二维连续型随机变量函数的分布

问题：已知二维随机变量(X,Y)的密度函数，$g(x,y)$为已知的二元函数，$Z=g(X,Y)$

求：Z的密度函数

方法：

- 先求Z的分布函数，将Z的分布函数转化为(X,Y)的事件

(1) 和的分布：$Z = X + Y$

设(X, Y)为连续型随机变量，联合密度函数为$f(x,y)$, 则

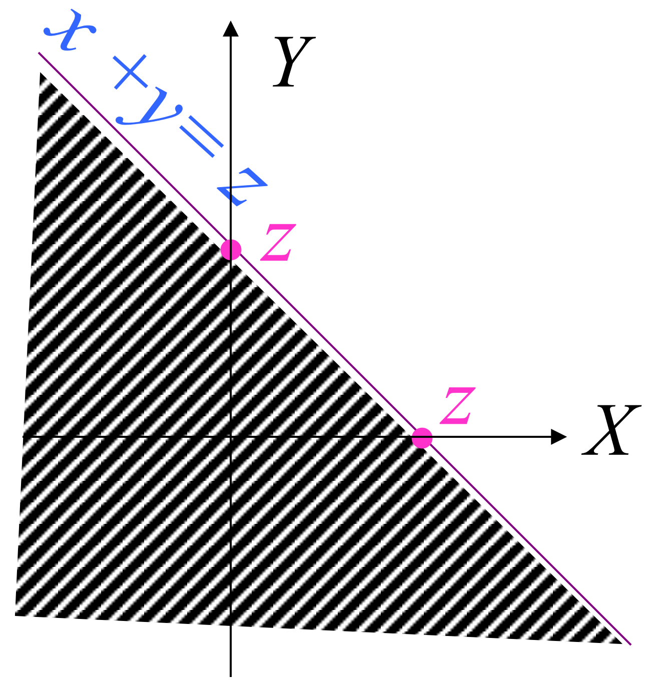
$$F_Z(z) = P(Z \le z)$$

$$= P(X + Y \le z)$$

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

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

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



$$-\infty < z < +\infty$$

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)dx \qquad -\infty<z<+\infty \quad (1)$$

$$=\int_{\overline{AB}} f(x,y)dx$$

$\overline{AB}$ 为有向直线 $x+y=z$

或 $$f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy \qquad -\infty<z<+\infty \quad (2)$$

$$=\int_{\overline{BA}} f(x,y)dy$$

特别地，若X,Y相互独立，则

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

$$-\infty<z<+\infty \quad (3)$$

或 $$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,dy \overset{\text{记作}}{=} f_X(z) * f_Y(z)$$

$$-\infty < z < +\infty \qquad (4)$$

称之为函数 $f_X(z)$ 与 $f_Y(z)$ 的卷积

正态随机变量的情形

- 若X,Y相互独立, $X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$

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

$$kX+c \sim N(k\mu_1+c,k^2\sigma_1^2)$$

$$aX+bY+c \sim N(a\mu_1+b\mu_2+c,a^2\sigma_1^2+b^2\sigma_2^2)$$

若 $X_1, X_2, \cdots, X_n$ 相互独立,

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

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

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

- 若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$

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

对于连续型随机变量，

设 X, Y 相互独立，$X \sim F_X(x)$, $Y \sim F_Y(y)$,

$U = \max\{X, Y\}$, $V = \min\{X, Y\}$,

求 U, V 的分布函数.

$$F_U(u) = P(\max\{X,Y\} \le u)$$

$$= P(X \le u, Y \le u) = P(X \le u)P(Y \le u)$$

$$= F_X(u)F_Y(u)$$

$$F_V(v) = P(\min\{X,Y\} \le v) = 1 - P(\min\{X,Y\} > v)$$

$$= 1 - P(X > v, Y > v) = 1 - P(X > v)P(Y > v)$$

$$= 1 - (1 - F_X(v))(1 - F_Y(v))$$

推广至相互独立的 n 个随机变量的情形：

设 $X_1, X_2, \cdots, X_n$ 相互独立，且

$$X_i \sim F_i(x_i), \quad i = 1, 2, \cdots, n$$

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

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

则

$$F_U(u) = \prod_{i=1}^{n} F_i(u)$$

$$F_V(v) = 1 - \prod_{i=1}^{n} (1 - F_i(v))$$

(2) 平方和的分布： $Z=X^2+Y^2$

设(X,Y)的联合密度函数为$f(x,y)$

则 $F_Z(z)=P(X^2+Y^2\le z)$

$$=\begin{cases} 0, & z<0, \\ \iint_{x^2+y^2\le z} f(x,y)\,dxdy & z\ge 0, \end{cases}$$

$$=\begin{cases} 0, & z<0, \\ \int_0^{2\pi} d\theta \int_0^{\sqrt{z}} f(r\cos\theta, r\sin\theta)\,r\,dr, & z\ge 0, \end{cases}$$

$$f_Z(z)=\begin{cases} 0, & z<0, \\ \frac{1}{2}\int_0^{2\pi} f(\sqrt{z}\cos\theta, \sqrt{z}\sin\theta)\,d\theta, & z\ge 0, \end{cases}$$

例如，$X \sim N(0,1)$, $Y \sim N(0,1)$, X, Y 相互独立，$Z = X^2 + Y^2$, 则

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ \dfrac{1}{2}\int_0^{2\pi} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z\cos^2\theta}{2}} \cdot \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z\sin^2\theta}{2}} d\theta, & z \geq 0, \end{cases}$$

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ \dfrac{1}{2} e^{-\frac{z}{2}}, & z \geq 0, \end{cases}$$ 称为自由度为2的χ^2分布

若 $X_1, X_2, \cdots, X_n$ 相互独立，且

$$X_i \sim N(0,1), \quad i = 1,2,\cdots,n$$

则 $Z = X_1^2 + X_2^2 + \cdots + X_n^2$ 所服从的分布称为

自由度为n 的 χ^2分布

它的概率密度函数为

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}}, & z \geq 0, \end{cases}$$

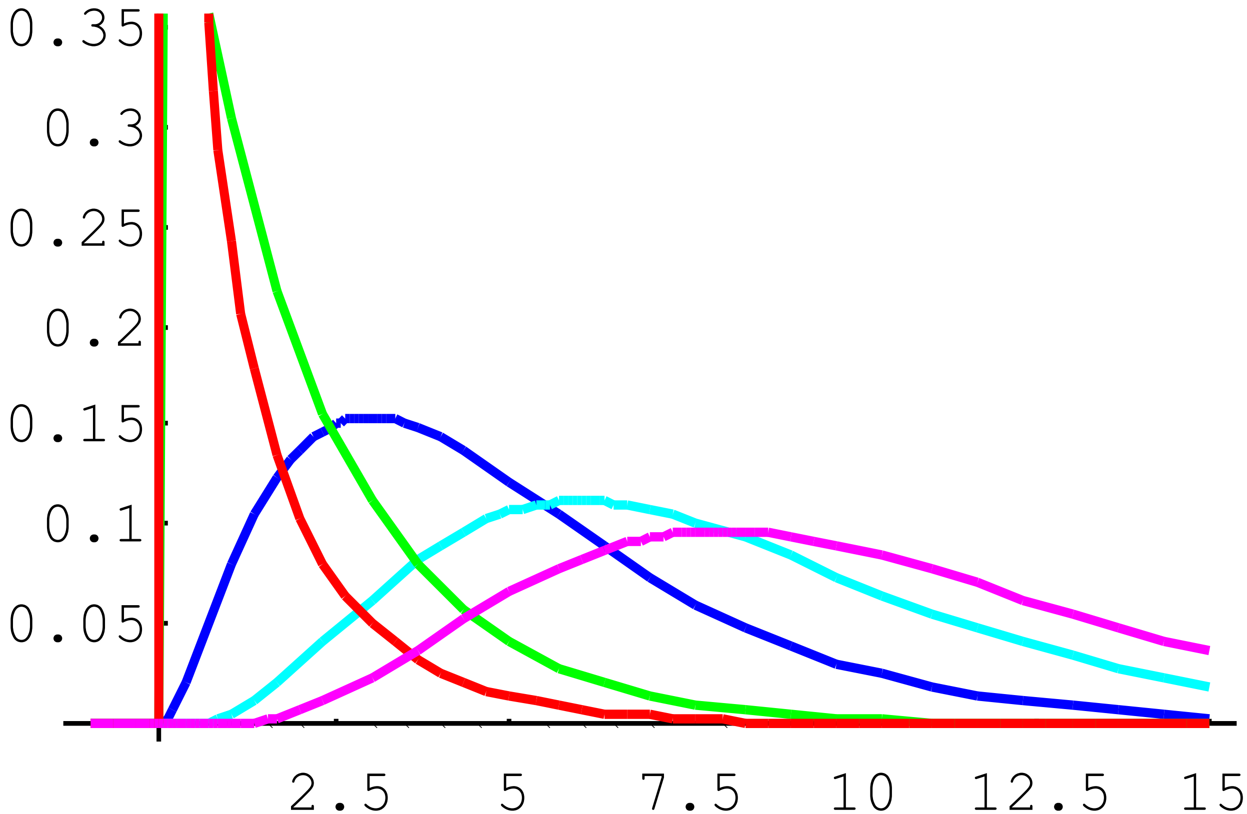
其中 $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt, \qquad x > 0$ —— 称为Γ函数

$$\Gamma(1)=1,\ \ \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi},$$

$$\Gamma(x+1)=x\Gamma(x),\ \ \Gamma(n+1)=n!$$

自由度分别为1,2,5,8,10的

χ^2分布的密度函数图形



自由度分别为1,2,5,8,10的χ^2分布的密度函数图形