概率统计

第三节 随机变量的函数及其分布(2)

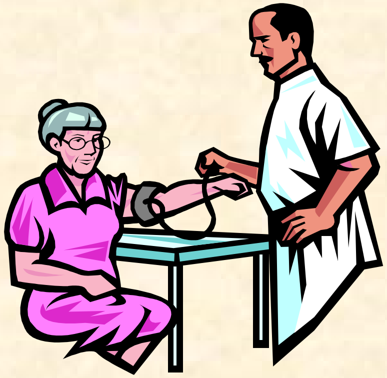
二维随机变量的函数的分布

- 一、问题的引出
- 二、离散型随机变量的函数的分布
- 三、连续型随机变量的函数的分布
- 四、内容小结

一、问题的引入

有一大群人,令 X 和 Y 分别表示一个人的年龄和体重, Z 表示该人的血压 ,并且已知 Z 与 X,Y 的函数关系 $Z=f(X,Y)$, 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.



二、离散型随机变量函数的分布

例1 设随机变量 (X,Y) 的分布律为

$X \backslash Y$	-2	-1	0
-1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{12}$
$\frac{1}{2}$	$\frac{2}{12}$	$\frac{1}{12}$	0
3	$\frac{2}{12}$	0	$\frac{2}{12}$

求 $(1) X+Y$, $(2) \left|X-Y\right|$ 的分布律.

解

$X \backslash Y$	-2	-1	0
-1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{12}$
$\frac{1}{2}$	$\frac{2}{12}$	$\frac{1}{12}$	0
3	$\frac{2}{12}$	0	$\frac{2}{12}$

等价于

概率	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$
(X,Y)	$(-1,-2)$	$(-1,-1)$	$(-1,0)$	$\left(\frac{1}{2},-2\right)$	$\left(\frac{1}{2},-1\right)$	$(3,-2)$	$(3,0)$

概率	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$
(X,Y)	$(-1,-2)$	$(-1,-1)$	$(-1,0)$	$\left(\frac{1}{2},-2\right)$	$\left(\frac{1}{2},-1\right)$	$(3,-2)$	$(3,0)$
$X+Y$	-3	-2	-1	$-\frac{3}{2}$	$-\frac{1}{2}$	1	3
$\lvert X-Y\rvert$	1	0	1	$\frac{5}{2}$	$\frac{3}{2}$	5	3

目录 上页 下页 返回 结束

所以 $X+Y, |X-Y|$ 的分布律分别为

$X+Y$	-3	-2	-1	$-\frac{3}{2}$	$-\frac{1}{2}$	1	3
P	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$

$\lvert X-Y \rvert$	0	1	$\frac{5}{2}$	$\frac{3}{2}$	5	3
P	$\frac{1}{12}$	$\frac{4}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$

结论

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 $Z=f(X,Y)$ 的分布律为

$$P\{Z=z_k\}=P\{f(X,Y)=z_k\}$$
$$=\sum_{z_k=f(x_i,y_j)} p_{ij} \qquad k=1,2,\cdots.$$

其中“$\sum\limits_{z_k=f(x_i,y_j)}$”是关于所有 $f(x_i,y_j)=z_k$ 的 (x_i,y_j) 求和.

例2 设两个独立的随机变量X 与Y 的分布律为

X	1	3
P_X	0.3	0.7

Y	2	4
P_Y	0.6	0.4

求随机变量 $Z=X+Y$ 的分布律.

解 因为 X 与 Y 相互独立, 所以

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\}P\{Y = y_j\},$$

得

$X \backslash Y$	2	4
1	0.18	0.12
3	0.42	0.28

$X \backslash Y$	2	4
1	0.18	0.12
3	0.42	0.28

可得

P	(X,Y)	$Z=X+Y$
0.18	(1,2)	3
0.12	(1,4)	5
0.42	(3,2)	5
0.28	(3,4)	7

所以

$Z=X+Y$	3	5	7
P	0.18	0.54	0.28

例3 设相互独立的两个随机变量 X, Y 具有同一分布律,且 X 的分布律为

X	0	1
P	0.5	0.5

试求 : $Z = \max(X, Y)$ 的分布律.

解 因为X与Y相互独立,

所以 $P\{X = i, Y = j\} = P\{X = i\}P\{Y = j\},$

于是

$X \backslash Y$	0	1
0	$1/2^2$	$1/2^2$
1	$1/2^2$	$1/2^2$

$$P\{\max(X,Y)=i\}$$
$$=P\{X=i, Y<i\}$$
$$+P\{X\le i, Y=i\}$$

$X \backslash Y$	0	1
0	$1/2^2$	$1/2^2$
1	$1/2^2$	$1/2^2$

$$\Rightarrow P\{\max(X,Y)=0\}=P\{0,0\}=\frac{1}{2^2},$$

$$P\{\max(X,Y)=1\}=P\{1,0\}+P\{0,1\}+P\{1,1\}$$
$$=\frac{1}{2^2}+\frac{1}{2^2}+\frac{1}{2^2}=\frac{3}{2^2}.$$

故$Z=\max(X,Y)$的分布律为

Z	0	1
P	$\frac{1}{4}$	$\frac{3}{4}$

三、连续型随机变量函数的分布

几种特殊形式的随机变量函数的分布

(1) $Z = X + Y$的分布

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

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

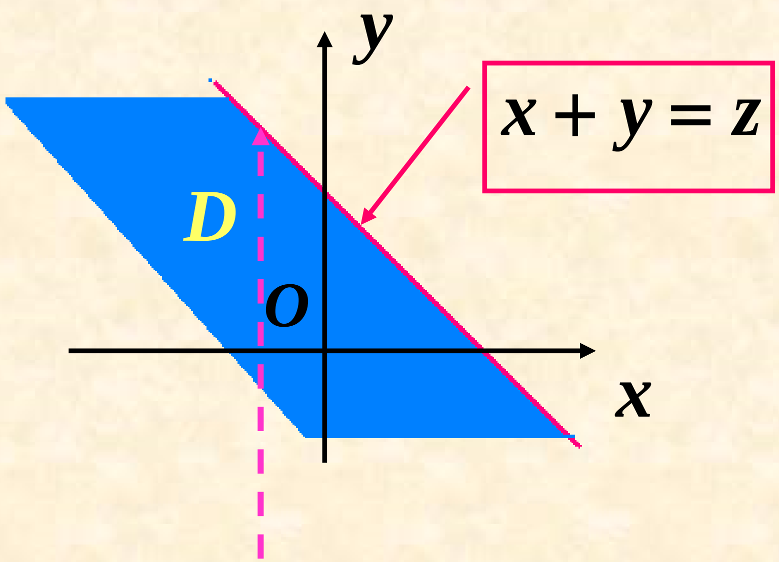
当X与Y独立时,

$$p_Z(z) = \int_{-\infty}^{+\infty} p_X(z-y) p_Y(y)\, dy$$

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

证 $\forall z \in R$

$$D = \{(x, y) \mid x + y \le z\}$$



$$F_Z(z) = P\{Z \le z\}$$

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

$$= \iint_D p(x, y)\,dx\,dy$$

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

$$\overline{\overline{\text{令} u = x + y}} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z} p(x, u - x)\,du$$

$$F_Z(z)=P\{Z\le z\}$$

积分时 z 视为常数

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

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

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

$$\therefore\quad p_Z(z)=\frac{dF_Z(z)}{dz}$$

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

例5 设两个独立的随机变量 X 与 Y 都服从标准正态分布,求 $Z=X+Y$ 的概率密度.

解 由于 $p_X(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad -\infty<x<+\infty,$

$$p_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}, \quad -\infty<y<+\infty,$$

由公式 $\quad p_Z(z)=\int_{-\infty}^{+\infty} p_X(x)p_Y(z-x)\,\mathrm{d}x.$

得 $$p_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{2\pi}e^{-\frac{x^2}{2}}e^{-\frac{(z-x)^2}{2}}\,\mathrm{d}x$$

$$=\frac{1}{2\pi}e^{-\frac{z^2}{4}}\int_{-\infty}^{+\infty}e^{-\left(x-\frac{z}{2}\right)^2}\,\mathrm{d}x$$

$$\overset{t=x-\frac{z}{2}}{=\!=\!=\!=}\ \frac{1}{2\pi}e^{-\frac{z^2}{4}}\int_{-\infty}^{+\infty}e^{-t^2}\,\mathrm{d}t=\frac{1}{2\sqrt{\pi}}e^{-\frac{z^2}{4}}.$$

即 Z 服从 $N(0,2)$ 分布.

说明

一般,设X,Y相互独立且$X \sim N(\mu_1,\sigma_1^2)$,$Y \sim N(\mu_2,\sigma_2^2)$.则$Z=X+Y$仍然服从正态分布,且有$Z \sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

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.

例如，设X、Y独立，都具有正态分布，则$3X+4Y+1$也具有正态分布.

(2) $Z=X-Y$的分布

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

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

当X与Y独立时,

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(z+y)p_Y(y)\,dy$$

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

(3) $Z=\dfrac{X}{Y}$的分布

$$p_Z(z)=\int_{-\infty}^{+\infty}|y|\,p(yz,y)\,dy$$

当X与Y独立时,

$$p_Z(z)=\int_{-\infty}^{+\infty}|y|\,p_X(yz)\,p_Y(y)\,dy$$

证 $\forall\, z\in R,\ D=\{(x,y)\,|\,\dfrac{x}{y}\le z\}$

$$F_Z(z)=P\{Z\le z\}=P\{\frac{X}{Y}\le z\}=\iint_D p(x,y)\,dx\,dy$$

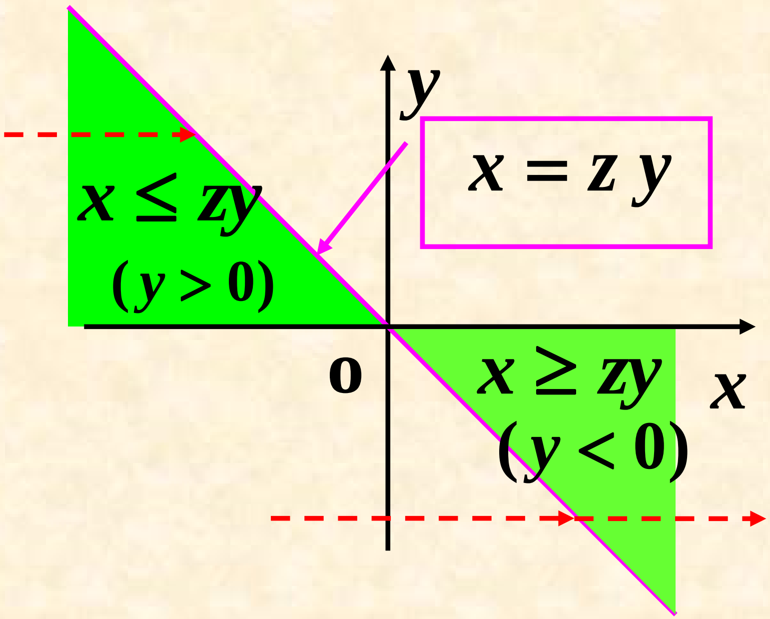
① 当$z \le 0$时，

$$D=\{(x,y) \mid \frac{x}{y} \le z\}$$

$$F_Z(z)=\iint\limits_D p(x,y)\,dx\,dy$$

$$=\int_{-\infty}^{0} dy \int_{zy}^{+\infty} p(x,y)dx+$$

$$+\int_{0}^{+\infty} dy \int_{-\infty}^{zy} p(x,y)dx$$



令$u=\dfrac{x}{y}$

$$=\int_{-\infty}^{0} dy \int_{z}^{-\infty} p(yu,y)\cdot y\,du+\int_{0}^{+\infty} dy \int_{-\infty}^{z} p(yu,y)\cdot y\,du$$

$$=\int_{-\infty}^{0}dy\int_{z}^{-\infty}p(yu,y)\cdot y\,du+\int_{0}^{+\infty}dy\int_{-\infty}^{z}p(yu,y)\cdot y\,du$$

$$=\int_{z}^{-\infty}du\int_{-\infty}^{0}p(yu,y)\cdot y\,dy+\int_{-\infty}^{z}du\int_{0}^{+\infty}p(yu,y)\cdot y\,dy$$

$$\therefore\quad p_Z(z)=\frac{dF_Z(z)}{dz}$$

$$=-\int_{-\infty}^{0}p(yz,y)\cdot y\,dy+\int_{0}^{+\infty}p(yz,y)\cdot y\,dy$$

$$=\int_{-\infty}^{+\infty}|y|\,p(yz,y)\,dy$$

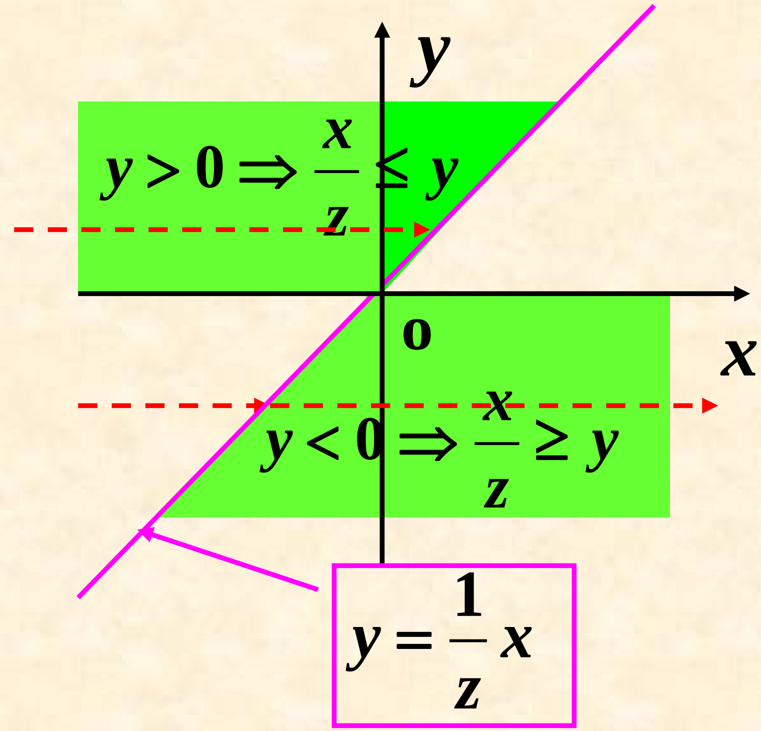
② 当$z>0$时，

$$D=\{(x,y)\mid \frac{x}{y}\le z\}$$

$$F_Z(z)=\iint_D p(x,y)\,dx\,dy$$

$$=\int_{-\infty}^{0}dy\int_{zy}^{+\infty}p(x,y)dx+$$

$$+\int_{0}^{+\infty}dy\int_{-\infty}^{zy}p(x,y)dx$$



令$u=\frac{x}{y}$

$$\underline{\underline{\qquad}}\int_{-\infty}^{0}dy\int_{z}^{-\infty}p(yu,y)\cdot y\,du+\int_{0}^{+\infty}dy\int_{-\infty}^{z}p(yu,y)\cdot y\,du$$

$$= \int_{-\infty}^{0} dy \int_{z}^{-\infty} p(yu, y) \cdot y du + \int_{0}^{+\infty} dy \int_{-\infty}^{z} p(yu, y) \cdot y du$$

$$= \int_{z}^{-\infty} du \int_{-\infty}^{0} p(yu, y) \cdot y dy + \int_{-\infty}^{z} du \int_{0}^{+\infty} p(yu, y) \cdot y dy$$

$$\therefore \quad p_Z(z) = \frac{dF_Z(z)}{dz}$$

$$= -\int_{-\infty}^{0} p(yz, y) \cdot y dy + \int_{0}^{+\infty} p(yz, y) \cdot y dy$$

$$= \int_{-\infty}^{+\infty} |y| p(yz, y)\, dy$$

(4) 极值分布，即 $M=\max\{X,Y\}$,

$N=\min\{X,Y\}$的分布.

① 当X，Y相互独立时，有

$$F_M(z)=F_X(z)F_Y(z)$$

$$F_N(z)=1-[1-F_X(z)][1-F_Y(z)]$$

② 当X，Y相互独立且同分布时，有

$$F_M(z)=F^2(z)$$

$$F_N(z)=1-[1-F(z)]^2$$

证 $F_M(z) = P\{M \le z\}$

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

$= P\{X \le z\} \cdot P\{Y \le z\}$ (X与Y独立)

$= F_X(z) F_Y(z)$

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

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

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

$= 1 - [1 - F_X(z)][1 - F_Y(z)]$

推广：一般地，设

$$M=\max\{X_1,X_2,\cdots X_n\},$$

$$N=\min\{X_1,X_2,\cdots X_n\},$$

则当 $X_1,X_2,\cdots,X_n$ 相互独立且同分布时，

有 $F_M(z)=F^n(z)$

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

其中 $F(z)=P\{X_1\le z\}$.

四、内容小结

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 $Z = f(X,Y)$ 的分布律为

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

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

2. 连续型随机变量函数的分布

(1) $Z = X + Y$ 的分布

(2) $Z = \dfrac{X}{Y}$ 的分布

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

备份题

例1 设随机变量 X 与 Y 相互独立,且其分布密度分别为

$$f_X(x)=\begin{cases}1, & 0\le x\le 1,\\ 0, & \text{其它}.\end{cases}\qquad f_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & \text{其它}.\end{cases}$$

求随机变量 $Z=2X+Y$ 的分布密度.

解 由于 X 与 Y 相互独立,所以 (X,Y) 的分布密度函数为

$$f(x,y)=f_X(x)\cdot f_Y(y)=\begin{cases}e^{-y}, & 0\le x\le 1, y>0,\\ 0, & \text{其它}.\end{cases}$$

随机变量 Z 的分布函数为

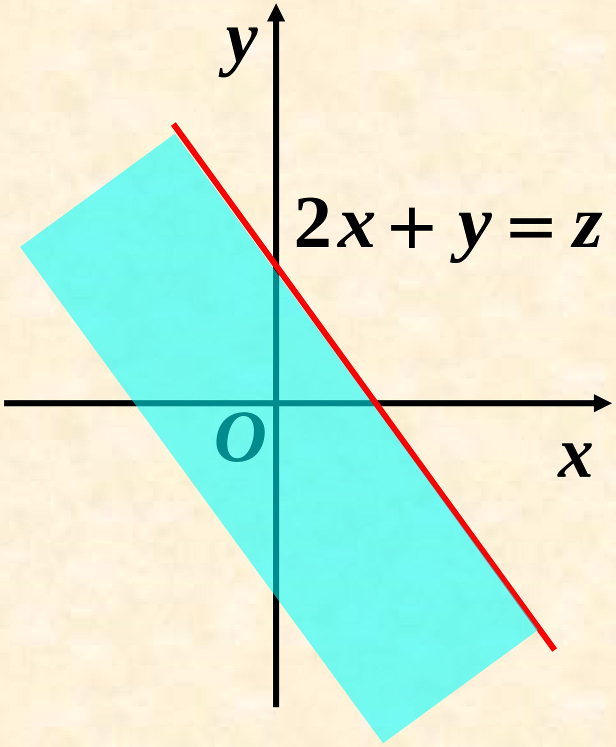
$$F_Z(z) = P\{Z \le z\}$$

$$= P\{2X + Y \le z\}$$

$$= \iint\limits_{2X+Y\le z} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \iint\limits_{2X+Y\le z} e^{-y}\,\mathrm{d}x\,\mathrm{d}y.$$

$$(0 \le x \le 1,\ y > 0)$$



$$F_Z(z)=\begin{cases}0, & z\le 0,\\ \int_0^{\frac{z}{2}}(1-e^{2x-z})\,\mathrm{d}x, & 0<z\le 2,\\ \int_0^{1}(1-e^{2x-z})\,\mathrm{d}x, & z>2.\end{cases}$$

所以随机变量 Z 的分布密度为

$$f_Z(z)=F_Z'(z)=\begin{cases}0, & z\le 0,\\ (1-e^{-z})/2, & 0<z\le 2,\\ (e^2-1)e^{-z}/2, & z\ge 2.\end{cases}$$

例2 若X和Y独立,具有共同的概率密度

$$p(x)=\begin{cases}1, & 0\le x\le 1\\ 0, & \text{其它}\end{cases}$$ 求$Z=X+Y$的概率密度．

解：由卷积公式

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

为确定积分限,先找出使被积函数不为0的区域

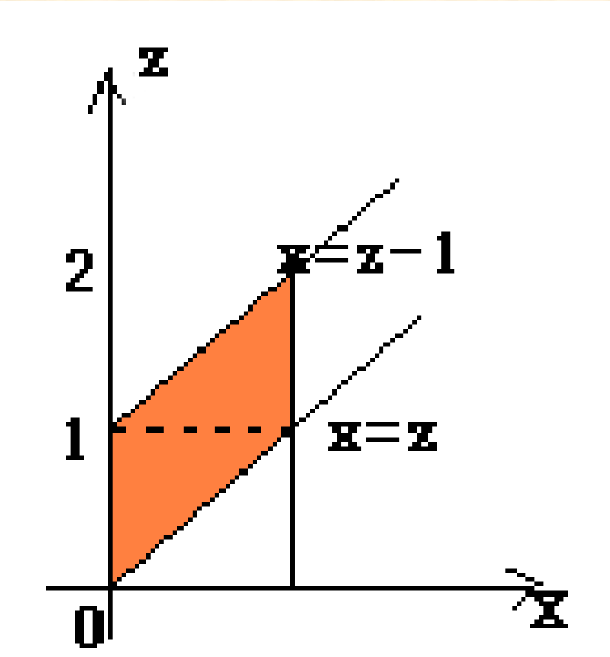
$$\begin{cases}0\le x\le 1\\ 0\le z-x\le 1\end{cases}\quad \text{也即}\quad \begin{cases}0\le x\le 1\\ z-1\le x\le z\end{cases}$$

$$p_Z(z)=\int_{-\infty}^{\infty} p_X(x)\,p_Y(z-x)dx$$

为确定积分限,先找出使被积函数不为0的区域

$$\begin{cases} 0\le x\le 1 \\ 0\le z-x\le 1 \end{cases} \quad \text{也即} \quad \begin{cases} 0\le x\le 1 \\ z-1\le x\le z \end{cases}$$

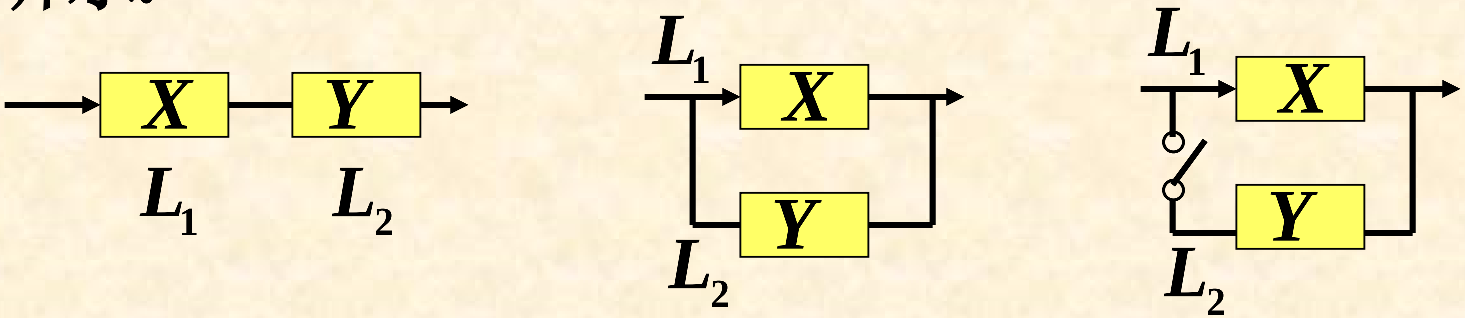
如图示:



于是

$$p_Z(z)=\begin{cases} \int_0^z dx = z, & 0\le z<1 \\ \int_{z-1}^1 dx = 2-z, & 1\le z<2 \\ 0, & \text{其它} \end{cases}$$

例3 设系统L由两个相互独立的子系统 L_1, L_2 联接而成,连接的方式分别为(i) 串联, (ii) 并联, (iii) 备用(当系统 L_1 损坏时,系统L_2开始工作),如图所示.



设 L_1, L_2 的寿命分别为X, Y,已知它们的概率密度分别为

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

其中 $\alpha>0, \beta>0$ 且 $\alpha\neq\beta$. 试分别就以上三种联接方式写出 L 的寿命 Z 的概率密度.

解 (i)串联情况

由于当 L_1, L_2 中有一个损坏时,系统 L 就停止工作,所以这时 L 的寿命为 $Z=\min(X,Y)$.

由 $f_X(x)=\begin{cases}\alpha e^{-\alpha x}, x>0,\\ 0, \quad x\le 0,\end{cases} \Rightarrow F_X(x)=\begin{cases}1-e^{-\alpha x}, x>0,\\ 0, \quad x\le 0,\end{cases}$

由 $f_Y(y)=\begin{cases}\beta e^{-\beta y}, y>0,\\ 0, \quad y\le 0;\end{cases} \Rightarrow F_Y(y)=\begin{cases}1-\beta e^{-\beta y}, y>0,\\ 0, \quad y\le 0.\end{cases}$

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

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

$$\Rightarrow f_{\min}(z)=\begin{cases}(\alpha+\beta)e^{-(\alpha+\beta)z}, z>0,\\ 0, \quad z\le 0.\end{cases}$$

(ii)并联情况

由于当且仅当 L_1, L_2 都损坏时,系统 L 才停止工作,

所以这时 L 的寿命为 $Z = \max(X, Y)$.

$Z = \max(X, Y)$的分布函数为

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

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

(iii)备用的情况

由于这时当系统 L_1 损坏时,系统 L_2 才开始工作,因此整个系统 L 的寿命 Z 是 L_1, L_2 两者之和,即

$$Z = X + Y$$

当 $z > 0$ 时, $Z = X + Y$ 的概率密度为

$$f(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) \mathrm{d}y = \int_0^z \alpha e^{-\alpha(z-y)} \beta e^{-\beta y} \mathrm{d}y$$

$$= \alpha\beta e^{-\alpha z} \int_0^z e^{-(\beta-\alpha)y} \mathrm{d}y$$

$$= \frac{\alpha\beta}{\beta - \alpha}[e^{-\alpha z} - e^{-\beta z}].$$

当 $z < 0$ 时, $f(z) = 0$,

于是 $Z = X + Y$ 的概率密度为

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