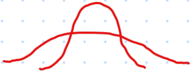
函数图像



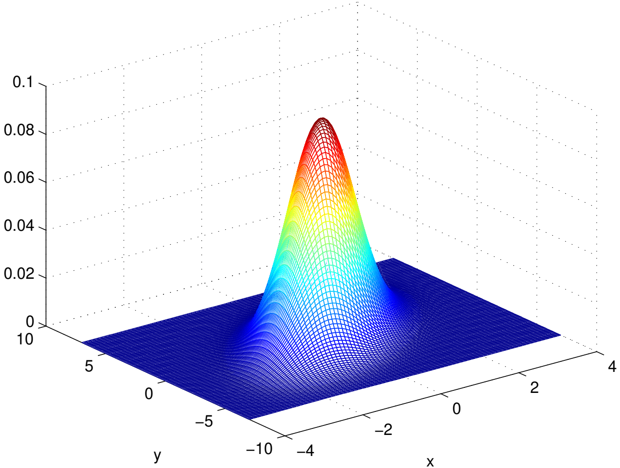


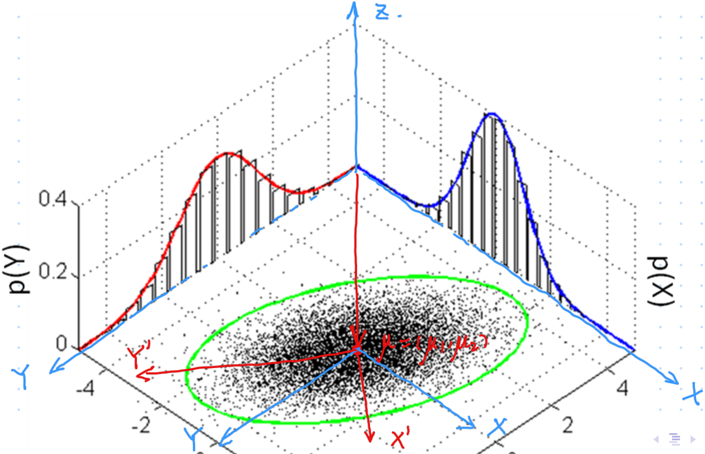
图: $\mu_1 = \mu_2 = 0,\ \sigma_1 = 1,\ \sigma_2 = 2,\ \rho = \frac{1}{2}$

边缘分布

二维正态分布的边缘分布为一维正态分布，即

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

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



边缘分布

$\rho = 0$ 是 X, Y 是相互独立随机变量的充要条件。

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left(\frac{[x-\mu_1)^2}{\sigma^2} + \frac{[y-\mu_2]^2}{\sigma^2}\right)}$$

$$f_{X|Y}(x|y) = f(x)$$

n 维正态分布（非退化情形）

$A: \lambda_i > 0 \quad \forall i=1,\dots,n.$

$A^{-1}: \lambda_i^{-1} > 0$

$X^T A X > 0.$

设 $\mu = (\mu_1, \mu_2, \cdots, \mu_n)'$，$\boldsymbol{\Sigma}$ 为 n 阶正定矩阵，记 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)'$，若

退化情形

$X \sim N(0,1)$

$(X,Y) \quad f_{X,Y}(x,y) = \begin{cases} f_X(x) & y=0 \\ 0 & y\neq 0 \end{cases}$

$$f_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

则称 $(X_1, X_2, \cdots, X_n)'$ 服从 n 维正态分布，记作 $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$。

$\Sigma^{-1} = (A^T \Lambda A)^{-1} = A^T \cdot \Lambda^{-1} \cdot A$

取 $y_i = x_i - \mu_i \quad \forall i = 1, \dots, n$

$$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) \overset{y = x-\mu}{=} \frac{1}{\left((2\pi)^n |\Sigma|\right)^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2} y^T \Sigma^{-1} y\right\}$$

（取 $\Sigma = A^T \Lambda A$ where $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

$\sigma_i > 0$ s.t. $\sigma_i^2 = \lambda_i$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \prod \sigma_i} \cdot \exp\left\{-\frac{1}{2} (Ay)^T \Lambda^{-1} (Ay)\right\}$$

n 维正态分布（非退化情形）

请大家验证，对于 $n = 2$ 的情形

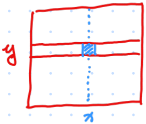
$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

二维连续型随机变量的条件密度函数

设 (X, Y) 为二维连续型的随机向量，其联合分布密度函数为 $f_{X,Y}(x, y)$，边缘分布密度函数分别为 $f_X(x)$ 和 $f_Y(y)$，若 $\boxed{f_Y(y) > 0}$，那么在 $\{Y = y\}$ 发生的条件下，X 的 条件密度函数 定义为

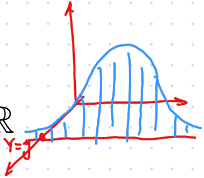
fixed.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \forall x \in \mathbb{R}$$



同样的，当 $f_X(x) > 0$，在 $\{X = x\}$ 发生的条件下 Y 的 条件密度函数 定义为

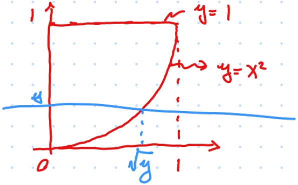
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \forall y \in \mathbb{R}$$



例

设 (X, Y) 的联合密度函数为

$$f_{X,Y}(x,y) = \begin{cases} 6xy, & x^2 \le y \le 1, 0 \le x \le 1 \\ 0, & \text{其他} \end{cases}$$



求 $f_{X|Y}(x|y=\frac{1}{2})$ 和 $f_{Y|X}(y|x=\frac{1}{3})$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 0 & y<0 \text{ 或 } y>1 \\ \frac{6xy}{3y^2} = \frac{2x}{y} & 0 \le y \le 1 \end{cases}$$

其中 $f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx$

$$= \begin{cases} 0, & y<0 \text{ 或 } y>1 \\ \int_0^{\sqrt{y}} 6xy\,dx & 0 \le y \le 1 \end{cases}$$

$$= 3y \int_0^{\sqrt{y}} dx^2$$

$$= 3y^2$$

例

大家可以自行验证，对于二维正态分布

$$(X,Y) \sim \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$

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

X 在 $\{Y=y\}$ 下和 Y 在 $\{X=x\}$ 下的条件分布分别为

$$N(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2))$$

和

$$N(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2))$$

Outline

简单情形

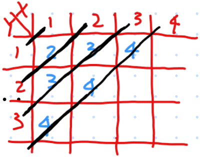
问题：根据 (X, Y) 的联合分布，求 $g(X, Y) = X + Y$ 的分布

离散型的简单情形

对于二维离散型随机向量的情形，设 (X, Y) 的分布列为

$$P(X = x_i, Y = y_j) = p_{ij}, i, j = 1, 2, \cdots$$

那么 $Z = X + Y$ 的分布列为



$$P(Z = z_k) = \sum_{x_i + y_j = z_k} p_{ij}, k = 1, 2, \cdots$$

特别的，若 $x_i = i$，$y_j = j$，那么

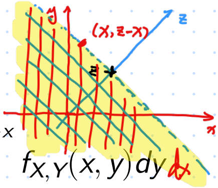
$$P(Z = k) = \sum_{i=0}^{k} P(X = i, Y = k - i) = \sum_{i=0}^{k} p_{i(k-i)}, k = 0, 1, 2, \cdots$$

进一步的，若 X 与 Y 独立，那么

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

连续型的简单情形

对于二维连续型随机向量，设 (X, Y) 的联合分布密度函数为 $f_{X,Y}$，那么 $Z = X + Y$ 的分布函数为



$$
\begin{aligned}
F_Z(z) &= P(Z \le z) = P(X + Y \le z) \\
&= \int_{x+y\le z} f_{X,Y}(x, y)dxdy = \int_{-\infty}^{\infty} x \int_{-\infty}^{z-x} f_{X,Y}(x, y)dy\ dx \\
&= \int_{-\infty}^{\infty} y \int_{-\infty}^{z-y} f_{X,Y}(x, y)dx\ dy
\end{aligned}
$$

连续型的简单情形

按照连续型随机变量函数的分布密度函数求法，对分布函数 $F_Z(z)$ 关于 z 求导，得到 Z 的分布密度函数为

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

特别的，若 X, Y 相互独立，那么

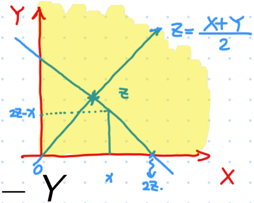
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x)dx = \int_{-\infty}^{\infty} f_X(z - y)f_Y(y)dy$$

即两个独立的随机变量和的分布密度函数为它们各自的分布密度函数的“卷积”。✓

练习

设 X, Y 的联合密度函数为

$$f_{X,Y}(x,y)=\begin{cases} e^{-(x+y)}, & x>0, y>0 \\ 0, & \text{其他} \end{cases}$$



求下面的随机变量的密度函数:

$$Z=(X+Y)/2, \quad W=X-Y$$

分布函数 $F_Z(z)=P(Z\le z)$

$=P(\frac{X+Y}{2}\le z)$ (得 $Y\le 2z-X$)

$=\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{2z-x} f_{X,Y}(x,y)dy\right)dx$

$=\int_0^{2z}\int_0^{2z-x} e^{-(x+y)}dy\,dx$

$=\begin{cases} 0 & z\le 0 \\ \int_0^{2z} e^{-x}\left(\int_0^{2z-x} e^{-y}dy\right)dx = \int_0^{2z} e^{-x}\cdot\underbrace{(-e^{-y})\Big|_0^{2z-x}}_{=1-e^{x-2z}}dx = \int_0^{2z} e^{-x}-e^{-2z}dx \end{cases}$

$=\int_0^{2z} e^{-x}dx - e^{-2z}\int_0^{2z}dx$

$=1-e^{-2z}-2z\cdot e^{-2z}$

多维正态分布的函数

$(X,Y) \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)$

X, Y 独立 $\Leftrightarrow \rho = 0$

- (X, Y) 是二维随机向量，X 与 Y 独立，$X \sim N(\mu_1, \sigma_1^2)$，$Y \sim (\mu_2, \sigma_2^2)$，那么 $Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- $(X_1, X_2, \cdots, X_n)$ 是 n 维正态分布，且 $X_1, X_2, \cdots, X_n$ 相互独立，$X_i \sim N(\mu_i, \sigma_i^2)$，$i = 1, \cdots, n$，那么 $\sum_{i=1}^n a_i X_i \sim N(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$

命题 3.4.1

随机向量 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ 服从 n 维正态分布的充要条件是对任意的 $\boldsymbol{k} = (k_1, k_2, \cdots, k_n)^T \in \mathbb{R}^n$，那么 $\boldsymbol{k}^T\boldsymbol{X}$ 服从一维正态分布

命题 3.4.2

设随机向量 $= (X_1, X_2, \cdots, X_n)'$ 服从 n 维正态分布，期望向量为 $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n)'$，协方差矩阵为 Σ，那么对于任意的实矩阵 $\boldsymbol{A}_{m\times n}$，有

$$\boldsymbol{A} \sim N(\boldsymbol{A}\boldsymbol{\mu}, \boldsymbol{A}\Sigma\boldsymbol{A}')$$

$\mathrm{Cov}(X,X).$

$$\begin{tikzcd} & \Omega \arrow[dl, "X"'] \arrow[dr, "Y"] & \\ \mathbb{R}^m \arrow[rr, "A"] & & \mathbb{R}^n \end{tikzcd}$$

练习

相互独立的随机变量 $X \sim \exp(\lambda_1)$，$Y \sim \exp(\lambda_2)$，求 $Z = X/Y$ 的密度函数。

Outline