4.5. The Bivariate Normal Distribution

1. *Definition.* Let X_1 and X_2 be continuous random variables. If X_1 and X_2 have the joint pdf

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 -2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\},$$

$$-\infty < x_1 < \infty,\ -\infty < x_2 < \infty,$$

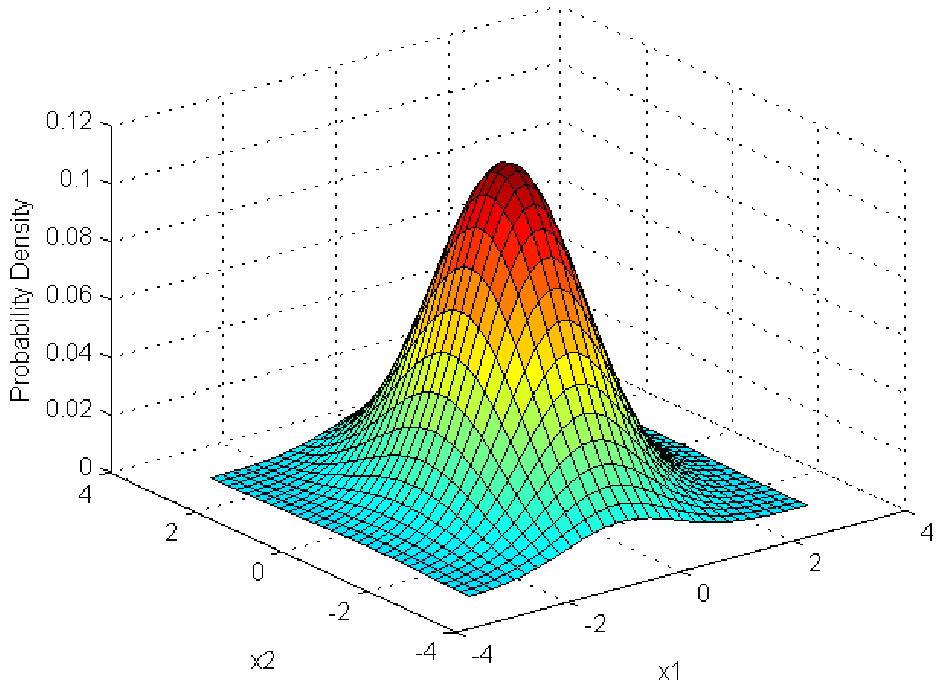
then we say X_1 and X_2 have a bivariate normal distribution, with parameters

$$\mu_1, \mu_2, \sigma_1, \sigma_2, \rho.$$

2. The plot of the joint pdf of the bivariate normal distribution with parameters

$$\mu_1 = 0,\ \mu_2 = 0,\ \sigma_1 = 1,\ \sigma_2 = \sqrt{2},\ \rho = \frac{\sqrt{2}}{10}$$

can be found on the next page.


Probability Density
0.12
0.1
0.08
0.06
0.04
0.02
0
x2
4
2
0
-2
-4
x1
-4
-2
0
2
4

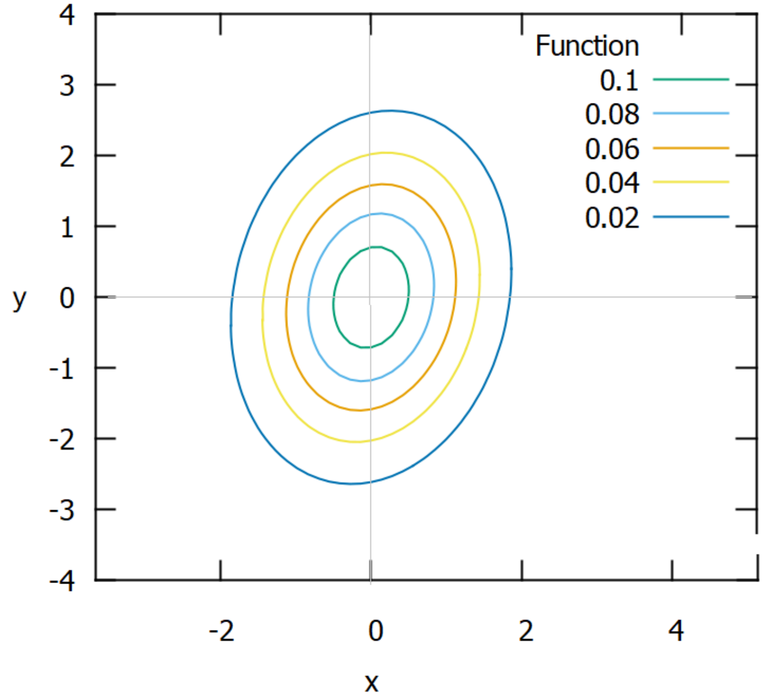
3. On the next page is the contour plot of joint pdf of the bivariate normal distribution with parameters

$$\mu_1 = 0,\ \mu_2 = 0,\ \sigma_1 = 1,\ \sigma_2 = \sqrt{2},\ \rho = \frac{\sqrt{2}}{10}.$$

   (a) The distribution is centered at the origin ($\mu_1 = \mu_2 = 0$).

   (b) The distribution is more spread-out in the x_2-direction than in the x_1 direction ($\sigma_2 > \sigma_1$).

   (c) The distribution is a little tilted to right, which means there is a positive weak correlation between X_1 and X_2 ($\rho = \sqrt{2}/10 \approx 0.14$).


Function
0.1
0.08
0.06
0.04
0.02
4
3
2
1
0
-1
-2
-3
-4
y
-2
0
2
4
x

4. In the special case when $\rho = 0$, the joint pdf of the bivariate normal distribution with parameters $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ reduces to

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot \exp\left\{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

If we let

$$f_1(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \cdot \exp\left(-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right),$$

$$f_2(x_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right),$$

then

$$f(x_1, x_2) = f_1(x_1) f_2(x_2).$$

In this case, $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, and X_1 and X_2 are independent.

5. *Formula.* If X_1 and X_2 have the bivariate normal distribution with parameters
$$\mu_1, \mu_2, \sigma_1, \sigma_2, \rho,$$
then the following hold:

   (a) The marginal distribution of X_1 is $N(\mu_1, \sigma_1^2)$.

   (b) The marginal distribution of X_2 is $N(\mu_2, \sigma_2^2)$.

   (c) The conditional distribution of X_1, given $X_2 = b$, is the normal distribution
$$N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(b - \mu_2), \sigma_1^2(1 - \rho^2)\right).$$

   (d) The conditional distribution of X_2, given $X_1 = a$, is the normal distribution
$$N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(a - \mu_1), \sigma_2^2(1 - \rho^2)\right).$$

   (e) The parameter ρ is the correlation of X_1 and X_2.

6. The expression for the joint pdf of bivariate normal distribution with parameters

$$\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$$

is rather cumbersome (see the definition on page 1). In fact, the joint pdf of the trivariate normal distribution is intolerably complicated. To simplify the notations, we need use matrix language.

7. *Bivariate normal in matrix notation.* Suppose that X_1 and X_2 have the bivariate normal distribution with parameters

$$\mu_1, \mu_2, \sigma_1, \sigma_2, \rho.$$

Denote

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},$$

$$V = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

Then the joint pdf of X_1 and X_2 is

$$f(x_1, x_2) = \frac{1}{2\pi |V|^{1/2}} \exp\left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T V^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right],$$

and we say the random vector $\mathbf{X}$ has the $N(\boldsymbol{\mu}, V)$ distribution.