多维随机变量及其分布03

多维随机变量的特征数

1、多维随机变量函数的数学期望

定理：若二维随机变量(X,Y)的分布用联合分布列$P(X=x_i,Y=y_i)$或用联合密度函数$p(x,y)$表示，则$Z=g(X,Y)$的数学期望为

$$E(Z)=\begin{cases}\sum_i\sum_j g(x_i,y_i)P(X=x_i,Y=y_i),\text{在离散场合}\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)p(x,y)\,dx\,dy,\text{在连续场合}\end{cases}$$

2、数学期望与方差的运算性质

性质 1：设(X,Y)是二维随机变量，则有$E(X+Y)=E(X)+E(Y)$

推广到n个场合为$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)$

性质 2：若随机变量X与Y互相独立，则有$E(XY)=E(X)E(Y)$

性质 3：若随机变量X与Y互相独立，则有$Var(X\pm Y)=Var(X)+Var(Y)$，推广到$n$个场合为$Var(X_1\pm X_2\pm\cdots\pm X_n)=Var(X_1)+Var(X_2)+\cdots+Var(X_n)$

3、协方差

定义：设(X,Y)是二维随机变量，若$E[(X-E(X))(Y-E(Y))]$存在，则称此数学期望为X与Y的协方差，或称X与Y的相关（中心）矩，并记为$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]$，特别有$Cov(X,Y)=Var(X)$

当$Cov(X,Y)>0$时，称为X与Y正相关

当$Cov(X,Y)<0$时，称为X与Y负相关

当$Cov(X,Y)=0$时，称为X与Y不相关

性质 1：$Cov(X,Y)=E(XY)-E(X)E(Y)$

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

性质 2：若随机变量X与Y互相独立，则$Cov(X,Y)=0$，反之不然。

性质 3：对任意的二维随机变量(X,Y)有$Var(X\pm Y)=Var(X)+Var(Y)\pm 2Cov(X,Y)$

性质 4：协方差$Cov(X,Y)=Cov(Y,X)$

性质 5：任意随便变量X与常数a的协方差为零，即$Cov(X,a)=0$

性质 6：对任意常数a,b,有$Cov(aX,bY)=abCov(X,Y)$

性质 7：设X,Y,Z是任意三个随机变量，则$Cov(X+Y,Z)=Cov(X,Z)+Cov(Y+Z)$

4、相关系数

为了消除量纲，对协方差除以相同量纲的量，就得到一个新的概念——相关系数。定义：设(X,Y)是一个二维随机变量，且$Var(X)={\sigma_X}^2>0,Var(Y)={\sigma_Y}^2$，则称$corr(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}$为$X$与$Y$的（线性）相关系数。

相关系数的另一种解释：他是标准化变量的协方差。若记X与Y的数学期望分别为μ_X,μ_Y，其标准化变量为：

$$X^*=\frac{X-\mu_X}{\sigma_X},\quad Y^*=\frac{Y-\mu_Y}{\sigma_Y}$$

$$Cov(X^*,Y^*)=Cov\left(\frac{X-\mu_X}{\sigma_X},\frac{Y-\mu_Y}{\sigma_Y}\right)=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=Corr(X,Y)$$

性质 1：（施瓦茨（Schwarz）不等式）对任意的二维随机变量(X,Y)，若X与Y的方差都存在，且记${\sigma_X}^2=Var(X),{\sigma_Y}^2=Var(Y)$，则有

$$[Cov(X,Y)]^2 \le {\sigma_X}^2{\sigma_Y}^2$$

性质 2：$-1 \le Corr(X,Y) \le 1$

性质3:$Corr(X,Y)=\pm 1$的充要条件是X与Y间几乎处处有线性关系，即存在$a(\ne 0)$与b，使得$P(Y=aX+b)=1$，其中，当$Corr(X,Y)=1$时，$a>0$；当$Corr(X,Y)=-1$时，$a<0$

注: 相关系数刻画了X与Y之间的线性关系强度，因此也称它为线性相关系数。

性质 4: 在二维正态分布$N(\mu_1,\mu_2,{\sigma_1}^2,{\sigma_2}^2,\rho)$场合，不相关与独立是等价的

5、随机向量的数学期望向量与协方差矩阵

定义：记n维随机变量$X=(X_1,X_2,\dots,X_n)'$，若其每个分量的数学期望都存在，则称$E(X)=(E(X_1),E(X_2),\dots,E(X_n))'$为$n$为随机向量$X$的数学期望向量，简称为$X$的数学期望，而称

$E\left\{\left(X-E(X)\right)\left(X-E(X)\right)'\right\}=$

$\begin{bmatrix} Var(X_1) & Cov(X_1,X_2) & \dots & Cov(X_1,X_n) \\ Cov(X_2,X_1) & Var(X_2) & \dots & Cov(X_2,X_n) \\ \dots & \dots & \dots & \dots \\ Cov(X_n,X_1) & Cov(X_n,X_2) & \dots & Var(X_n) \end{bmatrix}$为随机向量的方差-协方差矩阵，简称协方差矩阵，记为$Cov(X)$

条件分布与条件期望

6、条件分布

对二维随机变量(X,Y)而言，所谓随机变量X的条件分布，就是给定Y取某个值的条件下X的分布

（一）离散型随机变量的条件分布

设二维离散随机变量(X,Y)的联合分布列为

$$p_{ij}=P(X=x_i,Y=y_i),i=1,2,\dots\ \ j=1,2,\dots$$

定义：对一切使$P\left(Y=y_j\right)=p_{.j}=\sum_{i=1}^{m}p_{ij}>0$的$y_j$，称

$$p_{i|j}=P\left(X=x_i\middle|Y=y_j\right)=\frac{P(X=x_i,Y=y_i)}{P(Y=y_i)}=\frac{p_{ij}}{p_{.j}}$$

（二）连续型随机变量的条件分布

设二维离散随机变量(X,Y)的联合密度函数为$p(x,y)$，边际密度函数$p_X(x)$，$p_Y(y)$。由于连续型随机变量取某个值的概率为零，即$P(Y=y)=0$，所以无法用条件概率直接计算，因此，将$P(X\le x,Y\le y)$看成 $h\to 0$ 时 $P(X\le x|y\le Y\le y+h)$ 的极限，因此可以得到$P(X\le x|Y=y)=\int_{-\infty}^{x}\frac{p(u,y)}{P_Y(y)}du,$

对于一切使$P_Y(y)>0$的y, 给定$Y=y$条件下X的条件分布函数的条件密度函数分别为$F(x|y)=\int_{-\infty}^{x}\frac{p(u,y)}{P_Y(y)}du,p(x|y)=\frac{p(x,y)}{P_Y(y)}$, 同理, 对于一切使$P_X(x)>0$的$x$, 给定$X=x$条件下$Y$的条件分布函数的条件密度函数分别为$,p(y|x)=\frac{p(x,y)}{P_X(y)}$，**注：不同条件的条件分布函数不相同**

例题: 设二维离散随机变量(X,Y)服从$G=\{(x,y)|x^2+y^2\le 1\}$上均匀分布，试求给定$Y=y$条件下X的条件密度函数$p(x|y)$

解：因为

$$p(x,y)=\begin{cases}\frac{1}{\pi}，x^2+y^2\le 1\\ \quad 0，其他\end{cases}$$

由此，可得

$$P_Y(y)=\int_{-\infty}^{\infty}p(x,y)\,dx=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}p(x,y)\,dx$$

$$=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2},-1\le y\le 1\\ 0,其他\end{cases}$$

所以，当 $-1<y<1$ 时，有 $p(x|y)=\frac{p(x,y)}{P_Y(y)}=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2},-\sqrt{1-y^2}\le x\le\sqrt{1-y^2}\\ 0,其他\end{cases}$

将$y=0$和$y=0.5$分别代入上式可得（两个均匀分布）

$p(x|y=0)\begin{cases}\frac{1}{2},-1\le x\le 1\\ 0,其他\end{cases}$ $p(x|y=0.5)\begin{cases}\frac{1}{\sqrt{3}},-\frac{\sqrt{3}}{2}\le x\le\frac{\sqrt{3}}{2}\\ 0,其他\end{cases}$

当$-1<y<1$时，给定$Y=y$条件下，X服从$(-\sqrt{1-y^2},\sqrt{1-y^2})$上的均匀分布，同理$-1<x<1$，给定$X=x$条件下，$Y$服从$(-\sqrt{1-x^2},\sqrt{1-x^2})$上的均匀分布

（三）连续场合的全概率公式和贝叶斯公式

全概率公式的密度函数形式

$$p_Y(y)=\int_{-\infty}^{\infty}p_X(x)p(y|x)\,dx$$

$$p_X(x)=\int_{-\infty}^{\infty}p_Y(y)p(x|y)\,dy$$

贝叶斯公式的密度函数形式

$$p(x|y)=\frac{p_X(X)p(y|x)}{\int_{-\infty}^{\infty}p_X(x)p(y|x)\,dx}=\frac{p(x,y)}{p_Y(y)}$$

$$p(y|x)=\frac{p_X(X)p(y|x)}{\int_{-\infty}^{\infty}p_Y(y)p(x|y)\,dy}=\frac{p(x,y)}{p_X(x)}$$

7、条件数学期望

定义：条件分布的数学期望(若存在)称为条件期望，其定义如下：

$$E(X|Y=y)=\begin{cases}\displaystyle\sum_i x_iP(X=x_i|Y=y),(X,Y)\text{为二维离散随机变量}\\ \displaystyle\int_{-\infty}^{\infty}xp(x|y)dx,(X,Y)\text{为二维连续随机变量}\end{cases}$$

$$E(Y|X=x)=\begin{cases}\displaystyle\sum_j y_jP(Y=y_j|X=x),(X,Y)\text{为二维离散随机变量}\\ \displaystyle\int_{-\infty}^{\infty}yp(y|x)dy,(X,Y)\text{为二维连续随机变量}\end{cases}$$

注意:条件分布的期望与无条件期望的区别不仅在计算公式上，而且在于含以上，比如$E(X|Y=y)$是y的函数

定理:(重期望公式)设(X,Y)为二维随机变量,且E(X)存在,则$E(X)=E(E(X|Y))$，重期望公式的具体使用如下：

（1）如果Y是一个离散随机变量，则$E(X)=E(E(X|Y))$可以改写为

$$E(X)=\sum_j E(X|Y=y_j)P(Y=y_j)$$

（2）如果Y是一个连续随机变量，则$E(X)=E\big(E(X|Y)\big)$可以改写为

$$E(X)=\int_{-\infty}^{\infty}E(X|Y=y)p_Y(y)dy$$

例题：一矿工被困在有三个门的矿井里。第一个们通一坑道，沿此坑道走 3 小时可到达安全区，第二个门通一坑道，沿此坑道走 5 小时又回到原处，第三个门通一坑道，沿此坑走 7 小时也回到原处。假定此矿工总是等可能地在三个门中选择一个。试求他平均要用多少时间才能到达安全区？

解：设该矿工需要X小时达到安全区，则X的可能取值为

3，3+5，7+3，5+5+3，5+7+3，…

要写出X的分布是困难的，因此无法求出$E(X)$，记Y表示第一次所选的门，$|Y=i|$就是选择第i个门，由题设知

$$P(Y=1)=P(Y=2)=P(Y=3)=1/3$$

因此，选择第一个门后 3 小时可达到安全区，所以$E(X|Y=1)=3$

又因为选择第二个门 5 小时后又回到原处，所以$E(X|Y=2)=5+E(X)$

又因为选择第三个门 7 小时后又回到原处，所以$E(X|Y=3)=7+E(X)$

综上所述，$E(X)=\frac{1}{3}[3+5+E(X)+7+E(X)]=5+\frac{2}{3}E(X)$，得到$E(X)=15$，即该矿工平均要 15 小时才能达到安全区