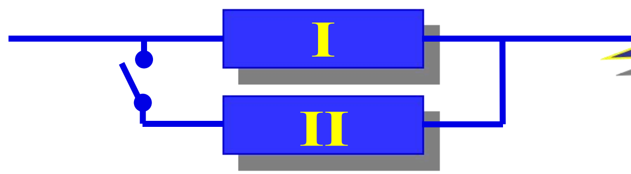
第三章 联合分布

§6 联合分布随机变量函数

实际背景

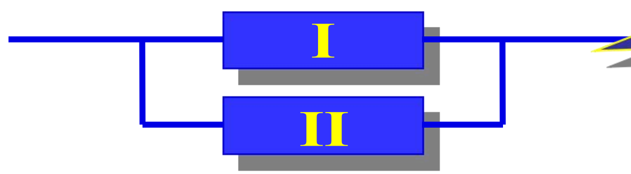
设有两个部件 I、II, 其工作寿命分别为 X, Y

冷冗余系统： 部件 I 坏了, 换上备用部件 II 继续工作



系统寿命 $X+Y$

热冗余系统： 部件I、II 并联同时工作, 仅当两个部件都损坏时，整个系统才失效



系统寿命 $\max\{X,Y\}$

串联系统： 部件I、II 串联同时工作, 只要有一个部件损坏，整个系统就失效

系统寿命 $\min\{X,Y\}$

问题 怎样确定上述各系统的寿命？

若 $(X,Y) \sim f(x,y)$, 怎样求

$X+Y$, $\max\{X,Y\}$, $\min\{X,Y\}$

的分布？

一般地 设 $z = g(x,y)$是一个二元函数,

怎样求 r.v $Z = g(X,Y)$的分布？

思路：

$$
\begin{aligned}
F_Z(z) &= P\{Z \le z\} \\
&= P\{g(X,Y) \le z\} \\
&= \iint\limits_{g(x,y)\le z} f(x,y)dxdy \\
&= \cdots = \int_{-\infty}^{z} f_Z(u)du
\end{aligned}
$$

$$\therefore \quad Z \sim f_Z(z)$$

(一) Z=X+Y 的分布(先讨论连续型)

设 $(X,Y) \sim f(x,y)$，则 $Z=X+Y$ 的分布函数为

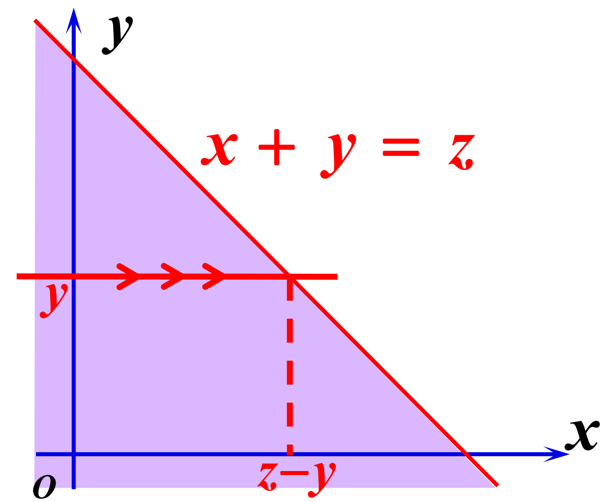
$$F_Z(z)=P\{Z\le z\}=P\{X+Y\le z\}$$

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

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

$$\overset{令x=u-y}{=\!=\!=}\int_{-\infty}^{\infty}\int_{-\infty}^{z}f(u-y,y)du\,dy$$

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



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

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

(一) $Z=X+Y$ 的分布 (先讨论连续型)

设 $(X,Y)\sim f(x,y)$，则 $Z=X+Y$ 的分布函数为

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

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

若 X,Y 相互独立, 则 $Z=X+Y$ 的密度函数为

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

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

称为**卷积(convolution)公式**, 记为

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

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

例 设r.v X,Y相互独立，且$X \sim N(0,1)$，$Y \sim N(0,1)$，求$Z = X + Y$的分布密度.

解 由独立性及卷积公式有

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

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

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

令 $t = x - \dfrac{z}{2}$

$$= \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} e^{-t^2} dt$$

$$= \frac{1}{2\pi} e^{-\frac{z^2}{4}} \sqrt{\pi} = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{z^2}{2(\sqrt{2})^2}}$$

$$\therefore \quad X + Y \sim N(0, 2).$$

什么结论？

独立正态r.v和的一般结果

1. 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

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

2. 一般地，若 $X_1, X_2, \cdots, X_n$ 相互独立, 且

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

则对于不全为零的常数 $a_1, a_2, \cdots, a_n$ 有

$$a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

独立正态r.v的非零线性组合仍服从正态分布

例 某电气设备中的两个部件存在接触电阻 R_1, R_2, 两个部件的工作状态是相互独立的, 概率密度均为

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

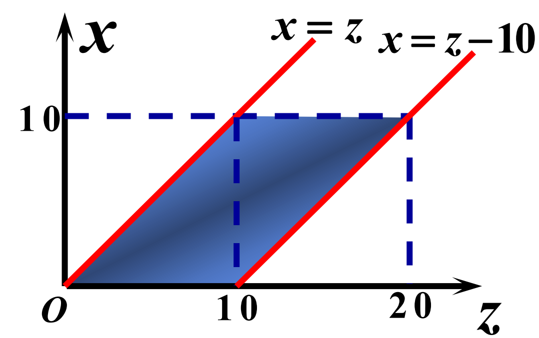
求 R_1, R_2 串联后的总电阻 $R=R_1+R_2$ 的概率密度.

解 由卷积公式有

$$f_R(z)=\int_{-\infty}^{\infty}f(x)f(z-x)dx$$

被积函数的非零区域是

$$\begin{cases}0<x<10\\ 0<z-x<10\end{cases}\Leftrightarrow\begin{cases}0<x<10\\ z-10<x<z\end{cases}$$



$$\therefore\quad f_R(z)=\begin{cases}\int_0^z f(x)f(z-x)dx, & 0\le z<10\\ \int_{z-10}^{10} f(x)f(z-x)dx, & 10\le z\le 20\\ 0, & \text{其它}\end{cases}$$

例 某电气设备中的两个部件存在接触电阻 R_1, R_2, 两个部件的工作状态是相互独立的, 概率密度均为

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

求 R_1, R_2 串联后的总电阻 $R=R_1+R_2$ 的概率密度.

解 由卷积公式有

$$f_R(z)=\int_{-\infty}^{\infty} f(x)f(z-x)dx$$

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

$$=\begin{cases}(600z-60z^2+z^3)/15000, & 0\le z<10\\ (20-z)^3/15000\ , & 10\le z<20\\ 0\ , & \text{其它}\end{cases}$$

例 设 X, Y 相互独立且都服从参数为 λ 的指数分布，求 r.v $Z = X + Y$ 的概率密度.

解 由卷积公式有，Z 的密度函数为

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

$$= \begin{cases} \int_0^z \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(z-x)} \mathrm{d}x, & z > 0 \\ 0, & z \le 0 \end{cases}$$

$$= \begin{cases} \lambda^2 z e^{-\lambda z}, & z > 0 \\ 0, & z \le 0 \end{cases} \sim \Gamma(2, \lambda)$$

研究问题
question

设 $X_1, X_2, \cdots, X_n$ 相互独立且都服从参数为 λ 的指数分布，求 $X_1 + X_2 + \cdots + X_n$ 的分布密度.

提示： 记 $X_1 + X_2 + \cdots + X_n \sim f_n(z)$，则 $f_2(z) = \lambda z f_1(z)$

设法导出递推公式，然后用归纳法证明.

(再讨论离散型)

设 X, Y 相互独立, 其频率函数分别为

$$P\{X = i\} = p_i \quad (i = 1, 2, \cdots)$$
$$P\{Y = j\} = q_j \quad (j = 1, 2, \cdots)$$

令 $Z = X + Y$, 则

$$P\{Z = k\} = \sum_{i=1}^{k-1} P\{X = i\} \cdot P\{Y = k - i\}$$
$$= \sum_{i=1}^{k-1} P\{X = k - i\} \cdot P\{Y = i\}$$
$$(k = 1, 2, \ldots)$$

(离散卷积公式)

比较一下连续型卷积公式

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

例 设X, Y独立,且$X \sim P(\lambda_1), Y \sim P(\lambda_2)$, 求$Z = X + Y$的分布.

解 由离散卷积公式有

$$P\{Z = k\} = \sum_{i=0}^{k} P\{X = k - i\} \cdot P\{Y = i\}$$

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

$$= e^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{k} \frac{1}{i!(k-i)!} \lambda_1^{k-i} \cdot \lambda_2^{i}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \lambda_1^{k-i} \cdot \lambda_2^{i}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!} \qquad (k = 0, 1, 2, \cdots)$$

$\therefore \quad Z \sim P(\lambda_1 + \lambda_2).$

(二) $Z=\dfrac{X}{Y}$ 的分布

例 设 X,Y 独立同分布，其密度函数为

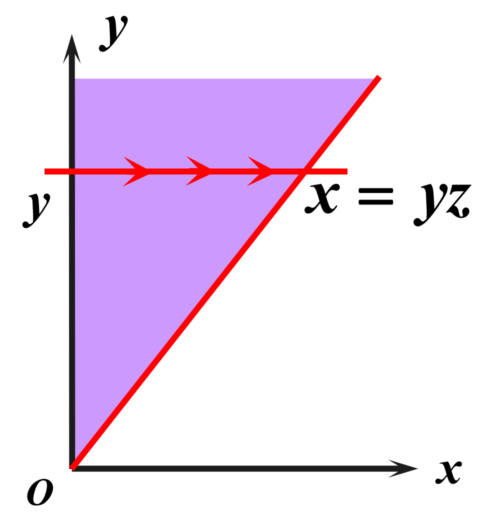
$$f(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & x\le 0, \end{cases}$$

直接法

求 $Z=X/Y$ 的概率密度.

解 当 $z\ge 0$ 时，Z 的分布函数为

$$F_Z(z)=P\{X/Y\le z\}$$
$$=\iint\limits_{\begin{cases} x/y\le z \\ x>0,\,y>0\end{cases}} e^{-(x+y)}dxdy$$
$$=\int_0^{\infty}dy\int_0^{yz}e^{-(x+y)}dx$$
$$=\int_0^{\infty}e^{-y}(1-e^{-yz})dy=1-\frac{1}{1+z}$$

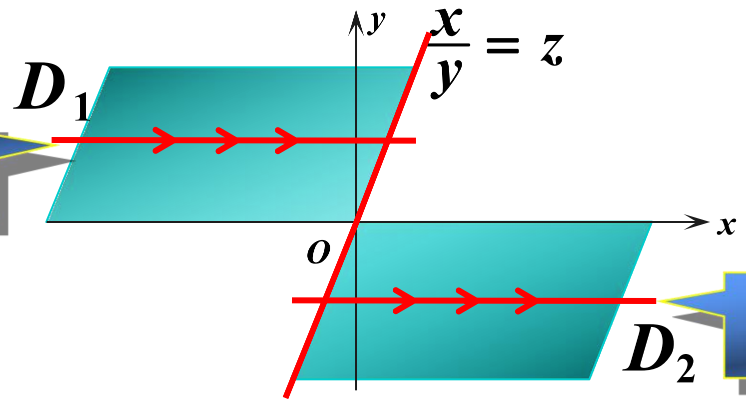


$$\therefore\quad f_Z(z)=\begin{cases}(1+z)^{-2}, & z>0 \\ 0, & z\le 0\end{cases}$$

(二) $Z=\dfrac{X}{Y}$ 的分布

设$(X,Y)\sim f(x,y)$, 则 $Z=X/Y$ 的分布函数为

$$F_Z(z)=P\{X/Y\le z\}=\iint_{\frac{x}{y}\le z} f(x,y)\mathrm{d}x\mathrm{d}y$$



当$y>0$, 沿这条直线积分

当$y<0$, 沿这条直线积分

$$F_Z(z)=\int_0^{+\infty}\mathrm{d}y\int_{-\infty}^{yz}f(x,y)\mathrm{d}x+\int_{-\infty}^{0}\mathrm{d}y\int_{yz}^{+\infty}f(x,y)\mathrm{d}x$$

由于积分区域不是矩形

这样计算积分比较繁!

回顾

【二重积分的变量替换（雅可比式）】

若连续可微分的函数

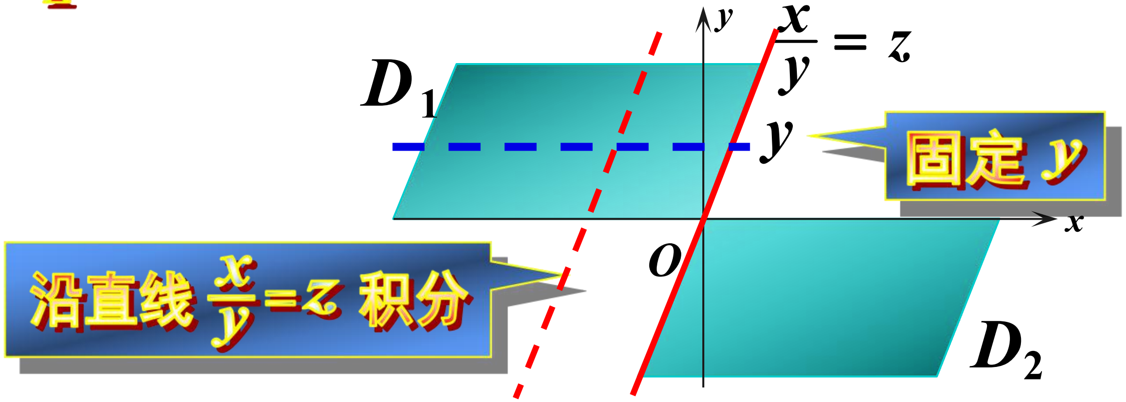
$$\begin{cases} x = x(u,v) \\ y = y(u,v) \end{cases}$$

把平面 Oxy 上的有界闭区域 Ω 单值映射到平面 $O'uv$ 上的闭区域 Ω'，其雅可比式为

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \neq 0$$

则 $$\iint_{\Omega} f(x,y)dxdy = \iint_{\Omega'} f[x(u,v),y(u,v)]\left|J\right|dudv$$

（二） $Z=\dfrac{X}{Y}$ 的分布



令 $\begin{cases} x/y=u \\ y=y \end{cases}$，则变换的Jacobi式为 $\dfrac{\partial(x,y)}{\partial(u,y)}=\begin{vmatrix} y & u \\ 0 & 1 \end{vmatrix}=y$

$$\therefore\quad F_Z(z)=\int_{-\infty}^{z}\left(\int_{-\infty}^{+\infty} f(uy,y)\,|y|\,\mathrm{d}y\right)\mathrm{d}u$$

$$\therefore\quad f_Z(z)=\int_{-\infty}^{+\infty} f(zy,y)\,|y|\,\mathrm{d}y$$

特别当 X,Y 独立时，则有

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(zy)f_Y(y)\,|y|\,\mathrm{d}y$$

例 设 X, Y 相互独立同服从正态分布 $N(0,1)$，求 r.v $Z = Y / X$ 的概率密度.

公式法

$$Z = \frac{Y}{X}$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, xz)\,|x|\,\mathrm{d}x$$

解
$$f_Z(z) = \int_{-\infty}^{\infty} \frac{|x|}{2\pi} e^{-x^2/2} e^{-x^2 z^2/2} \mathrm{d}x$$

$$= \frac{1}{\pi} \int_0^{\infty} x e^{-x^2((z^2+1)/2)} \mathrm{d}x$$

作变量替换 $u = x^2$，得到

$$f_Z(z) = \frac{1}{2\pi} \int_0^{\infty} e^{-u((z^2+1)/2)} \mathrm{d}u$$

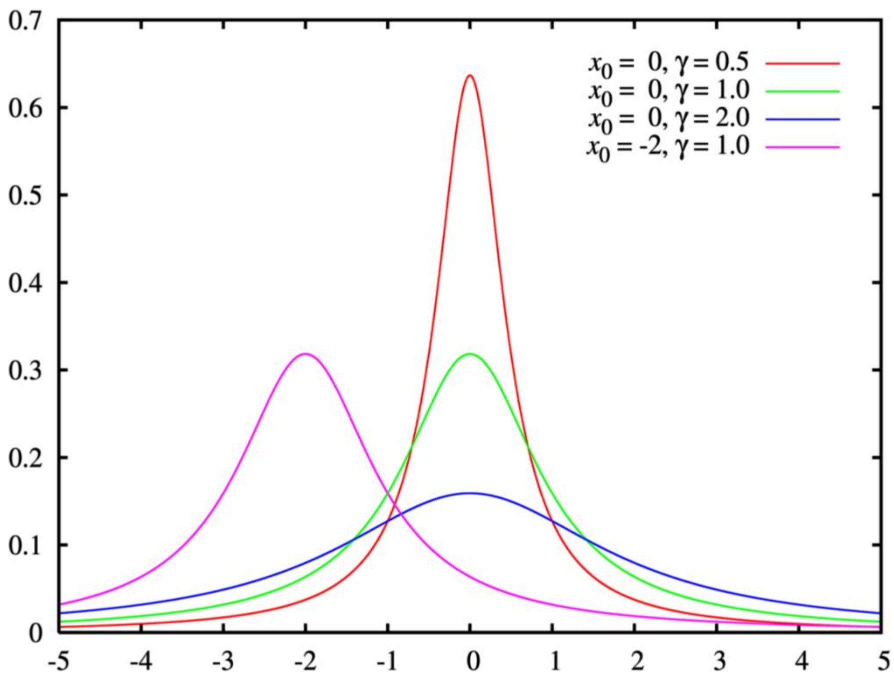
利用 $\int_0^{\infty} \lambda e^{-\lambda x} \mathrm{d}x = 1,\ \lambda = (z^2+1)/2$，得到

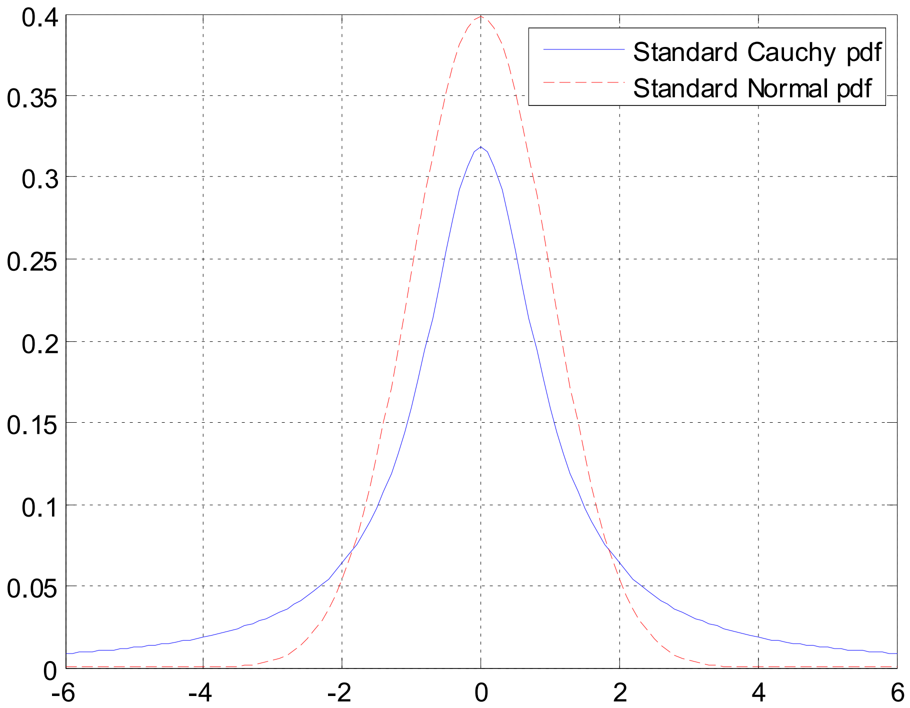
$$\therefore \quad f_Z(z) = \frac{1}{\pi(z^2+1)}, \quad -\infty < z < \infty.$$

(标准柯西Cauchy分布)

Cauchy分布

密度函数 $$f(x;x_0,\gamma)=\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]},\quad -\infty<x<\infty.$$





(二) $Z=\dfrac{X}{Y}$ 的分布

例 设 X, Y 独立同分布，其密度函数为

$$f(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & x\le 0, \end{cases}$$

公式法

求 $Z=X/Y$ 的概率密度.

解

当 X,Y 独立时，有

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(zy)f_Y(y)\,|y|\,\mathrm{d}y$$

$$f_Z(z)=\begin{cases} (1+z)^{-2}, & z>0 \\ 0, & z\le 0 \end{cases}$$

(三) 两个随机变量变换的分布

例 设 ξ, η 为独立同分布均服从指数分布的 r.v, 其密度为

$$f(x)=\begin{cases} e^{-x}, x \ge 0, \\ 0, \quad x<0. \end{cases}$$

试求$U=\xi+\eta, V=\xi/\eta$的联合密度，并证明U, V相互独立.

解 $F_{UV}(u,v)=P\{\xi+\eta \le u, \xi/\eta \le v\}$

$$=\iint_{\begin{cases} x+y\le u \\ x/y \le v \end{cases}} f(x)f(y)\mathrm{d}x\mathrm{d}y$$

令$x+y=\tilde{u}, x/y=\tilde{v}$, 则 $x=\dfrac{\tilde{u}\tilde{v}}{1+\tilde{v}}, y=\dfrac{\tilde{u}}{1+\tilde{v}}$,

变换的Jacobi式为

$$J=\frac{\partial(x,y)}{\partial(\tilde{u},\tilde{v})}=-\frac{\tilde{u}}{(1+\tilde{v})^2} \qquad (\tilde{u}\ge 0, \tilde{v}\ne -1)$$

(三) 两个随机变量变换的分布

例 设 ξ, η 为独立同分布均服从指数分布的 r.v, 其密度为

$$f(x)=\begin{cases} e^{-x}, x \ge 0, \\ 0, \quad x<0. \end{cases}$$

试求$U=\xi+\eta, V=\xi/\eta$的联合密度，并证明U, V相互独立.

解（续） $\therefore F_{UV}(u,v)=\iint\limits_{D:\begin{cases} x+y\le u \\ x/y\le v, x\ge 0, y\ge 0\end{cases}} e^{-(x+y)}\mathrm{d}x\mathrm{d}y$

$$=\int_0^v\int_0^u e^{-\tilde{u}}|J|\mathrm{d}\tilde{u}\mathrm{d}\tilde{v}$$

$$\therefore \quad f_{UV}(u,v)=\frac{ue^{-u}}{(1+v)^2} \quad (u\ge 0, v\ge 0)$$

$\because \quad f_{UV}(u,v)$可表为$g(u)h(v)$ $\quad \therefore U, V$相互独立.

例 假设X_1和X_2相互独立同服从标准正态分布N(0,1),

且 $Y_1=X_1$,

$Y_2=X_1+X_2$,

可以证明

$$(Y_1,Y_2)\sim N(0,0,1,2,\sqrt{\frac{1}{2}}).$$

推广: 两个独立标准正态r.v.的线性变换服从二元正态分布.

更一般地: 两个r.v.的联合分布是二元正态,则它们的非奇异线性变换还是二元正态分布.

(四)随机变量的其它函数

例 设 X,Y 相互独立同服从正态分布 $N(0,\sigma^2)$, 求 r.v $Z=\sqrt{X^2+Y^2}$ 的概率密度.

解 当 $z\geq 0$ 时, Z 的分布函数为

$$F_Z(z)=P\{\sqrt{X^2+Y^2}\leq z\}=\iint\limits_{\sqrt{x^2+y^2}\leq z} f_X(x)f_Y(y)dxdy$$

$$=\frac{1}{2\pi\sigma^2}\iint\limits_{\sqrt{x^2+y^2}\leq z} e^{-\frac{x^2+y^2}{2\sigma^2}}dxdy$$

$$\begin{cases} x=\rho\cos\varphi \\ y=\rho\sin\varphi \end{cases}$$

$(0\leq\rho\leq z, 0\leq\varphi\leq 2\pi)$

雅可比式: $J=\rho$

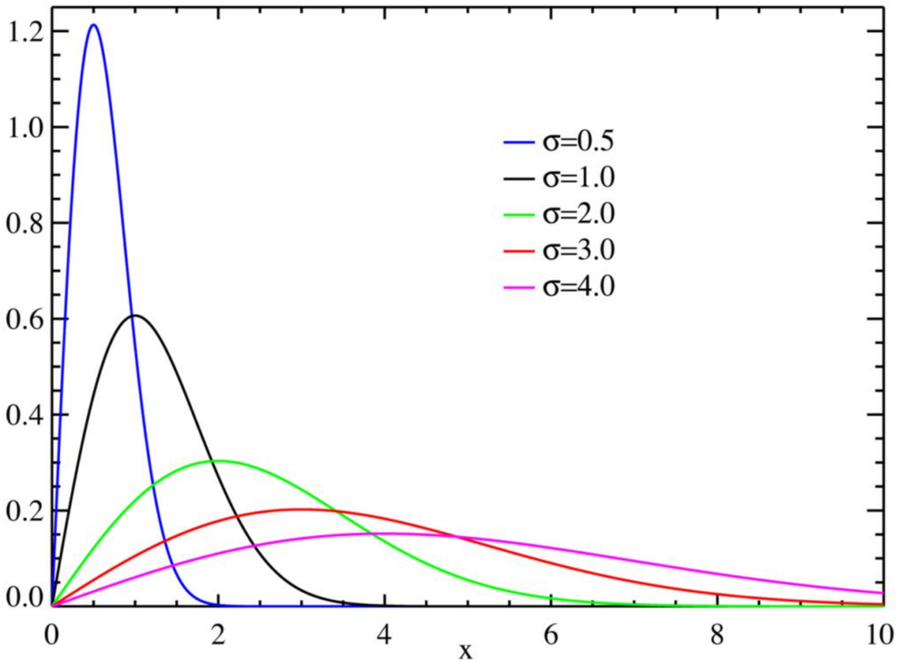
$$=\frac{1}{2\pi\sigma^2}\int_0^z d\rho\int_0^{2\pi}\rho e^{-\frac{\rho^2}{2\sigma^2}}d\varphi$$

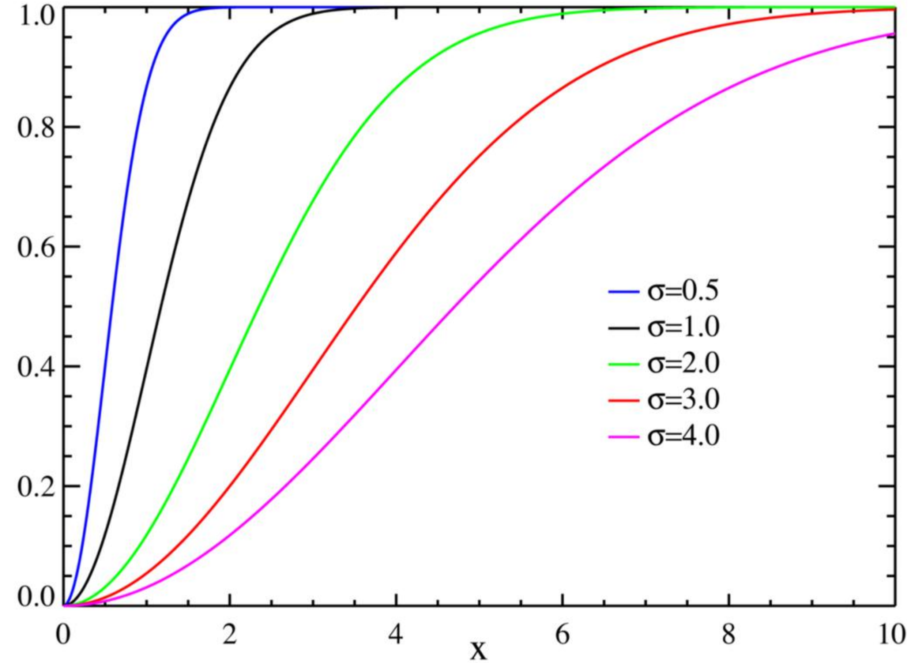
$$=\int_0^z e^{-\frac{\rho^2}{2\sigma^2}}d\frac{\rho^2}{2\sigma^2}=1-e^{-\frac{z^2}{2\sigma^2}} \quad (z\geq 0)$$

$$\therefore\quad f_Z(z)=\begin{cases}\dfrac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma^2}}, & z\geq 0\\ 0, & z<0\end{cases}$$

(瑞利Rayleigh分布)

Rayleigh分布





$$f_Z(z)=\begin{cases}\dfrac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma^2}}, & z\ge 0\\ 0\ , & z<0\end{cases}\qquad F_Z(z)=\begin{cases}1-e^{-\frac{z^2}{2\sigma^2}}, & z\ge 0\\ 0, & z<0\end{cases}$$

当一个二维随机向量的两个分量呈独立的、有着相同的方差的正态分布时，这个向量的模呈Rayleigh瑞利分布.

课后作业

P79：43、44、51、52、57，补充题1，2

1. 假设 X 和 Y 是两个独立的随机变量，服从标准正态分布 N(0,1)，令 U=X+Y, V=X-Y. 求 U 和 V 的边缘密度函数及联合密度函数，并讨论独立性。

2. 设二维连续随机变量(X,Y)的概率密度为

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

(1)求边缘密度函数;

(2)Z=2X-Y 的概率密度函数;

(3)P(Y<1/2|X<1/2).

END