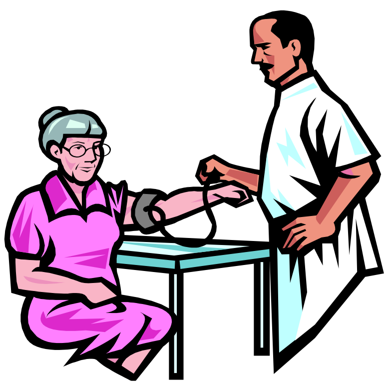
第三章 多维随机变量及其分布

§5 二维随机变量的函数的分布

问题的引入

有一大群人, 令 X 和 Y 分别表示一个人的身高和体重, Z 表示该人的BMI, 并且已知 Z 与 X,Y 的函数关系 $Z=g(X,Y)$,如何通过 X,Y 的分布确定 Z 的分布.

为了解决类似的问题下面我们讨论随机变量函数的分布.



§5 二维随机变量的函数的分布

• 离散型随机变量的函数的分布

例1 设(X,Y)的分布律为

$X \backslash Y$	0	1	2
-1	0.2	0.3	0.1
2	0.1	0.1	0.2

求 (1) $Z=X+Y$ (2) $Z=XY$ 的分布律.

解

(X,Y)	(-1,0)	(-1,1)	(-1,2)	(2,0)	(2,1)	(2,2)
$Z=X+Y$	-1	0	1	2	3	4
$Z=XY$	0	-1	-2	0	2	4
	0.2	0.3	0.1	0.1	0.1	0.2

$Z=XY$	-1	-2	0	2	4
	0.3	0.1	0.3	0.1	0.2

例　若X、Y独立，$P\{X=k\}=a_k$, $k=0,1,2,\ldots$,

$P\{Y=k\}=b_k$, $k=0,1,2,\ldots$,

求$Z=X+Y$的分布律.

解: $$P\{Z=r\}=P\{X+Y=r\}=\sum_{i=0}^{r}P\{X=i,Y=r-i\}$$

$$=\sum_{i=0}^{r}P\{X=i\}P\{Y=r-i\}=\sum_{i=0}^{r}a_i b_{r-i}$$

此即离散卷积公式

$=a_0b_r+a_1b_{r-1}+\ldots+a_rb_0$　　$r=0,1,2,\ldots$

例　若X和Y相互独立,它们分别服从参数为λ_1,λ_2的泊松分布,证明$Z=X+Y$服从参数为$\lambda_1+\lambda_2$的泊松分布.

解：依题意

$$P\{X=k\}=\frac{e^{-\lambda_1}\lambda_1^k}{k!} \qquad P\{Y=k\}=\frac{e^{-\lambda_2}\lambda_2^k}{k!}$$

k=0,1,2,… k=0,1,2,…

由卷积公式

$$P\{Z=r\}=\sum_{i=0}^{r}P\{X=i,Y=r-i\}=\sum_{i=0}^{r}e^{-\lambda_1}\frac{\lambda_1^i}{i!}\cdot e^{-\lambda_2}\frac{\lambda_2^{r-i}}{(r-i)!}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{r!}\sum_{i=0}^{r}\frac{r!}{i!(r-i)!}\lambda_1^i\lambda_2^{r-i}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{r!}(\lambda_1+\lambda_2)^r,$$ r=0,1， …

即Z服从参数为$\lambda_1+\lambda_2$的泊松分布.

- **连续型随机变量的函数的分布**

设(X,Y)的概率密度为$f(x,y)$，求$Z=g(X,Y)$的分布.

一般方法：分布函数法

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

$$f_Z(z)=F_Z'(z)$$

一、 $Z=X+Y$ 的分布

设(X,Y)的概率密度为$f(x,y)$，$Z=X+Y$的分布函数为

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

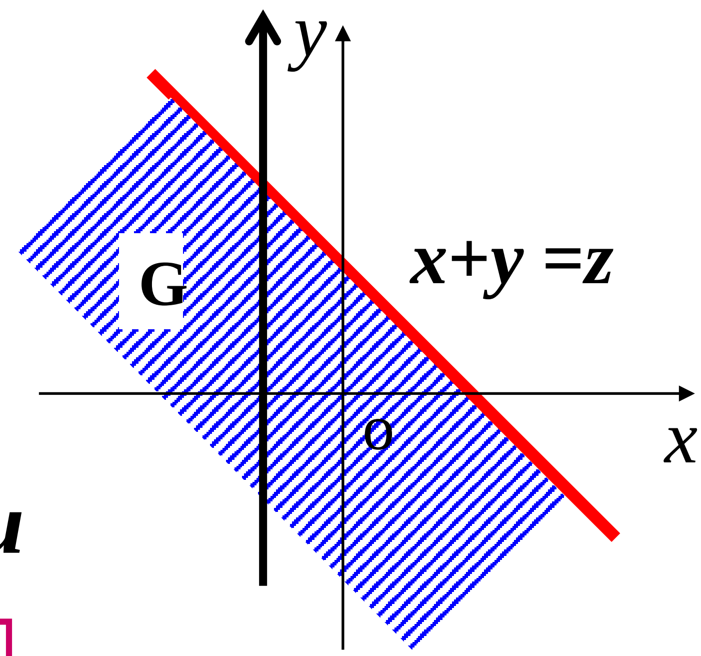
$$=\int_{-\infty}^{+\infty}dx\int_{-\infty}^{z-x} f(x,y)\,dy$$

$\boxed{y=u-x}$

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

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

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



$Z=X+Y$ 的概率密度:

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

$$f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy$$

当X,Y相互独立时, 卷积公式

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

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

$Z=X-Y$ 的概率密度:

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

$$\boxed{f_Z(z)=\int_{-\infty}^{+\infty} f(z+y,y)dy}$$

当X,Y相互独立时,

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

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

互相关

例1 设 $X\sim N(0,1)$, $Y\sim N(0,1)$且X与Y相互独立，求 $Z=X+Y$的概率密度。

解 $f_X(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$，$f_Y(y)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}},(-\infty<x,y<+\infty)$

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

$$\boxed{t=x-\frac{z}{2}}$$

$$=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\frac{1}{\sqrt{2\pi}}e^{\frac{-(z-x)^2}{2}}dx=\frac{1}{2\pi}e^{-\frac{z^2}{4}}\int_{-\infty}^{+\infty}e^{-(x-\frac{z}{2})^2}dx$$

$\Longrightarrow$ $f_Z(z)=\dfrac{1}{2\pi}e^{-\frac{z^2}{4}}\displaystyle\int_{-\infty}^{+\infty}e^{-t^2}dt=\dfrac{1}{\sqrt{2}\sqrt{2\pi}}e^{-\frac{z^2}{4}}\quad(-\infty<z<+\infty)$

$\Longrightarrow$ **$Z=X+Y\sim N(0,2)$.**

➢ 一般结论：

正态分布的可加性

(1) 若 $X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$ 且相互独立，则 $X+Y$ 仍服从正态分布，且

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

(2) 若 $X_i \sim N(\mu_i,\sigma_i^2), (i=1,2\cdots,n)$ 且相互独立，则

$$\sum_{i=1}^{n} X_i \sim N(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2)$$

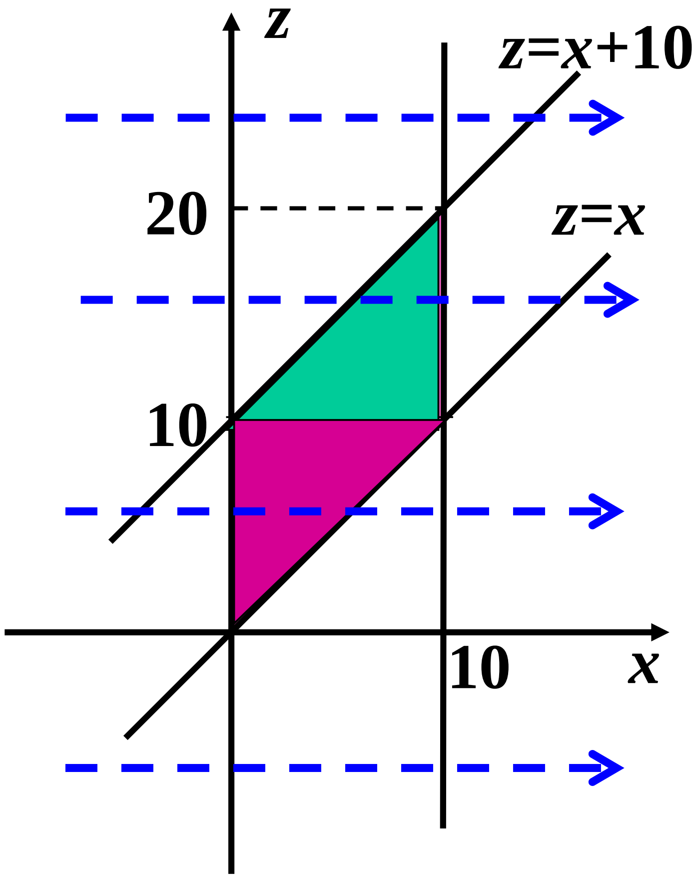
(3)有限个相互独立的正态随机变量的线性组合仍然服从正态分布

$$\sum_{i=1}^{n} c_i X_i \sim N(\sum_{i=1}^{n} c_i\mu_i, \sum_{i=1}^{n} c_i^2\sigma_i^2)$$

例2 在一简单电路中，两电阻R_1和R_2串联联接，设R_1, R_2相互独立，它们的概率密度均为

$$f(x)=\begin{cases}\dfrac{10-x}{50}, & 0\le x\le 10,\\ 0, & 其它.\end{cases}$$

求总电阻$R=R_1+R_2$的概率密度.



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

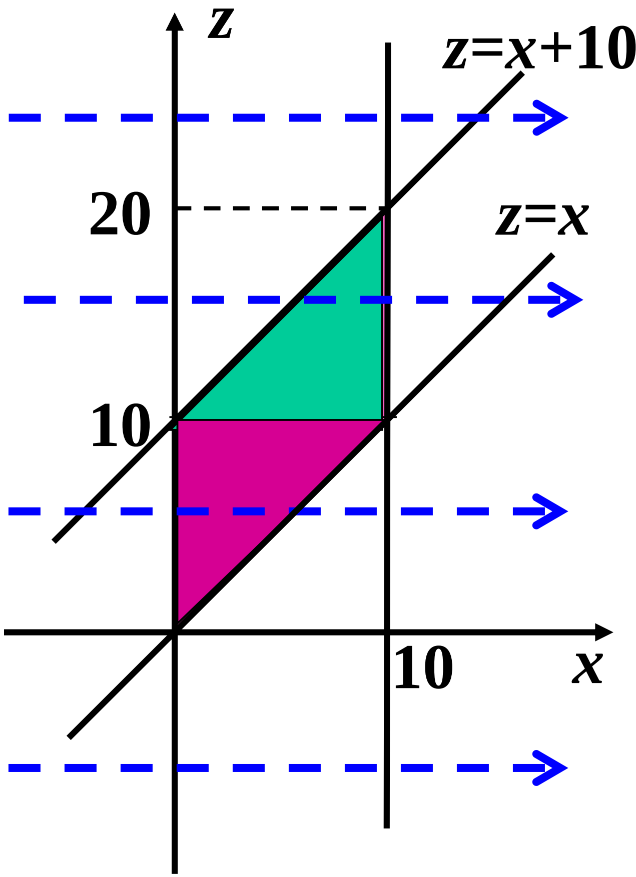
$$D:\begin{cases}0\le x\le 10,\\ 0\le z-x\le 10.\end{cases}$$

即 $D:\begin{cases}0\le x\le 10,\\ x\le z\le x+10.\end{cases}$

例2 在一简单电路中，两电阻R_1和R_2串联联接，设R_1, R_2相互独立，它们的概率密度均为

$$f(x)=\begin{cases}\dfrac{10-x}{50}, & 0\le x\le 10,\\ 0, & \text{其它.}\end{cases}$$

求总电阻$R=R_1+R_2$的概率密度.



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

$$f_Z(z)=\begin{cases}\displaystyle\int_0^z\frac{10-x}{50}\,\frac{10-(z-x)}{50}dx, & 0\le z\le 10,\\ \displaystyle\int_{z-10}^{10}\frac{10-x}{50}\,\frac{10-(z-x)}{50}dx, & 10<z\le 20,\\ 0, & \text{其它.}\end{cases}$$

◆ **Γ分布：** 若随机变量X的概率密度为

$$f(x)=\begin{cases}\dfrac{\beta}{\Gamma(\alpha)}(\beta x)^{\alpha-1}e^{-\beta x}, & x>0,\\ 0, & \text{其它.}\end{cases}\quad \alpha>0,\beta>0$$

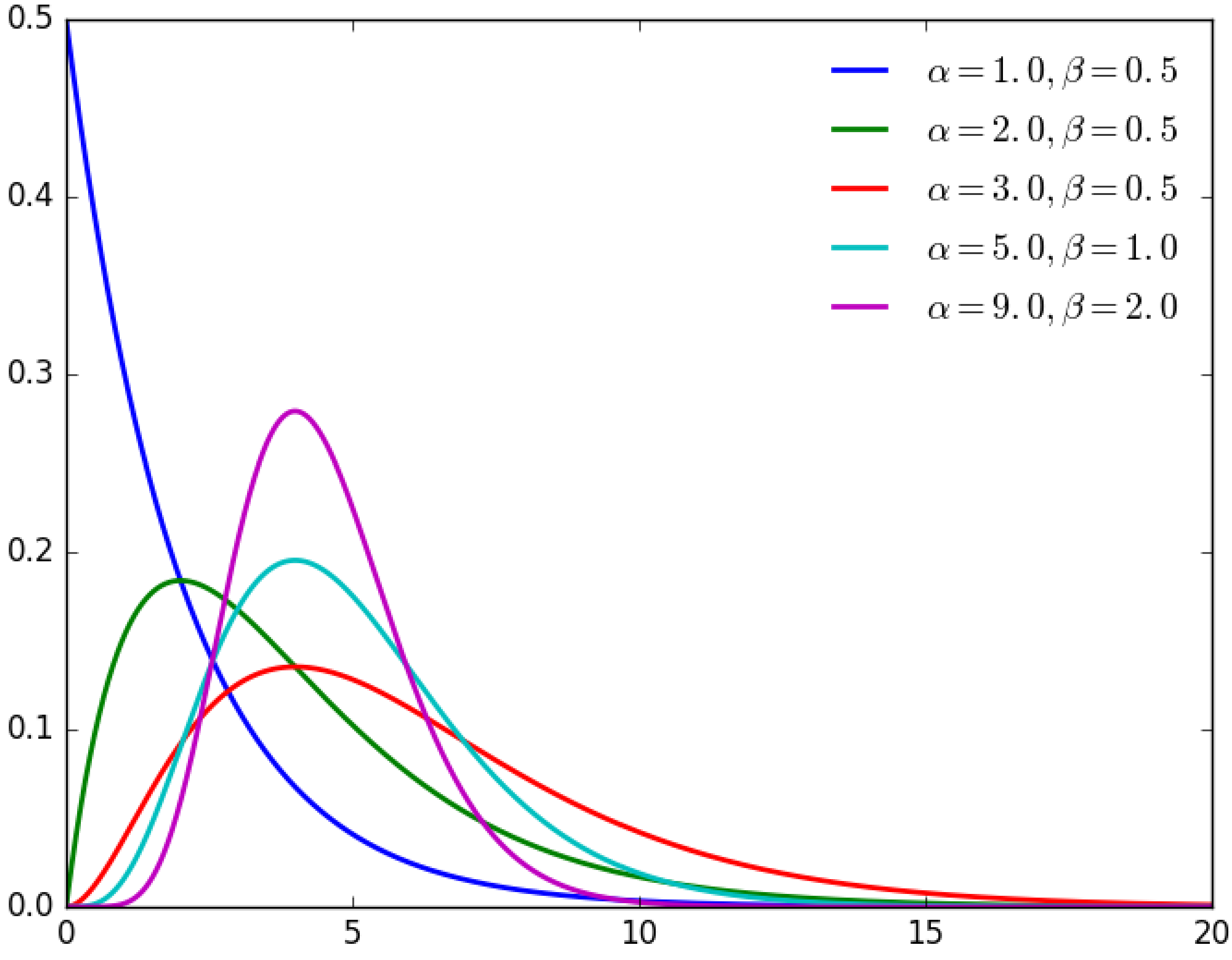
则称X服从参数为α, β的**Γ分布**.记为 $X\sim\Gamma(\alpha, \beta)$.

Γ分布的性质： 若$X_1\sim\Gamma(\alpha_1, \beta)$，$X_2\sim\Gamma(\alpha_2, \beta)$，且相互独立，则 $X_1+X_2\sim\Gamma(\alpha_1+\alpha_2, \beta)$.

一般结论： 若$X_1,X_2,\ldots X_n$相互独立，且X_i服从参数为$\alpha_i, \beta\ (i=1,2,\ldots n)$的Γ的分布，则$X_1+X_2+\ldots+X_n$服从参数为$\alpha_1+\alpha_2+\ldots+\alpha_n, \beta$的Γ分布.

［注］ Γ函数: $\Gamma(x)=\int_0^{+\infty}t^{x-1}e^{-t}dt.\ (x>0)$

Γ分布的概率密度函数



0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
$\alpha=1.0, \beta=0.5$
$\alpha=2.0, \beta=0.5$
$\alpha=3.0, \beta=0.5$
$\alpha=5.0, \beta=1.0$
$\alpha=9.0, \beta=2.0$

例3 设X_1, X_2相互独立分别服从参数为$\alpha_1, \beta; \alpha_2, \beta$ 的Γ分布, 即X_1, X_2的概率密度分别为

$$f_{X_1}(x)=\begin{cases}\dfrac{\beta}{\Gamma(\alpha_1)}(\beta x)^{\alpha_1-1}e^{-\beta x}, & x>0,\\ 0, & \text{其它}.\end{cases}\quad \alpha_1>0, \beta>0$$

$$f_{X_2}(y)=\begin{cases}\dfrac{\beta}{\Gamma(\alpha_2)}(\beta y)^{\alpha_2-1}e^{-\beta y}, & y>0,\\ 0, & \text{其它}.\end{cases}\quad \alpha_2>0, \beta>0,$$

试证：X_1+X_2服从参数为 $\alpha_1+\alpha_2, \beta$ 的Γ分布.

[注] Γ函数: $\Gamma(x)=\int_0^{+\infty} t^{x-1}e^{-t}dt. \ (x>0)$

证: $z<0,\ f_Z(z)=0.$ 当 $z>0$ 时,

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

$$=\int_0^z \frac{\beta}{\Gamma(\alpha_1)}(\beta x)^{\alpha_1-1}e^{-\beta x}\ \frac{\beta}{\Gamma(\alpha_2)}[\beta(z-x)]^{\alpha_2-1}e^{-\beta(z-x)}dx$$

$$=\frac{\beta^{\alpha_1+\alpha_2}e^{-\beta z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^z x^{\alpha_1-1}(z-x)^{\alpha_2-1}dx,\ \text{令}x=zt,$$

$$=(\beta z)^{\alpha_1+\alpha_2-1}e^{-\beta z}\ \boxed{\frac{\beta}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^1 t^{\alpha_1-1}(1-t)^{\alpha_2-1}dt} = A$$

$$=(\beta z)^{\alpha_1+\alpha_2-1}e^{-\beta z}\ \frac{\beta}{\Gamma(\alpha_1+\alpha_2)}$$

亦即$Z=X_1+X_2$服从参数为$\alpha_1+\alpha_2$, β的Γ分布.

A的计算： $A = \dfrac{\beta}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1} dt$

[注] Γ函数： $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt. \ (x > 0)$

$$1 = \int_{-\infty}^{+\infty} f_Z(z) dz$$

$$= \frac{A}{\beta} \int_0^{+\infty} (\beta z)^{\alpha_1 + \alpha_2 - 1} e^{-\beta z} d(\beta z)$$

$$= \frac{A}{\beta} \Gamma(\alpha_1 + \alpha_2) \longrightarrow A = \frac{\beta}{\Gamma(\alpha_1 + \alpha_2)}$$

一般结论：若$X_1, X_2, \ldots X_n$相互独立，且X_i服从参数为$\alpha_i, \beta\,(i=1,2,\ldots n)$的Γ的分布，则$X_1+X_2+\ldots+X_n$服从参数为$\alpha_1+\alpha_2+\ldots+\alpha_n, \beta$的Γ分布.

二、 Z=Y/X 的分布、 Z=XY的分布

设X,Y是二维连续型随机变量, 其概率密度为$f(x,y)$, 则$Z=Y/X$ 、$Z=XY$仍为连续型随机变量，其概率密度分别为

$$f_{Y/X}(z)=\int_{-\infty}^{+\infty}|x|\,f(x,zx)dx$$

$$f_{XY}(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f(x,\frac{z}{x})dx$$

当X,Y相互独立时,

$$f_{Y/X}(z)=\int_{-\infty}^{+\infty}|x|\,f_X(x)f_Y(zx)dx$$

$$f_{XY}(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f_X(x)f_Y(\frac{z}{x})dx$$

证: $F_{Y/X}(z)=P\{Y/X\le z\}=\iint\limits_{y/x\le z} f(x,y)dxdy$

$$=\iint\limits_{G_1} f(x,y)dxdy+\iint\limits_{G_2} f(x,y)dxdy$$

$$=\int_{-\infty}^{0}[\int_{zx}^{+\infty} f(x,y)dy]dx+\int_{0}^{+\infty}[\int_{-\infty}^{zx} f(x,y)dy]dx$$

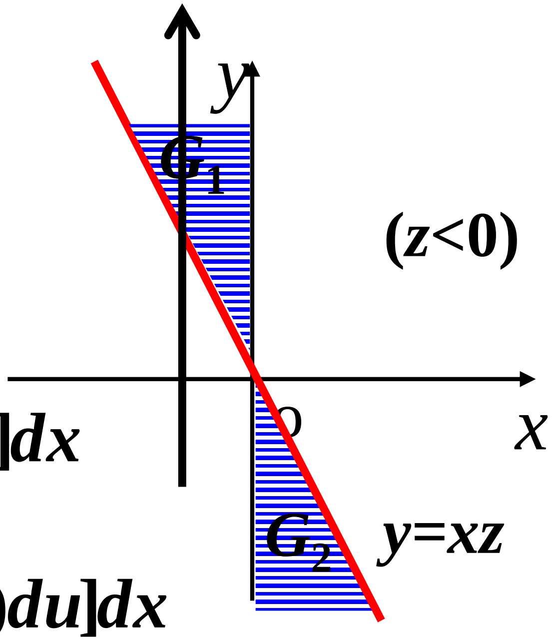
令$y=xu$

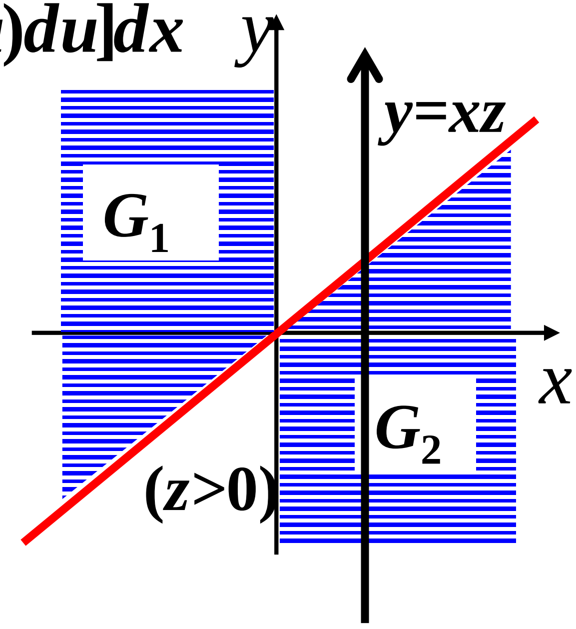
$$=\int_{-\infty}^{0}[\int_{z}^{-\infty} xf(x,xu)du]dx+\int_{0}^{+\infty}[\int_{-\infty}^{z} xf(x,xu)du]dx$$

$$=\int_{-\infty}^{0}[\int_{-\infty}^{z} -xf(x,xu)du]dx+\int_{0}^{+\infty}[\int_{-\infty}^{z} xf(x,xu)du]dx$$

$$=\int_{-\infty}^{+\infty}[\int_{-\infty}^{z} |x|\, f(x,xu)du]dx$$

$$=\int_{-\infty}^{z}[\boxed{\int_{-\infty}^{+\infty} |x|\, f(x,xu)dx}]du$$





例4 设X和Y分别表示两个不同电子元件的寿命，且相互独立, 服从同一分布，其概率密度为

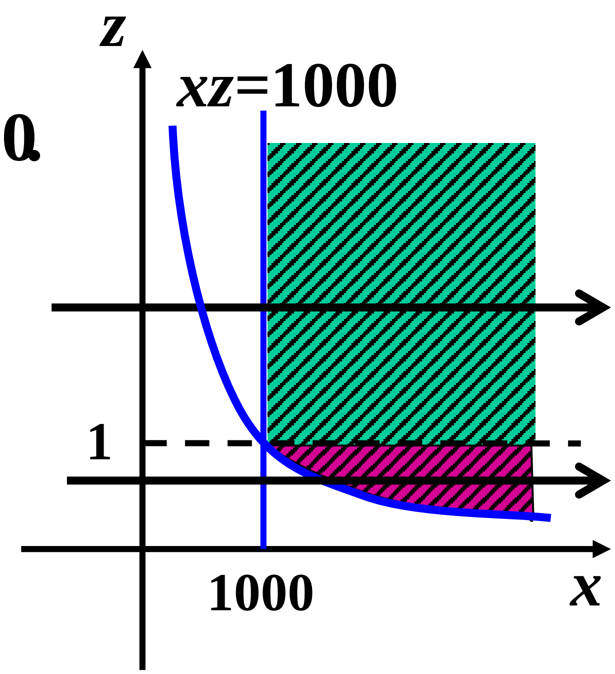
$$f(x)=\begin{cases}\dfrac{1000}{x^2}, & x>1000,\\ 0, & \text{其它}.\end{cases}$$

求$Z=Y/X$的概率密度.

解 $f_{Y/X}(z)=\int_{-\infty}^{+\infty}|x|f_X(x)f_Y(zx)dx$

被积函数的非零区域 $D: x>1000, zx>1000.$

$$=\begin{cases}0, & z\le 0\\ \int_{1000/z}^{+\infty}|x|\dfrac{1000}{x^2}\cdot\dfrac{1000}{(xz)^2}dx=\dfrac{1}{2}, & 0<z<1,\\ \int_{1000}^{+\infty}|x|\dfrac{1000}{x^2}\cdot\dfrac{1000}{(xz)^2}=\dfrac{1}{2z^2} & z\ge 1.\end{cases}$$



三、最大值、最小值的分布

设X,Y是两个相互独立的随机变量，它们的分布函数分别为$F_X(x),F_Y(y)$. 求 $M=\max\{X,Y\}$及 $N=\min\{X,Y\}$的分布函数.

对任意实数z,

$$F_{\max}(z)=P\{\max\{X,Y\}\le z\}=P\{X\le z,Y\le z\}$$

$$=P\{X\le z\}P\{Y\le z\}$$

$$\boxed{F_{\max}(z)=F_X(z)\cdot F_Y(z)}$$

$$F_{\min}(z)=P\{\min\{X,Y\}\le z\}=1-P\{\min\{X,Y\}\le z\}$$

$$=1-P\{X>z,Y>z\}=1-P\{X>z\}P\{Y>z\}$$

$$\boxed{F_{\min}(z)=1-[1-F_X(z)]\cdot[1-F_Y(z)]}$$

◆推广:

设$X_1,\ldots,X_n$相互独立,其分布函数分别为 $F_{X_i}(x_i)$,则 $M=\max\{X_1,X_2,\ldots,X_n\}$ 的分布函数为

$$F_{\max}(z)=\prod_{i=1}^{n}F_{X_i}(z)$$

$N=\min\{X_1,X_2,\ldots,X_n\}$的分布函数为

$$F_{\min}(z)=1-\prod_{i=1}^{n}[1-F_{X_i}(z)]$$

特别, 当$X_1,\ldots,X_n$相互独立且有相同分布函数$F(x)$时, 有

$$F_{\max}(z)=[F(z)]^n,$$

$$F_{\min}(z)=1-[1-F(z)]^n$$

例5 设系统L由两个相互独立的子系统L_1、L_2组成,其寿命分别为X, Y. 其概率密度分别为

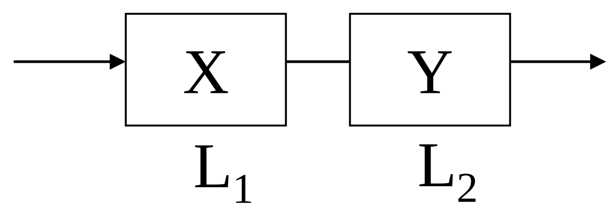
$$f_X(x)=\begin{cases}\alpha e^{-\alpha x}, & x>0,\\ 0, & x\le 0.\end{cases} \qquad f_Y(y)=\begin{cases}\beta e^{-\beta y}, & y>0,\\ 0, & y\le 0.\end{cases}$$

其中α>0,β>0,α≠β. 试求联接方式为： (1) 串联，(2) 并联,(3) 备用时系统L的寿命Z的概率密度.

解

$$F_X(x)=\begin{cases}1-e^{-\alpha x}, x>0,\\ 0, \quad x\le 0.\end{cases} \qquad F_Y(y)=\begin{cases}1-e^{-\beta y}, y>0,\\ 0, \quad y\le 0.\end{cases}$$

(1)串联系统：此时有 Z=min{X,Y}

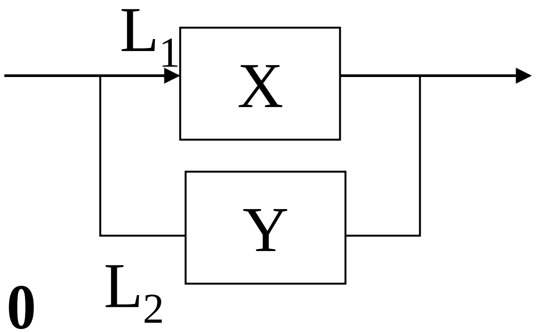


$$F_Z(z)=\begin{cases}1-e^{-(\alpha+\beta)z}, & z>0,\\ 0, & z\le 0.\end{cases}$$

$$F_{\min}(z)=1-[1-F_X(z)]\cdot[1-F_Y(z)]$$

$$f_Z(z)=F_Z'(z)=\begin{cases}(\alpha+\beta)e^{-(\alpha+\beta)z}, & z>0,\\ 0, & z<0.\end{cases}$$

(2)并联系统：

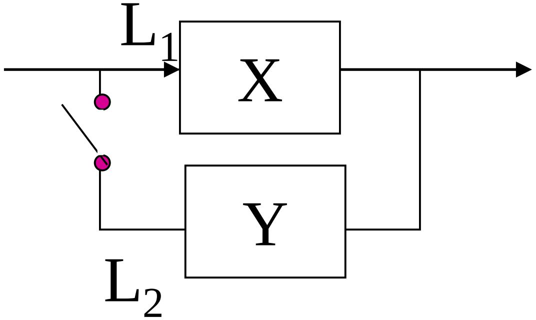


此时有 $Z=\max\{X, Y\}$

$$F_Z(z)=F_X(z)\cdot F_Y(z)=\begin{cases}(1-e^{-\alpha z})(1-e^{-\beta z}), & z>0\\ 0, & z\le 0\end{cases}$$

$$f_Z(z)=\begin{cases}\alpha e^{-\alpha z}+\beta e^{-\beta z}-(\alpha+\beta)e^{-(\alpha+\beta)z}, & z>0\\ 0, & z\le 0\end{cases}$$

(3)备用系统：　此时有 $Z=X+Y$



$$f(z)=\begin{cases}\dfrac{\alpha\beta}{\beta-\alpha}[\mathrm{e}^{-\alpha z}-\mathrm{e}^{-\beta z}], & z>0,\\ 0, & z\le 0.\end{cases}$$

注：若α=β，备用系统是Γ分布.

小 结

1. 离散型随机变量函数的分布律

若二维离散型随机变量 的联合分布律为

$$P\{X = x_i, Y = y_j\} = p_{ij}, i, j = 1,2,\cdots$$

则随机变量函数$Z = g(X,Y)$的分布律为

$$P\{Z = z_k\} = P\{g(X,Y) = z_k\}$$

$$= \sum_{z_k = g(x_i y_j)} p_{ij} \qquad k = 1,2,\cdots.$$

2. 连续型随机变量函数的分布

(1) $Z = X + Y$ 的分布

(2) $Z = \dfrac{Y}{X}$ 的分布

(3) $M = \max(X, Y)$ 及 $N = \min(X, Y)$ 的分布

作业

- **第三章习题20, 23, 24, 26, 29, 32**