Chapter 5
Correlation and Dependence

Correlation and dependence are statistical relationships that can be observed between distinct random variables or data samples. They are generally modeled by copulas which are used to describe the joint distribution of random variables. This chapter deals with uncertainty and dependence structures via the construction of copulas, starting from basic Bernoulli and Gaussian examples.

5.1 Joint Bernoulli Distribution ................. 89
5.2 Joint Gaussian Distribution .................. 90
5.3 Copulas and Dependence Structures ......... 92
5.4 Examples of Copulas .......................... 96
Exercises ........................................... 103

5.1 Joint Bernoulli Distribution

Given a choice of modeling based on the distributions of two random variables \( X \) and \( Y \), it is natural to consider a dependence structure between \( X \) and \( Y \).

Consider two Bernoulli random variables \( X \) and \( Y \), with

\[
p_X = \mathbb{P}(X = 1) = \mathbb{E}[\mathbb{1}_{\{X = 1\}}] \quad \text{and} \quad p_Y = \mathbb{P}(Y = 1) = \mathbb{E}[\mathbb{1}_{\{Y = 1\}}]
\]

and correlation

\[
\rho := \frac{\mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y]}{\sqrt{\text{Var}[X] \text{Var}[Y]}} = \frac{\mathbb{P}(X = 1 \text{ and } Y = 1) - p_X p_Y}{\sqrt{p_X(1 - p_X)p_Y(1 - p_Y)}}.
\]

* Correlation does not imply causation. Try Spurious Correlations.
We note that in that case the joint distribution \( P(X = i \text{ and } Y = j) \), \( i, j = 0, 1 \), is fully determined by \( P(X = 1) \), \( P(Y = 1) \) and the correlation \( \rho \in [-1, 1] \), as

\[
\begin{align*}
\mathbb{P}(X = 1 \text{ and } Y = 1) &= \mathbb{E}[XY] \\
&= pxpy + \rho \sqrt{pxpy(1-px)(1-py)},
\end{align*}
\]

\[
\begin{align*}
\mathbb{P}(X = 0 \text{ and } Y = 1) &= \mathbb{E}[(1-X)Y] = \mathbb{P}(Y = 1) - \mathbb{P}(X = 1 \text{ and } Y = 1) \\
&= (1-px)py - \rho \sqrt{pxpy(1-px)(1-py)},
\end{align*}
\]

\[
\begin{align*}
\mathbb{P}(X = 1 \text{ and } Y = 0) &= \mathbb{E}[X(1-Y)] = \mathbb{P}(X = 1) - \mathbb{P}(X = 1 \text{ and } Y = 1) \\
&= px(1-py) - \rho \sqrt{pxpy(1-px)(1-py)},
\end{align*}
\]

\[
\begin{align*}
\mathbb{P}(X = 0 \text{ and } Y = 0) &= \mathbb{E}[(1-X)(1-Y)] \\
&= (1-px)(1-py) + \rho \sqrt{pxpy(1-px)(1-py)},
\end{align*}
\]

see Exercise 5.2.

### 5.2 Joint Gaussian Distribution

Consider now two centered Gaussian random variables \( X \sim \mathcal{N}(0, \sigma^2) \) and \( Y \sim \mathcal{N}(0, \eta^2) \) with probability density functions

\[
f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)} \quad \text{and} \quad f_Y(x) = \frac{1}{\sqrt{2\pi\eta^2}} e^{-x^2/(2\eta^2)}, \quad x \in \mathbb{R}.
\]

Let

\[
\rho = \text{corr}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X] \text{Var}[Y]}} = \frac{\mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y]}{\sqrt{\text{Var}[X] \text{Var}[Y]}}.
\]

When the covariance matrix

\[
\Sigma := \begin{bmatrix}
\mathbb{E}[X^2] & \mathbb{E}[XY] \\
\mathbb{E}[XY] & \mathbb{E}[Y^2]
\end{bmatrix} = \begin{bmatrix}
\sigma^2 & \rho \sigma \eta \\
\rho \sigma \eta & \eta^2
\end{bmatrix}
\]

with determinant

\[
\det \Sigma = \mathbb{E}[X^2] \mathbb{E}[Y^2] - (\mathbb{E}[XY])^2
\]

\[
= \mathbb{E}[X^2] \mathbb{E}[Y^2] (1 - (\text{corr}(X, Y))^2)
\]

\[
\geq 0,
\]

is invertible, there exists a probability density function
Correlation and Dependence

\[
f_{\Sigma}(x, y) = \frac{1}{\sqrt{2\pi \det \Sigma}} \exp \left( \frac{-1}{2} \begin{bmatrix} x \\ y \end{bmatrix}^\top \Sigma^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \right) \tag{5.2}
\]

\[
= \frac{1}{\sqrt{2\pi \det \Sigma}} \exp \left( \frac{-1}{2} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} \mathbb{E}[X^2] & \mathbb{E}[XY] \\ \mathbb{E}[XY] & \mathbb{E}[Y^2] \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \right),
\]

with respective marginals \( \mathcal{N}(0, \sigma^2) \) and \( \mathcal{N}(0, \eta^2) \).

![Joint Gaussian probability density](image)

\textbf{Fig. 5.1: Joint Gaussian probability density.}

The probability density function (5.2) is called the centered joint (bivariate) Gaussian probability density with covariance matrix \( \Sigma \).

Note that when \( \rho = \text{corr}(X, Y) = \pm 1 \) we have \( \det \Sigma = 0 \) and the joint probability density function \( f_{\Sigma}(x, y) \) is not defined.

More generally, a random vector \((X_1, \ldots, X_n)\) has a multivariate centered Gaussian distribution if every linear combination \( Y = a_1X_1 + \ldots + a_nX_n \) is centered Gaussian, and in this case the probability density function of \((X_1, \ldots, X_n)\) takes the form

\[
f_{\Sigma}(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp \left( \frac{-1}{2} (x_1, \ldots, x_n)^T \Sigma^{-1} (x_1, \ldots, x_n) \right),
\]

\((x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\), where \( \Sigma \) is the covariance matrix

\[
\Sigma = \begin{bmatrix}
\text{Var}[X_1] & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_{n-1}) & \text{Cov}(X_1, X_n) \\
\text{Cov}(X_2, X_1) & \text{Var}[X_2^2] & \cdots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \text{Var}[X_{n-1}] & \text{Cov}(X_{n-1}, X_n) \\
\text{Cov}(X_1, X_n) & \text{Cov}(X_2, X_n) & \cdots & \text{Cov}(X_{n-1}, X_n) & \text{Var}[X_n^2]
\end{bmatrix}
\]
The next remark plays an important role in the modeling of joint default probabilities, see here for a detailed discussion.

**Remark 5.1.** There exist couples \((X,Y)\) with of random variables with Gaussian marginals \(\mathcal{N}(0,\sigma^2)\) and \(\mathcal{N}(0,\eta^2)\), such that

i) \((X,Y)\) does not have the bivariate Gaussian distribution with probability density function \(f_\Sigma(x,y)\), where \(\Sigma\) is the covariance matrix (5.1) of \((X,Y)\).

ii) the random variable \(X+Y\) is not even Gaussian.

*Proof.* See Exercise 5.5. □

### 5.3 Copulas and Dependence Structures

The word copula derives from the Latin noun for a “link” or “tie” that connects two different objects or concepts.

**Definition 5.2.** A two-dimensional copula is any joint cumulative distribution function

\[
C : [0, 1] \times [0, 1] \to [0, 1]
\]

\((u,v) \mapsto C(u,v)\)

with uniform \([0,1]\)-valued marginals.

In other words, any copula function \(C(u,v)\) can be written as

\[
C(u,v) = \mathbb{P}(U \leq u \text{ and } V \leq v), \quad 0 \leq u,v \leq 1,
\]

where \(U\) and \(V\) are uniform \([0,1]\)-valued random variables.

**Examples.**

i) The copula corresponding to independent uniform random variables \((U,V)\) is given by

\[
C(u,v) = \mathbb{P}(U \leq u \text{ and } V \leq v) \\
= \mathbb{P}(U \leq u)\mathbb{P}(V \leq v) \\
= uv, \quad 0 \leq u,v \leq 1.
\]

ii) The copula corresponding to the fully correlated case \(U = V\) is given by

\[
C(u,v) = \mathbb{P}(U \leq u \text{ and } V \leq v) \\
= \mathbb{P}(U \leq \min(u,v)) \\
= \min(u,v), \quad 0 \leq u,v \leq 1.
\]
Correlation and Dependence

iii) The copula corresponding to the fully anticorrelated case $U = 1 - V$ is given by

$$C(u, v) := P(U \leq u \text{ and } V \leq v)$$
$$= P(U \leq u \text{ and } 1 - U \leq v)$$
$$= P(1 - v \leq U \leq u)$$
$$= (u + v - 1)^+, \quad 0 \leq u, v \leq 1.$$  

The next lemma is well known and can be used to generate random samples of a cumulative distribution function $F_X$ based on uniformly distributed samples, see Proposition 3.1 in Embrechts and Hofert (2013) for its general statement.

**Lemma 5.3.** Consider $X$ a random variable with continuous and strictly increasing distribution function

$$F_X(x) := P(X \leq x), \quad x \in \mathbb{R}.$$  

a) The random variable

$$U := F_X(X)$$

is uniformly distributed on $[0, 1]$.

b) If $U$ is uniformly distributed on $[0, 1]$ then $F_X^{-1}(U)$ has same distribution as $X$.

**Proof.** We have

$$F_U(u) = P(U \leq u)$$
$$= P(F_X(X) \leq u)$$
$$= P(X \leq F_X^{-1}(u))$$
$$= F_X(F_X^{-1}(u))$$
$$= u, \quad 0 \leq u \leq 1,$$

and similarly

$$P(F_X^{-1}(U) \leq x) = P(U \leq F_X(x)) = F_X(x), \quad x \in \mathbb{R}.$$  

$\square$

As a consequence of Lemma 5.3, given $(X, Y)$ a couple of random variables with joint cumulative distribution function

$$F_{(X,Y)}(x, y) := P(X \leq x \text{ and } Y \leq y), \quad x, y \in \mathbb{R},$$

and cumulative distribution functions

$$F_X(x) = F_{(X,Y)}(x, \infty) = P(X \leq x) \text{ and } F_Y(y) = F_{(X,Y)}(\infty, y) = P(Y \leq y),$$

$\diamond$
we not the following points.

i) The random variables

\[ U := F_X(X) \quad \text{and} \quad V := F_Y(Y) \]

are uniformly distributed on \([0, 1]\).

ii) The copula function

\[ (u, v) \mapsto C_{(X,Y)}(u, v) := P(U \leq u \text{ and } V \leq v), \quad 0 \leq u, v \leq 1, \]

satisfies

\[ C_{(X,Y)}(u, v) := P(U \leq u \text{ and } V \leq v) = P(F_X(X) \leq u \text{ and } F_Y(Y) \leq v) = P(X \leq F_X^{-1}(u) \text{ and } Y \leq F_Y^{-1}(v)) = F_{(X,Y)}(F_X^{-1}(u), F_Y^{-1}(v)), \quad 0 \leq u, v \leq 1, \]

is a copula.

iii) The joint cumulative distribution function of \((X, Y)\) can be recovered as

\[ P(X \leq x \text{ and } Y \leq y) = P(F_X(X) \leq F_X(x) \text{ and } F_Y(Y) \leq F_Y(v)) = P(U \leq F_X(x) \text{ and } V \leq F_Y(y)) = C_{(X,Y)}(F_X(x), F_Y(y)), \quad x, y \in \mathbb{R}. \]

**Higher dimensional copulas**

**Definition 5.4.** An \(n\)-dimensional copula is any joint cumulative distribution function

\[ C : [0, 1] \times \cdots \times [0, 1] \rightarrow [0, 1] \quad (u_1, \ldots, u_n) \mapsto C(u_1, \ldots, u_n) \]

of \(n\) uniform \([0, 1]-valued\) random variables.

Consider the joint cumulative distribution function

\[ F_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n) := P(X_1 \leq x_1, \ldots, X_n \leq x_n) \]

of a family \((X_1, \ldots, X_n)\) of random variables with marginal cumulative distribution functions

\[ F_{X_i}(x) = F_{(X_1, \ldots, X_n)}(\pm \infty, \ldots, \pm \infty, x, \pm \infty, \ldots, \pm \infty), \quad x \in \mathbb{R}, \]
i = 1, 2, . . . , n. The copula defined in Sklar’s theorem encodes the dependence structure of the vector (X1, . . . , Xn).

**Theorem 5.5.** Sklar’s theorem (Sklar (1959)*, Sklar (2010)). Given a joint cumulative distribution function F(X1,...,Xn) there exists an n-dimensional copula C(u1,...,un) such that

\[ F(X_1,...,X_n)(x_1,x_2,...,x_n) = C(F_{X_1}(x_1),F_{X_2}(x_2),...,F_{X_n}(x_n)), \]

\[ x_1,x_2,...,x_n \in \mathbb{R}. \]

The following proposition is a consequence of Sklar’s Theorem 5.5.

**Proposition 5.6.** Assume that the marginal distribution functions FXi are continuous and strictly increasing. Then the joint cumulative distribution function F(X1,...,Xn) defines a n-dimensional copula

\[ C(u_1,...,u_n) := F(X_1,...,X_n)(F_{X_1}^{-1}(u_1),...,F_{X_n}^{-1}(u_n)), \]

\[ u_1,u_2,...,u_n \in [0,1], \text{ which encodes the dependence structure of the vector } (X_1,...,X_n). \]

**Proof.** Indeed, it can be checked as in Lemma 5.3 that C(u1,...,un) has uniform marginal distributions on [0,1], as

\[ C(1,...,1,u,1,...,1) \]
\[ = F(X_1,...,X_n)(F_{X_1}^{-1}(1),...,F_{X_{i-1}}^{-1}(1),F_{X_{i}}^{-1}(u),F_{X_{i+1}}^{-1}(1),...,F_{X_n}^{-1}(1)) \]
\[ = F(X_1,...,X_n)(+\infty,...,+\infty,F_{X_i}^{-1}(u),+\infty,...,+\infty) \]
\[ = F_{X_i}(F_{X_i}^{-1}(u)) \]
\[ = u, \quad 0 \leq u \leq 1. \]

\[ \square \]

**Proposition 5.7.** Given a family (X̂1,...,X̂n) of random variables with marginal cumulative distribution functions F̂X1,...,F̂Xn and a multidimensional copula C(u1,...,un), the function

\[ F^C_{X̂1,...,X̂n}(x_1,...,x_n) := C(F_{X̂1}(x_1),...,F_{X̂n}(x_n)), \quad x_1,x_2,...,x_n \in \mathbb{R}, \]

defines joint cumulative distribution function with marginals X̂1,...,X̂n.

* "The author considers continuous non-decreasing functions Cn on the n-dimensional cube [0,1]n with Cn(0,...,0) = 0, Cn(1,...,1,α,1,...,1) = α. Several theorems are stated relating n-dimensional distribution functions and their marginals in terms of functions Cn. No proofs are given." M. Loève, Math. Reviews MR0125600.
Proof. We note that the marginal distributions generated by \( F_{(\tilde{X}_1, \ldots, \tilde{X}_n)}(x_1, \ldots, x_n) \) coincide with the respective marginals of \((\tilde{X}_1, \ldots, \tilde{X}_n)\), as we have

\[
F_{(\tilde{X}_1, \ldots, \tilde{X}_n)}(+\infty, \ldots, +\infty, u, +\infty, \ldots, +\infty) = C(F_{\tilde{X}_1}(+\infty), \ldots, F_{\tilde{X}_i}(+\infty), F_{\tilde{X}_i}(u), F_{\tilde{X}_{i+1}}(+\infty), \ldots, F_{\tilde{X}_n}(+\infty)) = C(1, \ldots, 1, F_{\tilde{X}_i}(u), 1, \ldots, 1) = F_{\tilde{X}_i}(u), \quad 0 \leq u \leq 1.
\]

\[\square\]

5.4 Examples of Copulas

Gaussian copulas

The choice of (5.2) above as joint probability density function, see Figure 5.1, actually induces a particular dependence structure between the Gaussian random variables \(X\) and \(Y\), and corresponding to the joint cumulative distribution function

\[
\Phi_\Sigma(x, y) := \mathbb{P}(X \leq x \text{ and } Y \leq y)
\]

\[
= \frac{1}{\sqrt{2\pi} \det \Sigma} \int_{-\infty}^{x} \int_{-\infty}^{y} \exp \left( -\frac{1}{2} \left< \begin{bmatrix} u \\ v \end{bmatrix}, \Sigma^{-1} \begin{bmatrix} u \\ v \end{bmatrix} \right> \right) du dv,
\]

x, y \in \mathbb{R}. In case \((X, Y)\) are normalized centered Gaussian random variables with unit variance, \(\Sigma\) is given by

\[
\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},
\]

with correlation parameter \(\rho \in [-1, 1]\). Letting

\[
F_X(x) := \mathbb{P}(X \leq x) \quad \text{and} \quad F_Y(y) := \mathbb{P}(Y \leq y),
\]

denote the cumulative distribution functions of \(X\) and \(Y\), the random variables \(F_X(X)\) and \(F_Y(Y)\) are known to be uniformly distributed on \([0, 1]\), and \((F_X(X), F_Y(Y))\) is a \([0, 1] \times [0, 1]\)-valued random variable with joint cumulative distribution function

\[
C_{\Sigma}(u, v) := \mathbb{P}(F_X(X) \leq u \text{ and } F_Y(Y) \leq v)
\]

\[
= \mathbb{P}(X \leq F_X^{-1}(u) \text{ and } Y \leq F_Y^{-1}(v))
\]

\[
= \Phi_{\Sigma}(F_X^{-1}(u), F_Y^{-1}(v)), \quad 0 \leq u, v \leq 1. \quad (5.4)
\]
The function $C_{\Sigma}(u, v)$, which is the joint cumulative distribution function of a couple of uniformly distributed $[0, 1]$-valued random variables, is called the Gaussian copula generated by the jointly Gaussian distribution of $(X, Y)$ with covariance matrix $\Sigma$.

![Fig. 5.2: Different Gaussian copula graphs for $\rho = 0$, $\rho = 0.85$ and $\rho = 1$.](image)

The above leftmost figure corresponds to independent uniformly distributed $[0, 1]$-valued random variables $U, V$, i.e. to the copula

$$C(u, v) := \mathbb{P}(U \leq u \text{ and } V \leq v) = \mathbb{P}(U \leq u)\mathbb{P}(U \leq u) = uv, \quad 0 \leq u, v \leq 1.$$ 

On the other hand the rightmost figure corresponds to equal uniformly distributed $[0, 1]$-valued random variables $U = V$, i.e. to the copula

$$C(u, v) := \mathbb{P}(U \leq u \text{ and } V \leq v) = \mathbb{P}(U \leq u \text{ and } U \leq v) = \mathbb{P}(U \leq \min(u, v)) = \min(u, v), \quad 0 \leq u, v \leq 1,$$

The middle figure corresponds to an intermediate dependence level given by a Gaussian copula, cf. (5.4) below.
Fig. 5.3: Different Gaussian copula density graphs for $\rho = 0$, $\rho = 0.35$ and $\rho = 0.999$.

The leftmost figure above represents a uniform (product) probability density function on the square $[0, 1] \times [0, 1]$, which corresponds to two independent uniformly distributed $[0, 1]$-valued random variables $U, V$. The rightmost figure shows the probability distribution of the fully correlated couple $(U, U)$, which does not admit a probability density on the square $[0, 1] \times [0, 1]$.

The Gaussian copula $C_\Sigma(u, u)$ admits a probability density function on $[0, 1] \times [0, 1]$ given by

$$c_\Sigma(u, v) = \frac{\partial^2 C_\Sigma(u, v)}{\partial u \partial v}$$

$$= \frac{\partial^2}{\partial u \partial v} \Phi_\Sigma(F_X^{-1}(u), F_Y^{-1}(v))$$

$$= \frac{\partial}{\partial u} \left( \frac{1}{F_Y(F_Y^{-1}(v))} \frac{\partial \Phi_\Sigma}{\partial y}(F_X^{-1}(u), F_Y^{-1}(v)) \right)$$

$$= \frac{\partial}{\partial u} \left( \frac{1}{f_Y(F_Y^{-1}(v))} \frac{\partial \Phi_\Sigma}{\partial y}(F_X^{-1}(u), F_Y^{-1}(v)) \right)$$

$$= \frac{1}{f_X(F_X^{-1}(u)) f_Y(F_Y^{-1}(v))} \Phi_\Sigma(F_X^{-1}(u), F_Y^{-1}(v))$$

$$= \frac{f_\Sigma(F_X^{-1}(u), F_Y^{-1}(v))}{f_X(F_X^{-1}(u)) f_Y(F_Y^{-1}(v))},$$

hence the Gaussian copula $C_\Sigma(u, v)$ can be computed as

$$C_\Sigma(u, v) = \int_0^u \int_0^v c_\Sigma(a, b) \, da \, db$$

$$= \int_0^u \int_0^v \frac{f_\Sigma(F_X^{-1}(a), F_Y^{-1}(b))}{f_X(F_X^{-1}(a)) f_Y(F_Y^{-1}(b))} \, da \, db, \quad 0 \leq u, v \leq 1.$$
The joint cumulative distribution function $F_{(X,Y)}(x,y)$ of $(X,Y)$ can be recovered from Proposition 5.6 as

$$F_{(X,Y)}(x,y) = C_\Sigma(F_X(x), F_Y(y)), \quad x, y \in \mathbb{R}. \quad (5.5)$$

from the Gaussian copula $C_\Sigma(x,y)$ and the respective cumulative distribution functions $F_X(x)$, $F_Y(y)$ of $X$ and $Y$.

In that sense, the Gaussian copula $C_\Sigma(x,y)$ encodes the Gaussian dependence structure of the covariance matrix $\Sigma$. Moreover, the Gaussian copula $C_\Sigma(x,y)$ can be used to generate a joint distribution function $F_{(X,Y)}^C(x,y)$ by letting

$$F_{(X,Y)}^C(x,y) := C_\Sigma(F_X(x), F_Y(y)), \quad x, y \in \mathbb{R}, \quad (5.6)$$

based on other, possibly non-Gaussian cumulative distribution functions $F_X(x)$, $F_Y(y)$ of two random variables $X$ and $Y$. In this case we note that the marginals of the joint cumulative distribution function $F_{(X,Y)}^C(x,y)$ are $F_X(x)$ and $F_Y(y)$ because $C_\Sigma(x,y)$ has uniform marginals on $[0,1]$.

**Gumbel copula**

The Gumbel copula is given by

$$C(u,v) = \exp \left( - \left( (\log u)^\theta + (\log v)^\theta \right)^{1/\theta} \right), \quad 0 \leq u, v \leq 1,$$

with $\theta \geq 1$, and $C(u,v) = uv$ when $\theta = 1$.

**Uniform marginals with given copulas**

The following R code generates random samples according to the Gaussian, Student, and Gumbel copulas with uniform marginals, as illustrated in Figure 5.4.

```
1 install.packages("copula")
2 install.packages("gumbel")
3 library(copula);library(gumbel)
4 norm.cop <- normalCopula(0.35);norm.cop
5 persp(norm.cop, pCopula, n.grid = 51, xlab="u", ylab="v", zlab="C(u,v)", main="", sub="")
6 persp(norm.cop, dCopula, n.grid = 51, xlab="u", ylab="v", zlab="c(u,v)", main="", sub="")
7 norm <- rCopula(4000,normalCopula(0.7))
8 plot(norm[,1],norm[,2],cex=3,pch="",col="blue")
9 stud <- rCopula(4000,tCopula(0.5,dim=2,df=1))
10 points(stud[,1],stud[,2],cex=3,pch="",col="red")
11 gumb <- rgumbel(4000,4)
12 points(gumb[,1],gumb[,2],cex=3,pch="",col="green")
```
The following R code plots the histograms of Figure 5.4.

```r
joint_hist <- function(u) {x <- u[,1]; y <- u[,2]
xhist <- hist(x, breaks=40, plot=FALSE); yhist <- hist(y, breaks=40, plot=FALSE)
top <- max(c(xhist$counts, yhist$counts))
nf <- layout(matrix(c(2,0,1,3),2,2,byrow=TRUE), c(3,1), c(1,3), TRUE)
par(mar=c(3,3,1,1))
plot(x, y, xlab="", ylab="", col="blue", pch=19, cex=0.4)
points(x[1:50], -0.01 + rep(min(y),50), xlab="", ylab="", col="black", pch=18, cex=0.8)
points(0.01 + rep(min(x),50), y[1:50], xlab="", ylab="", col="black", pch=18, cex=0.8)
par(mar=c(0,3,1,1))
barplot(xhist$counts, axes=FALSE, ylim=c(0, top), space=0, col="purple")
par(mar=c(3,0,1,1))
barplot(yhist$counts, axes=FALSE, xlim=c(0, top), space=0, horiz=TRUE, col="purple")
joint_hist(norm); joint_hist(stud); joint_hist(gumb)
```

Fig. 5.4: Samples with histogram marginals and given copulas.

The next R code generates random samples according to the Gaussian, Student, and Gumbel copulas with Gaussian marginals, as illustrated in Figure 5.5.

Fig. 5.5: Samples with Gaussian marginals and given copulas.

This version: June 27, 2020
https://www.ntu.edu.sg/home/nprivault/index.html
The following R code plots joint densities with Gaussian marginals and given copulas, as illustrated in Figure 5.6.

```r
persp(gaussMVD, dMvdc, xlim = c(-3,3), ylim = c(-3,3), col = "lightblue")
persp(studentMVD, dMvdc, xlim = c(-3,3), ylim = c(-3,3), col = "lightblue")
persp(gumbelMVD, dMvdc, xlim = c(-3,3), ylim = c(-3,3), col = "lightblue")
```

Fig. 5.6: Joint densities with Gaussian marginals and given copulas.
The following R code generates contour plots with Gaussian marginals and given copulas, as illustrated in Figure 5.7.

```R
contour(gaussMVD,dMvdc,xlim=c(-3,3),ylim=c(-3,3),nlevels=10,xlab="X",ylab="Y",cex.axis=1.5)
contour(gaussMVD,dMvdc,xlim=c(-3,3),ylim=c(-3,3),nlevels=10,xlab="X",ylab="Y",cex.axis=1.5)
contour(gaussMVD,dMvdc,xlim=c(-3,3),ylim=c(-3,3),nlevels=10,xlab="X",ylab="Y",cex.axis=1.5)
```

**Gaussian marginals with given copulas**

The following R code generates random samples with exponential marginals according to the Gaussian, Student, and Gumbel copulas as illustrated in Figure 5.8.

```R
library(copula);set.seed(100);N=4000
gaussMVD<– mvdc(normalCopula(0.7), margins=c("exp","exp"),
paramMargins=list(list(rate=1),list(rate=1)))
studentMVD<– mvdc(tCopula(0.5,dim=2,df=1), margins=c("exp","exp"),
paramMargins=list(list(rate=1),list(rate=1)))
stud<– rMvdc(N,studentMVD)
gumbelMVD<– mvdc(gumbelCopula(param=4, dim=2), margins=c("exp","exp"),
paramMargins=list(list(rate=1),list(rate=1)))
gumb<– rMvdc(N,gumbelMVD)
plot(norm[,1],norm[,2],cex=3,pch=".",col="blue")
plot(stud[,1],stud[,2],cex=3,pch=".",col="blue")
plot(gumb[,1],gumb[,2],cex=3,pch=".",col="blue")
persp(gaussMVD, dMvdc, xlab = c(0,1), ylab = c(0,1))
persp(studentMVD, dMvdc, xlab = c(0,1), ylab = c(0,1))
persp(gumbelMVD, dMvdc, xlab = c(0,1), ylab = c(0,1))
contour(gaussMVD,dMvdc,xlim=c(0,1),ylim=c(0,1),nlevels=10,xlab="X",ylab="Y",cex.axis=1.5)
contour(studentMVD,dMvdc,xlim=c(0,1),ylim=c(0,1),nlevels=10,xlab="X",ylab="Y",cex.axis=1.5)
contour(gumbelMVD,dMvdc,xlim=c(0,1),ylim=c(0,1),nlevels=10,xlab="X",ylab="Y",cex.axis=1.5)
joint_hist(norm);joint_hist(stud);joint_hist(gumb)
```

![Figure 5.8: Samples with exponential marginals and given copulas.](https://www.ntu.edu.sg/home/nprivault/indext.html)
Exercises

Exercise 5.1 Copulas. In the sequel, $U$ denotes a uniformly distributed $[0, 1]$-valued random variable.

a) To which couple $(U, V)$ of uniformly distributed $[0, 1]$-valued random variables does the copula function

$$C_M(u, v) = \min(u, v), \quad 0 \leq u, v \leq 1,$$

correspond?

b) Show that the function

$$C_m(u, v) := (u + v - 1)^+, \quad 0 \leq u, v \leq 1,$$

is the copula on $[0, 1] \times [0, 1]$ corresponding to $(U, V) = (U, 1 - U)$.

c) Show that for any copula function $C(u, v)$ on $[0, 1] \times [0, 1]$ we have

$$C(u, v) \leq C_M(u, v), \quad 0 \leq u, v \leq 1. \quad (5.7)$$

d) Show that for any copula function $C(u, v)$ on $[0, 1] \times [0, 1]$ we also have

$$C_m(u, v) \leq C(u, v), \quad 0 \leq u, v \leq 1. \quad (5.8)$$

Hint: For fixed $v \in [0, 1]$, let $h(u) := C(u, v) - (u + v - 1)$ and show that $h(1) = 0$ and $h'(u) \leq 0$.

Exercise 5.2 Consider two Bernoulli random variables $X$ and $Y$, with $p_X = P(X = 1)$, $p_Y = P(Y = 1)$, correlation coefficient $\rho \in [-1, 1]$, and

$$\begin{align*}
P(X = 1 \text{ and } Y = 1) &= p_X p_Y + \rho \sqrt{p_X p_Y (1 - p_X)(1 - p_Y)}, \\
P(X = 0 \text{ and } Y = 1) &= (1 - p_X) p_Y - \rho \sqrt{p_X p_Y (1 - p_X)(1 - p_Y)}, \\
P(X = 1 \text{ and } Y = 0) &= p_X (1 - p_Y) - \rho \sqrt{p_X p_Y (1 - p_X)(1 - p_Y)}, \\
P(X = 0 \text{ and } Y = 0) &= (1 - p_X)(1 - p_Y) + \rho \sqrt{p_X p_Y (1 - p_X)(1 - p_Y)}.\end{align*}$$

Is it possible to have $\rho = 1$ without having $p_X = p_Y$ and
Exercise 5.3 Exponential copulas. Consider the random vector \((X, Y)\) of nonnegative random variables, whose joint distribution is given by the survival function
\[
P(X \geq x \text{ and } Y \geq y) := e^{-\lambda x - \mu y - \nu \max(x, y)}, \quad x, y \in \mathbb{R}_+,
\]
where \(\lambda, \mu, \nu > 0\).

a) Find the marginal distributions of \(X\) and \(Y\).
b) Find the joint cumulative distribution function \(F(x, y) := P(X \leq x \text{ and } Y \leq y)\) of \((X, Y)\).
c) Construct an “exponential copula” based on the joint cumulative distribution function of \((X, Y)\).

Exercise 5.4 Gumbel bivariate logistic distribution. Consider the random vector \((X, Y)\) of non-negative random variables, whose joint distribution is given by the joint cumulative distribution function (CDF)
\[
F_{(X, Y)}(x, y) := P(X \leq x \text{ and } Y \leq y) := \frac{1}{1 + e^{-x} + e^{-y}}, \quad x, y \in \mathbb{R}.
\]

a) Find the marginal distributions of \(X\) and \(Y\).
b) Construct the copula based on the joint CDF of \((X, Y)\).

Exercise 5.5 Consider the random vector \((X, Y)\) with the joint probability density function
\[
\tilde{f}(x, y) := \frac{1}{\pi \sigma \eta} \mathbb{1}_{\mathbb{R}_-^2}(x, y)e^{-x^2/(2\sigma^2) - y^2/(2\eta^2)} + \frac{1}{\pi \sigma \eta} \mathbb{1}_{\mathbb{R}_+^2}(x, y)e^{-x^2/(2\sigma^2) - y^2/(2\eta^2)},
\]
plotted as a heat map in Figure 5.9b.
Correlation and Dependence

Fig. 5.9: Truncated two-dimensional Gaussian density.

```r
library(MASS)
Sigma <- matrix(c(1,0,0,1),2,2);N=10000
u<-mvrnorm(N,rep(0,2),Sigma);j=1
for (i in 1:N){
  if (u[i,1]>0 & u[i,2]>0) {j<j+1;}
  if (u[i,1]<0 & u[i,2]<0) {j<j+1;}
  v<-matrix(nrow=j-1, ncol=2);j=1
  for (i in 1:N){
    if (u[i,1]>0 & u[i,2]>0) {v[i,]=u[i,];j<j+1;}
    if (u[i,1]<0 & u[i,2]<0) {v[i,]=u[i,];j<j+1;}
  }
joint_hist(v) # Function defined in the previous section
```

a) Show that $(X, Y)$ has the Gaussian marginals $\mathcal{N}(0,\sigma^2)$ and $\mathcal{N}(0,\eta^2)$.
b) Does the couple $(X, Y)$ have the bivariate Gaussian distribution with probability density function $f_{\Sigma}(x, y)$, where $\Sigma$ is the covariance matrix (5.1) of $(X, Y)$?
c) Show that the random variable $X + Y$ is not Gaussian (take $\sigma = \eta = 1$ for simplicity).
d) Show that under the rotation

$$
R = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix},
$$

of angle $\theta \in [0, 2\pi]$ the random vector $(X \cos \theta - Y \sin \theta, X \sin \theta + Y \cos \theta)$ can have an arbitrary covariance depending on the value of $\theta \in [0, 2\pi]$.

Exercise 5.6  Let $\tau_1$, $\tau_2$ and $\tau$ denote three independent exponentially distributed random times with respective parameters $\lambda_1, \lambda_2, \lambda > 0$. Consider two firms with respective default times $\tau_1 \wedge \tau = \min(\tau_1, \tau)$ and $\tau_2 \wedge \tau = \min(\tau_2, \tau)$, where $\tau$ represents the time of a macro-economic shock.

a) Find the (survival) distribution functions of $\tau_1 \wedge \tau$ and $\tau_2 \wedge \tau$.
b) Compute the joint survival probability

$$
P(\tau_1 \wedge \tau > s \text{ and } \tau_2 \wedge \tau > t), \quad s, t \in \mathbb{R}_+.
$$
Hint: Use the relation $\max(s, t) = s + t - \min(s, t)$, $s, t \in \mathbb{R}_+$.

c) Compute the joint cumulative distribution function

$$P(\tau_1 \land \tau \leq s \text{ and } \tau_2 \land \tau \leq t), \quad s, t \in \mathbb{R}_+.$$ 

d) Compute the resulting copula

$$C(u, v) := F_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v)), \quad 0 \leq u, v \leq 1.$$ 

e) Compute the resulting copula density function $\frac{\partial^2 C}{\partial u \partial v}(u, v), \ u, v \in [0, 1]$. 

This version: June 27, 2020
https://www.ntu.edu.sg/home/nprivault/index.html