概率论与数理统计

第12讲

连续型随机变量函数的分布

一. 二维连续型随机变量函数的分布

设(X,Y)是连续型随机变量，其概率密度函数为$f(x,y)$，$Z=g(X,Y)$是(X,Y)的函数，一般可以分为以下两种情况讨论。

第一，$Z=g(X,Y)$ 为离散型随机变量，此时，只需求Z的分布律，问题的实质是将 Z 取某个值 z 的概率转化为(X,Y)属于某个区域 D 的概率，即有

$$P\{Z=z\}=P\{(X,Y)\in D\}=\iint_D f(x,y)dxdy$$

第二，当Z不是离散型随机变量时，采用分布函数法求$Z=g(X,Y)$的分布函数

$$F_Z(z)=P\{Z\le z\}=P\{g(X,Y)\le z\}=\iint_{g(x,y)\le z} f(x,y)dxdy$$

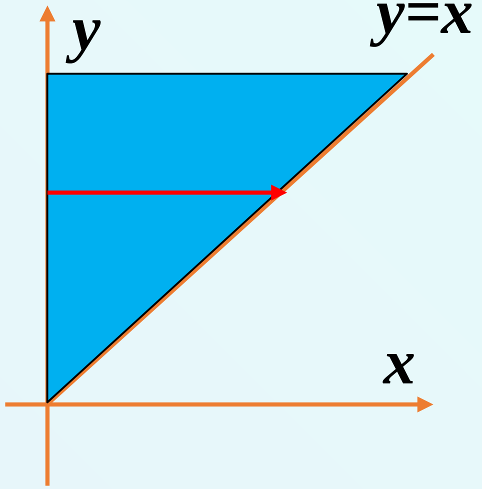
当 Z 为连续型随机变量时，其概率密度函数为 $f_Z(z)=F_Z'(z)$

例1. 设(X, Y)的联合概率密度为 $f(x,y)=\begin{cases}\lambda\mu e^{-(\lambda x+\mu y)}, & x>0, y>0\\ 0, & \text{其他}\end{cases}$

其中λ，μ为大于0的常数. 引入随机变量 $Z=\begin{cases}1, & X\le Y\\ 0, & X>Y\end{cases}$ 求Z的分布。

解：$P\{Z=1\}=P\{X\le Y\}=\iint\limits_{x\le y} f(x,y)dxdy=\int_0^{+\infty}dy\int_0^{y}\lambda\mu e^{-(\lambda x+\mu y)}dx$

$$=\int_0^{+\infty}\mu e^{-\mu y}dy\int_0^{y}\lambda e^{-\lambda x}dx=\int_0^{+\infty}\mu e^{-\mu y}(1-e^{-\lambda y})dy=\frac{\lambda}{\lambda+\mu}$$



一. 二维连续型随机变量函数的分布

$$P\{Z=0\}=1-P\{Z=1\}=1-\frac{\lambda}{\lambda+\mu}=\frac{\mu}{\lambda+\mu}$$

即 Z 的分布律为

Z	0	1
P	$\frac{\mu}{\lambda+\mu}$	$\frac{\lambda}{\lambda+\mu}$

例2. 已知X,Y相互独立，且均服从$N(0,\sigma^2)$ $(\sigma>0)$，求 $Z=\sqrt{X^2+Y^2}$ 的概率密度。

解：易知X和Y的密度函数分布为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}} \qquad f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}$$

又由独立性，得联合概率密度函数为

$$f(x,y)=f_X(x)\cdot f_Y(y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

设Z的分布函数和概率密度分别为$F_Z(z), f_Z(z)$

一. 二维连续型随机变量函数的分布

当 $z<0$ 时 $F_Z(z)=0$;

当 $z\geq 0$ 时

$$F_Z(z)=P\{Z\leq z\}=P\{\sqrt{X^2+Y^2}\leq z\}=\iint\limits_{\sqrt{x^2+y^2}\leq z} f(x,y)dxdy$$

$$=\iint\limits_{\sqrt{x^2+y^2}\leq z}\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}dxdy$$

$$=\iint\limits_{r\leq z}\frac{1}{2\pi\sigma^2}e^{-\frac{r^2}{2\sigma^2}}rdrd\theta=1-e^{-\frac{z^2}{2\sigma^2}}$$

$$\begin{cases} x=r\cos\theta \\ y=r\sin\theta \end{cases}$$

一. 二维连续型随机变量函数的分布

于是 Z 的分布函数为 $F_Z(z)=\begin{cases}1-e^{-\frac{z^2}{2\sigma^2}}, & z\ge 0\\ 0, & \text{其它}\end{cases}$

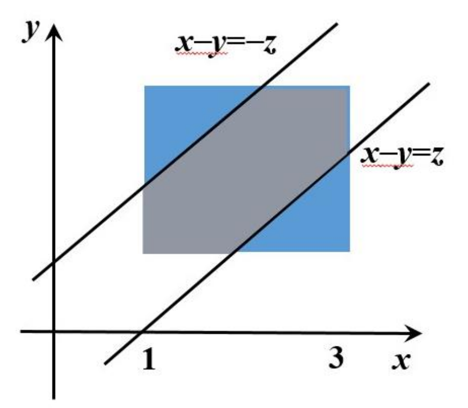
求导数得 Z 的密度函数为 $f_Z(z)=F_Z'(z)=\begin{cases}\dfrac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma^2}}, & z\ge 0\\ 0, & \text{其它}\end{cases}$

此分布称为**瑞利（*Rayleigh*）分布**。

例3.设 (X, Y) 在区域 $G=\{(x, y):1\le x\le 3,1\le y\le 3\}$ 上服从均匀分布，求 $Z=|X-Y|$ 的概率密度。

解：设 Z 的分布函数为 $F_Z(z)$ 。且易知 (X, Y) 的联合密度函数为

$$f(x,y)=\begin{cases}1/4, & 1\le x\le 3,1\le y\le 3\\ 0, & \text{其他}\end{cases}$$



当$z<0$时 $F_Z(z)=P\{Z\le z\}=P\{|X-Y|\le z\}=0$

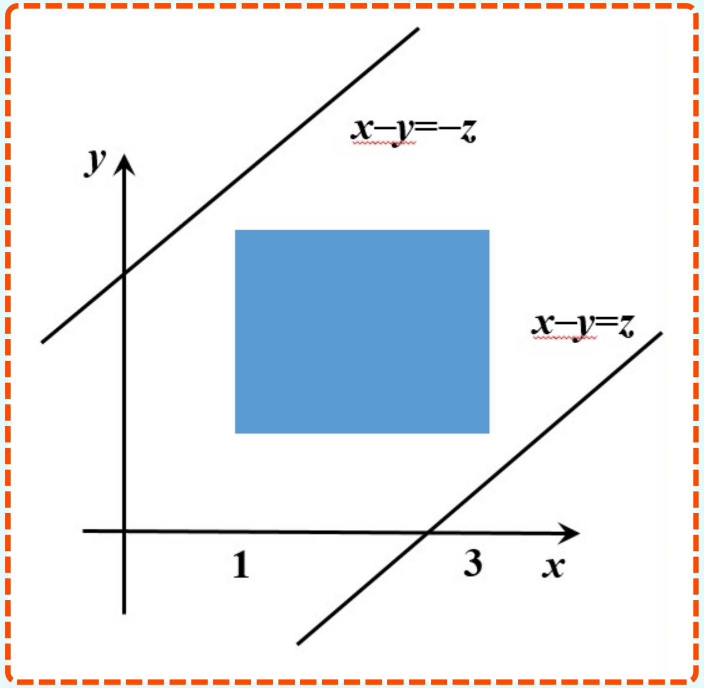
当$0\le z<2$时 $F_Z(z)=P\{Z\le z\}=P\{|X-Y|\le z\}=P\{-z\le X-Y\le z\}$

$$=\iint\limits_{|x-y|\le z} f(x,y)dxdy=\frac{1}{4}[4-(2-z)^2]=z-\frac{1}{4}z^2$$

一. 二维连续型随机变量函数的分布

当$z \geq 2$时

$$F_Z(z)=P\{Z\leq z\}=P\{|X-Y|\leq z\}=\iint\limits_{|x-y|\leq z} f(x,y)dxdy=1$$



故Z的分布函数为： $F_Z(z)=\begin{cases} 0, & z<0 \\ z-\frac{1}{4}z^2, & 0\leq z<2 \\ 1, & z\geq 2 \end{cases}$

则Z的密度函数为： $f_Z(z)=F_Z'(z)=\begin{cases} 1-\frac{z}{2}, & 0\leq z<2 \\ 0, & \text{其它} \end{cases}$

二. 两个随机变量和的分布

定理1 设二维随机变量(X, Y)的联合密度函数为$f(x,y)$，则$Z=X+Y$的概率密度为

$$f_Z(z)=\int_{-\infty}^{\infty} f(x,z-x)dx \quad \text{或} \quad f_Z(z)=\int_{-\infty}^{\infty} f(z-y,y)dy$$

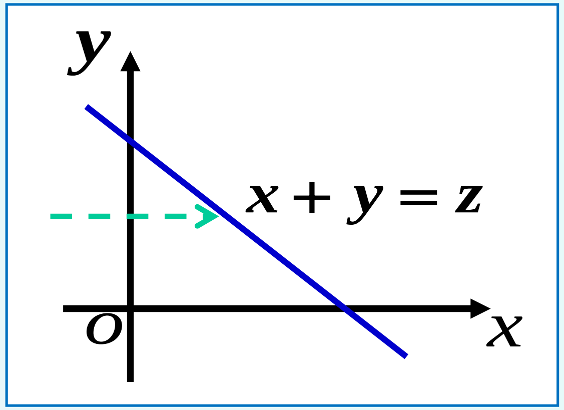
证明： 设Z的分布函数为$F_Z(z)$

$$F_Z(z)=P\{Z\le z\}=P\{X+Y\le z\}=\iint\limits_{x+y\le z} f(x,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z-y} f(x,y)dx\right)dy=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z} f(u-y,y)du\right)dy$$

$$=\int_{-\infty}^{z}\left(\int_{-\infty}^{+\infty} f(u-y,y)dy\right)du$$

令：$u=x+y$



$$D=\begin{cases}-\infty<y<+\infty\\ -\infty<x<z-y\end{cases}$$

二. 两个随机变量和的分布

$$F_Z(z)=\int_{-\infty}^{z}\left(\int_{-\infty}^{+\infty} f(u-y,y)dy\right)du$$

令：$f_Z(u)=\int_{-\infty}^{+\infty} f(u-y,y)dy$ 则有：$F_Z(z)=\int_{-\infty}^{z} f_Z(u)du$

从而得Z的密度函数为 $f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy$

同样的方法可得 $f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx$

二. 两个随机变量和的分布

设$(X,Y)\sim f(x,y)$，则 $Z=X+Y$ 为连续型随机变量，其概率密度为

$$f_Z(z)=\int_{-\infty}^{\infty} f(x,z-x)dx \quad \text{或} \quad f_Z(z)=\int_{-\infty}^{\infty} f(z-y,y)dy$$

又若X和Y独立，(X,Y) 的边缘密度分别为 $f_X(x)$，$f_Y(y)$，则

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx \qquad f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$$

二. 两个随机变量和的分布

例4.设随机变量X与Y相互独立，且都服从N(0,1)，求$Z=X+Y$的概率密度函数。

证明： 由于 X 和 Y 相互独立，所以 $Z=X+Y$ 的概率密度函数为

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(z-x)^2}{2}}dx=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\frac{x^2+(z-x)^2}{2}}dx$$

$$=\frac{1}{2\pi}e^{-\frac{z^2}{4}}\int_{-\infty}^{\infty}e^{-(x-\frac{z}{2})^2}dx=\frac{1}{2\sqrt{\pi}}e^{-\frac{z^2}{4}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\cdot\sqrt{1/2}}e^{-(x-\frac{z}{2})^2\Big/[2\cdot(\sqrt{\frac{1}{2}})^2]}dx$$

$$=\frac{1}{2\sqrt{\pi}}e^{-\frac{z^2}{4}}=\frac{1}{\sqrt{2\pi}\cdot\sqrt{2}}e^{-\frac{z^2}{2\cdot(\sqrt{2})^2}}$$

可见 $Z=X+Y$ 服从正态分布$N(0,2)$。

说明独立的正态分布具有可加性

二. 两个随机变量和的分布

推广： 若X和Y独立，$X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$

则有： $Z = X + Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

进一步，若随机变量 $X_1, X_2, \dots, X_n$ 相互独立，且满足 $X_i \sim N(\mu_i,\sigma_i^2), i=1,2,\cdots,n$

又$a_1, a_2, \dots, a_n$为n个实常数，令 $Z = \sum_{i=1}^{n} a_i X_i$

则有： $Z \sim N(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2)$

有限个独立正态变量的线性组合仍然服从正态分布

例5. 设(X,Y)的概率密度函数为 $f(x,y)=\begin{cases}1, & 0\le x\le 1, 0<y<2x\\ 0, & \text{其它}\end{cases}$

求$Z=X+Y$的概率密度

解: Z的密度函数为 $f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx$

其中 $f(x,z-x)=\begin{cases}1, & 0\le x\le 1, 0<z-x<2x\\ 0, & \text{其它}\end{cases}$

$f(x, z-x)$非0区域为 $\begin{cases}0\le x\le 1\\ 0<z-x<2x\end{cases}$ **即为** $\begin{cases}0\le x\le 1\\ x<z<3x\end{cases}$

二. 两个随机变量和的分布

当$z<0$或$z>3$时，对任意的x，$f(x, z-x)=0$

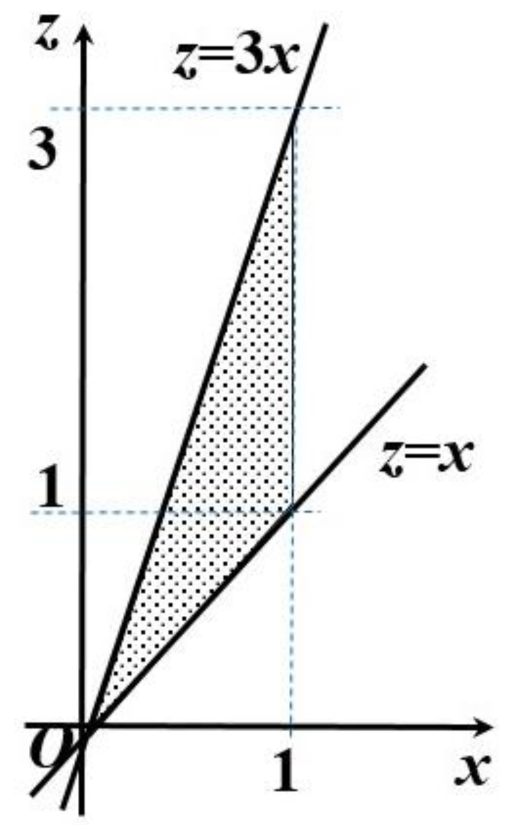
所以 $f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx=0$

当 $0\le z\le 1$ 时

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx=\int_{-\infty}^{z/3} 0dx+\int_{z/3}^{z} 1dx+\int_{z}^{+\infty} 0dx=\frac{2z}{3}$$

当 $1\le z\le 3$ 时

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx=\int_{-\infty}^{z/3} 0dx+\int_{z/3}^{1} 1dx+\int_{1}^{+\infty} 0dx=1-\frac{z}{3}$$



所以Z的密度函数为 $$f_Z(z)=\begin{cases}\dfrac{2z}{3}, & 0\le z\le 1\\ 1-\dfrac{z}{3}, & 1<z\le 3\\ 0, & \text{其他}\end{cases}$$

当联合密度函数 $f(x,y)$ 在某个区域不等于零而在其余的区域为零时，一般来说，$f_Z(z)$ 是一个分段函数，此时要注意 $f_Z(z)$ 的非零区间的确定，也要注意

$\int_{-\infty}^{\infty} f(x,z-x)dx$ 或 $\int_{-\infty}^{\infty} f(z-y,y)dy$ 的积分限的确定。

三. 最大值和最小值的分布

最大值和最小值的分布有广泛的应用，如在建筑高大建筑物时，要考虑若干年内的最大风压；建造桥梁时，要考虑若干年内洪水的最高水位等，这些问题的妥善解决，有着重要的实际意义。

设随机变量 X 和 Y 相互独立，其分布函数分别为$F_X(x)$ 和 $F_Y(y)$ ，现在来求$M=max(X,Y)$ 以及$N=min(X,Y)$的分布函数。

设M的分布函数为$F_M(z)$ 。$F_M(z)=P\{M\le z\} =P\{X\le z, Y\le z\}$

由于X和Y相互独立，于是$M=max(X, Y)$的分布函数为：

$F_M(z)=P\{X\le z\}P\{Y\le z\} =F_X(z)F_Y(z)$

即有：$F_M(z)= F_X (z)F_Y (z)$

$$M \le z \Leftrightarrow \begin{cases} X \le z \\ Y \le z \end{cases}$$

三. 最大值和最小值的分布

类似的，可求得$N=min(X, Y)$的分布函数

$F_N(z)=P\{N\le z\}=1-P\{N>z\}=1-P\{X>z, Y>z\}$

$$N>z \Leftrightarrow \begin{cases} X>z \\ Y>z \end{cases}$$

由于X和Y相互独立，于是$N=min(X,Y)$的分布函数为：

$F_N(z)=1-P\{X>z\}P\{Y>z\}$

即有： $F_N(z)=1-[1-F_X(z)][1-F_Y(z)]$

推广到n个随机变量的情形

设随机变量$X_1,\ldots,X_n$相互独立，其分布函数分别为 $F_{X_i}(x_i)\ (i=1,\ldots,n)$

则$M=max(X_1,\ldots,X_n)$和$N=min(X_1,\ldots,X_n)$的分布函数分别为

$F_M(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z)$ 和 $F_N(z)=1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]$

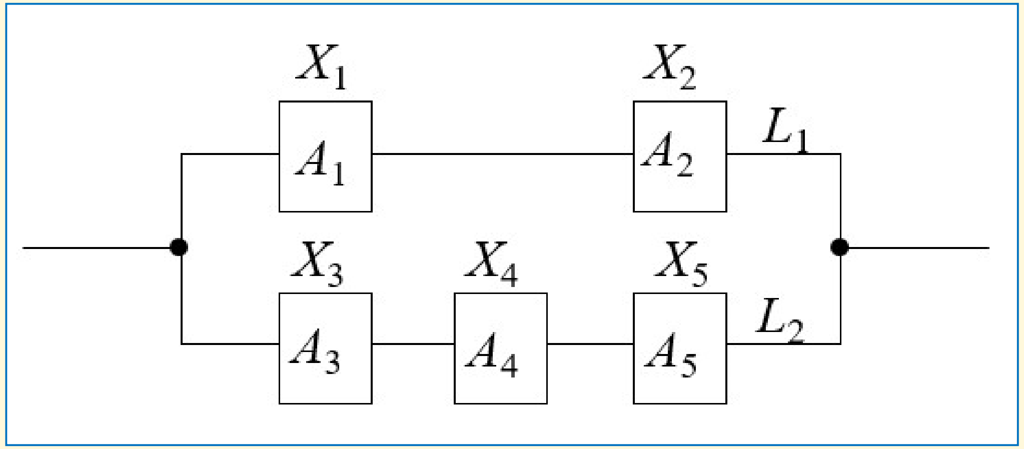
特别地，当$X_1,\ldots,X_n$相互独立且具有相同的分布函数$F(x)$时，有

$F_M(z)=[F(z)]^n \quad F_N(z)=1-[1-F(z)]^n$

例6. 系统L由两个相互独立的子系统L_1和L_2并联而成，而子系统又分别由互相独立的电子元件A_1，A_2，A_3，A_4，A_5按如图的方式串联而成，设每一电子元件的寿命X_i (i=1,2,3,4,5)都服从指数分布$E(0.1)$ 。

求（1）子系统L_1和L_2的寿命分布；

（2）系统L的寿命分布。



解：用Y_1和Y_2表示子系统L_1和L_2的寿命，Y表示系统L的寿命。

易知$Y_1=\min\{X_1, X_2\}$，$Y_2=\min\{X_3, X_4, X_5\}$，$Y=max\{Y_1, Y_2\}$且$Y_1$和$Y_2$相互独立。

易知X_i的分布函数为 $F(x)=\begin{cases}1-e^{-0.1x}, & x>0\\ 0, & x\le 0\end{cases}$

所以Y_1的分布函数为

$$F_{Y_1}(y_1)=1-[1-F(y_1)]^2=\begin{cases}1-[1-(1-e^{-0.1y_1})]^2, & y_1>0\\ 0, & y_1\le 0\end{cases}=\begin{cases}1-e^{-0.2y_1}, & y_1>0\\ 0, & y_1\le 0\end{cases}$$

三. 最大值和最小值的分布

所以Y_1的密度函数为

$$f_{Y_1}(y_1)=F'_{Y_1}(y_1)=\begin{cases}0.2e^{-0.2y_1}, & y_1>0\\ 0, & y_1\le 0\end{cases}$$

同理可得Y_2的分布函数和密度函数为

$$F_{Y_2}(y_2)=\begin{cases}1-e^{-0.3y_2}, & y_2>0\\ 0, & y_2\le 0\end{cases} \quad 和 \quad f_{Y_2}(y_2)=\begin{cases}0.3e^{-0.3y_2}, & y_2>0\\ 0, & y_2\le 0\end{cases}$$

三. 最大值和最小值的分布

所以Y的分布函数为

$$F_Y(y)=F_{Y_1}(y)F_{Y_2}(y)=\begin{cases}(1-e^{-0.2y})(1-e^{-0.3y}), & y>0\\ 0, & y\le 0\end{cases}$$

所以Y的密度函数为

$$f_Y(y)=F_Y'(y)=\begin{cases}0.2e^{-0.2y}+0.3e^{-0.3y}-0.5e^{-0.5y}, & y>0\\ 0, & y\le 0\end{cases}$$

三. 最大值和最小值的分布

以上介绍了X和Y独立时，最大值和最小值的分布的求法. 那么当X和Y不独立时，最大值和最小值的分布怎么求呢?

此时，设（X,Y）的联合密度函数为$f(x,y)$

则最大值$M=max(X,Y)$的分布函数为

$$F_M(z)=P\{M\leq z\} = P\{max(X,Y)\leq z\} =P\{X\leq z, Y\leq z\} = \iint_{x\leq z, y\leq z} f(x,y)dxdy$$

最小值$N=min(X,Y)$的分布函数为

$$F_N(z)=P\{N\leq z\} =P\{min(X,Y)\leq z\} =1-P\{N>z\}=1-P\{X>z, Y>z\} =1-\iint_{x>z, y>z} f(x,y)dxdy$$

这是一个积分问题，具体结果与$f(x,y)$的形式有关。

例7. 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1/2, & 0\le x\le 1, 0<y<2\\ 0, & \text{其它}\end{cases}$$

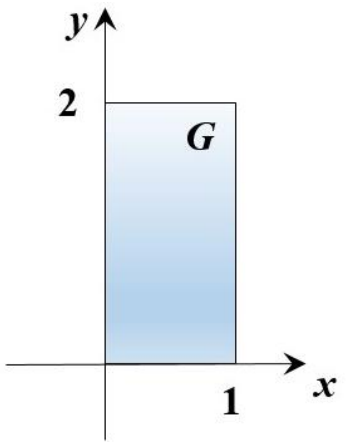
令Z=min{X,Y}. 求Z的密度函数。

解：联合密度函数的非0区域

$$G=\{0\le x\le 1, 0<y<2\}$$

$$F_Z(z)=1-\iint\limits_{x>z,y>z} f(x,y)dxdy$$

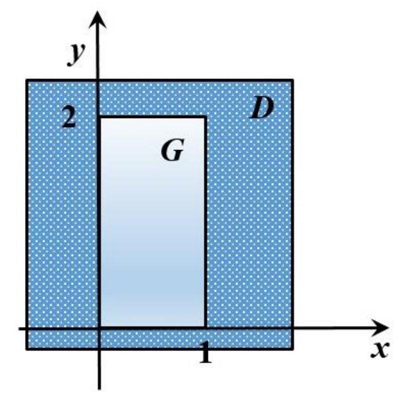
积分区域为 D={(x,y):x>z, y>z}



当$z<0$时

积分区域$D=\{(x,y):x>z, y>z\}$包含了整个非零区域

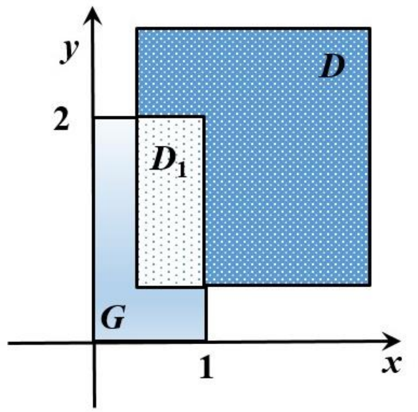
$$F_Z(z)=1-\iint\limits_{x>z,y>z} f(x,y)dxdy=1-1=0$$



当$0\le z<1$时

积分区域$D=\{(x,y):x>z, y>z\}$与非零区域G相交的区域为D_1，可表示为$D_1=\{(x,y): z\le x<1, z\le y<2\}$

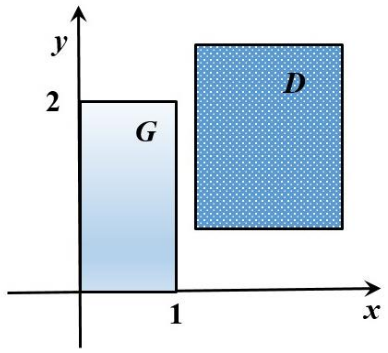
$$F_Z(z)=1-\int_z^1 dx\int_z^2 \frac{1}{2}dy=1-\frac{1}{2}(2-z)(1-z)$$



三. 最大值和最小值的分布

当$z\geq1$时

积分区域 $D=\{(x,y):x>z, y>z\}$ 与非零区域G不相交。



所以 $F_Z(z)=1-0=1$

所以 Z 的分布函数为
$$F_Z(z)=\begin{cases}0, & z<0\\ 1-\dfrac{1}{2}(2-z)(1-z), & 0\leq z<1\\ 1, & z\geq1\end{cases}$$

密度函数为
$$f_Z(z)=F_Z'(z)=\begin{cases}\dfrac{3}{2}-z, & 0\leq z<1\\ 0, & \text{其它}\end{cases}$$

例8.设随机变量 X 和 Y 相互独立，且 $X \sim U(0,1)$，$P\{Y=0\}=P\{Y=1\}=1/2$，令 $Z=X+Y$，求 Z 的分布。

解：首先易知X的密度函数和分布函数分别为

$$f_X(x)=\begin{cases}1, & 0<x<1\\ 0, & \text{其它}\end{cases} \quad 和 \quad F_X(x)=\begin{cases}0, & x\le 0\\ x, & 0<x<1\\ 0, & x\ge 1\end{cases}$$

则Z的分布函数为

$$F_Z(z)=P\{Z\le z\}=P\{X+Y\le z\}=P\Big\{\{X+Y\le z\}\cap\big\{\{Y=0\}\cup\{Y=1\}\big\}\Big\}$$

$$=P\{X+Y\le z, Y=0\}+P\{X+Y\le z, Y=1\}$$

$$=P\{X\le z\}P\{Y=0\}+P\{X\le z-1\}P\{Y=1\}$$

三. 最大值和最小值的分布

$$F_Z(z)=\frac{1}{2}[P\{X\le z\}+P\{X\le z-1\}]=\frac{1}{2}[F_X(z)+F_X(z-1)]$$

所以Z的密度函数为 $f_Z(z)=F_Z'(z)=\frac{1}{2}[f_X(z)+f_X(z-1)]$

当0<z<1时 $f_X(z)=1,\ f_X(z-1)=0\quad \therefore f_Z(z)=\frac{1}{2}[f_X(z)+f_X(z-1)]=\frac{1}{2}$

当1≤z<2时 $f_X(z)=0,\ f_X(z-1)=1\quad \therefore f_Z(z)=\frac{1}{2}[f_X(z)+f_X(z-1)]=\frac{1}{2}$

当z≤0或者z≥2时 $f_X(z)=0,\ f_X(z-1)=0\quad \therefore f_Z(z)=\frac{1}{2}[f_X(z)+f_X(z-1)]=0$

三. 最大值和最小值的分布

综上所述 $f_Z(z)=\begin{cases}\dfrac{1}{2}, & 0<z<2\\ 0, & 其它\end{cases}$

即 $Z \sim U(0,2)$

作业： 35,37,39,41,45,46

第 12 讲

谢谢聆听