
5.3.2 Bivariate Normal Distribution

Remember that the normal distribution is very important in probability theory and it shows up in many different applications. We have discussed a single normal random variable previously; we will now talk about two or more normal random variables. We recently saw in Theorem 5.2 that the sum of two independent normal random variables is also normal. However, if the two normal random variables are not independent, then their sum is not necessarily normal. Here is a simple counterexample:

Example 5.35

Let $X \sim N(0, 1)$ and $W \sim \text{Bernoulli}\left(\frac{1}{2}\right)$ be independent random variables. Define the random variable Y as a function of X and W :

$$Y = h(X, W) = \begin{cases} X & \text{if } W = 0 \\ -X & \text{if } W = 1 \end{cases}$$

Find the PDF of Y and $X + Y$.

Solution

Note that by symmetry of $N(0, 1)$ around zero, $-X$ is also $N(0, 1)$. In particular, we can write

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(Y \leq y | W = 0)P(W = 0) + P(Y \leq y | W = 1)P(W = 1) \\ &= \frac{1}{2}P(X \leq y | W = 0) + \frac{1}{2}P(-X \leq y | W = 1) \\ &= \frac{1}{2}P(X \leq y) + \frac{1}{2}P(-X \leq y) \quad (\text{since } X \text{ and } W \text{ are independent}) \\ &= \frac{1}{2}\Phi(y) + \frac{1}{2}\Phi(y) \quad (\text{since } X \text{ and } -X \text{ are } N(0, 1)) \\ &= \Phi(y). \end{aligned}$$

Thus, $Y \sim N(0, 1)$. Now, note that

$$Z = X + Y = \begin{cases} 2X & \text{with probability } \frac{1}{2} \\ 0 & \text{with probability } \frac{1}{2} \end{cases}$$

Thus, Z is a mixed random variable and its PDF is given by

$$\begin{aligned} f_Z(z) &= \frac{1}{2}\delta(z) + \frac{1}{2} \quad (\text{PDF of } 2X \text{ at } z), \\ f_Z(z) &= \frac{1}{2}\delta(z) + \frac{1}{2} \quad (\text{PDF of a } N(0, 4) \text{ at } z) \\ &= \frac{1}{2}\delta(z) + \frac{1}{4\sqrt{2\pi}}e^{-\frac{z^2}{8}}. \end{aligned}$$

In particular, note that X and Y are both normal but their sum is not. Now, we are ready to define **bivariate normal** or **jointly normal** random variables.

Definition 5.3

Two random variables X and Y are said to be **bivariate normal**, or **jointly normal**, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

In the above definition, if we let $a = b = 0$, then $aX + bY = 0$. We agree that the constant zero is a normal random variable with mean and variance 0. From the above definition, we can immediately conclude the following facts:

- If X and Y are bivariate normal, then by letting $a = 1, b = 0$, we conclude X must be normal.
- If X and Y are bivariate normal, then by letting $a = 0, b = 1$, we conclude Y must be normal.
- If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are independent, then they are jointly normal (Theorem 5.2).
- If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are jointly normal, then $X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2 + 2\rho(X, Y)\sigma_X\sigma_Y)$ (Equation 5.21).

But how can we obtain the joint normal PDF in general? Can we provide a simple way to generate jointly normal random variables? The basic idea is that we can start from several independent random variables and by considering their linear combinations, we can obtain bivariate normal random variables. Similar to our discussion on normal random variables, we start by introducing the **standard bivariate normal distribution** and then obtain the general case from the standard one. The following example gives the idea.

Example 5.36

Let Z_1 and Z_2 be two independent $N(0, 1)$ random variables. Define

$$\begin{aligned} X &= Z_1, \\ Y &= \rho Z_1 + \sqrt{1 - \rho^2} Z_2, \end{aligned}$$

where ρ is a real number in $(-1, 1)$.

- a. Show that X and Y are bivariate normal.
- b. Find the joint PDF of X and Y .
- c. Find $\rho(X, Y)$.

Solution

First, note that since Z_1 and Z_2 are normal and independent, they are jointly normal, with the joint PDF

$$\begin{aligned} f_{Z_1 Z_2}(z_1, z_2) &= f_{Z_1}(z_1) f_{Z_2}(z_2) \\ &= \frac{1}{2\pi} \exp \left\{ -\frac{1}{2} [z_1^2 + z_2^2] \right\}. \end{aligned}$$

- a. We need to show $aX + bY$ is normal for all $a, b \in \mathbb{R}$. We have

$$\begin{aligned}
aX + bY &= aZ_1 + b(\rho Z_1 + \sqrt{1 - \rho^2} Z_2) \\
&= (a + b\rho)Z_1 + b\sqrt{1 - \rho^2} Z_2,
\end{aligned}$$

which is a linear combination of Z_1 and Z_2 and thus it is normal.

- b. We can use the method of transformations (Theorem 5.1) to find the joint PDF of X and Y . The inverse transformation is given by

$$\begin{aligned}
Z_1 &= X = h_1(X, Y), \\
Z_2 &= -\frac{\rho}{\sqrt{1 - \rho^2}} X + \frac{1}{\sqrt{1 - \rho^2}} Y = h_2(X, Y).
\end{aligned}$$

We have

$$\begin{aligned}
f_{XY}(z_1, z_2) &= f_{Z_1 Z_2}(h_1(x, y), h_2(x, y))|J| \\
&= f_{Z_1 Z_2}\left(x, -\frac{\rho}{\sqrt{1 - \rho^2}} x + \frac{1}{\sqrt{1 - \rho^2}} y\right)|J|,
\end{aligned}$$

where

$$J = \det \begin{bmatrix} \frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial y} \\ \frac{\partial h_2}{\partial x} & \frac{\partial h_2}{\partial y} \end{bmatrix} = \det \begin{bmatrix} 1 & 0 \\ -\frac{\rho}{\sqrt{1 - \rho^2}} & \frac{1}{\sqrt{1 - \rho^2}} \end{bmatrix} = \frac{1}{\sqrt{1 - \rho^2}}.$$

Thus, we conclude that

$$\begin{aligned}
f_{XY}(x, y) &= f_{Z_1 Z_2}\left(x, -\frac{\rho}{\sqrt{1 - \rho^2}} x + \frac{1}{\sqrt{1 - \rho^2}} y\right)|J| \\
&= \frac{1}{2\pi} \exp \left\{ -\frac{1}{2} \left[x^2 + \frac{1}{1 - \rho^2} (-\rho x + y)^2 \right] \right\} \cdot \frac{1}{\sqrt{1 - \rho^2}} \\
&= \frac{1}{2\pi\sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} [x^2 - 2\rho xy + y^2] \right\}.
\end{aligned}$$

- c. To find $\rho(X, Y)$, first note

$$\begin{aligned}
\text{Var}(X) &= \text{Var}(Z_1) = 1, \\
\text{Var}(Y) &= \rho^2 \text{Var}(Z_1) + (1 - \rho^2) \text{Var}(Z_2) = 1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\rho(X, Y) &= \text{Cov}(X, Y) \\
&= \text{Cov}(Z_1, \rho Z_1 + \sqrt{1 - \rho^2} Z_2) \\
&= \rho \text{Cov}(Z_1, Z_1) + \sqrt{1 - \rho^2} \text{Cov}(Z_1, Z_2) \\
&= \rho \cdot 1 + \sqrt{1 - \rho^2} \cdot 0 \\
&= \rho.
\end{aligned}$$

We call the above joint distribution for X and Y **the standard bivariate normal distribution with correlation coefficient ρ** . It is the distribution for two jointly normal random variables when their variances are equal to one and their correlation coefficient is ρ .

Two random variables X and Y are said to have the **standard bivariate normal distribution with correlation coefficient** ρ if their joint PDF is given by

$$f_{XY}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2]\right\},$$

where $\rho \in (-1, 1)$. If $\rho = 0$, then we just say X and Y have the standard bivariate normal distribution.

Now, if you want two jointly normal random variables X and Y such that $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$, and $\rho(X, Y) = \rho$, you can start with two independent $N(0, 1)$ random variables, Z_1 and Z_2 , and define

$$\begin{cases} X &= \sigma_X Z_1 + \mu_X \\ Y &= \sigma_Y(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_Y \end{cases} \quad (5.23)$$

We can find the joint PDF of X and Y as above. While the joint PDF has a big formula, we usually do not need to use the formula itself. Instead, we usually work with properties of jointly normal random variables such as their mean, variance, and covariance.

Definitions 5.3 and 5.4 are equivalent in the sense that, if X and Y are jointly normal based on one definition, they are jointly normal based on the other definition, too. The proof of their equivalence can be concluded from Problem 10 in Section 6.1.6. In that problem, we show that the two definitions result in the same moment generating functions.

Definition 5.4

Two random variables X and Y are said to have a **bivariate normal distribution** with parameters $\mu_X, \sigma_X^2, \mu_Y, \sigma_Y^2$, and ρ , if their joint PDF is given by

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\} \quad (5.24)$$

where $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$ and $\rho \in (-1, 1)$ are all constants.

In the above discussion, we introduced bivariate normal distributions by starting from independent normal random variables, Z_1 and Z_2 . Another approach would have been to define the bivariate normal distribution using the joint PDF. The two definitions are equivalent mathematically. In particular, we can state the following theorem.

Theorem 5.3

Let X and Y be two bivariate normal random variables, i.e., their joint PDF is given by Equation 5.24. Then there exist independent standard normal random variables Z_1 and Z_2 such that

$$\begin{cases} X &= \sigma_X Z_1 + \mu_X \\ Y &= \sigma_Y(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_Y \end{cases}$$

Proof. (Sketch) To prove the theorem, define

$$\begin{cases} Z_1 &= \frac{X-\mu_X}{\sigma_X} \\ Z_2 &= -\frac{\rho}{\sqrt{1-\rho^2}} \frac{X-\mu_X}{\sigma_X} + \frac{1}{\sqrt{1-\rho^2}} \frac{Y-\mu_Y}{\sigma_Y} \end{cases}$$

Now find the joint PDF of Z_1 and Z_2 using the method of transformations (Theorem 5.1), similar to what we did above. You will find out that Z_1 and Z_2 are independent and standard normal and by definition satisfy the equations of Theorem 5.3. The reason we started our discussion on bivariate normal random variables from Z_1 and Z_2 is three fold. First, it is more convenient and insightful than the joint PDF formula. Second, sometimes the construction using Z_1 and Z_2 can be used to solve problems regarding bivariate normal distributions. Third, this method gives us a way to generate samples from the bivariate normal distribution using a computer program. Since most computing packages have a built-in command for independent normal random variable generation, we can simply use this command to generate bivariate normal variables using Equation 5.23.

Example 5.37

Let X and Y be jointly normal random variables with parameters μ_X , σ_X^2 , μ_Y , σ_Y^2 , and ρ . Find the conditional distribution of Y given $X = x$.

Solution

One way to solve this problem is by using the joint PDF formula (Equation 5.24). In particular, since $X \sim N(\mu_X, \sigma_X^2)$, we can use

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}.$$

Another way to solve this problem is to use Theorem 5.3. We can write

$$\begin{cases} X &= \sigma_X Z_1 + \mu_X \\ Y &= \sigma_Y(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_Y \end{cases}$$

Thus, given $X = x$, we have

$$Z_1 = \frac{x - \mu_X}{\sigma_X},$$

and

$$Y = \sigma_Y \rho \frac{x - \mu_X}{\sigma_X} + \sigma_Y \sqrt{1-\rho^2} Z_2 + \mu_Y.$$

Since Z_1 and Z_2 are independent, knowing Z_1 does not provide any information on Z_2 . We have shown that given $X = x$, Y is a linear function of Z_2 , thus it is normal. In particular

$$\begin{aligned}
E[Y|X=x] &= \sigma_Y \rho \frac{x - \mu_X}{\sigma_X} + \sigma_Y \sqrt{1 - \rho^2} E[Z_2] + \mu_Y \\
&= \mu_Y + \rho \sigma_Y \frac{x - \mu_X}{\sigma_X}, \\
\text{Var}(Y|X=x) &= \sigma_Y^2 (1 - \rho^2) \text{Var}(Z_2) \\
&= (1 - \rho^2) \sigma_Y^2.
\end{aligned}$$

We conclude that given $X = x$, Y is normally distributed with mean $\mu_Y + \rho \sigma_Y \frac{x - \mu_X}{\sigma_X}$ and variance $(1 - \rho^2) \sigma_Y^2$.

Theorem 5.4

Suppose X and Y are jointly normal random variables with parameters μ_X , σ_X^2 , μ_Y , σ_Y^2 , and ρ . Then, given $X = x$, Y is normally distributed with

$$\begin{aligned}
E[Y|X=x] &= \mu_Y + \rho \sigma_Y \frac{x - \mu_X}{\sigma_X}, \\
\text{Var}(Y|X=x) &= (1 - \rho^2) \sigma_Y^2.
\end{aligned}$$

Example 5.38

Let X and Y be jointly normal random variables with parameters $\mu_X = 1$, $\sigma_X^2 = 1$, $\mu_Y = 0$, $\sigma_Y^2 = 4$, and $\rho = \frac{1}{2}$.

- Find $P(2X + Y \leq 3)$.
- Find $\text{Cov}(X + Y, 2X - Y)$.
- Find $P(Y > 1|X = 2)$.

Solution

- Since X and Y are jointly normal, the random variable $V = 2X + Y$ is normal. We have

$$\begin{aligned}
EV &= 2EX + EY = 2, \\
\text{Var}(V) &= 4\text{Var}(X) + \text{Var}(Y) + 4\text{Cov}(X, Y) \\
&= 4 + 4 + 4\sigma_X\sigma_Y\rho(X, Y) \\
&= 8 + 4 \times 1 \times 2 \times \frac{1}{2} \\
&= 12.
\end{aligned}$$

Thus, $V \sim N(2, 12)$. Therefore,

$$P(V \leq 3) = \Phi\left(\frac{3-2}{\sqrt{12}}\right) = \Phi\left(\frac{1}{\sqrt{12}}\right) = 0.6136$$

- Note that $\text{Cov}(X, Y) = \sigma_X\sigma_Y\rho(X, Y) = 1$. We have

$$\begin{aligned}
\text{Cov}(X + Y, 2X - Y) &= 2\text{Cov}(X, X) - \text{Cov}(X, Y) + 2\text{Cov}(Y, X) - \text{Cov}(Y, Y) \\
&= 2 - 1 + 2 - 4 = -1.
\end{aligned}$$

c. Using Theorem 5.4, we conclude that given $X = 2$, Y is normally distributed with

$$E[Y|X = 2] = \mu_Y + \rho\sigma_Y \frac{2 - \mu_X}{\sigma_X} = 1$$
$$Var(Y|X = x) = (1 - \rho^2)\sigma_Y^2 = 3.$$

Thus

$$P(Y > 1|X = 2) = 1 - \Phi\left(\frac{1 - 1}{\sqrt{3}}\right) = \frac{1}{2}.$$

Remember that if two random variables X and Y are independent, then they are uncorrelated, i.e., $Cov(X, Y) = 0$. However, the converse is not true in general. In the case of jointly normal random variables, the converse is true. Thus, for jointly normal random variables, being independent and being uncorrelated are equivalent.

Theorem 5.5

If X and Y are bivariate normal and uncorrelated, then they are independent.

Proof. Since X and Y are uncorrelated, we have $\rho(X, Y) = 0$. By Theorem 5.4, given $X = x$, Y is normally distributed with

$$E[Y|X = x] = \mu_Y + \rho\sigma_Y \frac{x - \mu_X}{\sigma_X} = \mu_Y,$$
$$Var(Y|X = x) = (1 - \rho^2)\sigma_Y^2 = \sigma_Y^2.$$

Thus, $f_{Y|X}(y|x) = f_Y(y)$ for all $x, y \in \mathbb{R}$. Thus X and Y are independent. Another way to prove the theorem is to let $\rho = 0$ in Equation 5.24 and observe that $f_{XY}(x, y) = f_X(x)f_Y(y)$.