The Multivariate Normal Distribution and Copulas

Introduction

In this chapter we introduce the multivariate normal distribution and show how to generate random variables having this joint distribution. We also introduce copulas which are useful when choosing joint distributions to model random variables whose marginal distributions are known.

6.1 The Multivariate Normal

Let $Z_1, \ldots, Z_m$ be independent and identically distributed normal random variables, each with mean 0 and variance 1. If for constants $a_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, and μ_i, $i = 1, \ldots, n$,

$$\begin{aligned}
X_1 &= a_{11}Z_1 + a_{12}Z_2 + \cdots + a_{1m}Z_m + \mu_1 \\
\cdots &= \cdots \\
\cdots &= \cdots \\
X_i &= a_{i1}Z_1 + a_{i2}Z_2 + \cdots + a_{im}Z_m + \mu_i \\
&\cdots \\
&\cdots \\
X_n &= a_{n1}Z_1 + a_{n2}Z_2 + \cdots + a_{nm}Z_m + \mu_n
\end{aligned}$$

then the vector $X_1, \ldots, X_n$ is said to have a *multivariate normal distribution*. That is, $X_i, \ldots, X_n$ has a multivariate normal distribution if each is a constant plus a linear combination of the same set of independent standard normal random variables. Because the sum of independent normal random variables is itself normal, it follows that each X_i is itself a normal random variable.

Simulation. DOI: http://dx.doi.org/10.1016/B978-0-12-415825-2.00006-1

The means and covariances of multivariate normal random variables are as follows:

$$E[X_i] = \mu_i$$

and

$$\begin{aligned}\text{Cov}(X_i, X_j) &= \text{Cov}\left(\sum_{k=1}^{m} a_{ik} Z_k, \sum_{r=1}^{m} a_{jr} Z_r\right)\\ &= \sum_{k=1}^{m}\sum_{r=1}^{m} \text{Cov}(a_{ik} Z_k, a_{jr} Z_r)\\ &= \sum_{k=1}^{m}\sum_{r=1}^{m} a_{ik} a_{jr} \text{Cov}(Z_k, Z_r)\\ &= \sum_{k=1}^{m} a_{ik} a_{jk} \qquad (6.1)\end{aligned}$$

where the preceding used that

$$\text{Cov}(Z_k, Z_r) = \begin{cases} 1, \text{ if } r = k \\ 0, \text{ if } r \neq k \end{cases}$$

The preceding can be compactly expressed in matrix notation. Namely, if we let $\mathbf{A}$ be the $n \times m$ matrix whose row i column j element is a_{ij}, then the defining equation of the multivariate normal is

$$\mathbf{X}' = \mathbf{A}\mathbf{Z}' + \boldsymbol{\mu}' \qquad (6.2)$$

where $\mathbf{X} = (X_1, \ldots, X_n)$ is the multivariate normal vector, $\mathbf{Z} = (Z_1, \ldots, Z_m)$ is the row vector of independent standard normals, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ is the vector of means, and where $\mathbf{B}'$ is the transpose of the matrix $\mathbf{B}$. Because Equation (6.1) states that $\text{Cov}(X_i, X_j)$ is the element in row i column j of the matrix $\mathbf{AA}'$, it follows that if $\mathbf{C}$ is the matrix whose row i column j element is $c_{ij} = \text{Cov}(X_i, X_j)$, then Equation (6.1) can be written as

$$\mathbf{C} = \mathbf{AA}' \qquad (6.3)$$

An important property of multivariate normal vectors is that the joint distribution of $\mathbf{X} = (X_i, \ldots, X_n)$ is completely determined by the quantities $E[X_i]$ and $\text{Cov}(X_i, X_j)$, $i, j = 1, \ldots, n$. That is, the joint distribution is determined by knowledge of the mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and the covariance matrix $\mathbf{C}$. This result can be proved by calculating the joint moment generating function of $X_1, \ldots, X_n$, namely $E[\exp\{\sum_{i=1}^{n} t_i X_i\}]$, which is known to completely specify the joint distribution. To determine this quantity, note first that $\sum_{i=1}^{n} t_i X_i$ is itself a linear combination of the independent normal random variables

$Z_1, \ldots, Z_m$, and is thus also a normal random variable. Hence, using that $E\left[e^W\right] = \exp\{E[W] + \mathrm{Var}(W)/2\}$ when W is normal, we see that

$$E\left[\exp\left\{\sum_{i=1}^{n} t_i X_i\right\}\right] = \exp\left\{E\left[\sum_{i=1}^{n} t_i X_i\right] + \mathrm{Var}\left(\sum_{i=1}^{n} t_i X_i\right)\Big/ 2\right\}$$

As

$$E\left[\sum_{i=1}^{n} t_i X_i\right] = \sum_{i=1}^{n} t_i \mu_i$$

and

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} t_i X_i\right) &= \mathrm{Cov}\left(\sum_{i=1}^{n} t_i X_i, \sum_{j=1}^{n} t_j X_j\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} t_i t_j \mathrm{Cov}(X_i, X_j)\end{aligned}$$

we see that the joint moment generating function, and thus the joint distribution, of the multivariate normal vector is specified by knowledge of the mean values and the covariances.

6.2 Generating a Multivariate Normal Random Vector

Suppose now that we want to generate a multivariate normal vector $\mathbf{X} = (X_1, \ldots, X_n)$ having a specified mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$. Using Equations (6.2) and (6.3) along with the fact that the distribution of $\mathbf{X}$ is determined by its mean vector and covariance matrix, one way to accomplish this would be to first find a matrix $\mathbf{A}$ such that

$$\mathbf{C} = \mathbf{A}\mathbf{A}'$$

then generate independent standard normals $Z_1, \ldots, Z_n$ and set

$$\mathbf{X}' = \mathbf{A}\mathbf{Z}' + \boldsymbol{\mu}'$$

To find such a matrix $\mathbf{A}$ we can make use of a result known as the *Choleski decomposition*, which states that for any $n \times n$ symmetric and positive definite matrix $\mathbf{M}$, there is an $n \times n$ lower triangular matrix $\mathbf{A}$ such that $\mathbf{M} = \mathbf{A}\mathbf{A}'$, where by lower triangular we mean that all elements in the upper triangle of the matrix are equal to 0. (That is, a matrix is lower triangular if the element in row i column j is 0 whenever $i < j$.) Because a covariance matrix $\mathbf{C}$ will be symmetric (as $\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i)$) and as we will assume that it is positive definite (which is usually the case) we can use the Choleski decomposition to find such a matrix $\mathbf{A}$.

Example 6a The Bivariate Normal Distribution Suppose we want to generate the multivariate normal vector X_1, X_2, having means μ_i, variances σ_i^2, $i = 1, 2$, and covariance $c = \text{Cov}(X_1, X_2)$. (When $n = 2$, the multivariate normal vector is called a *bivariate normal*.) If the Choleski decomposition matrix is

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \tag{6.4}$$

then we need to solve

$$\begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & c \\ c & \sigma_2^2 \end{bmatrix}$$

That is,

$$\begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{11}a_{21} & a_{21}^2 + a_{22}^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & c \\ c & \sigma_2^2 \end{bmatrix}$$

This yields that

$$\begin{aligned} a_{11}^2 &= \sigma_1^2 \\ a_{11}a_{21} &= c \\ a_{21}^2 + a_{22}^2 &= \sigma_2^2 \end{aligned}$$

Letting $\rho = \frac{c}{\sigma_1\sigma_2}$ be the correlation between X_1 and X_2, the preceding gives that

$$\begin{aligned} a_{11} &= \sigma_1 \\ a_{21} &= c/\sigma_1 = \rho\sigma_2 \\ a_{22} &= \sqrt{\sigma_2^2 - \rho^2\sigma_2^2} = \sigma_2\sqrt{1-\rho^2} \end{aligned}$$

Hence, letting

$$\mathbf{A} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \tag{6.5}$$

we can generate X_1, X_2 by generating independent standard normals Z_1 and Z_2 and then setting

$$\mathbf{X}' = \mathbf{A}\mathbf{Z}' + \boldsymbol{\mu}'$$

That is,

$$\begin{aligned} X_1 &= \sigma_1 Z_1 + \mu_1 \\ X_2 &= \rho\sigma_2 Z_1 + \sigma_2\sqrt{1-\rho^2}\,Z_2 + \mu_2 \end{aligned}$$

The preceding can also be used to derive the joint density of the bivariate normal vector X_1, X_2. Start with the joint density function of Z_1, Z_2 :

$$f_{Z_1,Z_2}(z_1, z_2) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\left(z_1^2 + z_2^2\right)\right\}$$

and consider the transformation

$$x_1 = \sigma_1 z_1 + \mu_1 \tag{6.6}$$

$$x_2 = \rho\sigma_2 z_1 + \sigma_2\sqrt{1-\rho^2} \quad z_2 + \mu_2 \tag{6.7}$$

The Jacobian of this transformation is

$$\mathbf{J} = \begin{vmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{vmatrix} = \sigma_1\sigma_2\sqrt{1-\rho^2} \tag{6.8}$$

Moreover, the transformation yields the solution

$$z_1 = \frac{x_1 - \mu_1}{\sigma_1}$$

$$z_2 = \frac{x_2 - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)}{\sigma_2\sqrt{1-\rho^2}}$$

giving that

$$\begin{aligned} z_1^2 + z_2^2 &= \frac{(x_1-\mu_1)^2}{\sigma_1^2}\left(1 + \frac{\rho^2}{1-\rho^2}\right) + \frac{(x_2-\mu_2)^2}{\sigma_2^2(1-\rho^2)} \\ &\quad - \frac{2\rho}{\sigma_1\sigma_2(1-\rho^2)}(x_1-\mu_1)(x_2-\mu_2) \\ &= \frac{(x_1-\mu_1)^2}{\sigma_1^2(1-\rho^2)} + \frac{(x_2-\mu_2)^2}{\sigma_2^2(1-\rho^2)} - \frac{2\rho}{\sigma_1\sigma_2(1-\rho^2)}(x_1-\mu_1)(x_2-\mu_2) \end{aligned}$$

Thus, we obtain that the joint density of X_1, X_2 is

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= \frac{1}{|J|} f_{Z_1,Z_2}\left(\frac{x_1-\mu_1}{\sigma_1}, \frac{x_2 - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)}{\sigma_2\sqrt{1-\rho^2}}\right) \\ &= C\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \right.\right. \\ &\quad \left.\left. - \frac{2\rho}{\sigma_1\sigma_2}(x_1-\mu_1)(x_2-\mu_2)\right]\right\} \end{aligned}$$

where $C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$. □

It is generally easy to solve the equations for the Choleski decomposition of an $n \times n$ covariance matrix $\mathbf{C}$. As we take the successive elements of the matrix $\mathbf{AA}'$ equal to the corresponding values of the matrix $\mathbf{C}$, the computations are easiest if we look at the elements of the matrices by going down successive columns. That is, we equate the element in row i column j of $\mathbf{AA}'$ to c_{ij} in the following order of (i, j):

$$(1, 1), (2, 1), \ldots, (n, 1), (2, 2), (3, 2), \ldots, (n, 2),$$
$$(3, 3), \ldots, (n, 3), \ldots, (n-1, n-1), (n, n-1), (n, n)$$

By symmetry the equations obtained for (i, j) and (j, i) would be the same and so only the first to appear is given.

For instance, suppose we want the Choleski decomposition of the matrix

$$\mathbf{C} = \begin{bmatrix} 9 & 4 & 2 \\ 4 & 8 & 3 \\ 2 & 3 & 7 \end{bmatrix} \tag{6.9}$$

The matrix equation becomes

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} * \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ 0 & a_{22} & a_{32} \\ 0 & 0 & a_{33} \end{bmatrix} = \begin{bmatrix} 9 & 4 & 2 \\ 4 & 8 & 3 \\ 2 & 3 & 7 \end{bmatrix}$$

yielding the solution

$$\begin{aligned}
a_{11}^2 &= 9 & \Rightarrow \quad a_{11} &= 3 \\
a_{21}a_{11} &= 4 & \Rightarrow \quad a_{21} &= \frac{4}{3} \\
a_{31}a_{11} &= 2 & \Rightarrow \quad a_{31} &= \frac{2}{3} \\
a_{21}^2 + a_{22}^2 &= 8 & \Rightarrow \quad a_{22} &= \frac{\sqrt{56}}{3} \approx 2.4944 \\
a_{31}a_{21} + a_{32}a_{22} &= 3 & \Rightarrow \quad a_{32} &= \frac{3 - 8/9}{\sqrt{56}/3} = \frac{19}{3/\sqrt{56}} \approx 0.8463 \\
a_{31}^2 + a_{32}^2 + a_{33}^2 &= 7 & \Rightarrow \quad a_{33} &= \frac{1}{3}\sqrt{59 - (19)^2/56} \approx 2.4165
\end{aligned}$$

□

6.3 Copulas

A joint probability distribution function that results in both marginal distributions being uniformly distributed on $(0, 1)$ is called a *copula*. That is, the joint distribution function $C(x, y)$ is a copula if $C(0, 0) = 0$ and for $0 \le x, y \le 1$

$$C(x, 1) = x, \quad C(1, y) = y$$

Suppose we are interested in finding an appropriate joint probability distribution function $H(x, y)$ for random variables X and Y, whose marginal distributions are known to be the continuous distribution functions F and G, respectively. That is, knowing that

$$P(X \le x) = F(x)$$

and

$$P(Y \le y) = G(y)$$

and having some knowledge about the type of dependency between X and Y, we want to choose an appropriate joint distribution function $H(x, y) = P(X \le x, Y \le y)$. Because X has distribution F and Y has distribution G it follows that $F(X)$ and $G(Y)$ are both uniform on $(0, 1)$. Consequently the joint distribution function of $F(X), G(Y)$ is a copula. Also, because F and G are both increasing functions, it follows that $X \le x, Y \le y$ if and only if $F(X) \le F(x)$, $G(Y) \le G(y)$.Consequently, if we choose the copula $C(x, y)$ as the joint distribution function of $F(X), G(Y)$ then

$$\begin{aligned} H(x, y) &= P(X \le x, Y \le y) \\ &= P(F(X) \le F(x), G(Y) \le G(y)) \\ &= C(F(x), G(y)) \end{aligned}$$

The copula approach to choosing an appropriate joint probability distribution function for random variables X and Y is to first decide on their marginal distributions F and G, and then choose an appropriate copula to model the joint distribution of $F(X), G(Y)$. An appropriate copula to use would be one that models the presumed dependencies between $F(X)$ and $G(Y)$. Because F and G are increasing, the dependencies resulting from the resulting copula chosen should be similar to the dependency that we think holds between X and Y. For instance, if we believe that the correlation between X and Y is ρ, then we could try to choose a copula such that random variables whose distribution is given by that copula would have correlation equal to ρ. (Because correlation only measures the linear relationship between random variables, the correlation of X and Y is, however, not equal to the correlation of $F(X)$ and $G(Y)$.)

Example 6b The Gaussian Copula A very popular copula used in modeling is the Gaussian copula. Let Φ be the standard normal distribution function. If X and Y are standard normal random variables whose joint distribution is a bivariate normal distribution with correlation ρ, then the joint distribution of

$\Phi(X)$ and $\Phi(Y)$ is called the *Gaussian copula*. That is, the Gaussian copula C is given by

$$\begin{aligned} C(x,y) &= P(\Phi(X) \le x, \Phi(Y) \le y) \\ &= P(X \le \Phi^{-1}(x), Y \le \Phi^{-1}(y)) \\ &= \int_{-\infty}^{\Phi^{-1}(x)} \int_{-\infty}^{\Phi^{-1}(y)} \frac{1}{2\pi\sqrt{1-\rho^2}} \\ &\quad \times \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right\} dy\,dx \qquad \square \end{aligned}$$

Remark The terminology "Gaussian copula" is used because the normal distribution is often called the *Gaussian distribution* in honor of the famous mathematician J.F. Gauss, who made important use of the normal distribution in his astronomical studies. □

Suppose X, Y has a joint distribution function $H(x, y)$, and let

$$F(x) = \lim_{y\to\infty} H(x,y)$$

and

$$G(y) = \lim_{x\to\infty} H(x,y)$$

be the marginal distributions of X and Y. The joint distribution of $F(X), G(Y)$ is called the copula generated by X, Y, and is denoted as $C_{X,Y}$. That is,

$$\begin{aligned} C_{X,Y}(x,y) &= P(F(X) \le x, G(Y) \le y) \\ &= P(X \le F^{-1}(x), Y \le G^{-1}(y)) \\ &= H(F^{-1}(x), G^{-1}(y)) \end{aligned}$$

For instance, the Gaussian copula is the copula generated by random variables that have a bivariate normal distribution with means 0, variances 1, and correlation ρ.

We now show that if $s(x)$ and $t(x)$ are increasing functions, then the copula generated by the random vector $s(X), t(Y)$ is equal to the copula generated by X, Y.

Proposition *If s and t are increasing functions, then*

$$C_{s(X),t(Y)}(x,y) = C_{X,Y}(x,y)$$

Proof If F and G are the respective distribution functions of X and Y, then the distribution function of $s(X)$, call it F_s, is

$$\begin{aligned} F_s(x) &= P(s(X) \le x) \\ &= P(X \le s^{-1}(x)) \quad \text{(because } s \text{ is an increasing function)} \\ &= F(s^{-1}(x)) \end{aligned}$$

Similarly, the distribution function of $t(Y)$, call it F_t, is

$$F_t(y) = G(t^{-1}(y))$$

Consequently,

$$F_s(s(X)) = F(s^{-1}(s(X))) = F(X)$$

and

$$F_t(t(Y)) = G(Y)$$

showing that

$$\begin{aligned} C_{s(X),t(Y)}(x, y) &= P(F_s(s(X)) \leq x, F_t(t(Y)) \leq y) \\ &= P(F(X) \leq x, G(Y) \leq y) \\ &= C_{X,Y}(x, y) \end{aligned} \qquad \square$$

Suppose again that X, Y has a joint distribution function $H(x, y)$ and that the continuous marginal distribution functions are F and G. Another way to obtain a copula aside from using that $F(X)$ and $G(Y)$ are both uniform on $(0, 1)$ is to use that $1 - F(X)$ and $1 - G(Y)$ are also uniform on $(0, 1)$. Hence,

$$\begin{aligned} C(x, y) &= P(1 - F(X) \leq x, 1 - G(Y) \leq y) \\ &= P(F(X) \geq 1 - x, G(Y) \geq 1 - y) \\ &= P(X \geq F^{-1}(1 - x), Y \geq G^{-1}(1 - y)) \end{aligned} \tag{6.10}$$

is also a copula. It is sometimes called the copula generated by the tail distributions of X and Y.

Example 6c The Marshall–Olkin Copula A tail distribution generated copula that indicates a positive correlation between X and Y and which gives a positive probability that $X = Y$ is the *Marshall–Olkin copula*. The model that generated it originated as follows. Imagine that there are three types of shocks. Let T_i denote the time until a type i shock occurs, and suppose that T_1, T_2, T_3 are independent exponential random variables with respective means $E[T_i] = 1/\lambda_i$. Now suppose that there are two items, and that a type 1 shock causes item 1 to fail, a type 2 shock causes item 2 to fail, and a type 3 shock causes both items to fail. Let X be the time at which item 1 fails and let Y be the time at which item 2 fails. Because item 1 will fail either when a type 1 or a type 3 shock occurs, it follows from the fact that the minimum of independent exponential random variables is also exponential, with a rate equal to the sum of the rates, that X is exponential with rate $\lambda_1 + \lambda_3$. Similarly, Y is exponential with rate $\lambda_2 + \lambda_3$. That is, X and Y have respective distribution functions

$$F(x) = 1 - \exp\{-(\lambda_1 + \lambda_3)x\}, \quad x \geq 0 \tag{6.11}$$

$$G(y) = 1 - \exp\{-(\lambda_2 + \lambda_3)y\}, \quad y \geq 0 \tag{6.12}$$

Now, for $x \geq 0, y \geq 0$

$$
\begin{aligned}
P(X > x, Y > y) &= P(T_1 > x, T_2 > y, T_3 > \max(x, y)) \\
&= P(T_1 > x)P(T_2 > y)P(T_3 > \max(x, y)) \\
&= \exp\{-\lambda_1 x - \lambda_2 y - \lambda_3 \max(x, y)\} \\
&= \exp\{-\lambda_1 x - \lambda_2 y - \lambda_3(x + y - \min(x, y))\} \\
&= \exp\{-(\lambda_1 + \lambda_3)x\} \exp\{-(\lambda_2 + \lambda_3)y\} \exp\{\lambda_3 \min(x, y)\} \\
&= \exp\{-(\lambda_1 + \lambda_3)x\} \exp\{-(\lambda_2 + \lambda_3)y\} \\
&\quad \times \min(\exp\{\lambda_3 x\}, \exp\{\lambda_3 y\}) \qquad (6.13)
\end{aligned}
$$

Now, if $p(x) = 1 - e^{-ax}$, then $p^{-1}(x)$ is such that

$$x = p(p^{-1}(x)) = 1 - e^{-ap^{-1}(x)}$$

which yields that

$$p^{-1}(x) = -\frac{1}{a}\ln(1 - x) \qquad (6.14)$$

Consequently, setting $a = \lambda_1 + \lambda_3$ in Equation (6.14) we see from Equation (6.11) that

$$F^{-1}(1 - x) = -\frac{1}{\lambda_1 + \lambda_3}\ln(x), \quad 0 \leq x \leq 1$$

Similarly, setting $a = \lambda_2 + \lambda_3$ in Equation (6.14) yields from Equation (6.12) that

$$G^{-1}(1 - y) = -\frac{1}{\lambda_2 + \lambda_3}\ln(y), \quad 0 \leq y \leq 1$$

Consequently,

$$
\begin{aligned}
\exp\{-(\lambda_1 + \lambda_3)F^{-1}(1 - x)\} &= x \\
\exp\{-(\lambda_2 + \lambda_3)G^{-1}(1 - y)\} &= y \\
\exp\{\lambda_3 F^{-1}(1 - x)\} &= x^{-\frac{\lambda_3}{\lambda_1 + \lambda_3}} \\
\exp\{\lambda_3 G^{-1}(1 - y)\} &= y^{-\frac{\lambda_3}{\lambda_2 + \lambda_3}}
\end{aligned}
$$

Hence, from Equations (6.10) and (6.13) we obtain that the copula generated by the tail distribution of X and Y, referred to as the *Marshall–Olkin copula*, is

$$
\begin{aligned}
C(x, y) &= P(X \geq F^{-1}(1 - x), Y \geq G^{-1}(1 - y)) \\
&= xy \min\left(x^{-\frac{\lambda_3}{\lambda_1 + \lambda_3}}, y^{-\frac{\lambda_3}{\lambda_2 + \lambda_3}}\right) \\
&= \min(x^{\alpha} y, xy^{\beta})
\end{aligned}
$$

where $\alpha = \frac{\lambda_1}{\lambda_1 + \lambda_3}$ and $\beta = \frac{\lambda_2}{\lambda_2 + \lambda_3}$.

Multidimensional Copulas

We can also use copulas to model n-dimensional probability distributions. The n-dimensional distribution function $C(x_1, \ldots, x_n)$ is said to be a copula if all n marginal distributions are uniform on $(0, 1)$. We can now choose a joint distribution of a random vector $X_1, \ldots, X_n$ by first choosing the marginal distribution functions $F_i, i = 1, \ldots, n$, and then choosing a copula for the joint distribution of $F_1(X_1), \ldots, F_n(X_n)$. Again a popular choice is the Gaussian copula which takes C to be the joint distribution function of $\Phi(W_1), \ldots, \Phi(W_n)$ when $W_1, \ldots, W_n$ has a multivariate normal distribution with mean vector $\mathbf{0}$, and a specified covariance matrix whose diagonal (variance) values are all 1. (The diagonal values of the covariance matrix are taken equal to 1 so that the distribution of $\Phi(W_i)$ is uniform on $(0, 1)$.) In addition, so that the relationship between X_i and X_j is similar to that between W_i and W_j, it is usual to let $\mathrm{Cov}(W_i, W_j) = \mathrm{Cov}(X_i, X_j), i \neq j$.

6.4 Generating Variables from Copula Models

Suppose we want to generate a random vector $\mathbf{X} = (X_1, \ldots, X_n)$ with marginal distributions $F_1, \ldots, F_n$ and copula C. Provided we can generate a random vector whose distribution is C, and that we can invert the distribution functions $F_i, i = 1, \ldots, n$, it is easy to generate $\mathbf{X}$. Because the joint distribution of $F_1(X_1), \ldots, F_n(X_n)$ is C, we can generate $X_1, \ldots, X_n$ by first generating a random vector having distribution C and then inverting the generated values to obtain the desired vector $\mathbf{X}$. That is, if the generated values from the copula distribution function are $y_1, \ldots, y_n$, then the generated value of $X_1, \ldots, X_n$ are $F_1^{-1}(y_1), \ldots, F_n^{-1}(y_n)$.

Example 6d The following can be used to generate $X_1, \ldots, X_n$ having marginal distributions $F_1, \ldots, F_n$ and covariances $\mathrm{Cov}(X_i, X_j), i \neq j$, by using a Gaussian copula:

1. Use the Choleski decomposition method to generate $W_1, \ldots, W_n$ from a multivariate normal distribution with means all equal to 0, variances all equal to 1, and with $\mathrm{Cov}(W_i, W_j) = \mathrm{Cov}(X_i, X_j), i \neq j$.
2. Compute the values $\Phi(W_i), i = 1, \ldots, n$, and note that the joint distribution of $\Phi(W_1), \ldots, \Phi(W_n)$ is the Gaussian copula.
3. Let $F_i(Xi) = \Phi(W_i), i = 1, \ldots, n$.
4. Invert to obtain $X_i = F_i^{-1}(\Phi(W_i)), i = 1, \ldots, n$. □

Example 6e Suppose that we want to generate V, W having marginal distribution functions H and R using a Marshall–Olkin tail copula. Rather than generating directly from the copula, it is easier to first generate the Marshall–Olkin vector X, Y. With F and G denoting the marginal distribution functions of X and Y, we then take $1 - F(X) = e^{-(\lambda_1+\lambda_3)X}$, $1 - G(Y) = e^{-(\lambda_2+\lambda_3)Y}$ as the generated value of the vector having the distribution of the copula. We then set these values equal to $H(V)$ and to $R(W)$ and solve for V and W. That is, we use the following approach:

1. Generate T_1, T_2, T_3, independent exponential random variables with rates $\lambda_1, \lambda_2, \lambda_3$.
2. Let $X = \min(T_1, T_3)$, $Y = \min(T_2, T_3)$.
3. Set $H(V) = e^{-(\lambda_1+\lambda_3)X}$, $R(W) = e^{-(\lambda_2+\lambda_3)Y}$.
4. Solve the preceding to obtain V, W. □

Exercises

1. Suppose $Y_1, \ldots, Y_m$ are independent normal random variables with means $E[Y_i] = \mu_i$, and variances $\text{Var}(Y_i) = \sigma_i^2$, $i = 1, \ldots, m$. If

$$X_i = a_{i1}Y_1 + a_{i2}Y_2 + \cdots + a_{im}Y_m, \quad i = 1, \ldots, n$$

argue that $X_1, \ldots, X_n$ is a multivariate normal random vector.

2. Suppose that $X_1, \ldots, X_n$ has a multivariate normal distribution. Show that $X_1, \ldots, X_n$ are independent if and only if

$$\text{Cov}(X_i, X_j) = 0 \quad \text{when } i \neq j$$

3. If $\mathbf{X}$ is a multivariate normal n-vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$, show that $\mathbf{AX}'$ is multivariate normal with mean vector $\mathbf{A}\boldsymbol{\mu}'$ and covariance matrix $\mathbf{ACA}'$, when $\mathbf{A}$ is an $m \times n$ matrix.

4. Find the Choleski decomposition of the matrix

$$\begin{bmatrix} 4 & 2 & 2 & 4 \\ 2 & 5 & 7 & 0 \\ 2 & 7 & 19 & 11 \\ 4 & 0 & 11 & 25 \end{bmatrix}$$

5. Let X_1, X_2 have a bivariate normal distribution, with means $E[X_i] = \mu_i$, variances $\text{Var}(X_i) = \sigma_i^2$, $i = 1, 2$, and correlation ρ. Show that the conditional distribution of X_2 given that $X_1 = x$ is normal with mean $\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$ and variance $\sigma_2^2(1 - \rho^2)$.

6. Give an algorithm for generating random variables X_1, X_2, X_3 having a multivariate distribution with means $E[X_i] = i$, $i = 1, 2, 3$, and covariance matrix

$$\begin{bmatrix} 3 & -2 & 1 \\ -2 & 5 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

7. Find the copula $C_{X,X}$.

8. Find the copula $C_{X,-X}$.

9. Find the copula $C_{X,Y}$ when X and Y are independent.

10. If s is an increasing function, and t is a decreasing function, find $C_{s(X),t(Y)}$ in terms of $C_{X,Y}$.