多维随机变量及其分布02

1、边际分布函数

如果在二维随机变量(X,Y)的联合分布函数$F(x,y)$中令$y\to\infty$，由于$\{Y<\infty\}$为必然事件，故可得

$$\lim_{y\to\infty}F(x,y)=P(X\le x,Y\le\infty)=P(X\le x)$$

这是由(X,Y)的联合分布函数$F(x,y)$求得的X的分布函数，被称为X的边际分布，记为 $F_X(x)=F(x,\infty)$

同理，$F(x,y)$中令$x\to\infty$，可得Y的边际分布，$F_Y(y)=F(\infty,y)$

2、边际分布列

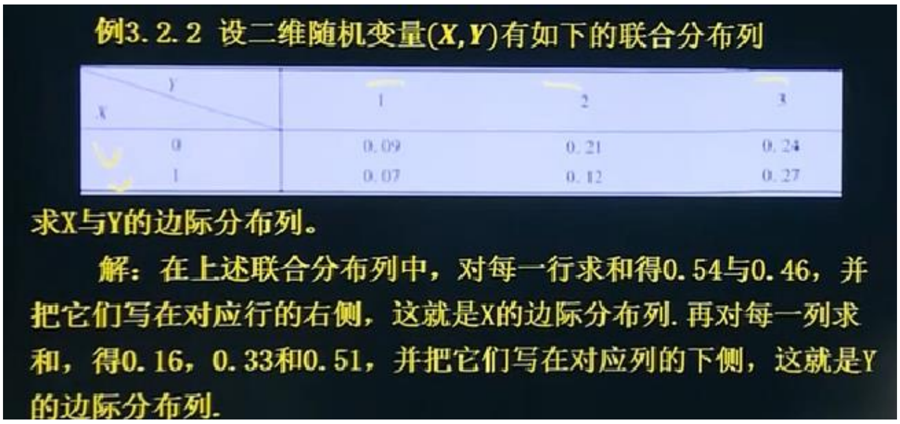
设二维随机变量(X,Y)的联合分布$\{P(X=x_i,Y=y_i)\}$中，对 j 求和所得的分布列

$$\sum_{j=1}^{\infty}P\left(X=x_i,Y=y_j\right)=P(X=x_i)\,,i=1,2,\ldots$$

被称为X的边际分布列，类似的，对i求和所得的分布列

$$\sum_{i=1}^{\infty}P\left(X=x_i,Y=y_j\right)=P(Y=y_j)\,,j=1,2,\ldots$$

被称为Y的边际分布列


例3.2.2 设二维随机变量(X,Y)有如下的联合分布列

X \ Y	1	2	3
0	0.09	0.21	0.24
1	0.07	0.12	0.27

求X与Y的边际分布列。

解：在上述联合分布列中，对每一行求和得0.54与0.46，并把它们写在对应行的右侧，这就是X的边际分布列. 再对每一列求和，得0.16，0.33和0.51，并把它们写在对应列的下侧，这就是Y的边际分布列.

3、设二维随机变量(X,Y)的联合密度函数为$p(x,y)$，因为

$$F_X(x)=F(x,\infty)=\int_{-\infty}^{x}(\int_{-\infty}^{\infty}p(u,v)dv)du=\int_{-\infty}^{x}p_X(u)\,du$$

$$F_Y(y)=F(\infty,y)=\int_{-\infty}^{y}(\int_{-\infty}^{\infty}p(u,v)du)dv=\int_{-\infty}^{y}p_Y(v)\,dv$$

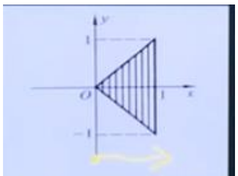
其中，$p_X(x)$和$p_Y(y)$分别为$p_X(x)=\int_{-\infty}^{\infty}p(x,y)dy$,$p_Y(y)=\int_{-\infty}^{\infty}p(x,y)dx$

他们恰好处于密度函数位置，故称上式给出的$p_X(x)$和$p_Y(y)$分别为X和Y的边际密度函数

例题：设二维随机变量(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases}1, 0<x<1,|y|<x\\0, & 其他\end{cases}$$

试求：边际密度函数，$p_X(x)$和$p_Y(y)$

解：首先识别非零区域
图3.2.1 $p(x,y)$的非零区域.

（1） 当$x\le 0$或$x\ge 1$时，有$p_X(x)=0$。 而当$0<x<1$时，有$p_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\int_{-y}^{y}1\ dy=2x$

所以X的边际密度函数为$P_X(x)=\begin{cases}2x, 0<x<1\\ 0,\ 其他\end{cases}$

（2） 当 $y\le -1$或$y\ge 1$时，有$p_Y(y)=0$。而当$-1<y<0$时，有$p_Y(y)=\int_{-\infty}^{\infty}p(x,y)dx=\int_{-y}^{1}1\,dx=1+y$ 。当$0<y<1$有，$p_Y(y)=\int_{-\infty}^{\infty}p(x,y)dx=\int_{y}^{1}1\,dx=1-y$

所以Y的边际密度函数为$P_Y(y)=\begin{cases}1+y, -1<y<0\\ 1-y, 0<y<1\\ 0, 其他\end{cases}$

4、随机变量间的独立性

当两个随机变量的取值互不影响时，就称他们是互相独立的，设 n 维随机变量$(X_1,X_2,\dots X_n)$的联合分布函数为 $F(X_1,X_2,\dots X_n)$，$F_i(x_i)$为X_i的边际分布函数，如果对任意n个实数$(x_1,x_2,\dots x_n)$，有

$F(X_1,X_2,\dots X_n)=\prod_{i=1}^{n}F_i(x_i)$，则称$X_1,X_2,\dots X_n$互相独立

离散型互相独立 $P(X_1=x_1,X_2=x_2,\dots X_n=x_n)=\prod_{i=1}^{n}P(X_i=x_i)$

连续性相互独立 $p(x_1,x_2,\dots x_n)=\prod_{i=1}^{n}p(x_i)$

例题：设二维随机变量(X,Y)的联合密度函数为$p(x,y)=\begin{cases}8xy,0\le x\le y\le 1\\ 0, \quad 其他\end{cases}$ 问X与Y是否相互独立？

解：当$x<0$或$x>1$时，$p_X(x)=0$，而当$0\le x\le 1$时，有$p_X(x)=\int_x^1 8xy\,dy=8x\left(\frac{1}{2}-\frac{x^2}{2}\right)=4x(1-x^2)$, 因此

$$p_X(x)=\begin{cases}4x(1-x^2),0\le x\le 1\\ 0, \quad 其他\end{cases}$$

当$y<0$或$y>1$时，$p_Y(y)=0$，而当$0\le y\le 1$时，有$p_Y(y)=\begin{cases}4y^3,0\le y\le 1\\ 0, 其他\end{cases}$,

由此可见$p(x,y)\neq p(x)p(y)$，故X与Y不独立，即相依

5、随机变量间的独立性

（1）X与Y是独立的其本质是：任对实数a,b,c,d,有$P(a<X<b,c<Y<d)=P(a<X<b)P(c<Y<d)$

（2）X与Y是独立的，则$g(X)$与$h(Y)$也是独立的

（3）(X,Y)服从矩形上的均匀分布，则X与Y独立

（4）(X,Y)服从单位圆上的均匀分布，则X与Y不独立

（5）联合密度$p(x,y)$的表达式中，若x的取值与y的取值有关系，则X与Y不独立

（6）若联合密度$p(x,y)$可分离变量，即$p(x,y)=g(x)h(y)$，则X与Y独立

（7）若(X,Y)服从二元正态$N(\mu_1,\mu_2,{\sigma_1}^2,{\sigma_2}^2,\rho)$则$X$与$Y$独立的充要条件是$\rho=0$

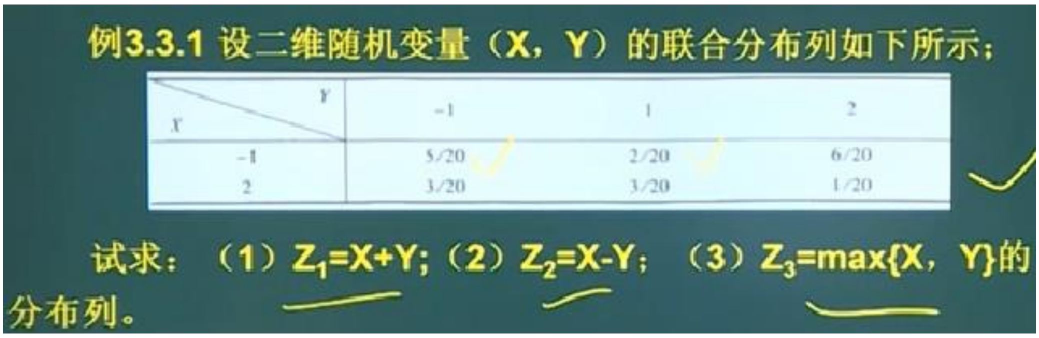
6、多维离散随机变量函数的分布

设$(X_1,X_2,\dots X_n)$是 n 维离散随机变量，则$Z=g(X_1,X_2,\dots X_n)$是一维离散随机变量

多维离散随机变量函数的分布是容易求的：

（1）写出$(X_1,X_2,\dots X_n)$的各种可能取值对，写出Z相应的取值

（2）对Z的相同的取值，合并其对应的概率



例3.3.1 设二维随机变量（X，Y）的联合分布列如下所示：

X \ Y	-1	1	2
-1	5/20	2/20	6/20
2	3/20	3/20	1/20

试求：（1）$Z_1=X+Y$；（2）$Z_2=X-Y$；（3）$Z_3=max\{X，Y\}$的分布列。

（一）最大值最小值分布

设$X_1,X_2,\dots X_n$**独立同分布**，其分布函数和密度函数分别为$F_X(x)$和$p_X(x)$，若记$Y=max(X_1,X_2,\dots X_n)$，$Z=min(X_1,X_2,\dots X_n)$，则

(1) Y的分布函数为：$F_Y(y)=[F_X(y)]^n$

证明：$F_Y(y)=P(Y\le y)=P(max(X_1,X_2,\dots X_n)\le y)=P(X_1\le y,X_2\le y,\dots X_n\le y)=[F_X(y)]^n$

(2) Y的密度函数为：$p_Y(y)=n[F_X(y)]^{n-1}p_X(y)$ 分布函数求导

(3) Z的分布函数为：$F_Z(z)=1-[1-F_X(Z)]^n$

证明：$F_Z(z)=P(Z\le z)=P(min(X_1,X_2,\dots X_n)\le z)=1-P(X_1\ge z,X_2\ge z,\dots X_n\ge z)=P(X_1\ge z)P(X_2\ge z)\dots P(X_n\ge z)=1-[1-F_X(Z)]^n$

(4) Z的密度函数为：$p_Z(z)=n[1-F_X(Z)]^{n-1}\,p_X(z)$

（二）卷积公式

（1）离散场合的卷积公式

设离散随机变量X与Y独立，则$Z=X+Y$的分布列为

$$P(Z=z_i)=\sum_{i=1}^{\infty}P(X=x_i)P(Y=z_i-x_i)=\sum_{j=1}^{\infty}P(X=z_j-y_j)P(Y=y_j)$$

(2) 连续场合的卷积公式

设连续随机变量X与Y独立，则$Z=X+Y$的密度函数为

$$p_z(z)=\int_{-\infty}^{\infty}p_X(x)p_Y(z-x)dx=\int_{-\infty}^{\infty}p_X(z-y)p_Y(y)dy$$

证明：$F_Z(z)=P(X+Y\le z)=\iint_{X+Y\le Z}P_X(x)P_Y(y)\,dx\,dy=\int_{-\infty}^{\infty}\int_{-\infty}^{z-y}P_X(x)dx\;P_Y(y)dy$

令$t=x+y$ $\int_{-\infty}^{\infty}\int_{-\infty}^{z}P_X(t-y)dt\;P_Y(y)dy$ 变换积分限 $\int_{-\infty}^{z}\int_{-\infty}^{\infty}P_X(t-y)\,dy\,P_Y(y)\,dt$

$$p_z(z)=F_Z(z)'=\int_{-\infty}^{\infty}P_X(t-y)\,P_Y(y)\,dy$$

7、分布可加性

若同一类分布的独立随机变量华人他的分布仍是此类分布，则称此类分布具有可加性。

（一）二项分布的可加性

设随机变量$X\sim b(n,p),Y\sim b(m,p)$，且$X$与$Y$相互独立，则$Z=X+Y\sim b(n+m,p)$

（二）泊松分布的可加性

设随机变量$X\sim P(\lambda_1),Y\sim P(\lambda_2)$，且$X$与$Y$相互独立，则$Z=X+Y\sim P(\lambda_1+\lambda_2)$

（三）正态分布的可加性

设随机变量$X\sim N(\mu_1,{\sigma_1}^2),Y\sim P(\mu_2,{\sigma_2}^2)$，且$X$与$Y$相互独立，则$Z=X+Y\sim N(\mu_1+\mu_2,{\sigma_1}^2+{\sigma_2}^2)$

独立正态变量的线性组合仍为正态变量

设$X\sim N(\mu,\sigma^2)$，则对任意非零实数 a 有$aX\sim N(a\mu,a^2\sigma^2)$，则任意 n 个

互相独立的正态变量的线性组合仍是正态变量，即

$$a_1X_1 + a_2X_2 + a_nX_n \sim N(\mu_0, {\sigma_0}^2)$$

若记$X_i \sim N(\mu_i, {\sigma_i}^2), i = 1,2,\dots,n$ 则参数μ_0和${\sigma_0}^2$分别为

$$\mu_0 = \sum_{i=1}^{n} a_i\mu_i \qquad , \qquad {\sigma_0}^2 = \sum_{i=1}^{n} {a_i}^2{\sigma_i}^2$$

（四）伽马分布的可加性

设随机变量$X \sim Ga(\alpha_1, \lambda), Y \sim Ga(\alpha_2, \lambda)$，且$X$与$Y$相互独立，则$Z = X + Y \sim Ga(\alpha_1 + \alpha_2, \lambda)$

注意:（1）$X - Y$不服从$Ga(\alpha_1 - \alpha_2, \lambda)$

（2）m个独立同分布的指数变量之和为伽马变量，即$Exp(\lambda) * Exp(\lambda) * \dots * Exp(\lambda) = Ga(m, \lambda)$

（3）独立的 0-1 分布随机变量之和服从二项分布

（4）m个独立的χ^2变量之和为χ^2变量，即$\chi^2(n_1) * \chi^2(n_2) * \dots * \chi^2(n_m) = \chi^2(n_1 + n_2 + \cdots + n_m)$