§3.4 随机变量的函数及其分布

问题的由来

很多实际问题中需要研究以随机变量为自变量的函数.

一般，若$(X_1, X_2, \ldots, X_n)$是已知联合分布的n维随机变量，则

$$Y = h(X_1, X_2, \cdots, X_n)$$

仍是随机变量，其中

$h(x_1, x_2, \cdots, x_n)$是 n 元连续函数.

问题 如何确定随机变量Y的分布?

基本思路 希望通过$(X_1, X_2, \ldots, X_n)$的**已知**分布去确定Y的分布.

例3.4.1 例3.4.2

一.离散型随机变量的函数及其分布律

离散型随机变量X的分布律为

$$P\{X = x_i\} = p_i, \qquad i = 1, 2, \ldots$$

$Y=g(X)$ 是随机变量, 则

$$P\{Y=y_j\}=P\{g(X)=y_j\}$$

$$=\sum_{x_i\in S_j}P\{X=x_i\},\qquad j=1,2,\dots$$

其中 $S_j=\{x_i \mid g(x_i)=y_j\}$

满足 $g(x_i)=y_j$ 的全体 x_i 构成的集合

二维离散型随机变量(X, Y)的联合分布律为

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

$Z = G(X, Y)$ **是随机变量, 则**

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

$$= \sum_{(x_i, y_j) \in T_k} P\{X = x_i, Y = y_j\}, \quad k = 1,2,\ldots$$

其中 $T_k = \{(x, y) \mid G(x_i, y_j) = z_k\}$

例 3.4.3

定理3.4.1 设随机变量(X, Y)是离散型随机变量，X, Y相互独立,其分布律分别为

$$P\{X=k\}=p(k) \quad k=0,1,2,\ldots$$

$$P\{Y=r\}=q(r) \quad r=0,1,2,\ldots$$

则$X+Y$的分布律为

$$P\{X+Y=m\}=\sum_{k=0}^{m} p(k)q(m-k) \quad m=0,1,2,\ldots$$

离散卷积公式

例3.4.4

结论 若$X_1, X_2, \ldots, X_n$相互独立,且$X_i \sim B(1, p)$则

$$X_1 + X_2 + \ldots + X_n \sim B(n, p)$$

反之若 $X \sim B(n, p)$, 则存在相互独立的

$X_i \sim B(1, p)$，使

$$X = X_1 + X_2 + \ldots + X_n$$

一般 1）随机变量$X_1, X_2, \ldots, X_n$相互独立;

2）具有相同类型的分布;

若

$$Y = \sum_{k=1}^{n} X_k$$

的分布除参数变化,而分布类型不变,称分布具有可加性.

二项分布具有可加性

泊松分布具有可加性

教材例3.4.3

二、连续型随机变量的函数及其概率密度

1. 设X 是连续型随机变量,若$Y= g(X)$也是连续型随机变量，则

$$F_Y(y) = P\{g(X) \le y\} = \int_{\{x \mid g(x) \le y\}} f_X(x)dx$$

$$f_Y(y) = \begin{cases} F_Y'(y), & f_Y(y)\text{的连续点;} \\ 0, & \text{其他.} \end{cases}$$

例3.4.5

例3.4.6

总结 从分布函数定义出发

解决问题的出发点

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

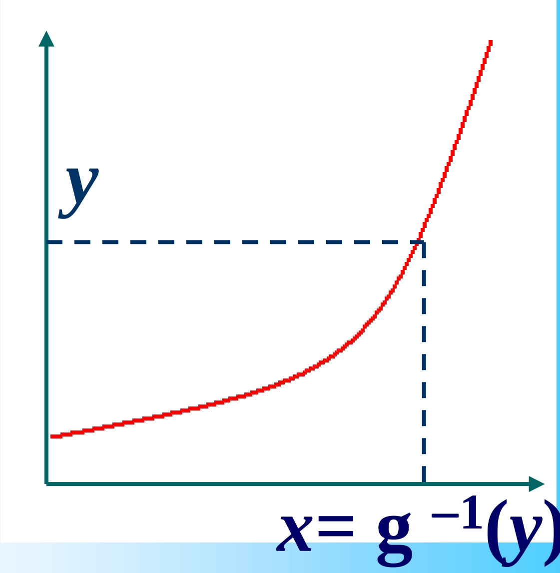
◆**求解的关键**

将$g(X) \le y$ 转换为关于 X 的不等式

当$g(x)$ 为单调递增函数时

$$\{ g(X) \le y \} = \{ X \le g^{-1}(y) \}$$

从而 $F_Y(y) = F_X[g^{-1}(y)]$



当 $g(x)$ 为单调递减函数

$$\{ g(X) \le y \} = \{ X \ge g^{-1}(y) \}$$

有 $F_Y(y) = 1 - F_X[g^{-1}(y)]$

2.求二维连续型随机变量 (X, Y) 的函数

$$Z = G(X, Y)$$

的概率密度 $f_z(z)$，

一般方法

1）**先求出 Z 的分布函数 $F_Z(Z)$；**

$$F_Z(z) = P\{ Z \le z \} = P\{ G(X, Y) \le z \}$$

$$= \iint\limits_{\{(x,y):G(x,y)\le z\}} f(x,y)dxdy$$

2）对$F_Z(Z)$微分得到 $f_z(z)$；

例3.4.7 例3.4.8

三.几种特殊函数的分布

1.$M = max\,(X,Y)$， $N = min\,(X\,,Y)$

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

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

若 X 与 Y 相互独立,有

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

又若 X 与 Y 有相同分布

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

从而 $$f_M(z) = 2F(z)f(z)$$

思考 已推得

$$F_N(z) = P\{\min(X,Y) \le z\} = P\{X \le z \text{或} Y \le z\}$$

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

若 X 与 Y 相互独立且具有相同分布

$$f_N(z) = 2[1 - F(z)]f(z)$$

?

见教材例3.4.9， 随机系统的串并联.

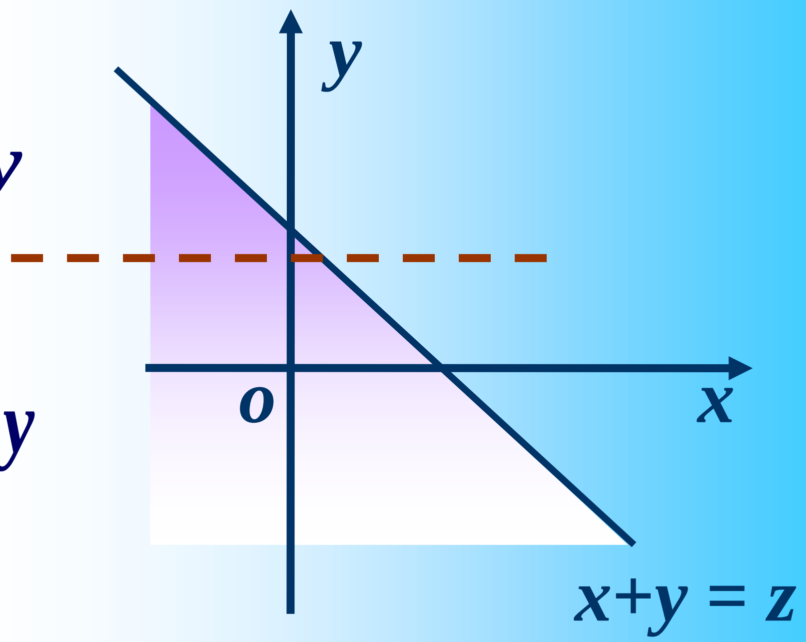
2. $Z = X + Y$ 的分布

设随机变量(X,Y)的联合概率密度为$f(x,y)$

$$F_Z(z) = P\{X + Y \le z\}$$

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

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



做积分变量变换,令

$$x = u - y$$

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

$x=u-y$

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

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

由连续型随机变量定义

$$F_Z(z)=\int_{-\infty}^{z} f_z(u)\,du$$

得到公式

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

若随机变量X, Y相互独立，则

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

类似可得

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

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

卷积公式

例 3.4.9

正态分布的可加性
见例3.4.11

解题步骤：

1）给出计算公式，并确定$f(x, z-x)$表达式；

2）在 XOZ平面上作出 $f(x, z-x)$ 的非零区域 G；；

3）在$f_Z(z)$非零区间内，逐段确定$f_Z(z)$的表达式；

4）写出 $f_Z(z)$的完整表达式.

例 3.4.10

3. $Z = X/Y$ 的分布

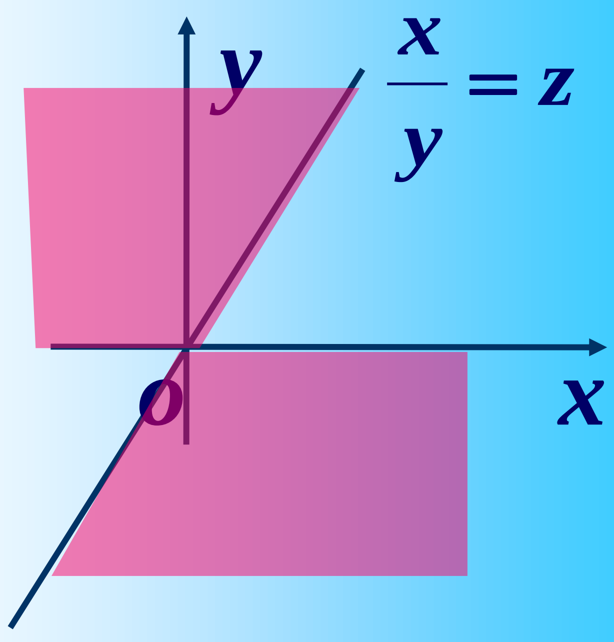
设随机变量(X, Y)的联合概率密度为$f(x,y)$

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

证 $F_Z(z) = P\{X/Y \le z\}$

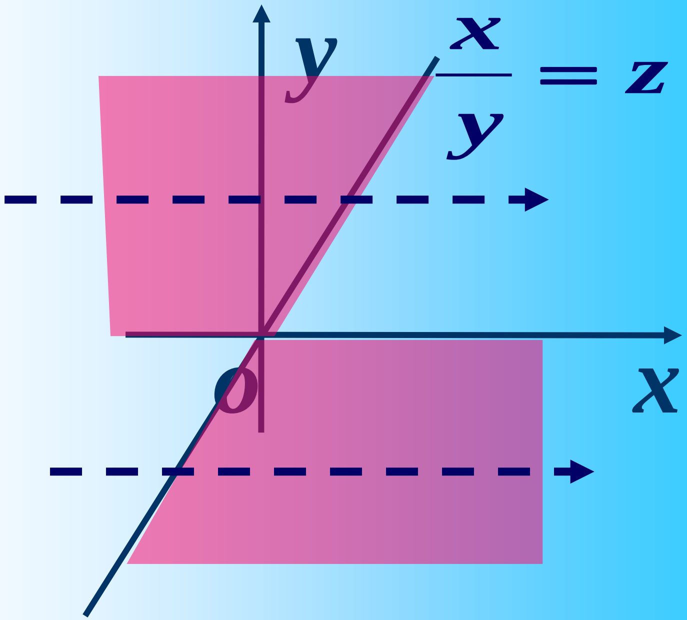
$$= \iint\limits_{x/y \le z} f(x,y)\,dxdy$$

$$x/y \le z \Leftrightarrow \begin{cases} x \le yz, & y>0; \\ x \ge yz, & y<0. \end{cases}$$



随机变量的函数

$$x/y \le z \Leftrightarrow \begin{cases} x \le yz, & y > 0; \\ x \ge yz, & y < 0. \end{cases}$$



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

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

做积分变量变换,令 $x = yu$,有

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

例 3.4.11

问题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?

问题2 由统计物理学,气体分子运动速率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?

例3.4.1 设随机变量 X 具有分布律

X	$-\frac{\pi}{2}$	0	$\frac{\pi}{2}$	π
$P\{X=x_i\}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

求：$Y=2X$ 以及 $Z=\sin X$ 的分布律。

解 首先由 X 的可能取值确定 Y 及 Z 的取值:

X	$-\frac{\pi}{2}$	0	$\frac{\pi}{2}$	π
$Y=2X$	$-\pi$	0	π	2π
$Z=\sin X$	-1	0	1	0

得到随机变量函数 Y 及 Z 的分布律为:

Y	$-\pi$	0	π	2π
$P\{Y = y_j\}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

Z	-1	0	1
$P\{Z = z_k\}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{3}{8}$

例3.4.2 设随机变量 X 具有分布函数 $F_X(x)$，试求$Y=X^2$ 的分布函数.

当 $y<0,\quad F_Y(y)=P\{Y\le y\}=0$

当 $y\ge 0$

$$F_Y(y)=P\{Y\le y\}=P\{X^2\le y\}$$

$$=P\{-\sqrt{y}\le X\le\sqrt{y}\}$$

$$=F_X(\sqrt{y})-F_X(-\sqrt{y}-0)$$

例3.4.3 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

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

XY	0	1
P	0.9	0.1

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

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

例3.4.4 设X ,Y相互独立,且$X\sim B(n_1, p)$, $Y\sim B(n_2, p)$ 则

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

证 $$P\{X=k\}=C_{n_1}^{k}p^{k}(1-p)^{n_1-k},\quad (k=0,1,\ldots,n_1)$$

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

$$P\{X+Y=m\}=\sum_{k=0}^{m}p(k)q(m-k)$$

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

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

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

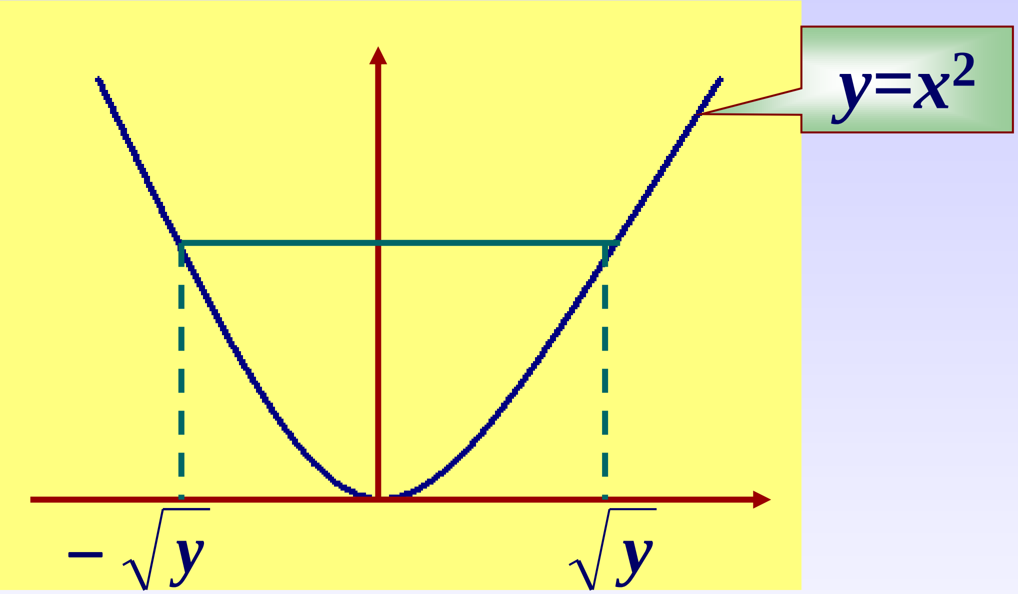
求和为 $C_{n_1+n_2}^{m}$

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

二项分布具有可加性

例3.4.5 设$X \sim N(0,1)$，求$Y=X^2$的概率密度

解



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

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

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

$$f_Y(y) = F_Y'(y)$$

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

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

称Y服从自由度为1的χ^2分布

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

例3.4.6 已知随机变量 X 的概率密度为连续函数 $f_X(x)$，求: $Y = aX + b \ (a \neq 0)$ 的概率密度 $f_Y(y)$。

解 当$a > 0$ 时,

$$F_Y(y) = P\{aX + b \leq y\} = P\{X \leq \frac{y-b}{a}\} = F_X(\frac{y-b}{a})$$

当$a < 0$ 时,

$$F_Y(y) = P\{aX + b \leq y\} = P\{X \geq \frac{y-b}{a}\} = 1 - F_X(\frac{y-b}{a})$$

对 y 求导得到

$$f_Y(y) = F_Y'(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

特别当 $X \sim N(\mu, \sigma^2)$, **则**X**的线性函数**$Y = aX + b \ (a \neq 0)$ **服从正态分布** $N(a\mu + b, \ a^2\sigma^2)$.

证

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

$$= \frac{1}{|a|\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}}$$

$$= \frac{1}{|a|\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{[y-(a\mu+b)]^2}{2a^2\sigma^2}}, \quad x \in R$$

即$Y \sim N(a\mu + b,\ a^2\sigma^2)$.

标准化变换

特别 取 $a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$, 则 $Y = \frac{X-\mu}{\sigma}$

有Y 服从标准正态分布 $N(0, 1)$，其概率密度为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, x\in R$$

- **结论：正态分布的线性函数仍然服从正态分布;**

例3.4.7 设随机变量X的概率密度为

$$f_X(x)=\begin{cases}\frac{1}{2}, & -1<x<0;\\ \frac{1}{4}, & 0\le x<2;\\ 0, & \text{其他。}\end{cases}$$

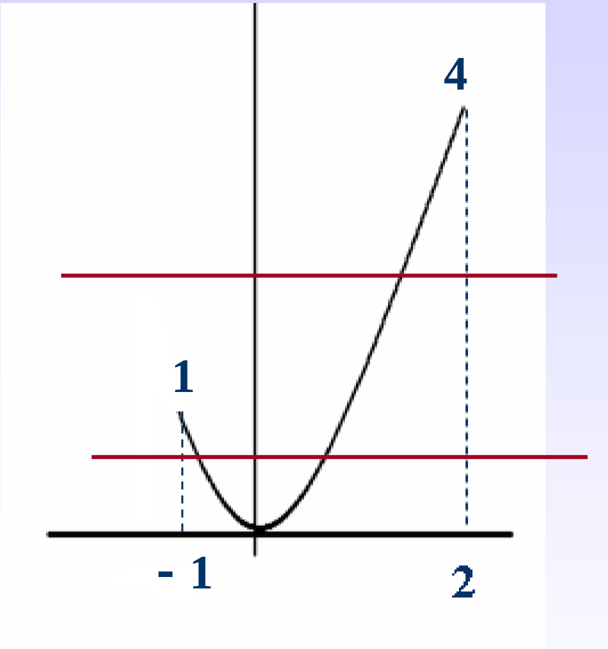
令$Y=X^2$, $F(x,y)$为二维随机变量(X,Y)的联合分布函数.

(I) 求Y的概率密度$f_Y(y)$;

(II) $F(-\frac{1}{2},4)$

[分析] 该题本质上是求一个随机变量的分布函数和概率计算问题。

[解] （I） 设Y的分布函数为



$$F_Y(y)=P\{Y\le y\}=P\{X^2\le y\}$$

$$=P\{-\sqrt{y}\le X\le\sqrt{y}\}$$

1）当 $y<0$, $F_Y(y)=0$;

2）当 $0\le y<1$,

$$F_Y(y)=\int_{-\sqrt{y}}^{0}\frac{1}{2}dx+\int_{0}^{\sqrt{y}}\frac{1}{4}dx=\frac{3}{4}\sqrt{y}$$

3）当 $1 \le y < 4$,

$$F_Y(y) = \int_{-1}^{0} \frac{1}{2} dx + \int_{0}^{\sqrt{y}} \frac{1}{4} dx = \frac{1}{4}\sqrt{y} + \frac{1}{2}$$

4）当 $4 \le y$, $F_Y(y) = 1$;

最后

$$f_Y(y) = F_Y'(y) = \begin{cases} \dfrac{3}{8\sqrt{y}}, & 0 < y < 1; \\ \dfrac{1}{8\sqrt{y}}, & 1 \le y \le 4; \\ 0, & \text{其他}. \end{cases}$$

（II） $F(-\frac{1}{2},4)$

$$= P\{X \le -\frac{1}{2}, Y \le 4\} = P\{X \le -\frac{1}{2}, X^2 \le 4\}$$

$$= P\{X \le -\frac{1}{2}, -2 \le X \le 2\} = P\{-2 \le X \le -\frac{1}{2}\}$$

$$= \int_{-1}^{-\frac{1}{2}} \frac{1}{2} dx = \frac{1}{4}$$

例3.4.8 设二维随机变量(X,Y)的联合概率密度为:

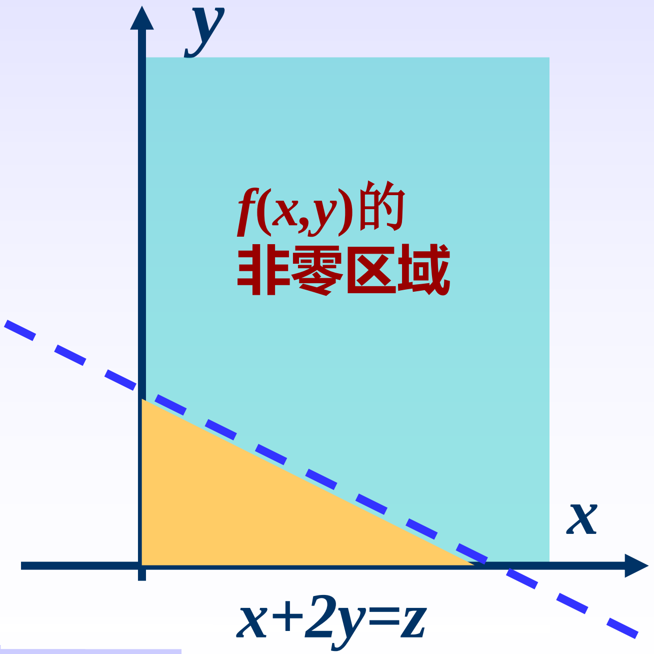
$$f(x,y)=\begin{cases}2e^{-(x+2y)}, & x,y>0;\\ 0, & \text{其它}.\end{cases}$$

求随机变量$Z=X+2Y$的分布函数和概率密度.

解 $F_Z(z)=P\{Z\le z\}$

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

$$=\iint\limits_{x+2y\le z} f(x,y)\,dx\,dy$$



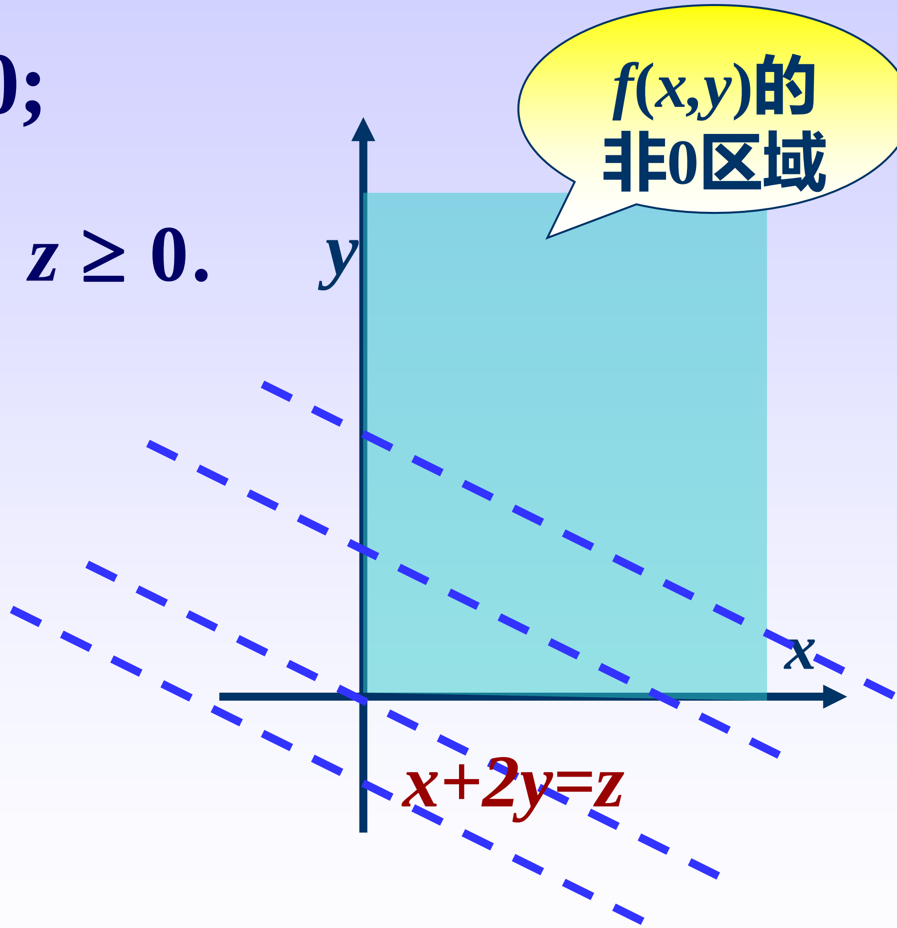
$$\iint\limits_{x+2y\le z} f(x,y)\,dxdy$$

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

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

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

$$=\begin{cases} 0, & z<0; \\ ze^{-z}, & z\ge 0. \end{cases}$$



例3.4.9 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$f_Z(z)$.

解 $\because$ 随机变量X,Y相互独立，

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

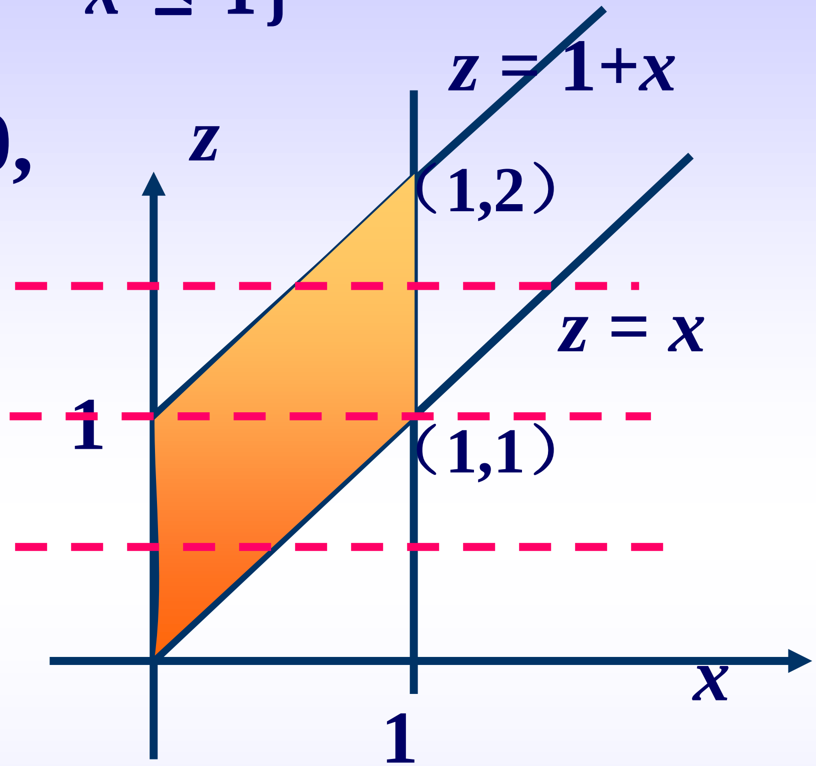
则 $$f_X(x)f_Y(z-x)=\begin{cases}0 & (x,z)\notin G\\ 1 & (x,z)\in G\end{cases}$$

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

在XOZ平面上作出区域G

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

当$z\le 0$ 或 $z>2$ 时 $f_Z(z)=0$,



当$0<z\le 1$时

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

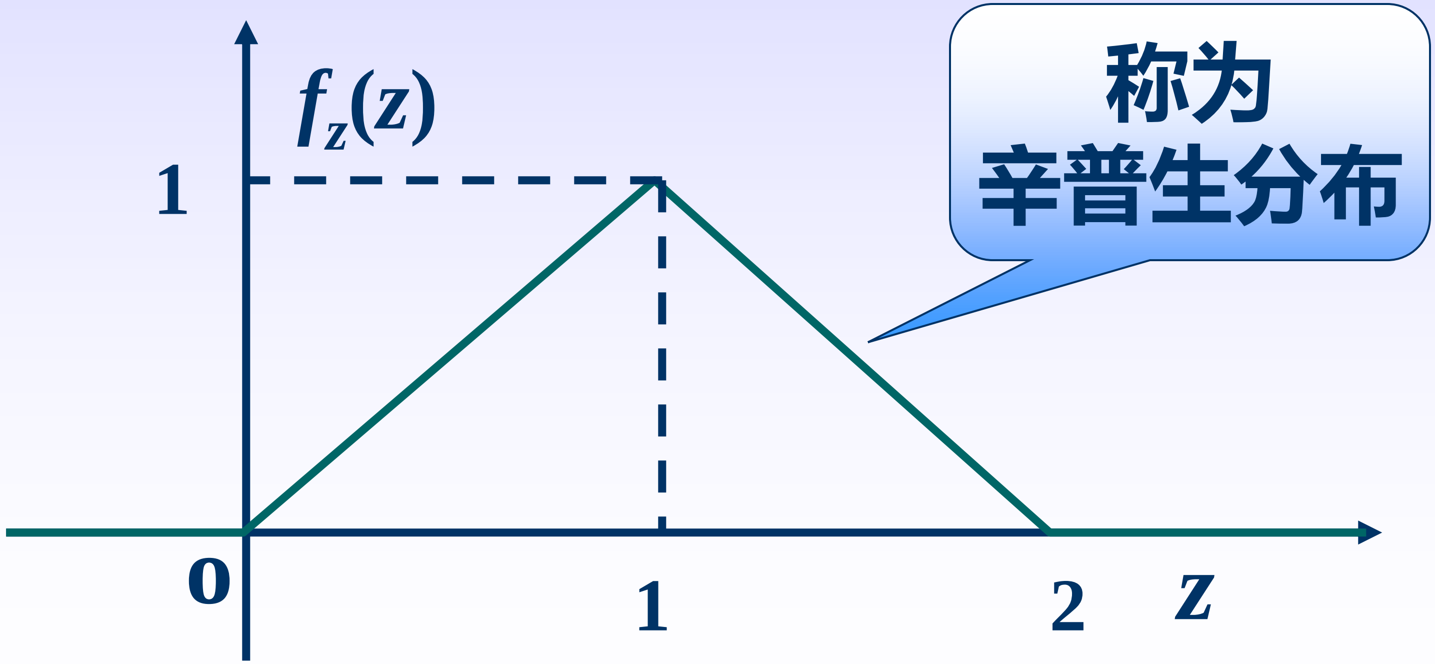
当$1<z\le 2$时

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

综上得$Z=X+Y$的概率密度为：

$$f_z(z)=\begin{cases} z, & 0<z\le 1;\\ 2-z, & 1<z\le 2;\\ 0, & \text{其他}. \end{cases}$$

概率密度曲线为



例3.4.10 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

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

其中:

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

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

在XOZ平面上作出G区域

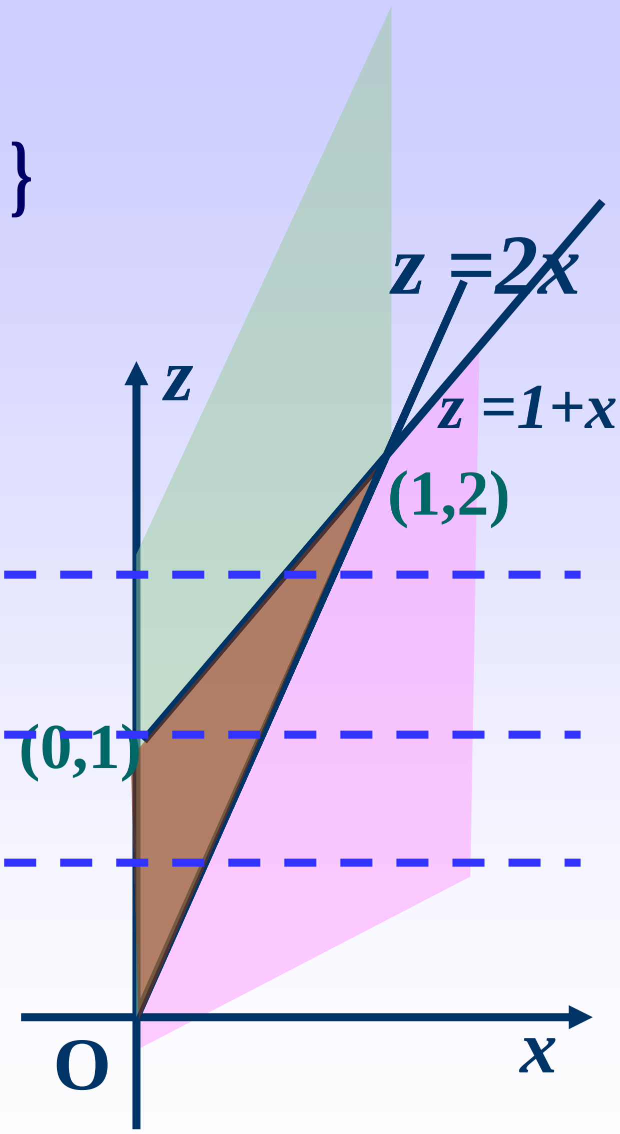
当$z\le 0$ 或 $z>2$ 时 $f_Z(z)=0$

当$0<z\le 1$时

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

当$1<z\le 2$时

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



$$= 2z - z^2,$$

综上得 $Z = X + Y$ 的概率密度为:

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

例3.4.11 **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Z=X/Y的分布.

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

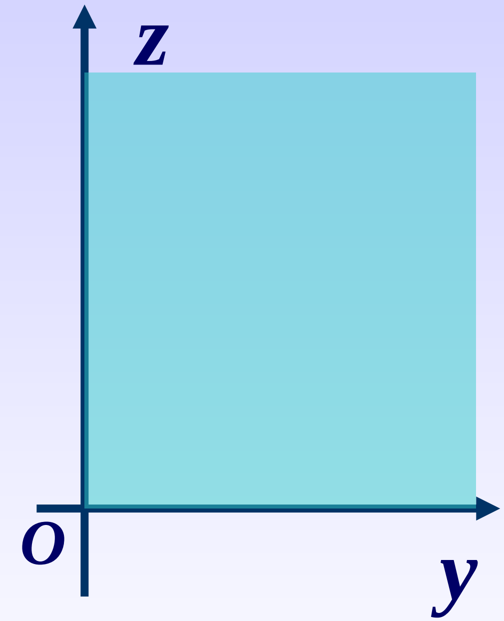
因为X,Y相互独立同分布，所以

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

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

令 $G=\{(y,z):yz>0,\ y>0\}$

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



在YOZ平面上作出G区域

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

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

$$=\begin{cases}\int_0^{+\infty}ye^{-yz-y}dy, & z>0;\\ 0, & \text{其它}.\end{cases}$$

$$= \begin{cases} \int_0^{+\infty} y e^{-yz-y} dy, & z > 0; \\ 0, & 其它. \end{cases}$$

分部积分

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