Lyon PhD Course Actuarial Science

Chapter 6 - Introduction to the theory of copulas

Professor : Etienne Marceau

May 16, 2017

Contents

1 Introduction **14**

2 Motivation **15**

2.1 Estimation for continuous univariate distribution 15

2.2 Estimation for continuous multivariate distribution 17

2.3 Definitions 20

2.4 Sklar's Theorem 22

2.5 Definition of a bivariate copula 24

2.6 Definition for a multivariate copula 27

3 Theorem of Sklar 31

3.1 The theorem itself 31

3.2 Proof of the theorem 32

4 Practical consequences of Sklar's theorem 36

4.1 Identify a copula using Sklar's theorem 37

4.1.1 Example - EFGM's copula 38

4.2 Building a multivariate distribution using Sklar's Theorem . . . 39

4.3 Advantages of the approach based on copulas 40

5 Fréchet's Bounds 41

6 Copula and pdf 43

7 Simulation 45

8 Associated copulas **47**

9 Survival copula and Sklar's Theorem **52**

10 Invariance **55**

11 Conditional cdf **57**

11.1 Definition 57

11.2 Simulation and conditional cdf 60

12 Famillies of copulas **61**

13 Basic copulas **62**

13.1 Independence copula 63

13.1.1 Definition 63

13.1.2 Simulation 63

13.2 Frechet Upper's bound copula 65

13.2.1 Definition 65

13.2.2 Simulation 65

13.3 Frechet's lower bound copula 67

13.3.1 Definition 67

13.3.2 Simulation 67

13.4 Complete copula 70

13.5 Eyraud-Farlie-Gumbel-Morgenstern (EFGM) bivariate copula . . 71

13.5.1 Definition 71

13.5.2 Simulation 73

14 Archimedean Copulas 76

14.1 Introduction 77

14.2 Construction of archimedean copulas by common frailty - Step #1 79

14.3 Construction of archimedean copulas by common frailty - Step #2 81

14.4 A simulation procedure for Archimedean copulas based on common mixture 84

14.5 Bivariate Ali-Mikhail-Haq (AMH) Copula 88

14.5.1 Definition 88

14.5.2 Simulation procedures 91

14.5.3 Simulation based on a general method for Archimedean copulas 91

14.5.4 Simulation based on the conditional cdf 93

14.6 Multivariate copula 96

14.7 Bivariate Clayton Copula 97

14.7.1 Definition 97

14.7.2 Simulation procedures 99

14.7.3 Simulation based on a general method for Archimedean copulas 100

14.7.4 Simulation based on the conditional cdf 104

14.7.5 Multivariate copula 107

14.8 Bivariate Frank copula 109

14.8.1 Definition 109

14.8.2 Simulation procedures 113

14.8.3 Simulation based on a general method for Archimedean copulas 114

14.8.4 Simulation based on the conditional cdf 118

14.8.5 Multivariate copula 121

14.9 Bivariate Gumbel copula 123

14.9.1 Definition 123

14.9.2 Simulation procedures 127

14.9.3 Simulation based on a general method for Archimedean copulas 128

14.9.4 Simulation based on the conditional cdf 133

14.9.5 Multivariate copula 134

15 Elliptic copulas 137

15.1 Bivariate Normal copula 138

15.1.1 Definition 138

15.1.2 Simulation procedure for the bivariate standard normal distribution 142

15.1.3 Simulation procedure for the bivariate copula 147

15.1.4 Multivariate copula 151

15.2 Bivariate Student copula 153

16 Challenges for research in actuarial science 154

1 Introduction

Nowadays, copulas are vevy well known in actuarial science, quantitive risk management, and related field.

2 Motivation

2.1 Estimation for continuous univariate distribution

Let X be a continuous rv with cdf F_X and pdf f_X.

We have a set of observations denoted by

$$\{x_1, ..., x_m\}.$$

The parameters of the distribution of X is represented by the vector

$$\underline{\theta} = (\theta_1, ..., \theta_l).$$

We can use the ML method in order to estimate the parameters.

The likelihood function is given by

$$L\left(\underline{\theta}\right) = \prod_{j=1}^{m} f_X\left(x_j; \underline{\theta}\right).$$

The loglikelihood function is given by

$$l\left(\underline{\theta}\right) = \sum_{j=1}^{m} \ln\left(f_X\left(x_j; \underline{\theta}\right)\right).$$

Generally, we need to use numeral optimization tools in order to find the estimator of $\widehat{\underline{\theta}}$ of $\underline{\theta}$.

Remark: usually the number of parameters are small (1,2,3) and the task is not to difficult to perform.

2.2 Estimation for continuous multivariate distribution

Let $\underline{X}$ be a vector of continuous rvs with joint cdf $F_{\underline{X}}$ and joint pdf $f_{\underline{X}}$.

We have a set of observations denoted by

$$\{\underline{x}_1, ..., \underline{x}_m\}.$$

The parameters of the joint distribution of $\underline{X}$ is represented by the vector

$$\underline{\theta} = (\theta_1, ..., \theta_l).$$

We can use the ML method in order to estimate the parameters.

The likelihood function is given by

$$L\left(\underline{\theta}\right) = \prod_{j=1}^{m} f_{\underline{X}}\left(\underline{x}_j; \underline{\theta}\right).$$

The loglikelihood function is given by

$$l\left(\underline{\theta}\right) = \sum_{j=1}^{m} \ln\left(f_{\underline{X}}\left(\underline{x}_j; \underline{\theta}\right)\right).$$

Again, we need to use numeral optimization tools in order to find the estimator of $\widehat{\underline{\theta}}$ of $\underline{\theta}$. However, the procedure can become very difficult to perform due the large number of parameters to estimate.

The theory of copulas allows (notably) to split the estimation procedure in two steps:

- estimate the parameters of each marginals
- estimate the function that links these marginals
- this function is called a copula.

2.3 Definitions

A copula C is a bivariate function defined over $[0,1]^2$, which takes values in $[0,1]$ and which satisfies the following properties :

1. $C(u_1,u_2) \uparrow$ as $u_1 \uparrow$ and $u_2 \uparrow$;
2. $C(u_1,1) = u_1$;
3. $C(1,u_2) = u_2$;
4. $C(1,1) = 1$;
5. $C(0,u_2) = 0$;
6. $C(u_1,0) = 0$; and
7. Rectangle property:

$$C(b_1,b_2) - C(a_1,b_2) - C(b_1,a_2) + C(a_1,a_2) \geq 0$$

for any rectangle defined in $[0,1]^2$ with vertices (a_1,a_2), (b_1,a_2), (a_1,b_2), (b_1,b_2), $0 \leq a_1 < b_1 \leq 1$ and $0 \leq a_2 < b_2 \leq 1$.

In fact, the copula corresponds to the cdf of the vector of rvs $\underline{U} = (U_1, U_2)$ where $U_1 \sim U_2 \sim Unif\,(0,1)$ over $[0,1]^2$ i.e.

$$F_{\underline{U}}\,(u_1, u_2) = C\,(u_1, u_2)$$

for $(u_1, u_2) \in [0,1]^2$.

Now, we are ready to provide Sklar's Theorem.

2.4 Sklar's Theorem

The theorem has two parts:

Theorem. Let $F_{\underline{X}} \in \Gamma\left(F_{X_1}, F_{X_2}\right)$.

1. **(Derivation of the copula from a bivariate cdf)** Then, it is existed a copula C such that
$$F_{X_1,X_2}(x_1, x_2) = C\left(F_{X_1}(x_1), F_{X_2}(x_2)\right).$$
If the marginals are continuous, then the copula C is unique.
2. **(Construction of bivariate distribution)** Let F_{X_1} and F_{X_2} be two marginals. Let C be a copula. Then, we can use the copula to link these marginals and the resulting bivariate function is a proper cdf $\in \Gamma\left(F_{X_1}, F_{X_2}\right)$ i.e
$$C\left(F_{X_1}(x_1), F_{X_2}(x_2)\right) = F_{X_1,X_2}(x_1, x_2)$$

where $F_{X_1,X_2} \in \Gamma\left(F_{X_1}, F_{X_2}\right)$. The marginals can be continuous, discrete or mixed.

2.5 Definition of a bivariate copula

Definition 1 *A copula* $C(u_1, u_2)$ *is an application* $[0,1]^2 \longrightarrow [0,1]$ *with the following properties (bivariate cdf of a pair of rvs with standard uniform marginals) :*

- $C(u_1, u_2)$ *is non decreasing on* $[0,1]^2$ *;*
- $C(u_1, u_2)$ *is right continuous on* $[0,1]^2$ *;*
- $\lim_{u_i \longrightarrow 0} C(u_1, u_2) = 0$ *for* $i = 1, 2$ *;*
- $\lim_{u_1 \longrightarrow 1} C(u_1, u_2) = u_2$ *and* $\lim_{u_2 \longrightarrow 1} C(u_1, u_2) = u_1$ *;*
- **Inequality of the rectangle** *: for all* $a_1 \leq b_1$ *and* $a_2 \leq b_2$*, we have*

$$C(b_1, b_2) - C(b_1, a_2) - C(a_1, b_2) + C(a_1, a_2) \geq 0.$$

Interpretation, the bivariate function C *is a copula if it allocates a positive probability mass over all rectangles* $(a_1, b_1] \times (a_2, b_2]$ *included in* $[0,1]^2$.

Notation 2 *Let C be a copula and let $\underline{U} = (U_1, U_2)$ ($U_1 \sim U_2 \sim Unif\,(0,1)$ with*

$$F_{\underline{U}}\,(u_1, u_2) = C\,(u_1, u_2)$$

for $u_i \in [0,1]$, $i = 1, 2$. Then, we denote $\underline{U} \sim C$.

Remark 3 *Let $\underline{U} \sim C$. Then, the inequality of the rectangle means*

$$\begin{aligned} C\,(b_1, b_2) - C\,(b_1, a_2) - C\,(a_1, b_2) + C\,(a_1, a_2) &= \Pr\,(a_1 < U_1 \leq b_1, a_2 < U_2 \leq \\ &= \Pr\,(U_1 \in (a_1, b_1]\,, U_2 \in (a_2, b \\ &= \Pr\left(\bigcap_{i=1}^{2} \{U_i \in (a_i, b_i]\}\right) \\ &= \Pr\,(\underline{U} \in (a_1, b_1] \times (a_2, b_2])\,, \end{aligned}$$

which corresponds to the probability that $\underline{U} = (U_1, U_2)$ takes values with the rectangle $(a_1, b_1] \times (a_2, b_2]$.

Remark 4 *Inequality of the rectangle and continuous copula. Let C be a continuous copula. The inequality of the rectangle is satisfied if*

$$\frac{\partial^2}{\partial u_1 \partial u_2} C\left(u_1, u_2\right) \geq 0,$$

for $u_i \in [0,1]$, $i = 1, 2$.

2.6 Definition for a multivariate copula

A copula C is the joint cdf on $[0,1]^n$ of a vector of rvs $\underline{U} = (U_1, ..., U_n)$ where the components U_i $(i = 1, ..., n)$ follow the (univariate) standard uniform distribution $(0,1)$.

Definition 5 *A multivariate copula* $C(u_1, ..., u_n)$ *is application* $[0,1]^n \rightarrow [0,1]$ *with the following properties:*

- $C(u_1, ..., u_n)$ *is non decreasing on* $[0,1]^n$ *;*
- $C(u_1, ..., u_n)$ *is right continuous on* $[0,1]^n$ *;*
- $\lim_{u_i \longrightarrow 0} C(u_1, ..., u_n) = 0$ *pour* $i = 1, ..., n$ *;*
- $\lim_{u_j \longrightarrow 1, j \neq i} C(u_1, ..., u_n) = u_i$ *;*

- *Inequality of the hyperrectangle: for all* $a_1 \leq b_1$, *...,* $a_n \leq b_n$*, we have*

$$\sum_{i_1=1}^{2} \dots \sum_{i_n=1}^{2} (-1)^{i_1+\dots+i_n} C\left(u_{1,i_1}, \dots, u_{n,i_n}\right) \geq 0,$$

with $u_{j,1} = a_j$ *and* $u_{j,2} = b_j$ *for* $j \in \{1, 2, \dots, n\}$.
Interpretatuon: the multivariate function C *is a copula if it allocated a positive probability mass on every hyperrectangle* $(a_1, b_1] \times \dots \times (a_n, b_n]$ *included in* $[0, 1]^n$.

Definition 6 *Hyperrectangle = généralisation à une dimension supérieure (*$n > 2$*) d'un rectangle de dimension deux*

Notation 7 *Let* C *be a copula and let* $\underline{U} = (U_1, \dots, U_n)$ *(*$U_1 \sim \dots \sim U_n \sim Unif\,(0, 1)$ *with*

$$F_{\underline{U}}\,(u_1, \dots, u_n) = C\,(u_1, \dots, u_n)\,,$$

for $u_i \in [0, 1]$, $i = 1, ..., n$. *Then, we denote* $\underline{U} \sim C$.

Remark 8 *Let* $\underline{U} \sim C$. *Then, the inequality of the hyperrectangle means*

$$\begin{aligned} \sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1+\ldots+i_n} C\left(u_{1,i_1}, \ldots, u_{n,i_n}\right) &= \Pr\left(\bigcap_{i=1}^{n} \{U_i \in (a_i, b_i]\}\right) \\ &= \Pr\left(\underline{U} \in (a_1, b_1] \times \ldots \times (a_n, b_n]\right) \end{aligned}$$

for every hyperrectangle $(a_1, b_1] \times \ldots \times (a_n, b_n] \subseteq [0, 1]^n$.

If $n = 3$, we have

$$\begin{aligned} \Pr\left(\bigcap_{i=1}^{3} \{U_i \in (a_i, b_i]\}\right) &= C(b_1, b_2, b_3) - C(a_1, b_2, b_3) - C(b_1, a_2, b_3) - C(b_1, \\ &\quad + C(a_1, a_2, b_3) + C(a_1, b_2, a_3) + C(b_1, a_2, a_3) \\ &\quad - C(a_1, a_2, a_3). \end{aligned}$$

Remark 9 *inequality of the hyperrectangle and continuous copula. Let C be a continuous copula. The inequality of the hyperrectangle is satisfied if*

$$\frac{\partial^n}{\partial u_1 ... \partial u_n} C\left(u_1, ..., u_n\right) \geq 0,$$

forr $u_i \in [0, 1]$, $i = 1, ..., n$.

3 Theorem of Sklar

3.1 The theorem itself

The fundamental result of the theory of copulas is the theorem of Sklar (Sklar's Theorem)

Theorem 10 Théorème de Sklar

- *Let* $F \in \Gamma\left(F_1, ..., F_n\right)$ *with marginal cdf* F_1, ..., F_n.

- *Part #1 (Identification of the copula). Then, there is a copula C such that for all $\underline{x} \in \mathbb{R}^n$, we have*

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)).$$

If F_1, ..., F_n are continuous, then C is unique. Otherwise, C uniquely determined on $RanF_1 \times ... \times RanF_n$.

- *Part #2 (Construction of a multivariate cdf). Ii C is a copula and F_1 ... F_2 are marginal cdfs. Then the function defined by*

$$F(x_1, ..., x_2) = C(F_1(x_1), ..., F_n(x_n))$$

is a multivarate cdf with marginal univariate cdf F_1 ... F_n.

3.2 Proof of the theorem

We provide the provide the proof when F_1, ..., F_n are continuous.

For a detailed proof, see e.g. [**?**] or [**?**].

- Part #1. Let $\underline{X} = (X_1, ..., X_n)$ be a vector of continous rvs with a multivariate $F_{\underline{X}}$ for which F_1, ..., F_n are continuous. Then, we have

$$\begin{aligned} F_{\underline{X}}(x_1, ..., x_n) &= \Pr(X_1 \leq x_1, ..., X_n \leq x_n) \\ &= \Pr\left(F_{X_1}(X_1) \leq F_{X_1}(x_1), ..., F_{X_n}(X_n) \leq F_{X_n}(x_n)\right), \end{aligned}$$

for $x_i \in \mathbb{R}$, $i = 1, 2, ..., n$. Using the probability integral transform Theorem (or probability transformation theorem), we know that

$$F_{X_i}(X_i) = U_i \sim Unif(0, 1)$$

for $i = 1, 2, ..., n$. Then,

$$\Pr\left(F_{X_1}(X_1) \leq F_{X_1}(x_1), ..., F_{X_n}(X_n) \leq F_{X_n}(x_n)\right)$$

corresponds to the multivariate cdf of

$$\left(F_{X_1}(X_1), .., F_{X_n}(X_n)\right) = (U_1, ..., U_n).$$

We introduce the multivariate function C such that

$$\begin{aligned} F_{\underline{X}}(x_1, ..., x_n) &= \Pr\left(F_{X_1}(X_1) \leq F_{X_1}(x_1), ..., F_{X_n}(X_n) \leq F_{X_n}(x_n)\right) \\ &= C\left(F_{X_1}(x_1), ..., F_{X_n}(x_n)\right) \end{aligned}$$

From Definition 5, the function C is a copula.

- Part #2. Let $\underline{U} \sim C$. We define a vector of rvs. $\underline{X} = (X_1, ..., X_n)$, where

$$X_i = F_i^{-1}(U_i)$$

for $i = 1, 2, ..., n$. Then, the multivariate cdf of $\underline{X}$ is given by

$$\begin{aligned}
F_{\underline{X}}(x_1, ..., x_n) &= \Pr(X_1 \leq x_1, ..., X_n \leq n_n) \\
&= \Pr\left(F_1^{-1}(U_1) \leq x_1, ..., F_n^{-1}(U_n) \leq x_n\right) \\
&= \Pr\left(U_1 \leq F_{X_1}(x_1), ..., U_n \leq F_{X_n}(x_n)\right) \\
&= C(F_1(x_1), ..., F_n(x_n)).
\end{aligned}$$

4 Practical consequences of Sklar's theorem

There is two practical consequences to Sklar's theorem:

1. identify (extract) the copula associated to the joint cdf on any continuous multivariate distribution;
2. build a multivariate distribution by linking univariate marginals (continuous or discrete) with a copula.

4.1 Identify a copula using Sklar's theorem

Corollary 11 *Let $F_{\underline{X}} \in \Gamma(F_1, ..., F_n)$, with continuous univariate marginal cdfs $F_1, ..., F_n$. Then, the copula C associated to $F_{\underline{X}}$ is*

$$C(u_1, ..., u_n) = F_{\underline{X}}\left(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)\right).$$

Preuve. From the proof of Part#1 of Sklar's Theorem, the copula C associated to $F_{\underline{X}}$ is

$$\begin{aligned} C(u_1, ..., u_n) &= \Pr(U_1 \leq u_1, ..., U_n \leq u_n) \\ &= \Pr(F_1(X_1) \leq u_1, ..., F_n(X_n) \leq u_n) \\ &= \Pr\left(X_1 \leq F_1^{-1}(u_1), ..., X_n \leq F_n^{-1}(u_n)\right) \\ &= F_{\underline{X}}\left(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)\right). \end{aligned}$$

■

4.1.1 Example - EFGM's copula

Let F_{X_1,X_2} be the joint cdf of EFGM's bivariate exponential distribution given by

$$\begin{aligned} F_{X_1,X_2}(x_1,x_2) &= \left(1-\mathrm{e}^{-\beta_1 x_1}\right)\left(1-\mathrm{e}^{-\beta_2 x_2}\right) \\ &\quad + \theta\left(1-\mathrm{e}^{-\beta_1 x_1}\right)\left(1-\mathrm{e}^{-\beta_2 x_2}\right)\mathrm{e}^{-\beta_1 x_1}\mathrm{e}^{-\beta_2 x_2}, \quad (1) \end{aligned}$$

with a dependence paramter $-1 \leq \theta \leq 1$ and with $\beta_i > 0$.

Replacing x_i by $-\frac{1}{\beta_i}\ln(1-u_i)$ $(i=1,2)$ in (1), we find

$$C(u_1,u_2) = u_1u_2 + \theta u_1u_2(1-u_1)(1-u_2)$$

for $0 \leq u_1, u_2 \leq 1$.

4.2 Building a multivariate distribution using Sklar's Theorem

Due to Part #2 of Sklar's Theroem, a copula allows us to combine different marginals (contiuous, discrete) in the aim to create a multivariate distribution.

The copula $C(u_1, ..., u_n)$ describe the dependence relation between the X_1, ..., X_n while the marginals describe the stochastic behavior of each rvs.

When the rvs are continuous, the copula contains all the information with regards to the dependence relation between the rvs.

Therefore, it is possible to analyze the marginals and the copula separately.

The case with discrete marginals is treated in Section **??**.

4.3 Advantages of the approach based on copulas

- The dependence relation is contained within the copulas : this is true if the marginals are continuous
- Flexbility: we can build any multivariate distributions with copulas and marginals (either discrete or continuous)
- It helps a lot to understand the dependence relation between rvs.
- Eventually: it helps when we need to estimate the parameters of a multivariate distributions
- Simulation is "very easy".

5 Fréchet's Bounds

Let C be a copula defined over $[0,1]^n$. Since C is a cdf, we have

$$W\left(u_1,...,u_n\right) \leq C\left(u_1,...,u_n\right) \leq M\left(u_1,...,un\right),$$

where

$$W\left(u_1,...,u_n\right) = \max\left(u_1+...+u_n-(n-1);0\right)$$

and

$$M\left(u_1,...,u_n\right) = \min\left(u_1;...;u_n\right)$$

correspond to lower and upper Fréchet's bounds, respectively.

For all $n = 2, 3, ...,$

$$C^+\left(u_1,...,u_n\right) = M\left(u_1,...,u_n\right)$$

corresponds to the copula Fréchet's upper bound.

For $n = 2$ (**only**),

$$C^{-}(u_1, u_2) = W(u_1, u_2)$$

corresponds to the copula Fréchet's upper bound.

6 Copula and pdf

Let C be a continuous copula.

The joint pdf associated to the copula C is given by

$$c(u_1, ..., u_n) = \frac{\partial^n}{\partial u_1 ... \partial u_n} C(u_1, ..., u_n).$$

Let $\underline{X} = (X_1, ..., X_n)$ be a vector of continuous rvs with joint cdf given by

$$F_{\underline{X}}(x_1, ..., x_n) = C\left(F_{X_1}(x_1), ..., F_{X_n}(x_n)\right),$$

Then, the joint pdf of $\underline{X} = (X_1, ..., X_n)$ is given

$$\begin{aligned} f_{\underline{X}}(x_1, ..., x_n) &= \frac{\partial^n}{\partial x_1 ... \partial x_n} F_{\underline{X}}(x_1, ..., x_n) \\ &= c\left(F_{X_1}(x_1), ..., F_{X_n}(x_n)\right) f_{X_1}(x_1) ... f_{X_n}(x_n), \end{aligned}$$

where $f_{X_i}(x_i)$ is the pdf of the rv X_i i.e. $f_{X_i}(x_i) = \frac{\mathrm{d}F_{X_i}(x_i)}{\mathrm{d}x_i}$ for $i = 1, ..., n$.

7 Simulation

Simulation with copulas are easy.

Let $\underline{X} = (X_1, ..., X_n)$ be a vector of rvs with

$$F_{\underline{X}}(x_1, ..., x_n) = C\left(F_{X_1}(x_1), ..., F_{X_n}(x_n)\right),$$

We want to simulate a sampled value $\underline{X}^{(j)} = \left(X_1^{(j)}, ..., X_n^{(j)}\right)$ of $\underline{X}$.

The marginals of $F_{\underline{X}}$ may be continuous or discrete.

Algorithm 12 *General algorithm.*

1. *Simulate a sampled value* $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ *of* $\underline{U} \sim C$.
2. *Simulate a sampled value* $\underline{X}^{(j)}$ *of* $\underline{X}$ *with*

$$X_1^{(j)} = F_{X_1}^{-1}\left(U_1^{(j)}\right) \quad ... \quad X_n^{(j)} = F_{X_n}^{-1}\left(U_n^{(j)}\right).$$

8 Associated copulas

Let C be a copula and $\underline{U} = (U_1, U_2) \sim C$.

We consider 3 copulas for the 3 following pairs of rvs :

$$(U_1, 1 - U_2), (1 - U_1, U_2), (1 - U_1, 1 - U_2).$$

The copula for $(U_1, 1 - U_2)$ is

$$\begin{aligned} F_{U_1, 1-U_2}(u_1, u_2) &= \Pr(U_1 \leq u_1, 1 - U_2 \leq u_2) \\ &= \Pr(U_1 \leq u_1, U_2 > 1 - u_2) \\ &= u_1 - C(u_1, 1 - u_2). \end{aligned}$$

The copula for $(1-U_1, U_2)$ is

$$\begin{aligned} F_{1-U_1,U_2}(u_1,u_2) &= \Pr(1-U_1 \leq u_1, U_2 \leq u_2) \\ &= \Pr(U_1 > 1-u_1, U_2 \leq u_2) \\ &= u_2 - C(1-u_1, u_2). \end{aligned}$$

The copula for $(1-U_1, 1-U_2)$ is called survival copula and denoted $\widehat{C}(u_1,u_2)$.

It is given by

$$\begin{aligned} F_{1-U_1,1-U_2}(u_1,u_2) &= \widehat{C}(u_1,u_2) \\ &= \Pr(1-U_1 \leq u_1, 1-U_2 \leq u_2) \\ &= \Pr(U_1 > 1-u_1, U_2 > 1-u_2) \\ &= 1-(1-u_1)-(1-u_2)+C(1-u_1,1-u_2) \\ &= C(1-u_1,1-u_2)+u_1+u_2-1. \end{aligned}$$

The forementioned copulas satisfy the following properties:

$$C^{-}(u_1,u_2) \leq F_{U_1,1-U_2}(u_1,u_2) = u_1 - C(u_1, 1-u_2) \leq C^{+}(u_1,u_2)$$

$$C^{-}(u_1,u_2) \leq F_{1-U_1,U_2}(u_1,u_2) = u_2 - C(1-u_1, u_2) \leq C^{+}(u_1,u_2)$$

$$C^{-}(u_1,u_2) \leq F_{1-U_1,1-U_2}(u_1,u_2) = \widehat{C}(u_1,u_2) = C(1-u_1, 1-u_2) + u_1 + u_2 - 1$$

Let c be the pdf associated to the continuous copula C.

- The pdf associated to copula $u_1 - C(u_1, 1-u_2)$ is
$$c(u_1, 1-u_2).$$
- The pdf associated to copula $u_2 - C(1-u_1, u_2)$ is
$$c(1-u_1, u_2).$$

- The pdf associated to $\widehat{C}(u_1, u_2)$ is

$$\widehat{c}(u_1, u_2) = c(1 - u_1, 1 - u_2).$$

Let (X_1, X_2) be a pair of rvs with marginals F_{X_1} and F_{X_2} :

- Let $X_1 = F_{X_1}^{-1}(U_1)$ and $X_2 = F_{X_2}^{-1}(1 - U_2)$. Then, we have

$$F_{X_1,X_2}(x_1, x_2) = F_{U_1,1-U_2}\left(F_{X_1}(x_1), F_{X_2}(x_2)\right) = F_{X_1}(x_1) - C\left(F_{X_1}(x_1), 1\right.$$

- Let $X_1 = F_{X_1}^{-1}(1 - U_1)$ and $X_2 = F_{X_2}^{-1}(U_2)$. Then, we have

$$F_{X_1,X_2}(x_1, x_2) = F_{1-U_1,U_2}\left(F_{X_1}(x_1), F_{X_2}(x_2)\right) = F_{X_2}(x_2) - C\left(1 - F_{X_1}(x\right.$$

- Let $X_1 = F_{X_1}^{-1}(1-U_1)$ and $X_2 = F_{X_2}^{-1}(1-U_2)$. Then, we have

$$\begin{aligned} F_{X_1,X_2}(x_1,x_2) &= F_{1-U_1,1-U_2}\left(F_{X_1}(x_1), F_{X_2}(x_2)\right) \\ &= \widehat{C}\left(F_{X_1}(x_1), F_{X_2}(x_2)\right) \\ &= C\left(1-F_{X_1}(x_1), 1-F_{X_2}(x_2)\right) + F_{X_1}(x_1) + F_{X_2}(x_2) - \end{aligned}$$

9 Survival copula and Sklar's Theorem

Let C be a copula of dimension 2 and $\underline{U} = (U_1, U_2) \sim C$.

Let (X_1, X_2) be a pair of rvs with marginals F_{X_1} and F_{X_2}.

Then, we have

$$\begin{aligned}
\overline{F}_{X_1,X_2}(x_1, x_2) &= 1 - F_{X_1}(x_1) - F_{X_2}(x_2) + F_{X_1,X_2}(x_1, x_2) \\
&= C\left(F_{X_1}(x_1), F_{X_2}(x_2)\right) + 1 - F_{X_1}(x_1) - F_{X_2}(x_2) \\
&= C\left(1 - \overline{F}_{X_1}(x_1), 1 - \overline{F}_{X_2}(x_2)\right) + \overline{F}_{X_1}(x_1) + \overline{F}_{X_2}(x_2) - 1
\end{aligned}$$

Since

$$\widehat{C}(u_1, u_2) = C(1 - u_1, 1 - u_2) + u_1 + u_2 - 1,$$

we find

$$\overline{F}_{X_1,X_2}(x_1,x_2) = \widehat{C}\left(\overline{F}_{X_1}(x_1), \overline{F}_{X_2}(x_2)\right).$$

Then, it is possible to adapt Sklar's Theorem in the context of multivariate survival functions.

For $n = 2, 3, 4, \ldots$, we have

$$\overline{F}_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \widehat{C}\left(\overline{F}_{X_1}(x_1), \ldots, \overline{F}_{X_n}(x_n)\right)$$

allowing to identify a copula $\widehat{C}$ from the joint survival function of the multivariate continuous distribution with

$$\widehat{C}(u_1,\ldots,u_n) = \overline{F}_{X_1,\ldots,X_n}\left(\overline{F}_{X_1}^{-1}(u_1), \ldots, \overline{F}_{X_n}^{-1}(u_n)\right).$$

Since $\widehat{C}$ is a copula, we will drop the decoration "$\widehat{\ }$" when it will not be useful.

We will use this method to generate copulas belonging to the class of Archimedean copulas (including notably the Clayton copula, the AMH copula, the Frank copula and the Gumbel copula).

10 Invariance

The invariance property is an attracting property of the copula.

Let (X_1, X_2) be a pair of continuous rvs.

The dependence structure between X_1 and X_2 is defined by the copula C.

Let ϕ_1 and ϕ_2 be continuous monotone functions.

We have the following properties :

- If ϕ_1 and ϕ_2 are non decreasing, the dependence structure of $(\phi_1(X_1), \phi_2(X_2))$ corresponds to the copula $C(u_1, u_2)$.

- If ϕ_1 is non decreasing and ϕ_2 is non increasing, the dependence structure of $(\phi_1(X_1), \phi_2(X_2))$ corresponds to the copula $u_1 - C(u_1, 1 - u_2)$.
- If ϕ_1 is non increasing and ϕ_2 is non decreasing, the dependence structure of $(\phi_1(X_1), \phi_2(X_2))$ corresponds to the copula $u_2 - C(1 - u_1, u_2)$.
- If ϕ_1 is non increasing and ϕ_2 is non increasing, the dependence structure of $(\phi_1(X_1), \phi_2(X_2))$ corresponds to the copula $\widehat{C}(u_1, u_2)$.

The form of the copula does not depend on ϕ_1 and ϕ_2, which it is important for estimation.

11 Conditional cdf

11.1 Definition

Let C be a copula for which the partial derivatives with regards to u_1 and u_2 exist.

Let $\underline{U} = (U_1, U_2) \sim C$.

The conditional cdf of U_2 given $U_1 = u_1$ is

$$\begin{aligned} C_{2|1}(u_2|u_1) &= \Pr(U_2 \leq u_2 | U_1 = u_1) \\ &= \lim_{h\to 0} \frac{C(u_1+h, u_2) - C(u_1, u_2)}{h} \\ &= \frac{\partial}{\partial u_1} C(u_1, u_2). \end{aligned}$$

Similarly, the conditional cdf of U_1 given $U_2 = u_2$ is

$$\begin{aligned} C_{1|2}(u_1|u_2) &= \Pr(U_1 \leq u_1 | U_2 = u_2) \\ &= \lim_{h\to 0} \frac{C(u_1, u_2+h) - C(u_1, u_2)}{h} \\ &= \frac{\partial}{\partial u_2} C(u_1, u_2). \end{aligned}$$

Let (X_1, X_2) be a pair of continuous rvs with

$$F_{X_1,X_2}(x_1, x_2) = C\left(F_{X_1}(x_1), F_{X_2}(x_2)\right)$$

for $x_1, x_2 \in \mathbb{R}$.

Then, we have

$$\begin{aligned} F_{X_2|X_1=x_1}(x_2) &= \Pr(X_2 \leq x_2 | X_1 = x_1) \\ &= C_{2|1}\left(F_{X_2}(x_2) | F_{X_1}(x_1)\right) \end{aligned}$$

and

$$\begin{aligned} F_{X_1|X_2=x_2}(x_1) &= \Pr(X_1 \leq x_1 | X_2 = x_2) \\ &= C_{1|2}\left(F_{X_1}(x_1) | F_{X_2}(x_2)\right). \end{aligned}$$

11.2 Simulation and conditional cdf

Let $\underline{U} = (U_1, U_2) \sim C$.

One could apply an approach based on the conditional cdf to simulate sampled values $\left(U_1^{(j)}, U_2^{(j)}\right)$ $(j = 1, 2, ..., m)$ of (U_1, U_2).

Let $C_{2|1}^{-1}(v|u_1)$ be the inverse of $C_{2|1}(u_2|u_1)$, which is obtained by finding the solution to $C_{2|1}(u_2|u_1) = v$

Algorithm 13 General procedure based on conditional cdf

1. *Simulate the sampled values $V_1^{(j)}$ and $V_2^{(j)}$ of the iid V_1 and V_2 where $V_i \sim U(0,1)$ for $i = 1, 2$.*
2. *Calculate $U_1^{(j)} = V_1^{(j)}$ and $U_2^{(j)} = C_{2|1}^{-1}\left(V_2^{(j)}|U_1^{(j)}\right)$.*

12 Famillies of copulas

To simplify the presentation, we have put togother copulas with similar features :

- basic copulas
- archimedean copulas
- elliptic copulas
- copulas with singularities
- etc.

13 Basic copulas

13.1 Independence copula

13.1.1 Definition

The independence copula is defined by

$$C^{\perp}(u_1, ..., u_n) = u_1 ... u_n,$$

for $u_i \in [0,1]$, $i = 1, 2, ..., n$.

13.1.2 Simulation

Algorithm 14 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C^{\perp}$.

1. *Simulate* $V_1^{(j)}$, ..., $V_n^{(j)}$ *of the iid rvs* V_1, ..., V_n *where* $V_i \sim U(0,1)$ *for* $i = 1, 2, ..., n$.
2. *Calculate* $U_i^{(j)} = V_i^{(j)}$, $i = 1, 2, ..., n$.

See Figure 1) for a typical scatter plot of sampled values of $\underline{U} = (U_1, U_2) \sim C^{\perp}$: there are uniformly distributed over $[0,1] \times [0,1]$.

13.2 Frechet Upper's bound copula

13.2.1 Definition

The copula is defined by

$$C^{+}(u_1, ..., u_n) = \min(u_1; ...; u_n),$$

for $u_i \in [0,1]$, $i = 1, 2, ..., n$.

13.2.2 Simulation

Algorithm 15 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C^{+}$.

1. *Simulate a sampled value* $V^{(j)}$ *of the rv* $V \sim U\,(0,1)$.
2. *Calculate* $U_i^{(j)} = V^{(j)}$, $i = 1, 2, ..., n$.

See Figure 1) for a typical scatter plot of sampled values of $\underline{U} = (U_1, U_2) \sim C^+:$ there are uniformly distributed on the diagonale $u_1 = u_2$ in $[0,1] \times [0,1]$.

13.3 Frechet's lower bound copula

13.3.1 Definition

The copula is defined by

$$C^{-}(u_1, u_2) = \max(u_1 + u_2 - 1; 0).$$

13.3.2 Simulation

Algorithm 16 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C^{-}$.

1. *Simulate a sampled value* $V^{(j)}$ *of the rv* $V \sim U\,(0,1)$.
2. *Calculate* $U_1^{(j)} = V^{(j)}$ *and* $U_2^{(j)} = 1 - V^{(j)}$.

See Figure 1) for a typical scatter plot of sampled values of $\underline{U} = (U_1, U_2) \sim C^-$: there are uniformly distributed on the diagonale $u_1 = 1 - u_2$ in $[0,1] \times [0,1]$.

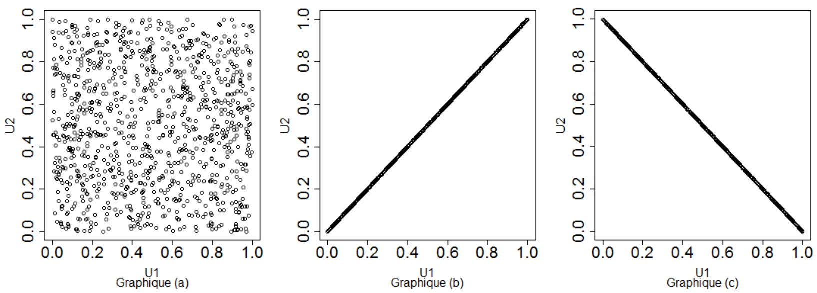

Figure 1: Graphs (a), (b) et (c) are the scatterplots of 1000 sampled values of (U_1, U_2) from independence copula, Frechet's upper bound copula, and Frechet's lowser bound copula.

13.4 Complete copula

When $n = 2$, a copula C is said to be *complete* if it includes Frechet's upper bound copula, independence copula, and Frechet's lower bound copula, either as a special case or a limit case.

13.5 Eyraud-Farlie-Gumbel-Morgenstern (EFGM) bivariate copula

13.5.1 Definition

The EFGM bivariate copula is obtained from EFGM's bivariate exponential ditribution using the relation (Part#1 of Sklar's Theorem)

$$C(u_1, u_2) = F_{X_1,X_2}\left(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)\right),$$

for $u_1, u_2 \in [0,1]$.

The EFGM bivariate copula can be seen as a perturbation to the independence copula

$$C_\alpha (u_1, u_2) = u_1 u_2 + \alpha u_1 u_2 (1 - u_1) (1 - u_2), \tag{2}$$

for $\alpha \in [-1, 1]$.

Special case: $C_0 (u_1, u_2) = C^\perp (u_1, u_2)$.

The pdf is

$$c (u_1, u_2) = 1 + \alpha (1 - 2u_1) (1 - 2u_2).$$

The EFGM induces both negative and positive dependence relations between two rvs.

However, it is not complete.

13.5.2 Simulation

L'algorithme suivant de simulation est basée sur la méthode de la fonction de répartition conditionnelle.

Algorithm 17 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C_\alpha^{EFGM}$.

1. *Simulate sampled values $V_1^{(j)}$ and $V_2^{(j)}$ of the iid rvs V_1 and V_2 where $V_i \sim U(0,1)$ for $i = 1, 2$.*
2. *Let $U_1^{(j)} = V_1^{(j)}$.*

3. *Define* $W_1^{(j)} = \alpha\left(2U_1^{(j)} - 1\right) - 1$ *and*

$$W_2^{(j)} = \left(1 - \alpha\left(2U_1^{(j)} - 1\right)\right)^2 + 4\alpha V_2^{(j)}\left(2U_1^{(j)} - 1\right).$$

4. *Calculate* $U_2^{(j)} = \frac{2V_2^{(j)}}{\left(\sqrt{W_2^{(j)}} - W_1^{(j)}\right)}$.

See Figure 2 for a typical scatter plot of sampled values of $\underline{U} = (U_1, U_2) \sim C_\alpha^{EFGM}$.

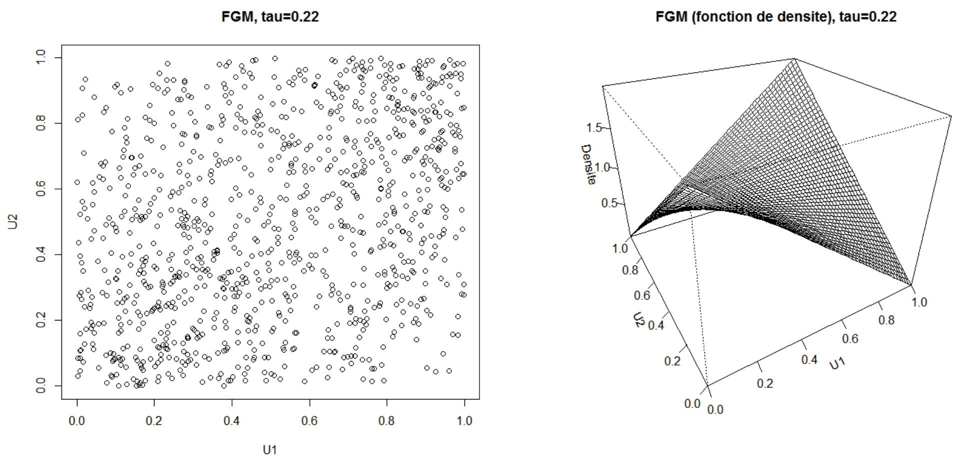

Figure 2: EFGM copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.22).

14 Archimedean Copulas

14.1 Introduction

A copula C is said to be archimedean if it can be written in the following way :

$$C(u_1, ..., u_n) = \psi\left(\psi^{-1}(u_1) + ... + \psi^{-1}(u_n)\right)$$

where the function ψ is called a generator.

The function ψ must satisfies the following properties :

- $\psi : [0, \infty) \to [0, 1]$ with $\psi(0) = 0$ and $\lim_{x \to \infty} \psi(x) = 0$
- ψ is a continuous function;
- ψ is strictly decreasing over $\left[0, \psi^{-1}(0)\right]$
- ψ^{-1} is the inverse of ψ, i.e.

$$\psi^{-1}(x) = \inf\{u, \psi(u) \leq x\}.$$

What is the function ψ ?

Which functions do meet these properties ?

Important: The function ψ will lead to an archimedean copula for any dimension $n = 2, 3, ...$ if and only if ψ is completely monotone i.e.

$$(-1)^k \frac{\mathrm{d}^k \psi(x)}{\mathrm{d}x^k} \geq 0.$$

A function is completely monotone if and only if it is a Laplace-Stieltjes transform of a strictly positive rv.

We present below a general method to construct archimedean copulas via Laplace-Stieltjes transform of a strictly positive rv.

This method is called by **common frailty**.

14.2 Construction of archimedean copulas by common frailty - Step #1

Let Θ be **a strictely positive** rv with cdf F_Θ and Laplace-Stietljes transform $\mathcal{L}_\Theta$.

Strictly positive : $F_\Theta(0) = 0$.

The rv Θ can be discrete or continuous.

Let $\underline{Y} = (Y_1, ..., Y_n)$ be a vector of rvs where

- $(Y_1|\Theta = \theta)$, ..., $(Y_n|\Theta = \theta)$ are conditionally independent;
- $(Y_i|\Theta = \theta) \sim Exp(\theta)$, $i = 1, 2, ..., n$.

The survival function of the rv Y_i is

$$\overline{F}_{Y_i}(x_i) = E\left[\mathrm{e}^{-\Theta x_i}\right] = \mathcal{L}_{\Theta}(x_i)$$

for $i = 1, 2, ..., n$.

The joint survival function of $\underline{Y}$ is

$$\begin{aligned}
\overline{F}_{\underline{Y}}(x_1, ..., x_n) &= \int_0^{\infty} \overline{F}_{\underline{Y}|\Theta=\theta}(x_1, ..., x_n)\, \mathrm{d}F_{\Theta}(\theta) \\
&= \int_0^{\infty} \overline{F}_{Y_1|\Theta=\theta}(x_1) \times ... \times \overline{F}_{Y_n|\Theta=\theta}(x_n)\, \mathrm{d}F_{\Theta}(\theta) \\
&= \int_0^{\infty} \mathrm{e}^{-\theta x_1} \times ... \times \mathrm{e}^{-\theta x_n} \mathrm{d}F_{\Theta}(\theta) \\
&= E\left[\mathrm{e}^{-\Theta x_1} \times ... \times \mathrm{e}^{-\Theta x_n}\right] \\
&= E\left[\mathrm{e}^{-\Theta(x_1 + ... + x_n)}\right] \\
&= \mathcal{L}_{\Theta}(x_1 + ... + x_n)
\end{aligned}$$

pour $i = 1, 2, ..., n$.

14.3 Construction of archimedean copulas by common frailty - Step #2

Now, we apply Part#1 of Sklar's Theorem to identify the copula C.

The expression of the copula is given by

$$C(u_1, ..., u_n) = \overline{F}_{\underline{Y}}\left(\overline{F}_{Y_1}^{-1}(u_1), ..., \overline{F}_{Y_i}^{-1}(u_n)\right).$$

We know that

$$x_i = \overline{F}_{Y_1}^{-1}(u_1) = \mathcal{L}_{\Theta}^{-1}(u_i)$$

forr $i = 1, 2, ..., n$.

Then, the expression of the copula is

$$C(u_1,...,u_n) = \mathcal{L}_\Theta\left(\mathcal{L}_\Theta^{-1}(u_1) + ... + \mathcal{L}_\Theta^{-1}(u_n)\right)$$

for $0 \leq u_1,...,u_n \leq 1$.

The function ψ corresponds to the Laplace-Stieltjes transform of the mixing rv Θ.

Example 18 *Let* $\Theta \sim Gamma\left(\frac{1}{\alpha},1\right)$ *with*

$$\mathcal{L}_\Theta(t) = \left(\frac{1}{1+t}\right)^{\frac{1}{\alpha}}.$$

The expression for $\mathcal{L}_\Theta^{-1}$ *is the solution of*

$$\mathcal{L}_\Theta(x) = \left(\frac{1}{1+x}\right)^{\frac{1}{\alpha}} = u.$$

We obtain

$$x = \mathcal{L}_{\Theta}^{-1}(u) = u^{-\alpha} - 1$$

for $u \in [0, 1]$.

We conclude

$$\begin{aligned} C(u_1, ..., u_n) &= \left(\frac{1}{1 + \sum_{i=1}^{n} \left(u_i^{-\alpha} - 1 \right)} \right)^{\frac{1}{\alpha}} \\ &= \left(\sum_{i=1}^{n} u_i^{-\alpha} - (n-1) \right)^{-\frac{1}{\alpha}}, \end{aligned}$$

for $0 \leq u_1, ..., u_n \leq 1$. *The result is called the Clayton Copula.* $\square$

14.4 A simulation procedure for Archimedean copulas based on common mixture

Recall: we know that

$$C(u_1, ..., u_n) = \overline{F}_{\underline{Y}}\left(\overline{F}_{Y_1}^{-1}(u_1), ..., \overline{F}_{Y_i}^{-1}(u_n)\right).$$

Objective: simulate sampled values $\underline{U}^{(j)}$ of $\underline{U} = (U_1, ..., U_n)$ where

$$F_{\underline{U}}(u_1, ..., u_n) = C(u_1, ..., u_n).$$

General method :

1. Simulate$\underline{Y}^{(j)}$ of $\underline{Y} = (Y_1, ..., Y_n)$.

2. Simulatie $\underline{U}^{(j)}$ of $\underline{U} = (U_1, ..., U_n)$ with
$$U_i^{(j)} = \overline{F}_{Y_i}\left(Y_i^{(j)}\right)$$
with $i = 1, 2, ..., n$.

General method of simulation for archimedan:

1. Simulate $\underline{Y}^{(j)}$ of $\underline{Y} = (Y_1, ..., Y_n)$ (derived from the construction of the distribution of$\underline{Y}$) :
   (a) Simulate the sampled value $\Theta^{(j)}$ of Θ.
   (b) Simulate the sampled value $\underline{V}^{(j)}$ of $\underline{V} = (V_1, ..., V_n)$ which is a vector of iid rvs with $V_i \sim Unif\,(0, 1)$.
   (c) Simulate $Y_i^{(j)}$ with
$$\left(Y_i|\Theta = \Theta^{(j)}\right) = F^{-1}_{Y_i|\Theta=\Theta^{(j)}}\left(V_i^{(j)}\right) = -\frac{1}{\Theta^{(j)}}\ln\left(1 - V_i^{(j)}\right)$$

forr $i = 1, 2, ..., n$.

2. Simlate the sampled value $\underline{U}^{(j)}$ of $\underline{U} = (U_1, ..., U_n)$ with

$$U_i^{(j)} = \overline{F}_{Y_i}\left(Y_i^{(j)}\right) = \mathcal{L}_\Theta\left(Y_i^{(j)}\right),$$

for $i = 1, 2, ..., n$.

We apply this algorithm for the Clayton copula:

1. Simulate $\underline{Y}^{(j)}$ of $\underline{Y} = (Y_1, ..., Y_n)$ (derived from the construction of the distribution of$\underline{Y}$) :
   (a) Simulate the sampled value $\Theta^{(j)}$ of $\Theta \sim Gamma\left(\frac{1}{\alpha}, 1\right)$.
   (b) Simulate the sampled value $\underline{V}^{(j)}$ of $\underline{V} = (V_1, ..., V_n)$ which is a vector of iid rvs with $V_i \sim Unif\,(0, 1)$.

(c) Simulate $Y_i^{(j)}$ with

$$\left(Y_i|\Theta=\Theta^{(j)}\right)=F^{-1}_{Y_i|\Theta=\Theta^{(j)}}\left(V_i^{(j)}\right)=-\frac{1}{\Theta^{(j)}}\ln\left(1-V_i^{(j)}\right)$$

forr $i=1,2,...,n$.

2. Simulate the sampled value $\underline{U}^{(j)}$ of $\underline{U}=(U_1,...,U_n)$ with

$$U_i^{(j)}=\overline{F}_{Y_i}\left(Y_i^{(j)}\right)=\mathcal{L}_{\Theta}\left(Y_i^{(j)}\right)=\left(\frac{1}{1+Y_i^{(j)}}\right)^{\frac{1}{\alpha}},$$

for $i=1,2,...,n$.

14.5 Bivariate Ali-Mikhail-Haq (AMH) Copula

14.5.1 Definition

The Ali-Mikhail-Haq (AMH) copula is an Archimedean copula.

Let Θ be strictly positive discrete rv which follows a geometric of type 2 distribuion.

- Notation : $\Theta \sim GeomType2(q)$
- Parameter : $q \in]0,1[$
- Support : $k \in \mathbb{N}^{+}$
- Pmf : $\Pr(\Theta = k) = q\,(1-q)^{k-1}$
- Cdf : $F_{\Theta}(k) = 1-(1-q)^{k}$

- Expectation : $E[\Theta] = \frac{1}{q}$
- Variance : $\mathrm{Var}(\Theta) = \frac{1-q}{q^2}$
- LST : $\mathcal{L}_\Theta(t) = \frac{q\mathrm{e}^t}{(1-(1-q)\mathrm{e}^t)}$
- Inverse of LST: $\mathcal{L}_\Theta^{-1}(u)$
- Pgf : $P_\Theta(t) = \frac{qt}{(1-(1-q)t)}$

Let $\alpha = 1 - q$.

The expression for the copula is

$$\begin{aligned} C_\alpha(u_1, u_2) &= \frac{u_1 u_2}{1 - \alpha(1-u_1)(1-u_2)} \\ &= u_1 u_2 \sum_{k=0}^{\infty} (\alpha(1-u_1)(1-u_2))^k . \end{aligned} \quad (3)$$

Given the representation by common frailty, one may expect that the dependence parameter $\alpha \in [0, 1]$.

However, (3) is also valid for $\alpha \in [-1, 1]$.

Special case: $C_0(u_1, u_2) = C^{\perp}(u_1, u_2)$.

The copula is not complete.

The copula introduces at most a moderate level of dependence relation between the rvs.

The pdf is

$$c(u_1, u_2) = \frac{1 - \alpha + 2\alpha \frac{u_1 u_2}{1-\alpha(1-u_1)(1-u_2)}}{(1 - \alpha(1 - u_1)(1 - u_2))^2}.$$

14.5.2 Simulation procedures

We provide two simulation algorithms:

- The first algorithm works for $a > 0$. It is based on Sklar's Theorem. It is application of a general method for Archimedean copulas.
- The second algorithm works for $\alpha > 0$ and $\alpha < 0$. It is application of the method based on conditional cdf.

14.5.3 Simulation based on a general method for Archimedean copulas

Algorithm 19 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_\alpha^{AMH}$ ($\alpha > 0$).

1. **(Simulate from the bivariate mixed exponential-geometric distribution).**
   (a) *Simulate the sampled value* $\Theta^{(j)}$ *of* $\Theta \sim GeometricType2\,(1-\alpha)$.
   (b) *Simulate the sampled values* $\left(Y_1^{(j)}, Y_2^{(j)}\right)$ *of* (Y_1, Y_2) *with*

$$Y_1^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right)$$

$$Y_2^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_2^{(j)}\right)$$

   where $V_1^{(j)}$ *and* $V_2^{(j)}$ *are the sampled values of the independent rvs* $V_1 \sim Unif\,(0,1)$ *and* $V_2 \sim Unif\,(0,1)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values* $U_1^{(j)}$ *and* $U_2^{(j)}$ *with*

$$U_1^{(j)} = \overline{F}_{Y_1}\left(Y_1^{(j)}\right) = \mathcal{L}_{\Theta}\left(Y_1^{(j)}\right) = \frac{(1-\alpha)}{\left(e^{Y_1^{(j)}} - \alpha\right)}$$

$$U_2^{(j)} = \overline{F}_{Y_2}\left(Y_2^{(j)}\right) = \mathcal{L}_{\Theta}\left(Y_2^{(j)}\right) = \frac{(1-\alpha)}{\left(e^{Y_2^{(j)}} - \alpha\right)}.$$

14.5.4 Simulation based on the conditional cdf

Algorithm 20 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_{\alpha}^{AMH}$ ($-1 \leq \alpha < 0$ and $0 < \alpha \leq 1$)

1. *Simulate the sampled values $V_1^{(j)}$ et $V_2^{(j)}$ of the iid rvs V_1 and V_2 where $V_i \sim U(0,1)$ for $i = 1, 2$.*
2. *Let $U_1^{(j)} = V_1^{(j)}$.*

3. *Define* $A = 1 - U_1^{(j)}$, $B = -\alpha\left(2AV_2^{(j)} + 1\right) + 2\alpha^2 A^2 \times V_2^{(j)} + 1$ *and*

$$C = 1 + \alpha^2\left(4A^2V_2^{(j)} - 4AV_2^{(j)} + 1\right) + \alpha\left(4AV_2^{(j)} - 4V_2^{(j)} + 2\right).$$

4. *Calculate* $U_2^{(j)} = \frac{2 \times V_2^{(j)}(A\alpha - 1)^2}{\left(B + \sqrt{C}\right)}$.

See Figure 3 for typical scatter plots of sampled values of $\underline{U} = (U_1, U_2) \sim C_\alpha^{AMH}$ and typical graphs of pdf of the copula.

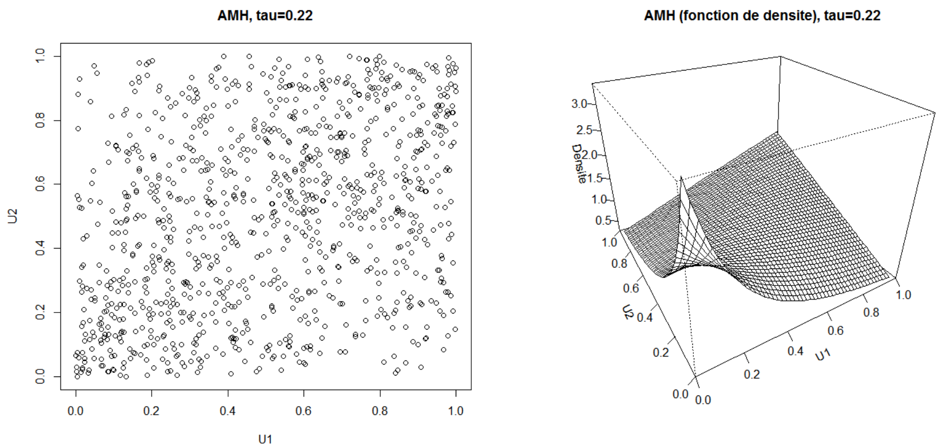

Figure 3: AMH copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.22).

14.6 Multivariate copula

The expression of the multivariate copula is

$$C_\alpha(u_1, ..., u_n) = \mathcal{L}_\Theta\left(\mathcal{L}_\Theta^{-1}(u_1) + ... + \mathcal{L}_\Theta^{-1}(u_n)\right),$$

for $u_i \in [0,1]$, $i = 1, ..., n$ and $\alpha \in [0,1]$.

14.7 Bivariate Clayton Copula

14.7.1 Definition

The Clayton copula is an Archimedean copula.

Let Θ be a strictly positive continuous rv with $\Theta \sim Gamma\left(\frac{1}{\alpha}, 1\right)$:

- LST:

$$\mathcal{L}_{\Theta}(t) = \left(\frac{1}{1+t}\right)^{\frac{1}{\alpha}}$$

- Inverse of LST:

$$x = \mathcal{L}_{\Theta}^{-1}(u) = u^{-\alpha} - 1$$

  for $u \in [0, 1]$.

The expression for the copula is

$$C_\alpha(u_1, u_2) = \left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-\frac{1}{\alpha}},$$

for $u_i \in [0,1]$, $i = 1,2$ is $\alpha > 0$.

Limit cases:

- Independence copula: $\lim_{\alpha\to 0} C_\alpha(u_1, u_2) = C^{\perp}(u_1, u_2)$
- Frechet's upper bound copula: $\lim_{\alpha\to\infty} C_\alpha(u_1, u_2) = C^{+}(u_1, u_2)$.

The pdf is

$$c(u_1, u_2) = \frac{1+\alpha}{(u_1 u_2)^{\alpha+1}} \left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-2-\frac{1}{\alpha}}.$$

The conditional cdf of U_2 given $U_1 = u_1$ is

$$C_{2|1}(u_2|u_1) = \frac{1}{u_1^{\alpha+1}} \left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-1-\frac{1}{\alpha}}.$$

14.7.2 Simulation procedures

We provide two simulation algorithms:

- The first algorithm works for $a > 0$. It is based on Sklar's Theorem. It is application of the general method for Archimedean copulas.
- The second algorithm works for $\alpha > 0$ and $\alpha < 0$. It is application of the method based on conditional cdf.

14.7.3 Simulation based on a general method for Archimedean copulas

Algorithm 21 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_\alpha^{Clay}$ ($\alpha > 0$).

1. **(Simulate from the bivariate mixed exponential-gamma distribution).**
(a) Simulate the sampled value $\Theta^{(j)}$ de la v.a. $\Theta \sim Ga\left(\frac{1}{\alpha}, 1\right)$.
(b) Simulate the sampled values $\left(Y_1^{(j)}, Y_2^{(j)}\right)$ de (Y_1, Y_2) où $\left(Y_i|\Theta = \Theta^{(j)}\right) \sim Exp\left(\Theta^{(j)}\right)$, pour $i = 1, 2$, i.e.

$$Y_1^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right)$$

$$Y_2^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_2^{(j)}\right)$$

où $V_1^{(j)}$ *et* $V_2^{(j)}$ *sont des réalisations indépendantes de* $V_1 \sim Unif\,(0,1)$ *et* $V_2 \sim Unif\,(0,1)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values* $U_1^{(j)}$ *et* $U_2^{(j)}$ *avec*

$$U_1^{(j)} = \overline{F}_{Y_1}\left(Y_1^{(j)}\right) = \left(1 + Y_1^{(j)}\right)^{-\frac{1}{\alpha}}$$

$$U_2^{(j)} = \overline{F}_{Y_2}\left(Y_2^{(j)}\right) = \left(1 + Y_2^{(j)}\right)^{-\frac{1}{\alpha}}$$

See Figure 4 and 5 for typical scatter plots of sampled values of $\underline{U} = (U_1, U_2) \sim C_\alpha^C$ and typical graphs of pdf of the copula.

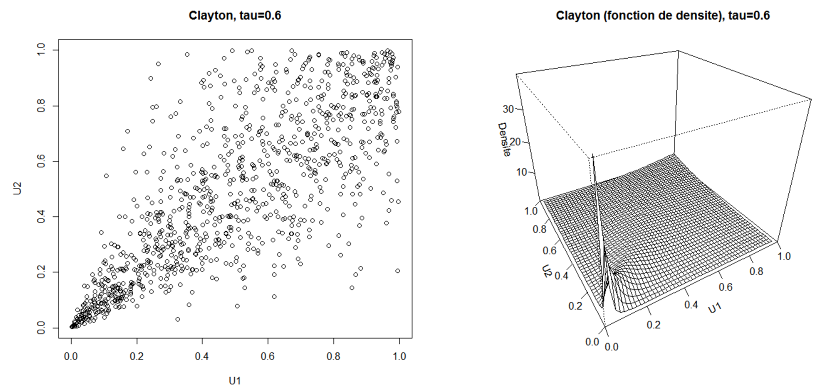

Figure 4: Clayton copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.22).

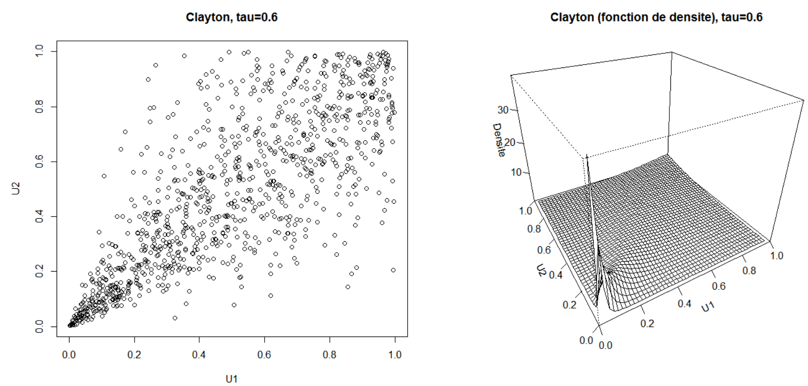

Figure 5: Clayton copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.6).

14.7.4 Simulation based on the conditional cdf

To apply the simulation method, we need the expression of $C_{2|1}^{-1}(v|u_1)$.

Recall

$$C_{2|1}(u_2|u_1) = \frac{1}{u_1^{\alpha+1}}\left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-1-\frac{1}{\alpha}}.$$

Let $v \in]0,1[$. We need to identify to the expression of u_2 such that

$$C_{2|1}(u_2|u_1) = \frac{1}{u_1^{\alpha+1}}\left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-1-\frac{1}{\alpha}} = v.$$

We have

$$\left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-\frac{(1+\alpha)}{\alpha}} = v \times u_1^{\alpha+1}$$

which becomes

$$\left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right) = \left(v \times u_1^{\alpha+1}\right)^{-\frac{\alpha}{\alpha+1}}.$$

We obtain

$$C_{2|1}^{-1}(v|u_1) = \left(\left(v \times u_1^{\alpha+1}\right)^{-\frac{\alpha}{\alpha+1}} - u_1^{-\alpha} + 1\right)^{-\frac{1}{\alpha}}.$$

Algorithm 22 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_\alpha^{Clay}$ ($\alpha > 0$).

1. *Simulate the sampled values $V_1^{(j)}$ and $V_2^{(j)}$ of the iid rvs V_1 and V_2 where $V_i \sim U(0,1)$ for $i = 1, 2$.*

2. Let $U_1^{(j)} = V_1^{(j)}$.
3. Calculate

$$U_2^{(j)} = C_{2|1}^{-1}\left(V_2^{(j)}|u_1\right) = \left(\left(V_2^{(j)} \times \left(U_1^{(j)}\right)^{\alpha+1}\right)^{-\frac{\alpha}{\alpha+1}} - \left(U_1^{(j)}\right)^{-\alpha} + 1\right)^{-\frac{1}{\alpha}}.$$

14.7.5 Multivariate copula

The expression of the multivariate copula is

$$C_\alpha(u_1, ..., u_n) = \mathcal{L}_\Theta\left(\mathcal{L}_\Theta^{-1}(u_1) + ... + \mathcal{L}_\Theta^{-1}(u_n)\right) = \left(u_1^{-\alpha} + ... + u_n^{-\alpha} - (n-1)\right)$$

for $u_i \in [0,1]$, $i = 1, ..., n$ and $\alpha > 0$.

Algorithm 23 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C_\alpha^{Clay}$ ($\alpha > 0$).

1. **(Simulate from the multivariate mixed exponential-gamma distribution).**

(a) Simulate the sampled value $\Theta^{(j)}$ of the rv $\Theta \sim Ga\left(\frac{1}{\alpha}, 1\right)$.

(b) *Simulate the sampled values* $\left(Y_1^{(j)}, ..., Y_n^{(j)}\right)$ *of* $(Y_1, ..., Y_n)$ *where* $\left(Y_i|\Theta = \Theta^{(j)}\right)$ $\sim Exp\left(\Theta^{(j)}\right)$, *i.e.*

$$Y_i^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right)$$

for $i = 1, ..., n$, *where* $V_1^{(j)}$ ... $V_n^{(j)}$ *are sampled values of the iid rvs* $V_1 \sim Unif\,(0,1)$ ... $V_n \sim Unif\,(0,1)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values* $U_1^{(j)}$ ... $U_n^{(j)}$ *with*

$$U_1^{(j)} = \overline{F}_{Y_1}\left(Y_1^{(j)}\right) = \mathcal{L}_\Theta\left(Y_1^{(j)}\right) = \left(1 + Y_1^{(j)}\right)^{-\frac{1}{\alpha}}$$

...

$$U_n^{(j)} = \overline{F}_{Y_n}\left(Y_n^{(j)}\right) = \mathcal{L}_\Theta\left(Y_n^{(j)}\right) = \left(1 + Y_n^{(j)}\right)^{-\frac{1}{\alpha}}$$

14.8 Bivariate Frank copula

14.8.1 Definition

The Clayton copula is an Archimedean copula.

Let Θ be a strictly positive discrete rv which follows a logarithmic distribution.

- Notation : $\Theta \sim Log(\gamma)$
- Parameter : $\gamma \in]0,1[$
- Support : $k \in \mathbb{N}^{+}$
- Pmf : $\Pr(\Theta = k) = \frac{-1}{\ln(1-\gamma)} \frac{\gamma^k}{k}$
- Expectation : $E[\Theta] = \frac{-1}{\ln(1-\gamma)} \frac{\gamma}{1-\gamma}$

- Variance : $\mathsf{Var}(\Theta) = \frac{\gamma+\ln(1-\gamma)}{(1-\gamma)^2(\ln(1-\gamma))^2}$
- LST : $\mathcal{L}_\Theta(t) = \frac{\ln(1-\gamma e^t)}{\ln(1-\gamma)}$
- Inverse of the LST: $\mathcal{L}_\Theta^{-1}(u) =$
- Pgf : $P_\Theta(t) = \frac{\ln(1-\gamma t)}{\ln(1-\gamma)}$

The expression for the copula is

$$C_\alpha(u_1, u_2) = -\frac{1}{\alpha}\ln\left(1 + \frac{\left(\mathrm{e}^{-\alpha u_1} - 1\right)\left(\mathrm{e}^{-\alpha u_2} - 1\right)}{\left(\mathrm{e}^{-\alpha} - 1\right)}\right), \tag{4}$$

for $u_i \in [0, 1]$, $i = 1, 2$.

Given the representation by common frailty, one may expect that the dependence parameter $\alpha \in \mathbb{R}^+$ (>0).

However, (??) is also valid for $\alpha \neq 0$.

Limit case:

- Frechet's lower bound copula: $\lim_{\alpha \to --\infty} C_\alpha(u_1, u_2) = C^-(u_1, u_2)$
- Independence copula: $\lim_{\alpha \to 0} C_\alpha(u_1, u_2) = C^\perp(u_1, u_2)$
- Frechet's upper bound copula: $\lim_{\alpha \to \infty} C_\alpha(u_1, u_2) = C^+(u_1, u_2)$.

The pdf is

$$c(u_1, u_2) = \frac{\alpha \mathrm{e}^{-\alpha(u_1+u_2)} \left(1 - \mathrm{e}^{-\alpha}\right)}{\left(\mathrm{e}^{-\alpha(u_1+u_2)} - \mathrm{e}^{-\alpha u_1} - \mathrm{e}^{-\alpha u_2} + \mathrm{e}^{-\alpha}\right)^2}.$$

The conditional cdf of U_2 given $U_1 = u_1$ is

$$C_{2|1}(u_2|u_1) = \frac{\mathrm{e}^{-\alpha u_1}\left(\mathrm{e}^{-\alpha u_2} - 1\right)}{\left(\mathrm{e}^{-\alpha} - 1\right) + \left(\mathrm{e}^{-\alpha u_1} - 1\right)\left(\mathrm{e}^{-\alpha u_2} - 1\right)}.$$

The Frank copula is complete.

14.8.2 Simulation procedures

We provide two simulation algorithms:

- The first algorithm works for $a > 0$. It is based on Sklar's Theorem. It is application of the general method for Archimedean copulas.
- The second algorithm works for $\alpha > 0$ and $\alpha < 0$. It is application of the method based on conditional cdf.

14.8.3 Simulation based on a general method for Archimedean copulas

Algorithm 24 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_\alpha^F$ ($\alpha > 0$).

1. **(Simulate from the bivariate mixed exponential-logarithmic distribution).**

   (a) Simulate the sampled value $\Theta^{(j)}$ of the rv $\Theta \sim Log(\gamma)$ with $\gamma = 1 - \mathrm{e}^{-\alpha}$.

   (b) Simulate the sampled values $\left(Y_1^{(j)}, Y_2^{(j)}\right)$ of (Y_1, Y_2) where $\left(Y_i | \Theta = \Theta^{(j)}\right) \sim Exp\left(\Theta^{(j)}\right)$, for $i = 1, 2$, i.e.

$$Y_1^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right)$$

$$Y_2^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_2^{(j)}\right)$$

where $V_1^{(j)}$ and $V_2^{(j)}$ are sampled values of the iid rvs $V_1 \sim Unif\,(0,1)$ and $V_2 \sim Unif\,(0,1)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values $U_1^{(j)}$ et $U_2^{(j)}$ with*

$$U_1^{(j)} = \overline{F}_{Y_1}\left(Y_1^{(j)}\right) = \mathcal{L}_{\Theta}\left(Y_1^{(j)}\right) = \frac{\ln(1-\gamma e^{-Y_1^{(j)}})}{\ln(1-\gamma)}$$

$$U_2^{(j)} = \overline{F}_{Y_2}\left(Y_2^{(j)}\right) = \mathcal{L}_{\Theta}\left(Y_2^{(j)}\right) = \frac{\ln(1-\gamma e^{-Y_2^{(j)}})}{\ln(1-\gamma)}$$

See Figure 6 and 7 for typical scatter plots of sampled values of $\underline{U} = (U_1, U_2) \sim C_\alpha^F$ and typical graphs of pdf of the copula.

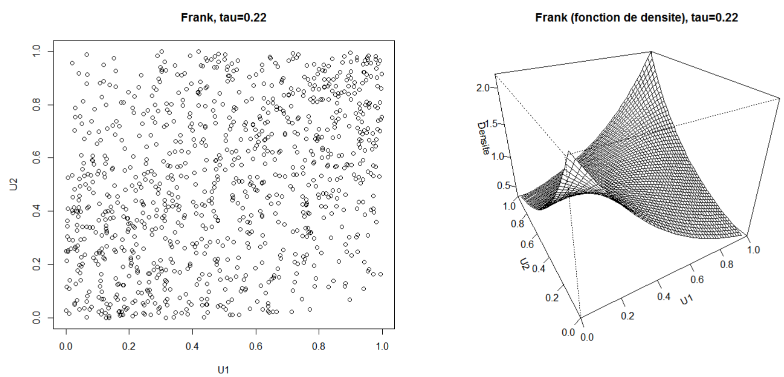

Figure 6: Frank copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau $= 0.22$).

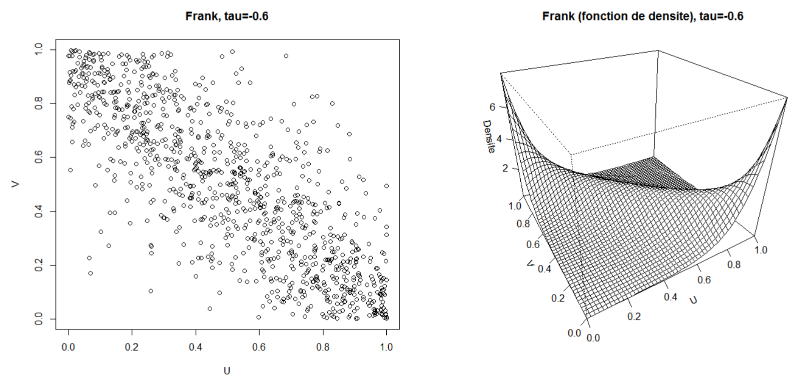

Figure 7: Frank copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.6).

14.8.4 Simulation based on the conditional cdf

To apply the simulation method, we need the expression of $C_{2|1}^{-1}(v|u_1)$.

Recall

$$C_{2|1}(u_2|u_1) = \frac{\mathrm{e}^{-\alpha u_1}\left(\mathrm{e}^{-\alpha u_2}-1\right)}{\left(\mathrm{e}^{-\alpha}-1\right)+\left(\mathrm{e}^{-\alpha u_1}-1\right)\left(\mathrm{e}^{-\alpha u_2}-1\right)}.$$

Let $v \in]0,1[$. We need to identify to the expression of u_2 such that

$$C_{2|1}(u_2|u_1) = \frac{\mathrm{e}^{-\alpha u_1}\left(\mathrm{e}^{-\alpha u_2}-1\right)}{\left(\mathrm{e}^{-\alpha}-1\right)+\left(\mathrm{e}^{-\alpha u_1}-1\right)\left(\mathrm{e}^{-\alpha u_2}-1\right)} = v.$$

We have

$$\mathrm{e}^{-\alpha u_1}\left(\mathrm{e}^{-\alpha u_2}-1\right)=v\times\left(\mathrm{e}^{-\alpha}-1\right)+v\times\left(\mathrm{e}^{-\alpha u_1}-1\right)\left(\mathrm{e}^{-\alpha u_2}-1\right)$$

which becomes

$$\mathrm{e}^{-\alpha u_1}\left(\mathrm{e}^{-\alpha u_2}-1\right)-v\times\left(\mathrm{e}^{-\alpha u_1}-1\right)\left(\mathrm{e}^{-\alpha u_2}-1\right)=v\times\left(\mathrm{e}^{-\alpha}-1\right)$$

and

$$\left(\mathrm{e}^{-\alpha u_2}-1\right)=\frac{v\times\left(\mathrm{e}^{-\alpha}-1\right)}{\mathrm{e}^{-\alpha u_1}(1-v)+v}.$$

We obtain

$$u_2=C_{2|1}^{-1}(v|u_1)=-\frac{1}{\alpha}\ln\left(1+\frac{v\times\left(\mathrm{e}^{-\alpha}-1\right)}{\mathrm{e}^{-\alpha u_1}(1-v)+v}\right).$$

Algorithm 25 Simulation des réalisations de (U_1, U_2).

1. *Simulate the sampled values* $V_1^{(j)}$ *and* $V_2^{(j)}$ *of the iid rvs* V_1 *and* V_2 *where* $V_i \sim U(0,1)$ *for* $i = 1, 2$.
2. *Let* $U_1^{(j)} = V_1^{(j)}$.
3. *Calculate*

$$U_2^{(j)} = -\frac{1}{\alpha} \ln \left(1 + \frac{V_2^{(j)} \times \left(\mathrm{e}^{-\alpha} - 1 \right)}{\mathrm{e}^{-\alpha U_1^{(j)}} \left(1 - V_2^{(j)} \right) + V_2^{(j)}} \right).$$

14.8.5 Multivariate copula

The expression of the multivariate copula is

$$C_\alpha(u_1, ..., u_n) = \mathcal{L}_\Theta\left(\mathcal{L}_\Theta^{-1}(u_1) + ... + \mathcal{L}_\Theta^{-1}(u_n)\right) = -\frac{1}{\alpha}\ln\left(1 + \frac{\prod_{i=1}^{n}\left(\mathrm{e}^{-\alpha u_i} - 1\right)}{\left(\mathrm{e}^{-\alpha} - 1\right)^{n-1}}\right)$$

for $u_i \in [0, 1]$, $i = 1, 2, ..., n$ and $\alpha > 0$.

Algorithm 26 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C_\alpha^F$ ($\alpha > 0$).

1. **(Simulate from the bivariate mixed exponential-logarithmic distribution).**

(a) *Simulate the sampled value* $\Theta^{(j)}$ *of the rv* $\Theta \sim Log(\gamma)$ *with* $\gamma = 1 - e^{-\alpha}$.

(b) *Simulate the sampled values* $\left(Y_1^{(j)}, ..., Y_n^{(j)}\right)$ *of* $(Y_1, ..., Y_n)$ *where* $\left(Y_i | \Theta = \Theta^{(j)}\right)$ $\sim Exp\left(\Theta^{(j)}\right)$, *i.e.*

$$Y_i^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right),$$

for $i = 1, ..., n$, *where* $V_1^{(j)}$ ... $V_n^{(j)}$ *are sampled values of the iid rvs* $V_1 \sim Unif\,(0,1)$ ... $V_n \sim Unif\,(0,1)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values* $U_1^{(j)}$ ... $U_n^{(j)}$ *with*

$$U_1^{(j)} = \overline{F}_{Y_1}\left(Y_1^{(j)}\right) = \mathcal{L}_\Theta\left(Y_1^{(j)}\right) = \frac{\ln(1-\gamma e^{-Y_1^{(j)}})}{\ln(1-\gamma)}$$

...

$$U_n^{(j)} = \overline{F}_{Y_n}\left(Y_n^{(j)}\right) = \mathcal{L}_\Theta\left(Y_n^{(j)}\right) = \frac{\ln(1-\gamma e^{-Y_n^{(j)}})}{\ln(1-\gamma)}$$

14.9 Bivariate Gumbel copula

14.9.1 Definition

The Gumbel copula is an Archimedean copula.

Let Θ be a strictly positive continous rv which follows a stable distribution.

As mentioned in [**?**], there is several parametrisation for the stable distribution.

Here, in the present context, the $Stable\,(\varphi, 1, \gamma, 0)$ distribution is defined over $\mathbb{R}^+$:

- $\varphi^{-1} > 1$ (or $0 < \varphi < 1$)

- $\gamma = \left(\cos\left(\frac{\pi\varphi}{2}\right)\right)^{\frac{1}{\varphi}}$
- LST: $\mathcal{L}_{\Theta}(t) = \mathrm{e}^{-t^{\varphi}}$
- Inverse of the LST: $\mathcal{L}_{\Theta}^{-1}(u) = (-\ln(u))^{\frac{1}{\varphi}}$

Here is the simulation algorithm for the stable distribution with this specific parametrization.

Algorithm 27 Simulation of a sampled value $W^{(j)}$ of a rv. $W \sim Stable(\varphi, 1, \gamma, 0)$ with $\gamma = \left(\cos\left(\frac{\pi\varphi}{2}\right)\right)^{\frac{1}{\varphi}}$

1. *Simulate a sample value of $R^{(j)}$ of the rv $R \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.*
2. *Simulate a sample value of $V^{(j)}$ of the rv $V \sim Exp(1)$.*

3. *Calculate*

$$A = \frac{\sin\left(\varphi\left(R^{(j)} + \frac{\pi}{2}\right)\right)}{\left(\cos\left(R^{(j)}\right)\right)^{\frac{1}{\varphi}}}$$

$$B = \left(\frac{\cos\left(\varphi\frac{\pi}{2} + (\varphi - 1)\, R^{(j)}\right)}{V^{(j)}}\right)^{\frac{1-\varphi}{\varphi}}.$$

4. *Calculate* $W^{(j)} = A \times B$.

We can represent $W = \gamma Y$ *where* $Y \sim Stable\,(\varphi, 1, 1, 0)$ *and* $\gamma = \left(\cos\left(\frac{\pi\varphi}{2}\right)\right)^{\frac{1}{\varphi}}$.

The algorithm provided above is an adaptation of an algorithm presented in [?] allowing to simulate sampled values of the rv Y.

Let $\alpha = \frac{1}{\varphi}$. The expression for the copula is

$$C_\alpha(u_1, u_2) = \mathcal{L}\left(\mathcal{L}^{-1}(u_1) + \mathcal{L}^{-1}(u_2)\right) = \mathrm{e}^{-\{(-\ln u_1)^\alpha + (-\ln u_2)^\alpha\}^{(1/\alpha)}},$$

for $u_i \in [0,1]$, $i = 1, 2$ and $\alpha \geq 1$.

Limit cases :

- Independence copula: $\lim_{\alpha \to 1} C_\alpha(u_1, u_2) = C^\perp(u_1, u_2)$
- Frechet's upper bound copula: $\lim_{\alpha \to \infty} C_\alpha(u_1, u_2) = C^+(u_1, u_2)$.

The copula is not complete.

The pdf is

$$
\begin{aligned}
& c\left(u_1, u_2\right) \\
= \; & C_\alpha(u_1, u_2) \times \frac{(-\ln u_1)^{\alpha-1}(-\ln u_2)^{\alpha-1}}{u_1 u_2} \\
& \times \left((-\ln u_1)^\alpha + (-\ln u_2)^\alpha\right)^{\frac{1}{\alpha}-2}\left(\alpha - 1 + \left((-\ln u_1)^\alpha + (-\ln u_2)^\alpha\right)^{\frac{1}{\alpha}}\right).
\end{aligned}
$$

The conditional cdf of U_2 given $U_1 = u_1$ is

$$
C_{2|1}\left(u_2|u_1\right) = C_\alpha(u_1, u_2)\frac{(-\ln u_1)^{\alpha-1}}{u_1}\left((-\ln u_1)^\alpha + (-\ln u_2)^\alpha\right)^{\frac{1}{\alpha}-1}.
$$

14.9.2 Simulation procedures

We provide two simulation algorithms:

- The first algorithm works for $a > 0$. It is based on Sklar's Theorem. It is application of the general method for Archimedean copulas.
- The second algorithm works for $\alpha > 0$ and $\alpha < 0$. It is application of the method based on conditional copula.

14.9.3 Simulation based on a general method for Archimedean copulas

Algorithm 28 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_\alpha^G$ ($\alpha > 0$).

1. **(Simulate from the bivariate mixed exponential-stable distribution, also called bivariate Weibull distribution)**

(a) *Simulate the sampled value* $\Theta^{(j)}$ *of the rv* $\Theta \sim Stable\left(\varphi, 1, \gamma, 0\right)$ *where* $\varphi = \frac{1}{\alpha}$ *and* $\gamma = \left(\cos\left(\frac{\pi}{2\alpha}\right)\right)^{\alpha}$ $(\alpha > 1)$.

(b) *Simulate the sampled values* $\left(W_1^{(j)}, W_2^{(j)}\right)$ *of* (W_1, W_2) *where* $\left(W_i | \Theta = \Theta^{(j)}\right) \sim Exp\left(\Theta^{(j)}\right)$, *for* $i = 1, 2$, *i.e.*

$$W_1^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right)$$

$$W_2^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_2^{(j)}\right)$$

where $V_1^{(j)}$ *and* $V_2^{(j)}$ *are sampled values of the iid rvs* $V_1 \sim Unif\left(0, 1\right)$ *and* $V_2 \sim Unif\left(0, 1\right)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values* $U_1^{(j)}$ *et* $U_2^{(j)}$

with

$$U_1^{(j)} = \overline{F}_{W_1}\left(W_1^{(j)}\right) = \mathcal{L}_{\Theta}\left(W_1^{(j)}\right) = \exp\left(-\left(W_1^{(j)}\right)^{\frac{1}{\alpha}}\right)$$

$$U_2^{(j)} = \overline{F}_{W_2}\left(W_2^{(j)}\right) = \mathcal{L}_{\Theta}\left(W_2^{(j)}\right) = \exp\left(-\left(W_2^{(j)}\right)^{\frac{1}{\alpha}}\right)$$

See Figures 8 and 9 for typical scatter plots of sampled values of $\underline{U} = (U_1, U_2) \sim C_{\alpha}^{C}$ and typical graphs of pdf of the copula.

The copula is also called «Gumbel-Hougard».

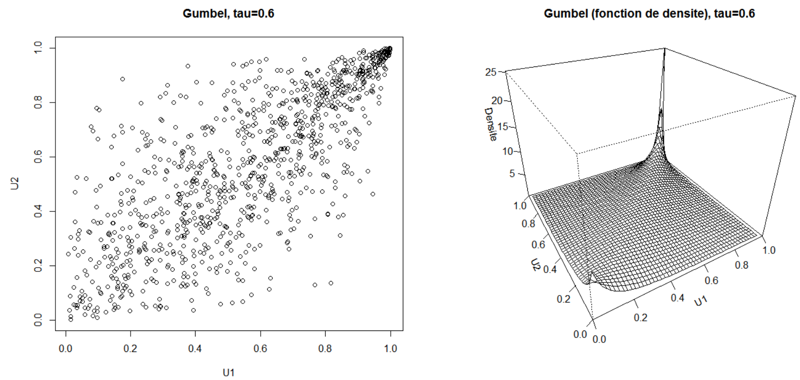

Figure 8: Gumbel copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.22).

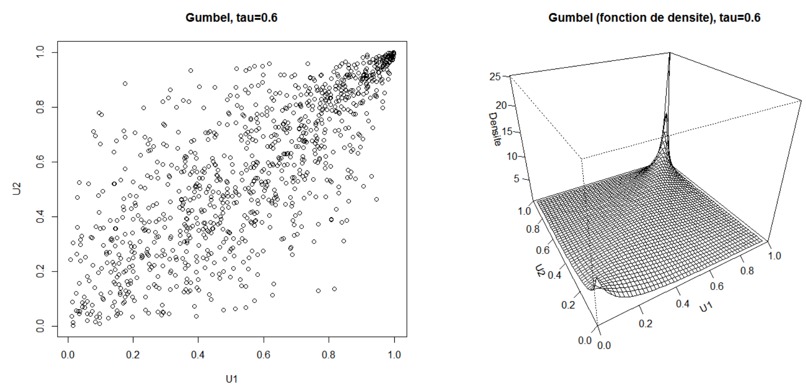

Figure 9: Gumbel copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.6).

14.9.4 Simulation based on the conditional cdf

To apply the simulation method, we need the expression of $C_{2|1}^{-1}(v|u_1)$.

Recall

$$C_{2|1}(u_2|u_1) = C_\alpha(u_1,u_2)\frac{(-\ln u_1)^{\alpha-1}}{u_1}((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{\frac{1}{\alpha}-1}.$$

Let $v \in \,]0,1[$. We need to identify to the expression of u_2 such that

$$C_{2|1}(u_2|u_1) = C_\alpha(u_1,u_2)\frac{(-\ln u_1)^{\alpha-1}}{u_1}((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{\frac{1}{\alpha}-1} = v.$$

We need to numerical optimization to compute values of $C_{2|1}^{-1}(v|u_1)$.

14.9.5 Multivariate copula

The expression of the multivariate copula is

$$C_\alpha(u_1, ..., u_n) = \mathcal{L}_\Theta\left(\mathcal{L}_\Theta^{-1}(u_1) + ... + \mathcal{L}_\Theta^{-1}(u_n)\right) = \exp\left(-\left\{\sum_{i=1}^{n}(-\ln u_i)^\alpha\right\}\right),$$

for $u_i \in [0, 1]$, $i = 1, 2, ..., n$ and $\alpha \geq 1$.

Algorithm 29 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, ..., U_n^{(j)}\right)$ of $\underline{U} \sim C_\alpha^G$ ($\alpha > 0$).

1. **(Simulate from the multivariate mixed exponential-stable distribution, also called multivariate Weibull distribution)**

(a) *Simulate the sampled value* $\Theta^{(j)}$ *of the rv* $\Theta \sim Stable\left(\varphi, 1, \gamma, 0\right)$ *where* $\varphi = \frac{1}{\alpha}$ *and* $\gamma = \left(\cos\left(\frac{\pi}{2\alpha}\right)\right)^{\alpha}$ *(*$\alpha > 1$*).*

(b) *Simulate the sampled values* $\left(W_1^{(j)}, W_2^{(j)}\right)$ *of* (W_1, W_2) *where* $\left(W_i | \Theta = \Theta^{(j)}\right) \sim Exp\left(\Theta^{(j)}\right)$, *for* $i = 1, ..., n$, *i.e.*

$$W_1^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_1^{(j)}\right)$$
$$\ldots$$
$$W_n^{(j)} = -\frac{1}{\Theta^{(j)}} \ln\left(1 - V_n^{(j)}\right)$$

where $V_1^{(j)}$, ..., $V_n^{(j)}$ *are sampled values of the iid rvs* $V_1 \sim Unif\left(0, 1\right)$, ..., $V_n \sim Unif\left(0, 1\right)$.

2. **(Part#1 of Sklar's Theorem).** *Simulate the sampled values* $U_1^{(j)}$ *et* $U_2^{(j)}$

with

$$U_1^{(j)} = \overline{F}_{W_1}\left(W_1^{(j)}\right) = \mathcal{L}_{\Theta}\left(W_1^{(j)}\right) = \exp\left(-\left(W_1^{(j)}\right)^{\frac{1}{\alpha}}\right)$$

$$U_n^{(j)} = \overline{F}_{W_n}\left(W_n^{(j)}\right) = \mathcal{L}_{\Theta}\left(W_n^{(j)}\right) = \exp\left(-\left(W_n^{(j)}\right)^{\frac{1}{\alpha}}\right)$$

15 Elliptic copulas

The family of elliptic copulas includes the normal copula and the student copula

Common characteristic: the scatterplots of the sampled values from copulas of this family have elliptic forms

For the moment, the detailed study of this family is not considered in the present version of these Lecture Notes.

15.1 Bivariate Normal copula

15.1.1 Definition

The normal copula is obtained by using Part#1 of Sklar's Theorem.

Let $\underline{Z} = (Z_1, Z_2)$ follows a bivariate standard normal distribution with

- $\mu_1 = \mu_2 = 0$;
- $\sigma_1 = \sigma_2 = 1$;
- $\rho_{12} = \rho_{21} = \rho \in [0, 1]$.

It means that

$$F_{Z_i}(x_i) = \Phi(x_i)$$

for $x_i \in \mathbb{R}$, $i = 1, 2$.

It means that

$$F_{Z_i}^{-1}(u_i) = \Phi^{-1}(x_i)$$

for $u_i \in (0, 1)$, $i = 1, 2$.

It means that

$$F_{Z_1,Z_2}(x_1, x_2) = \underline{\Phi}_\rho(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x2} f_{Z_1,Z_2}(x_1, x_2)\, dx_1 dx_2$$

for $x_i \in \mathbb{R}$, $i = 1, 2$.

Now, we apply Part #1 of Sklar's Theorem.

Indeed, the Normal copula is given by

$$C(u_1, u_2) = F_{Z_1,Z_2}\left(F_{Z_1}^{-1}(u_1), F_{Z_2}^{-1}(u_2)\right)$$

which is denoted by

$$C_\alpha^N(u_1, u_2) = \underline{\Phi}_{\underline{\alpha}}\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right),$$

for $u_i \in [0, 1]$, $i = 1, 2$, and with

$$\underline{\alpha} = \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix},$$

where $\alpha \in [-1, 1]$.

Now, $\rho = \alpha$ corresponds to the dependence parameter.

Notice that the expression of C_α^N is not closed-form.

Special cases are :

- $C_0(u_1, u_2) = C^{\perp}(u_1, u_2)$
- $C_1(u_1, u_2) = C^{+}(u_1, u_2)$
- $C_{-1}(u_1, u_2) = C^{-}(u_1, u_2)$.

The pdf $c(u_1, u_2)$ is

$$\frac{1}{\sqrt{1-\alpha^2}} \mathrm{e}^{-\frac{\left(\Phi^{-1}(u_1)^2 - 2\alpha\Phi^{-1}(u_1)\Phi^{-1}(u_2) + \Phi^{-1}(u_2)^2\right)}{2\left(1-\alpha^2\right)}} \mathrm{e}^{\frac{\left(\Phi^{-1}(u_1)^2 + \Phi^{-1}(u_2)^2\right)}{2}}.$$

The conditional cdf of U_2 given $U_1 = u_1$ is

$$C_{2|1}(u_2|u_1) = \Phi\left(\frac{\Phi^{-1}(u_2) - \alpha\Phi^{-1}(u_1)}{\sqrt{1-\alpha^2}}\right).$$

Lorsque $\alpha > 0$ ($\alpha < 0$), la copule introduit une relation de dépendance positive (négative) entre les composantes de $\underline{U}$.

The normal copula is complete.

15.1.2 Simulation procedure for the bivariate standard normal distribution

Let $\underline{Z} = (Z_1, Z_2)$ follows a bivariate standard normal distribution with

- $\mu_1 = \mu_2 = 0$;
- $\sigma_1 = \sigma_2 = 1$;
- $\rho_{12} = \rho_{21} = \rho \in [0, 1]$.

Let $\underline{Z} = (Z_1, Z_2)$ follows a bivariate standard normal distribution with

- $\mu_1 = \mu_2 = 0$;

- $\sigma_1 = \sigma_2 = 1$;
- $\rho_{12} = \rho_{21} = \rho \in [0, 1]$.

Let Y_1, Y_2 be two iid rvs with $Y_1 \sim Y_2 \sim Norm\,(0, 1)$.

We want to represent the components of (Z_1, Z_2) in terms of Y_1 and Y_2.

We propose the following representation:

- First, $Z_1 = Y_1$.
- Second,

$$Z_2 = \rho \times Y_1 + \sqrt{\left(1 - \rho^2\right)} \times Y_2.$$

We need to verify that (Z_1, Z_2) follow a bivariate normal distribution.

- Clearly, $Z_1 \sim Norm\,(0,1)$.
- Then, we have

$$
\begin{aligned}
M_{Z_2}(t) &= E\left[\mathrm{e}^{Z_2 t}\right] \\
&= E\left[\mathrm{e}^{\left(\rho\times Y_1+\sqrt{(1-\rho^2)}\times Y_2\right)t}\right] \\
&= E\left[\mathrm{e}^{\rho\times Y_1 t}\right]\times E\left[\mathrm{e}^{\left(\sqrt{(1-\rho^2)}\times Y_2\right)t}\right]
\end{aligned}
$$

- It becomes

$$
\begin{aligned}
M_{Z_2}(t) &= \mathrm{e}^{\frac{1}{2}\rho^2 t^2}\times \mathrm{e}^{\frac{1}{2}\left(\sqrt{1-\rho^2}\right)^2 t^2} \\
&= \mathrm{e}^{\frac{1}{2}\rho^2 t^2}\times \mathrm{e}^{\frac{1}{2}(1-\rho^2)t^2} \\
&= \mathrm{e}^{\frac{1}{2}t^2}
\end{aligned}
$$

- Also, we have

$$
\begin{aligned}
Cov(Z_1, Z_2) &= Cov\left(Y_1, \rho \times Y_1 + \sqrt{\left(1-\rho^2\right)} \times Y_2\right) \\
&= Cov(Y_1, \rho \times Y_1) + Cov\left(Y_1, \sqrt{\left(1-\rho^2\right)} \times Y_2\right) \\
&= \rho Cov(Y_1, Y_1) + \sqrt{\left(1-\rho^2\right)} \times Cov(Y_1, Y_2) \\
&= \rho.
\end{aligned}
$$

- Finally, we obtain

$$\begin{aligned} M_{Z_1,Z_2}(t_1,t_2) &= E\left[\mathrm{e}^{Y_1\times t_1}\mathrm{e}^{\left(\rho\times Y_1+\sqrt{(1-\rho^2)}\times Y_2\right)t_2}\right] \\ &= E\left[\mathrm{e}^{Y_1\times(t_1+\rho t_2)}\mathrm{e}^{\sqrt{(1-\rho^2)}\times Y_2 t_2}\right] \\ &= E\left[\mathrm{e}^{Y_1\times(t_1+\rho t_2)}\right]E\left[\mathrm{e}^{\sqrt{(1-\rho^2)}\times Y_2 t_2}\right] \\ &= \mathrm{e}^{\frac{1}{2}\left(t_1^2+2\rho t_1 t_2+\rho^2 t_2^2\right)}\mathrm{e}^{\frac{1}{2}\left(1-\rho^2\right)t_2^2} \\ &= \mathrm{e}^{\frac{1}{2}\left(t_1^2+2\rho t_1 t_2+t_2^2\right)}, \end{aligned}$$

which is the mgf of the bivariate Normal distribution.

This representation will help us to simulate sampled values of (Z_1, Z_2).

Algorithm 30 Simulate a sampled value $\underline{Z}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{Z} = (Z_1, Z_2)$.

1. *Simulate a sampled value $\underline{Y}^{(j)} = \left(Y_1^{(j)}, Y_2^{(j)}\right)$ of a pair of iid standard normal rvs $\underline{Y} = (Y_1, Y_2)$.*
2. *Let $Z_1^{(j)} = Y_1^{(j)}$.*
3. *Calculate $Z_2^{(j)} = \rho \times Y_1^{(j)} + \sqrt{\left(1 - \rho^2\right)} \times Y_2^{(j)}$.*

15.1.3 Simulation procedure for the bivariate copula

The simulation method is based on the application of Part#1 of Sklar's Theorem and the simulation procedure for the bivariate standard normal distribution.

Algorithm 31 Simulate a sampled value $\underline{U}^{(j)} = \left(U_1^{(j)}, U_2^{(j)}\right)$ of $\underline{U} \sim C_\alpha^N$.

1. *Simulate a sampled value $\underline{Z}^{(j)} = \left(Z_1^{(j)}, Z_2^{(j)}\right)$ of $\underline{Z}$.*
2. **Apply Part#1 of Sklar's Theorem.** *Calculate $U_i^{(j)} = \Phi\left(Z_i^{(j)}\right)$, $i = 1, 2$.*

See Figure 10 and 11 for typical scatter plots of sampled values of $\underline{U} = (U_1, U_2) \sim C_\alpha^N$.

The scatterplots take an elliptic form.

- If $\alpha > 0$, the elliptic form is around the diagonale line from the point (0,0) to the point (1,1) in $[0, 1] \times [0, 1]$.
- If $\alpha < 0$, the elliptic form is around the diagonale line from the point (0,1) to the point (1,0) in $[0, 1] \times [0, 1]$.

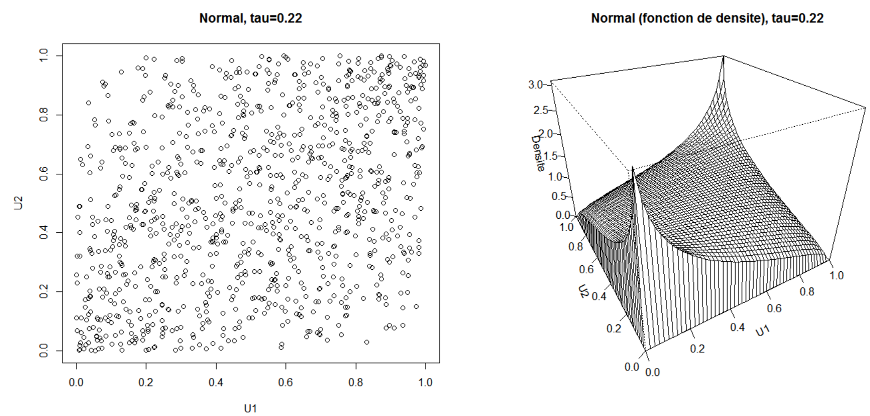

Figure 10: Normal copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = 0.22).

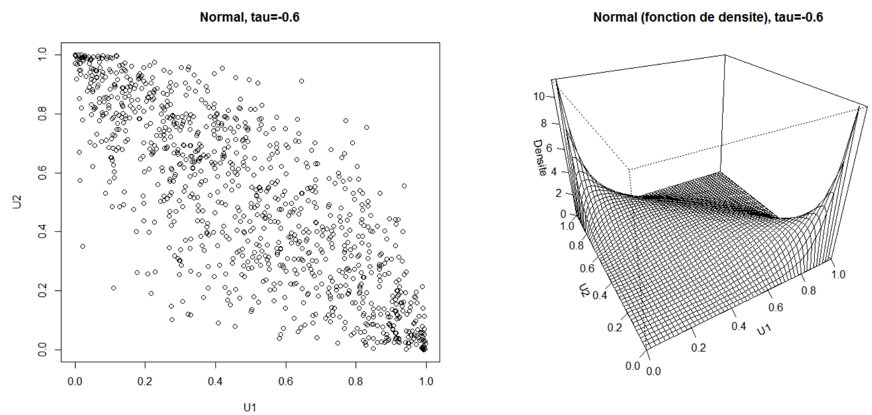

Figure 11: Normal copula: scatterplot of 1000 sampled values and 3-D plot of its pdf (Kendall's tau = -0.6)..

15.1.4 Multivariate copula

The multivariate copula is

$$C_{\underline{\alpha}}(u_1, ..., u_n) = \underline{\Phi}_{\underline{\alpha}}\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)\right),$$

for $u_i \in [0, 1]$, $i = 1, ..., n$ and with

$$\underline{\alpha} = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{12} & 1 & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & 1 \end{pmatrix},$$

where $\alpha_{ij} \in [0, 1]$, $i < j \in \{1, 2, ..., n\}$.

Let $\underline{Z} = (Z_1, ..., Z)$ be a vector of rvs which follows a multivariate standard normal distribution with a matrix of Pearson's correlation coefficients equal to $\underline{\alpha}$.

Algorithm 32 Simulation des réalisations de $(U_1, ..., U_n)$

1. *Simulate a sampled value* $\left(Z_1^{(j)}, ..., Z_n^{(j)}\right)$ *of* $\underline{Z}$.
2. *Calculate* $U_i^{(j)} = \Phi\left(Z_i^{(j)}\right)$, $i = 1, ..., n$.

15.2 Bivariate Student copula

(Later)

16 Challenges for research in actuarial science

- Advanced statistical methods for copulas.
- Development of new copulas.
- Construction methods for copulas.
- Hierarchical Archimedean copulas.
- Applications and innovations in non-life insurance.
- Using copulas to define joint distributions for vectors of counting rvs.
- Using copulas to define joint distributions for vectors of continuous rvs.
- Using copulas to define joint distributions for vectors of mixed rvs.
- Aggregation of dependent risks with joint distribution defined with copulas.
- Etc.