sqrt{Var(X)Var(Y)}.
\end{aligned}
$$

下面证明等号成立的充要条件. 如果$Y=aX+b$, 则

$$Cov(X,Y)=Cov(X,aX+b)=aVar(X),\quad Var(Y)=a^2Var(X),$$

所以
$$Cov^2(X,Y)=a^2Var^2(X)=Var(X)a^2Var(X)=Var(X)Var(Y).$$
另一方面, 若$(Cov(X,Y))^2=Var(X)Var(Y)$, 则
$$(E[(X-EX)(Y-EY)])^2=E(X-EX)^2E(Y-EY)^2,$$
设
$$\begin{aligned} f(t) &= E[t(X-EX)-(Y-EY)]^2 \\ &= t^2E[X-E(X)]^2-2tE[(X-E(X))(Y-E(Y))]+E[Y-E(Y)]^2 \end{aligned}$$
根据一元二次方程的性质
$$\Delta=4(E[(X-EX)(Y-EY)])^2-4E(X-EX)^2E(Y-EY)^2=0,$$
得到$f(t)=0$恰有一重根t_0. 由此得到
$$f(t_0)=0=E[(t_0(X-EX)-(Y-EY))^2]$$
因为$[t_0(X-EX)-(Y-EY)]^2=0$, 所以$Y=t_0(X-E(X))+E(Y)=aX+b$. □

习题. *随机变量X与Y独立, 且$Var(X)=6$和$Var(Y)=3$, 求$Var(2X\pm Y)=$?*

习题. *随机变量$X\sim P(2)$, $Y\sim\mathcal{N}(-2,4)$, X与Y独立, 则$E[(X-Y)^2]=$?*

根据性质 5.5可知
$$\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}\le 1.$$
等号成立的充要条件是X与Y存在线性相关. 上式一定程度上反应了随机变量X和Y的线性相关程度, 由此引入一个新概念: 相关系数.

定义5.10. *设X和Y为二维随机变量, 如果$Var(X), Var(Y)$存在且不为0, 则称*
$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$
为X与Y的相关系数, 简记ρ.

注1) 这里使用相关系数而不是$Cov(X,Y)$, 主要是规范$|\rho_{XY}|\le 1$, $Cov(X,Y)$受数值大小影响.

注2) 相关系数$|\rho_{XY}|\le 1$: 若$\rho>0$, X与Y正相关; 若$\rho<0$, X与Y负相关; $|\rho_{XY}|=1$的充要条件为X与Y有线性关系$Y=aX+b$. 本质上ρ_{XY}刻画了X,Y的线性相关程度, 又称为“线性相关系数”.

注3) 相关系数$\rho=0$称X与Y不相关(线性不相关). 独立$\Rightarrow$不相关, 不相关$\nRightarrow$ 独立.

注4) 随机变量X与Y不相关, 仅表示X与Y之间无线性关系, 还可能存在其他关系. 例如: $X\sim U[-\frac{1}{2},\frac{1}{2}]$, $Y=cosX$. 易有$E(X)=0$,
$$\begin{aligned} & Cov(X,Y)=E[(X-E(X))(Y-E(Y))] \\ =\ & E[X\cdot\cos(X)-XE(\cos(X))]=E[X\cdot\cos(X)]=\int_{-\frac{1}{2}}^{\frac{1}{2}}x\cdot\cos(x)dx=0. \end{aligned}$$

定理5.15. 随机变量X,Y不相关的等价性如下:

$$\rho_{XY}=0\Longleftrightarrow Cov(X,Y)=0\Longleftrightarrow E(X,Y)=E(X)E(Y)\Longleftrightarrow Var(X\pm Y)=Var(X)\pm Var(Y).$$

二维正态分布:

$$\begin{pmatrix}X_1\\X_2\end{pmatrix}\sim\mathcal{N}\left(\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)$$

1) $X_1\sim\mathcal{N}(\mu_1,\sigma_1^2)$, $X_2\sim\mathcal{N}(\mu_2,\sigma_2^2)$.

2) $Cov(X_1,X_2)=\rho\sigma_1\sigma_2$, 参数$\rho$为$X_1$与$X_2$的相关系数.

3) X_1与X_2独立$\Longleftrightarrow\rho=0$(不相关).

习题. 随机变量(X,Y)联合概率密度为

$$f(x,y)=\begin{cases}\frac{1}{8}(x+y) & 0\le x\le 2,0\le y\le 2\\0 & 其它\end{cases}$$

求$Cov(X,Y),Var(X+Y)$.

习题. 随机变量$X\sim\mathcal{N}(\mu,\sigma^2)$, $Y\sim\mathcal{N}(\mu,\sigma^2)$, 且$X$与$Y$相互独立. 求$Z_1=\alpha X+\beta Y$和$Z_2=\alpha X-\beta Y$的相关系数($\alpha,\beta$不全为0).

5.5.3 随机向量的数学期望与协方差阵

定义5.11. 设$X=(X_1,X_2,\cdots,X_n)'$, 则随机向量的期望

$$E(X)=(E(X_1),E(X_2),\cdots,E(X_n))',$$

称随机变量X的协方差矩阵为

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

定理5.16. 随机变量X的协方差阵是对称非负定的矩阵.

多维正态分布$X=(X_1,X_2,\ldots,X_n)'\sim N(\mu,\Sigma)$, 则有

$$\mu=(E[X_1],E[X_2],\ldots,E[X_n])'\qquad\Sigma=[Cov(X_i,X_j)]_{n\times n}$$

其概率密度为:

$$f(x)=\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)).$$

多维随机变量具有如下性质:

1) $X = (X_1, X_2, \cdots, X_n)' \sim \mathcal{N}(\mu, \Sigma)$, 则每个变量$X_i$的边缘分布是正态分布.

2) $X = (X_1, X_2, \cdots, X_n)' \sim \mathcal{N}(\mu, \Sigma)$, 则$\sum_{i=1}^{n} a_i X_i$是正态分布.

3) $X = (X_1, X_2, \cdots, X_n)' \sim \mathcal{N}(\mu, \Sigma)$, 则$X_1, X_2, \cdots, X_n$相互独立$\Longleftrightarrow X_1, X_2, \cdots, X_n$相互不相关.