第四章 正态分布 附录
正态分布性质的证明

刘 春 光

暨南大学数学系

2018年4月

目录

二维正态分布

Definition (二元正态分布(Bivariate normal distribution))

以以下函数为密度的分布称为二元正态分布，

简记为$N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

其中μ_1, μ_2为实数，$\sigma_1, \sigma_2 > 0$，$|\rho| < 1$。

二维正态分布

Theorem (二维正态分布的数字特征)

设随机向量$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，则

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

且

$$R(X, Y) = \rho.$$

证明-二维正态分布的边缘分布-1/3

令

$$u(x,y)=\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]$$

则二维正态分布的联合密度函数可写为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot e^{-u(x,y)}.$$

X的边缘密度为

$$f_X(x)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}e^{-u(x,y)}dy.$$

证明-二维正态分布的边缘分布-2/3

对$u(x,y)$中含有y的项进行配方得

$$
\begin{aligned}
u(x,y) &= \frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right] \\
&= \frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{1}{2(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2} - \rho\frac{x-\mu_1}{\sigma_1}\right]^2.
\end{aligned}
$$

对积分$\int_{-\infty}^{+\infty} e^{-u(x,y)}dy$做变量替换

$$
t = \frac{1}{\sqrt{1-\rho^2}}\left[\frac{y-\mu_2}{\sigma_2} - \rho\frac{x-\mu_1}{\sigma_1}\right],
$$

证明-二维正态分布的边缘分布-3/3

对积分$\int_{-\infty}^{+\infty} e^{-u(x,y)}dy$做变量替换

$$t = \frac{1}{\sqrt{1-\rho^2}}\left[\frac{y-\mu_2}{\sigma_2} - \rho\frac{x-\mu_1}{\sigma_1}\right],$$

得

$$f_X(x) = \frac{1}{2\pi\sigma_1}\cdot\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}\int_{-\infty}^{+\infty} e^{-t^2/2}dt$$
$$= \frac{1}{\sqrt{2\pi}\sigma_1}\cdot\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}.$$

证明-二维正态分布的相关系数-1/3

令

$$u(x,y)=\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]$$

则X,Y的协方差可写为

$$\mathrm{cov}(X,Y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\iint_{\mathbb{R}^2}(x-\mu_1)(y-\mu_2)e^{-u(x,y)}dxdy,$$

相关系数为

$$R(X,Y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\iint_{\mathbb{R}^2}\frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2}e^{-u(x,y)}dxdy.$$

证明-二维正态分布的相关系数-2/3

注意到

$$u(x,y)=\frac{(x-\mu_1)^2}{2\sigma_1^2}+\frac{1}{2(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2,$$

对积分

$$R(X,Y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\iint_{\mathbb{R}^2}\frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_2}e^{-u(x,y)}dxdy,$$

做变量替换$x_1=\frac{x-\mu_1}{\sigma_1}$, $y_1=\frac{1}{\sqrt{1-\rho^2}}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]$，得

$$\begin{aligned}R(X,Y)&=\frac{1}{2\pi}\iint_{\mathbb{R}^2}x_1(\rho x_1+\sqrt{1-\rho^2}y_1)e^{-x_1^2/2-y_1^2/2}dx_1dy_1\\&=I_1+I_2\end{aligned}$$

证明-二维正态分布的相关系数-3/3

其中

$$I_2 = \frac{1}{2\pi}\iint_{\mathbb{R}^2}\sqrt{1-\rho^2}x_1y_1e^{-x_1^2/2-y_1^2/2}dx_1dy_1 = 0,$$

$$\begin{aligned} I_1 &= \frac{1}{2\pi}\iint_{\mathbb{R}^2}\rho x_1^2e^{-x_1^2/2-y_1^2/2}dx_1dy_1 \\ &= \rho\left[\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}x_1^2e^{-x_1^2/2}dx_1\right]\cdot\left[\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-y_1^2/2}dy_1\right] \\ &= \rho. \end{aligned}$$

故有$R(X,Y)=\rho$。

目录

相互独立正态随机变量的和的分布

Theorem

设随机变量X与Y相互独立，且都服从正态分布：

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

则它们的和也服从正态分布，且有

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

证明-相互独立正态随机变量的和的分布-1/4

随机变量X与Y的概率密度分别为：

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma_1}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\},$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\cdot\sigma_2}\exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\},$$

由卷积公式，$Z = X + Y$的概率密度为：

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(z-x-\mu_2)^2}{\sigma_2^2}\right]\right\}dx \\ &= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2}(ax^2-2bx+c)\right\}dx \end{aligned}$$

证明-相互独立正态随机变量的和的分布-2/4

$Z = X + Y$的概率密度为：

$$f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}(ax^2 - 2bx + c)\right\} dx,$$

其中

$$a = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2},$$
$$b = \frac{\mu_1}{\sigma_1^2} + \frac{z-\mu_2}{\sigma_2^2},$$
$$c = \frac{\mu_1^2}{\sigma_1^2} + \frac{(z-\mu_2)^2}{\sigma_2^2}.$$

证明-相互独立正态随机变量的和的分布-3/4

对积分

$$f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}(ax^2 - 2bx + c)\right\} dx,$$

做换元法$t = \sqrt{a}(x - \frac{b}{a})$，得：

$$f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot \frac{1}{\sqrt{a}} \exp\left\{-\frac{ac - b^2}{2a}\right\} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}t^2\right\} dt$$

$$= \frac{1}{\sqrt{2\pi a}\sigma_1\sigma_2} \exp\left\{-\frac{1}{2} \cdot \frac{ac - b^2}{a}\right\}.$$

证明-相互独立正态随机变量的和的分布-4/4

由

$$\frac{1}{\sqrt{2\pi a}\sigma_1\sigma_2} = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}},$$

$$\frac{ac-b^2}{a} = \frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2},$$

得

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} \exp\left\{-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right\},$$

即$Z \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$。