问题的由来

例1: 炮击某一目标O，已知弹着点(X,Y)服从二维正态分布. 点(X,Y)与目标O的距离$Z=\sqrt{X^2+Y^2}$服从什么分布?

例2: 由统计物理学, 气体分子运动速率 v 服从马克斯维尔分布.

$$f(x)=\begin{cases}\dfrac{4x^2}{\alpha^3\sqrt{\pi}}e^{\frac{-x^2}{\alpha^2}} & x>0,\alpha>0\\ 0 & x\le 0\end{cases}$$

分子运动动能$\eta=\dfrac{1}{2}mv^2$服从什么分布?

引例1 博彩问题

一个庄家在一个签袋中放有8个白、8个黑的围棋子。规定：每个摸彩者交一角钱作“手续费”，然后一个从袋中摸出五个棋子，按下面“摸子中彩表”给“彩金”。

摸到	五个白	四个白	三个白	其它
彩金	2元	2角	5分	共乐一次

设X表示摸出的白围棋子个数，Y表示一次得到的彩金，则X，Y的分布律为

X	0	1	2	3	4	5
P	0.0128	0.1282	0.3590	0.3590	0.1282	0.0128
Y	0	0	0	0.05	0.2	2
P	0.0128	0.1282	0.3590	0.3589	0.1282	0.0128

已知X的分布律，如何确定$Y=g(X)$的分布律?

思路:

① 确定Y的所有可能取值

② 找出与Y取值 所对应的X取值

③ 确定Y的取值所对应的概率

引例2 骰子点数和

抛两颗骰子，分析点数之和的分布.

设X表示第一颗骰子的点数，Y表示第二颗骰子的点数，Z表示点数之和，则

思路:

① 先考虑Z的所有可能取值

② 找出与Z取值对应的(X,Y)组合

③ 计算Z所有取值对应的概率

Z	2	3	4	5	6	7
(X,Y)	1种	2种	3种	4种	5种	6种
Z	8	9	10	11	12	
(X,Y)	5种	4种	3种	2种	1种	

从而Z的分布律为:

Z	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

函数的分布律

设(X,Y)的联合分布律为:

$X \backslash Y$	0	1
0	3/10	3/10
1	3/10	1/10

试求 1) $\sin X$　2) $X+Y$　3) XY　4) $\max(X,Y)$的分布律.

解:由(X,Y)的分布律得

P	3/10	3/10	3/10	1/10
(X,Y)	(0,0)	(0,1)	(1,0)	(1,1)
X	0		1	
$\sin X$	0		$\sin 1$	
$X+Y$	0	1	1	2
XY	0	0	0	1
$\max(X,Y)$	0	1	1	1

P	3/10	3/10	3/10	1/10
X	0		1	
$\sin X$	0		$\sin 1$	
(X,Y)	(0,0)	(0,1)	(1,0)	(1,1)
$X+Y$	0	1	1	2
XY	0	0	0	1
$\max(X,Y)$	0	1	1	1

$\sin X$	0	$\sin 1$
P	0.6	0.4

$X+Y$	0	1	2
P	0.3	0.6	0.1

XY	0	1
P	0.9	0.1

$\max(X,Y)$	0	1
P	0.3	0.7

二项分布之和

设X ,Y相互独立,且$X\sim B(n_1, p)$, $Y\sim B(n_2, p)$　则

$$X+Y\sim B(n_1+n_2, p)$$

推导中是否存在错误？如何理解？

证：

$$P\{X=k\}=C_{n_1}^{k}p^{k}(1-p)^{n_1-k}\qquad k=0,1,..n_1$$

$$P\{Y=r\}=C_{n_2}^{r}p^{r}(1-p)^{n_2-r}\qquad r=0,1,..n_2$$

$$P\{X+Y=m\}=\sum_{k=0}^{m}C_{n_1}^{k}p^{k}(1-p)^{n_1-k}C_{n_2}^{m-k}p^{m-k}(1-p)^{n_2-m+k}$$

$$=p^{m}(1-p)^{n_1+n_2-m}\sum_{k=0}^{m}C_{n_1}^{k}C_{n_2}^{m-k}$$

利用 $\sum_{k=0}^{m}C_{n_1}^{k}C_{n_2}^{m-k}=C_{n_1+n_2}^{m}$

$$=p^{m}(1-p)^{n_1+n_2-m}C_{n_1+n_2}^{m}\qquad m=0,1,\ldots,n_1+n_2$$

二项分布具有可加性

设$X\sim N(0,1)$，求$Y=X^2$的概率密度.

分析:求随机变量的函数的概率密度,我们一般是通过求分布函数得到的.

解: 当$y\le 0$ $F_Y(y)=P\{X^2\le y\}=0$

当$y>0$ $F_Y(y)=P\{X^2\le y\}=P\{-\sqrt{y}\le X\le \sqrt{y}\}$

$$=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

继续积分?

$$F_Y(y)=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

此时 $f_Y(y)=F_Y'(y)=\frac{1}{\sqrt{2\pi}}\left[e^{-\frac{(\sqrt{y})^2}{2}}(\sqrt{y})'-e^{-\frac{(-\sqrt{y})^2}{2}}(-\sqrt{y})'\right]$

$$=\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}$$

$$f_Y(y)=\begin{cases}\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}} & y>0\\ o & y\le 0\end{cases}$$

称Y服从自由度为1的χ^2分布

$X+2Y$的概率密度

设二维随机变量(X,Y)的联合概率密度为:

$$f(x,y)=\begin{cases}2e^{-(x+2y)} & x,y>0\\ 0 & \text{其它}\end{cases}$$

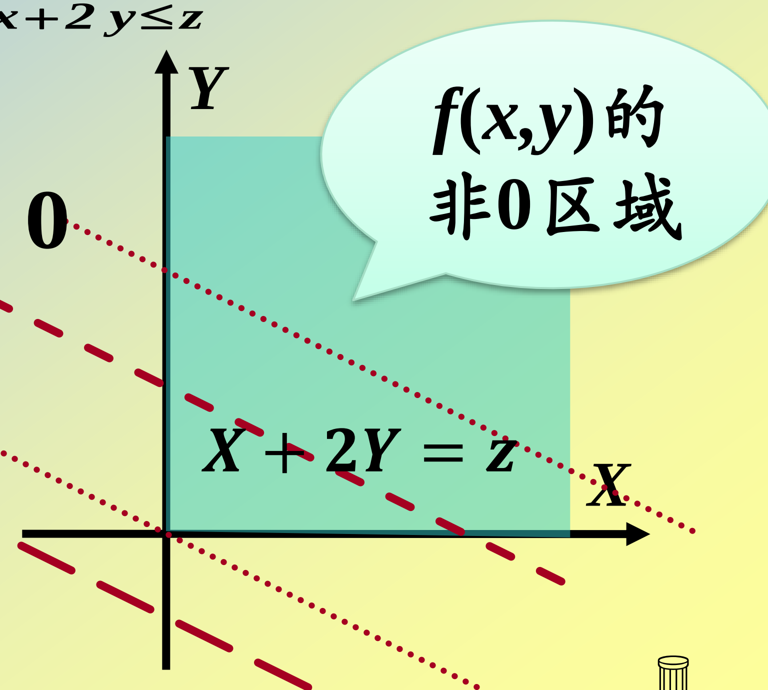
求随机变量$Z=X+2Y$的分布函数和概率密度.

解: $F_Z(z)=P\{Z\le z\}=P\{X+2Y\le z\}=\iint\limits_{x+2y\le z} f(x,y)d\sigma$

$$=\begin{cases}0 & z<0\\ \int_0^z\left[\int_0^{\frac{z-x}{2}} 2e^{-(x+2y)}dy\right]dx & z\ge 0\end{cases}$$

$$=\begin{cases}0 & z<0\\ 1-e^{-z}-ze^{-z} & z\ge 0\end{cases}$$

$$f_Z(z)=F_Z'(z)=\begin{cases}0 & z<0\\ ze^{-z} & z\ge 0\end{cases}$$



再求平方的概率密度

设$X \sim N(0,1)$，求$Y = X^2$的概率密度.

分析:注意到$y > 0$时，$y = x^2$分段单调连续，可用定理

解：由于$y = x^2 \geq 0$，故此$y < 0$时，$f_Y(y) = 0$

$y > 0$时，$y = x^2$分段单调连续

I_1：$x \in (-\infty, 0], x = h_1(y) = -\sqrt{y}$

I_2：$x \in (0, +\infty), x = h_2(y) = \sqrt{y}$

$$f_Y(y) = f_X[h_1(y)]|h_1'(y)| + f_X[h_2(y)]|h_2'(y)|$$

$$= f_X[-\sqrt{y}]|(-\sqrt{y}\,)'| + f_X[\sqrt{y}]|(\sqrt{y}\,)'|$$

$$= \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(-\sqrt{y})^2} + \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\sqrt{y})^2}\right]$$

$$= \frac{1}{2\sqrt{\pi y}}e^{-\frac{1}{2}y}$$

自己写出完整概率密度函数

练习

设$X \sim N(0,1)$，求$Y = |X|$的概率密度.

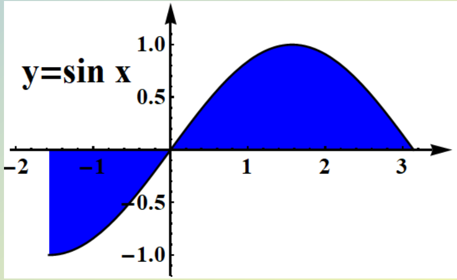
参考答案:

$$f_Y(y) = \begin{cases} \sqrt{\dfrac{2}{\pi}} e^{-\frac{y^2}{2}}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

往年考题

设X的概率密度为

$$f_X(x)=\begin{cases}\dfrac{8}{9\pi^2}\left(x+\dfrac{\pi}{2}\right), & -\dfrac{\pi}{2}\le x\le \pi\\ 0, & \text{其他}\end{cases}$$



求$Y=\sin X$的概率密度函数。

分析：一般步骤

(1) 确定$f_Y(y)$的非零区域$[a,b]$，在其外$f_Y(y)=0$

(2) 当$y\in[a,b]$时，作出$Y=g(X)$的图形，确定以$y=y_0$为水平线截得的图形，将$[a,b]$分为不同区间；

(3) 在各区间上结合图形求$F_Y(y)$，有时候无需具体积分，然后求导得到$f_Y(y)$；

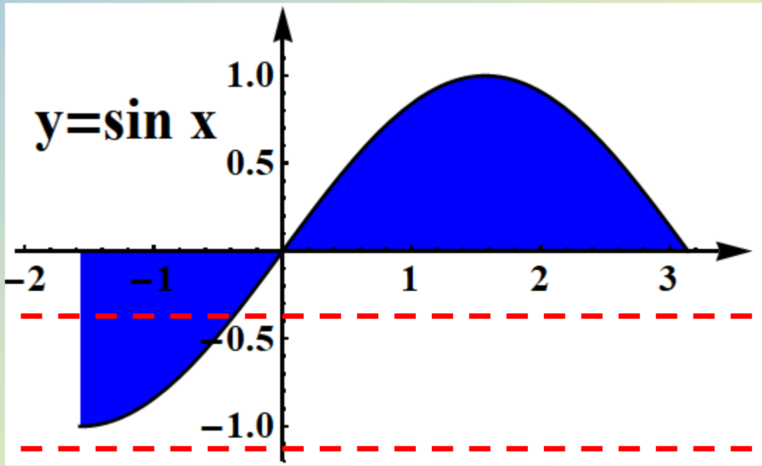
(4) 合并各区间得到综合表达式。

往年考题

设X的概率密度为

$$f_X(x)=\begin{cases}\dfrac{8}{9\pi^2}\left(x+\dfrac{\pi}{2}\right), & -\dfrac{\pi}{2}\le x\le\pi\\ 0, & \text{其他}\end{cases}$$

求$Y=\sin X$的概率密度函数。



解：由于$-1\le y\le 1$

结合图形可知非零区域分两段

当$-1\le y<0$时，

$$F_Y(y)=P\{Y\le y\}=P\{\sin X\le y\}$$

$$=P\left\{-\frac{\pi}{2}<X<\arcsin y\right\}$$

$$=\int_{-\frac{\pi}{2}}^{\arcsin y}\frac{8}{9\pi^2}\left(x+\frac{\pi}{2}\right)dx$$

往年考题

求导得

$$f_Y(y)=\frac{8}{9\pi^2}\left(\arcsin y+\frac{\pi}{2}\right)\frac{1}{\sqrt{1-y^2}}$$

当$0\le y\le 1$时，

$$F_Y(y)=P\{Y\le y\}=P\{\sin X\le y\}$$

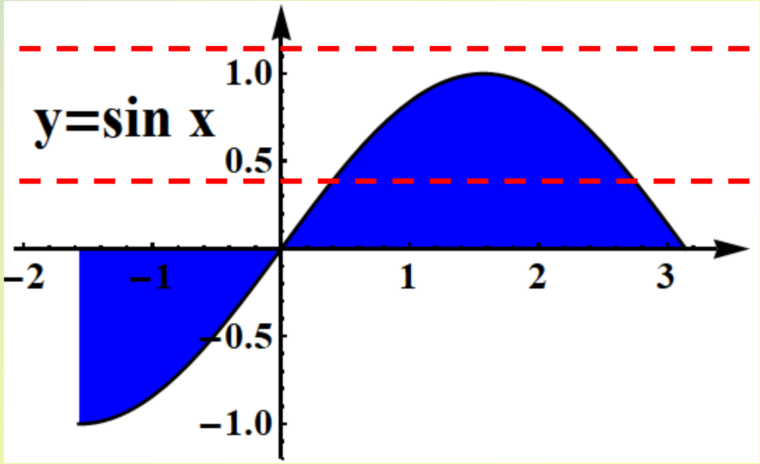
$$=P\left\{-\frac{\pi}{2}<X<\arcsin y\right\}+P\{\pi-\arcsin y<X<\pi\}$$

$$=\int_{-\frac{\pi}{2}}^{\arcsin y}\frac{8}{9\pi^2}\left(x+\frac{\pi}{2}\right)dx+\int_{\pi-\arcsin y}^{\pi}\frac{8}{9\pi^2}\left(x+\frac{\pi}{2}\right)dx$$

求导得

$$f_Y(y)$$

$$=\frac{8}{9\pi^2}\left[\left(\arcsin y+\frac{\pi}{2}\right)+\left(\frac{3\pi}{2}-\arcsin y\right)\right]\frac{1}{\sqrt{1-y^2}}$$



往年考题

求导得

$$f_Y(y)=\frac{8}{9\pi^2}\left[\left(\arcsin y+\frac{\pi}{2}\right)+\left(\frac{3\pi}{2}-\arcsin y\right)\right]\frac{1}{\sqrt{1-y^2}}$$

$$=\frac{16}{9\pi}\frac{1}{\sqrt{1-y^2}}$$

从而可得

$$f_Y(y)=\begin{cases}\dfrac{16}{9\pi}\dfrac{1}{\sqrt{1-y^2}}, & 0<y\le 1\\ \dfrac{8}{9\pi^2}\left(\arcsin y+\dfrac{\pi}{2}\right)\dfrac{1}{\sqrt{1-y^2}}, & -1<y\le 0\\ 0, & |y|>1\end{cases}$$

尝试：采用逐段确定概率密度方法求解该问题

均匀分布之和

设随机变量X, Y相互独立，均服从区间(0，1) 上的均匀分布，求：$Z = X + Y$ 的概率密度$f_Z(z)$。

分析： 1. X, Y 相互独立，所以有：

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

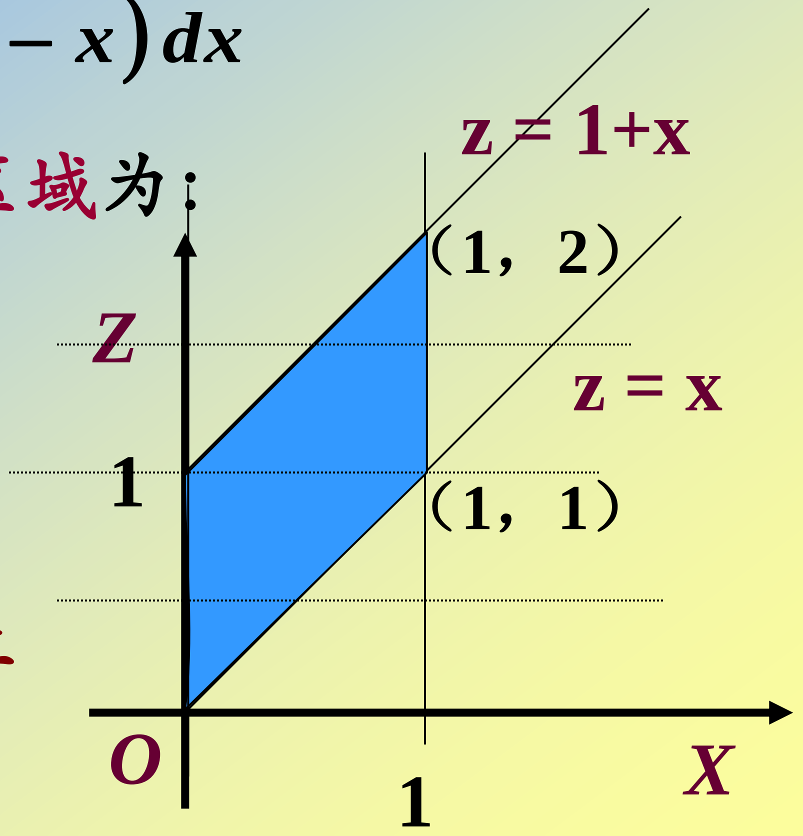
2. 使 $f_X(x)\, f_Y(z-x)$ 为非零的区域为：

$$0 \le x \le 1 \quad 0 \le z - x \le 1$$

3. $f_Z(z)$的非零区域为：

$$0 \le z \le 2$$

4.在不同的区间段 积分的上下限是不相同的。



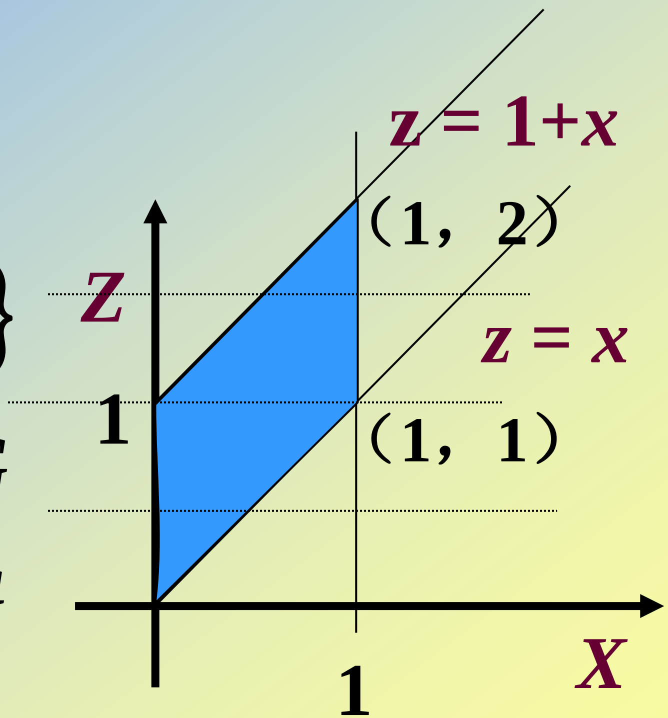
解: $\because$ 随机变量 X, Y 相互独立

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

在XOZ平面上作出区域G

$$G = \{(x,z) \mid 0 \le x \le 1, 0 \le z - x \le 1\}$$

$$f_X(x) f_Y(z-x) = \begin{cases} 0 & (x,z) \notin G \\ 1 & (x,z) \in G \end{cases}$$



当$z<0$ 或 $z\geq 2$ 时 $f_Z(z)=0$

当$0\leq z<1$时

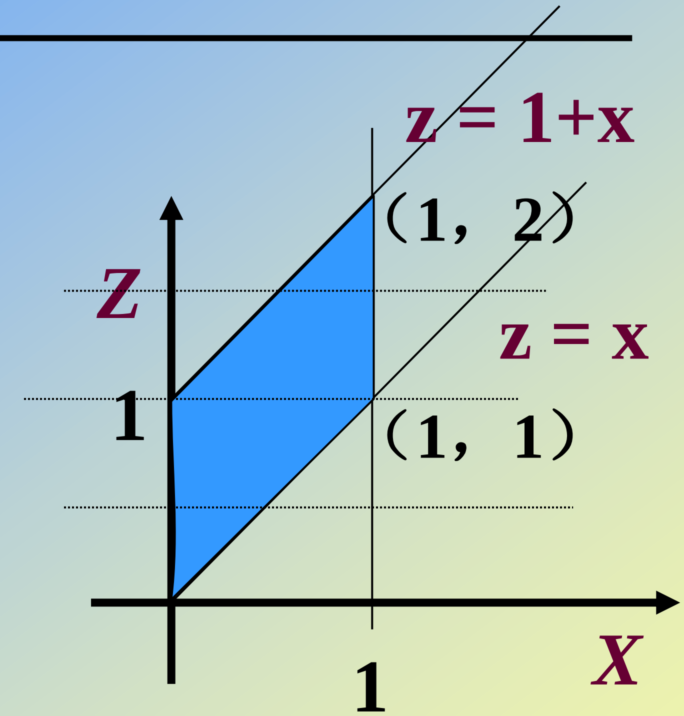
$$f_Z(z)=\int_0^z 1\,dx$$

$$=z$$

当$1\leq z<2$时

$$f_Z(z)=\int_{z-1}^1 1\,dx$$

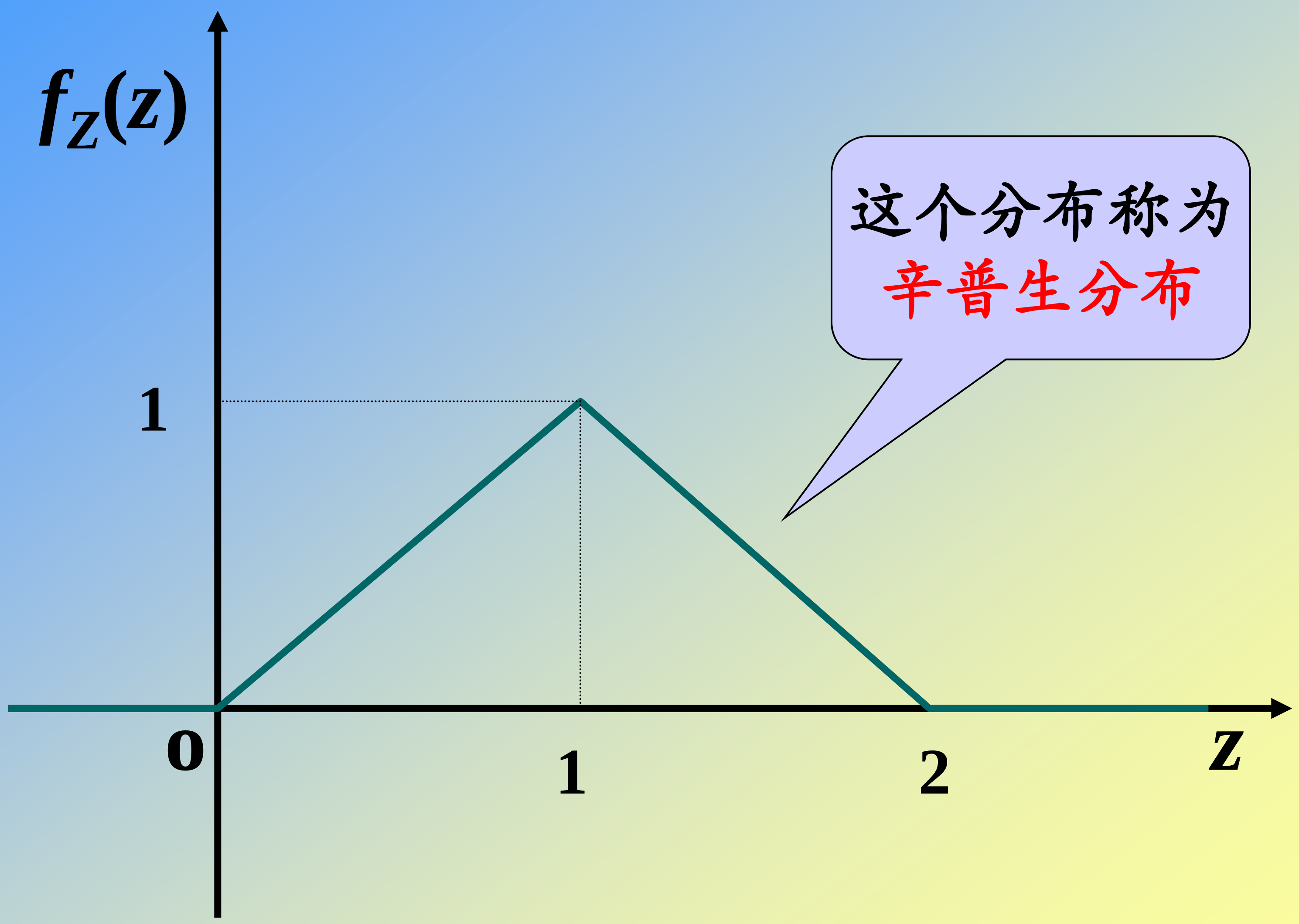
$$=2-z$$



综上得$Z=X+Y$的概率密度为：

$$f_Z(z)=\begin{cases} z & 0\leq z<1 \\ 2-z & 1\leq z<2 \\ 0 & \text{其他} \end{cases}$$

$Z = X + Y$ 的概率密度曲线为

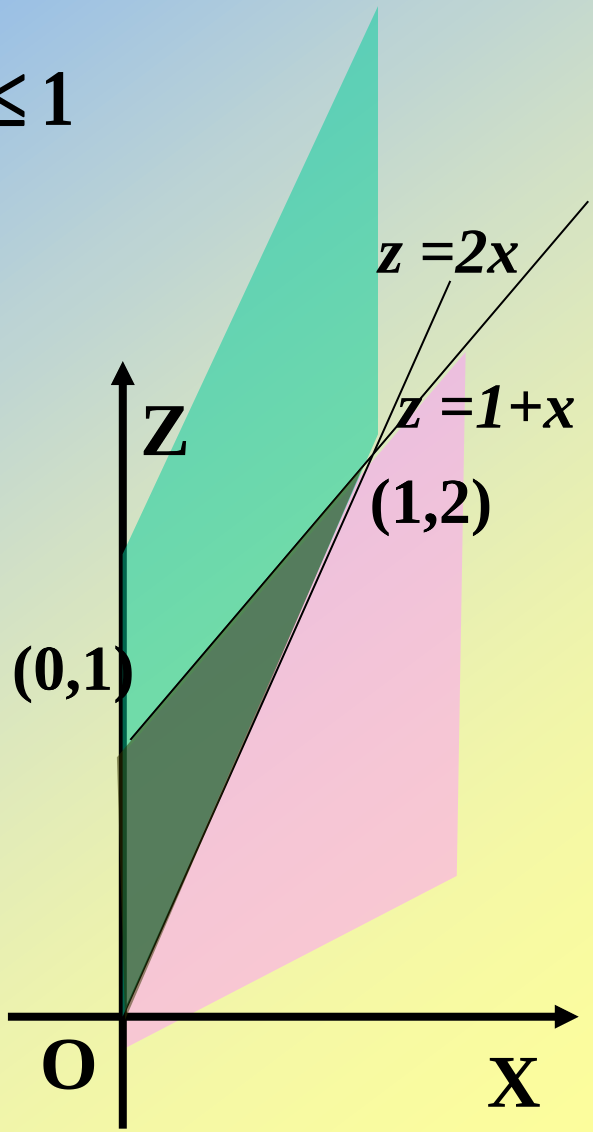


$X+Y$ 的概率密度

已知二维随机变量（X,Y）的联合概率密度为

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

求： $Z=X+Y$ 的概率密度。



解： 在XOZ平面上作出区域

$$G=\{(x,z)|0\le x\le z-x\le 1\}$$

$$=\{(x,z)|0\le x\le 1,\ 2x\le z\le 1+x\}$$

$$f(x,y)=\begin{cases}2z & (x,z)\in G\\ 0 & \text{其他}\end{cases}$$

当 $z<0$ 或 $z\geq 2$ 时 $f_Z(z)=0$

当 $0\leq z<1$ 时

$$f_Z(z)=\int_0^{z/2}2z\,dx=z^2$$

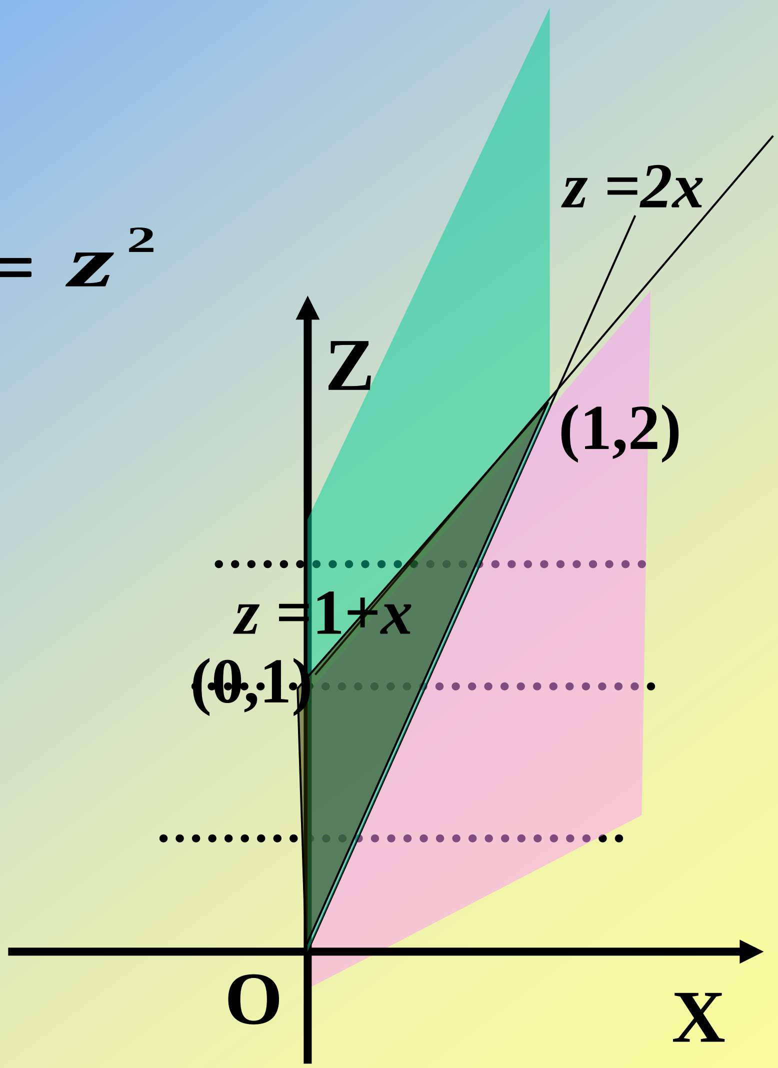
当 $1\leq z<2$ 时

$$f_Z(z)=\int_{z-1}^{z/2}2z\,dx$$

$$=2z-z^2$$

综上得 $Z=X+Y$ 的概率密度为:

$$f_Z(z)=\begin{cases} z^2 & 0\leq z<1 \\ 2z-z^2 & 1\leq z<2 \\ 0 & \text{其他} \end{cases}$$



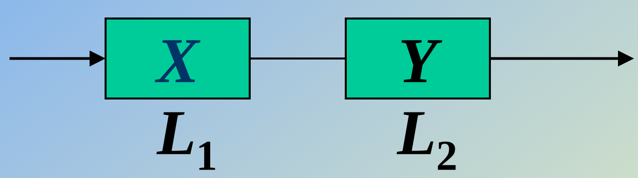
系统寿命

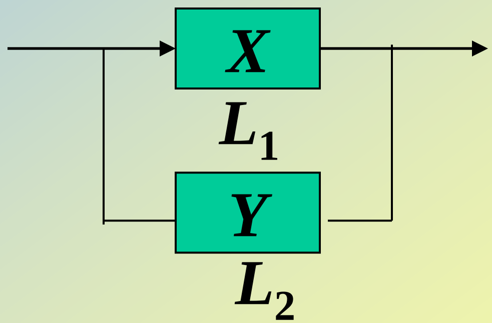
设系统L由两个功能相似且相互独立的子系统L_1, L_2连接而成,连接的方式分别为:1) 串联 2) 并联 3) 备用,如图所示. 设L_1, L_2的寿命分别为X, Y 它们的概率密度分别为:

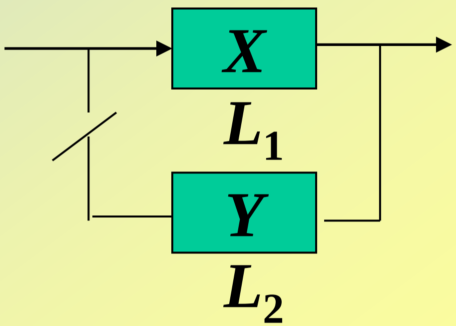
$$f_X(x)=\begin{cases}\alpha e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases} \qquad f_Y(y)=\begin{cases}\beta e^{-\beta y} & y>0\\ 0 & y\le 0\end{cases}$$

$$\alpha,\beta>0 \quad \alpha\ne\beta$$

试求出在以上3种连接方式下系统寿命T的概率密度

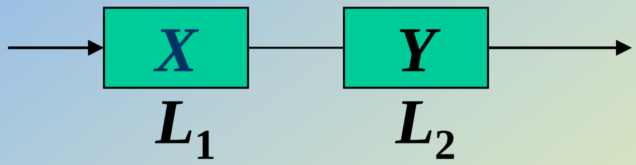






解：1) 串联系统

由于当L_1,L_2有一个损坏时，系统L就停止工作，此时系统寿命$T=\min(X,Y)$



X,Y 的分布函数分别为：

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

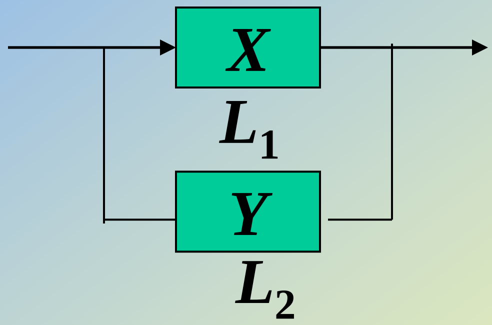
T 的概率密度函数为：

$$f_T(t)=f_X(t)[1-F_Y(t)]+f_Y(t)[1-F_X(t)]$$

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

电子科技大学数学科学学院 杜鸿飞 hongfeidu@qq.com

2) 并联系统



由于当L_1,L_2 都损坏时,系统L才停止工作,此时系统寿命

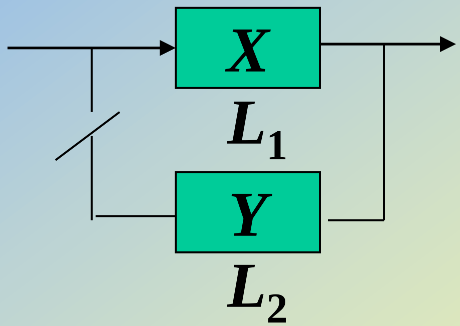
$$T = \max(X, Y)$$

T 的概率密度函数为:

$$f_T(t) = f_X(t)F_Y(t) + f_Y(t)F_X(t)$$

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

3) 备用系统



由于当L_1损坏时,L_2才开始工作

此时系统寿命$T = X + Y$

由卷积公式当 $t > 0$ 时, T 的概率密度函数为:

$$f_T(t) = \int_{-\infty}^{+\infty} f_X(t-y) f_Y(y) dy = \int_0^t \alpha e^{-\alpha(t-y)} \cdot \beta e^{-\beta y} dy$$

$$= \frac{\alpha\beta}{\beta - \alpha}\left(e^{-\alpha t} - e^{-\beta t}\right)$$

当 $t \le 0$ 时: $f_T(t) = 0$ 从而T 的概率密度为

$$f_T(t) = \begin{cases} \dfrac{\alpha\beta}{\beta - \alpha}\left(e^{-\alpha t} - e^{-\beta t}\right) & t > 0 \\ 0 & t \le 0 \end{cases}$$

商的分布

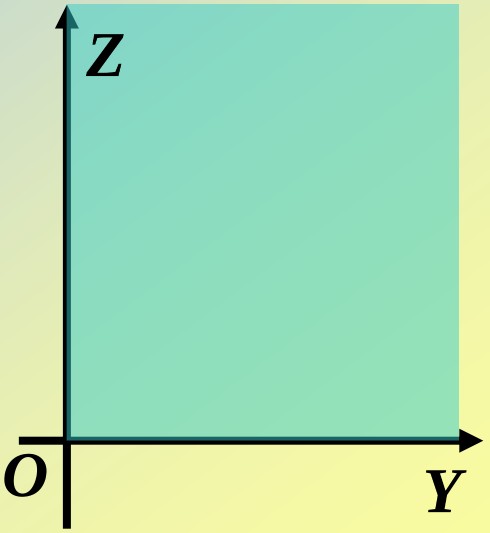
已知随机变量X,Y相互独立同分布.

$$f_X(x)=f_Y(x)=\begin{cases} e^{-x} & x>0 \\ 0 & 其它 \end{cases}$$

求：X/Y 的分布.

解：令 $G=\{(y,z): yz>0, y>0\}$

$$=\{(y,z): y>0, z>0\}$$



$$f(yz,y)=f_X(yz)f_Y(y)$$

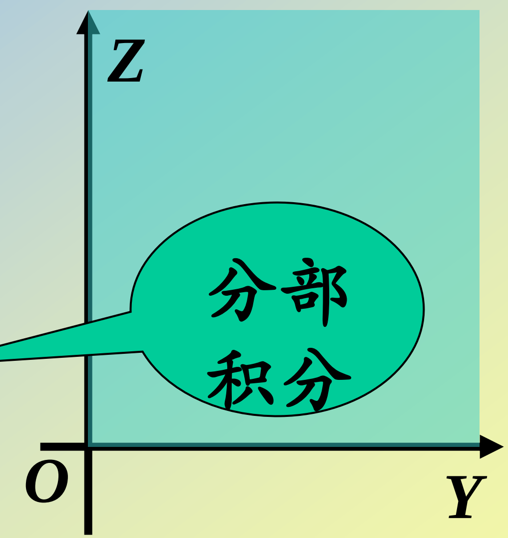
$$=\begin{cases} e^{-yz-y} & (y,z)\in G \\ 0 & 其它 \end{cases}$$

$$f(yz,y)=f_X(yz)f_Y(y)$$

$$=\begin{cases} e^{-yz-y} & (y,z)\in G \\ 0 & \text{其它} \end{cases}$$

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

$$=\int_0^{+\infty} yf(yz,z)dy$$



$$=\begin{cases} \dfrac{1}{(z+1)^2} & z>0 \\ 0 & \text{其它} \end{cases}$$