第九周 协方差与相关系数

9.1. 随机变量函数的期望

随机变量函数的期望

n维随机变量 $X=(X_1,X_2,\cdots,X_n)$，若 $Z=g(X_1,X_2,\cdots,X_n)$，则

离散情形：$E(Z)=\sum_{i_1}\cdots\sum_{i_n}g(x_1,x_2,\cdots,x_n)P(X_1=x_1,\cdots,X_n=x_n)$

连续情形：$E(Z)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,x_2,\cdots,x_n)f(x_1,x_2,\cdots,x_n)dx_1\cdots dx_n$

$Var(Z)=E\left[(Z-E(Z))^2\right]=E(Z^2)-E(Z)^2$

****

例 9.1.1 从$1,2,3,4$中等可能地取 1 个数记为 X，再从$1,2,\cdots,X$中等可能地取 1 个数记为 Y。求 $E(X+2Y)$ 与 $Var(X+2Y)$。

解： (X,Y)的联合分布列为

$X\backslash Y$	1	2	3	4
1	1/4	0	0	0
2	1/8	1/8	0	0
3	1/12	1/12	1/12	0
4	1/16	1/16	1/16	1/16

$$E(X+2Y)=\sum_{i=1}^{4}\sum_{k=1}^{4}(x_i+2y_k)P(X=x_i,Y=y_k)$$

$$=\frac{1}{4}\cdot(1+2\cdot 1)+\frac{1}{8}(2+2\cdot 1+2+2\cdot 2)+\frac{1}{12}\sum_{k=1}^{3}(3+2\cdot k)+\frac{1}{16}\sum_{k=1}^{4}(4+2\cdot k)=6$$

$$E\left[(X+2Y)^2\right]=\sum_{i=1}^{4}\sum_{k=1}^{4}(x_i+2y_k)^2P(X=x_i,Y=y_k)=\sum_{i=1}^{4}\frac{1}{4i}\sum_{k=1}^{i}(x_i+2y_k)^2=\frac{259}{6}$$

$$Var(X+2Y)=E\left[(X+2Y)^2\right]-E(X+2Y)^2=\frac{43}{6}。$$

(X,Y)的联合与边缘分布列为

$X\backslash Y$	$Y=1$	$Y=2$	$Y=3$	$Y=4$	$P(X=i)$
$X=1$	1/4	0	0	0	1/4
$X=2$	1/8	1/8	0	0	1/4
$X=3$	1/12	1/12	1/12	0	1/4
$X=4$	1/16	1/16	1/16	1/16	1/4
$P(Y=k)$	25/48	13/48	7/48	1/16	1

$$E(X)=\frac{1}{4}(1+2+3+4)=\frac{5}{2}, \quad E(Y)=\frac{25}{48}+2\cdot\frac{13}{48}+3\cdot\frac{7}{48}+4\cdot\frac{3}{48}=\frac{7}{4}$$

$$E(X+2Y)=E(X)+E(2Y)=\frac{5}{2}+2\cdot\frac{7}{4}=6, \qquad Var(X+2Y)\neq Var(X)+Var(2Y)$$

****

随机变量和的期望等于期望的求和

$$E(X_1+X_2)=\sum_i\sum_j(i+j)\cdot P(X_1=i,X_2=j)$$

$$=\sum_i\sum_j i\cdot P(X_1=i,X_2=j)+\sum_i\sum_j j\cdot P(X_1=i,X_2=j)$$

$$=\sum_i i\cdot\sum_j P(X_1=i,X_2=j)+\sum_i j\cdot\sum_j P(X_1=i,X_2=j)$$

$$=\sum_i i\cdot P(X_1=i)+\sum_i j\cdot P(X_2=j)=E(X_1)+E(X_2)$$

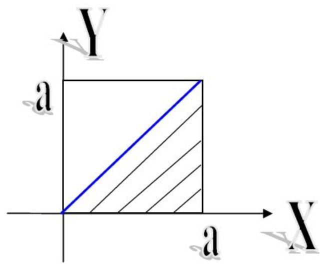
$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)\cdots+E(X_n)$$

若 $X_1,X_2,\cdots,X_n$ 相互独立，$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n)$，

若 $X_1,X_2,\cdots,X_n$ 相互独立，$Var(X_1+X_2+\cdots+X_n)=Var(X_1)+Var(X_2)\cdots+Var(X_n)$

****

例 9.1.2 在长度为 a 的线段上随机任取两点，求两点距离的期望与方差。



解： 设两个点到线段一个固定端点的距离分别为随机变量 X 与 Y，则 (X,Y) 的密度函数为

$$f_{XY}(x,y)=\begin{cases}1/a^2, & 0<x,y<a\\ 0, & 其他\end{cases}$$

$$E(|X-Y|)=\int_0^a\int_0^a|x-y|\cdot\frac{1}{a^2}dxdy=\int_0^a dx\int_0^x 2\cdot(x-y)\cdot\frac{1}{a^2}dy=\frac{a}{3}$$

$$E(|X-Y|^2)=\int_0^a\int_0^a(x-y)^2\cdot\frac{1}{a^2}dxdy=\frac{a^2}{6}, \quad Var(|X-Y|)=\frac{a^2}{18}。$$

****