Multivariate probability Distributions: Copulas

Alexander Ritz

University of Göttingen

1. Dezember 2019

Overview

Section 1

Idea

Preliminaries

In general, a multivariate probability distribution is described by its cumulative distribution function (cdf). For simplicity's sake, we will mostly restrict ourselves to real valued and continuous RVs in this introdution. Let $\mathbf{X} = (X_1, ..., X_n)$ be an n-dimensional random variable such that:

$$\mathbf{X} : \Omega \to \mathbb{R}^n : \mathbf{X}(\omega) = (X_1(\omega), ..., X_n(\omega))'$$

with probability space

$$(\Omega, \mathcal{F}, P)$$

or specifically

$$(\mathbb{R}^n, \mathcal{B}^n, P_X)$$

Preliminaries

Then $\mathbf{X}$ can be described by the joint cdf of $X_1, ..., X_n$:

$$F_{\mathbf{X}}(x_1, ..., x_n) = P_{\mathbf{X}}(X_1 \leq x_1, ..., X_n \leq x_n)$$

with marginal cdf's

$$F_{X_1}(x_1) = P_{X_1}(X_1 \leq x_1), ..., F_{X_n}(x_n) = P_{X_n}(X_n \leq x_n)$$

Preliminaries

Recall the properties of the following transformations:

1. If $U \sim \mathcal{U}(0,1)$ and F is a cdf, then: $\quad X = F^{-1}(U) \sim F$
2. If $X \sim F$ and F is a continuous cdf, then: $\quad U = F(X) \sim \mathcal{U}(0,1)$

Remark: $U = F(X)$ is often called the grade of X.

Motivation

Try to decompose the multivariate distribution into:

- n univariate margins $F_{X_1}, ..., F_{X_n}$
- a Copula C

Section 2

Definition

What is a Copula?

What is a Copula?

What is a Copula?

What is a Copula?

A copula is the joint distribution of the grades of our variables of interest. We "erase" the information on the marginals by constructing the grades and only information on the structure of the dependence remains.

Definition

A copula is a function

$$C : [0,1]^n \to [0,1] :$$

$$(u_1, ..., u_n) \mapsto C(u_1, ..., u_n)$$

such that:

- $\forall u \in [0,1]^n : \quad \exists u_j = 0 : \quad C(u) = 0, \quad j = 1, ..., n$
- $\forall u_j : \qquad C(1, ..., u_j, ..., 1) = u_j \quad (u_1, ..., u_{j-1}, u_{j+1}, ...u_n = 1)$
- C is n-increasing

Bivariate Case

In the bivariate case, the properties are easier to parse:

Definition

A bivariate copula is a function

$$C : [0,1]^2 \to [0,1] :$$

$$(u_1, u_2) \mapsto C(u_1, u_2)$$

such that:

- $\forall u \in [0,1]^2 : \quad C(u_1, 0) = C(0, u_2) = 0$
- $C(u_1, 1) = u_1$ and $C(1, u_2) = u_2$
- C is 2-increasing, meaning: $C(u_1, u_2) - C(u_1, v_2) - C(v_1, u_2) + C(v_1, v_2) \geq 0$, for $v_1 \leq u_1$ and $v_2 \leq u_2$

Section 3

Sklar's Theorem

What exactly does the relationship between joint cdf, marginals and copula look like?

Sklar's Theorem (bivariate)

Let $F_{\mathbf{X}}$ be a bivariate joint cdf for the RV $\mathbf{X} = (X_1, X_2)$ with marginal cdfs F_{X_1} and F_{X_2}. Then there exists a copula C such that:

$$F_{\mathbf{X}}(x_1, x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)) = C(u_1, u_2),$$

with $x_1, x_2 \in \bar{\mathbb{R}}$ and $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$

Relating the joint density

Assuming $F_{\mathbf{X}}$ and C are differentiable, calculating the joint density in combination with Sklar's Theorem gives:

$$\frac{\partial^2 F_{\mathbf{X}}(x_1, x_2)}{\partial x_1 \partial x_2} = f_{\mathbf{X}}(x_1, x_2) = c(F_{X_1}(x_1), F_{X_2}(x_2)) f_{X_1}(x_1) f_{X_2}(x_2)$$

with

$$c(F_{X_1}(x_1), F_{X_2}(x_2)) = \frac{\partial^2 C(F_{X_1}(x_1), F_{X_2}(x_2))}{\partial F_{X_1}(x_1) \partial F_{X_2}(x_2)},$$

the copula pdf.

Interpretation

The copula density can therefore be interpreted as a multiplicative adjustment of the joint density under independence. The concept easily transfers to higher dimensions in an analogous manner.

Section 4

Archimedean Copulas

Classification of Copulas

Different kinds of copulas can be distinguished, e.g.

- Elliptical Copulas, such as the Gaussian
- Archimedean Copulas
- ...

Archimedean Copulas

A copula C is called an Archimedean copula, if it can be written in terms of a continuous generator function $\phi : [0, \infty] \to [0, 1]$:

$$C(u) = \phi(\phi^{-1}(u_1) + ... + \phi^{-1}(u_n)),$$

with $u \in [0, 1]^n$.
The generator function ϕ holds the properties:

- $\phi(0) = 1$
- $\phi(\infty) = 0$
- $(-1)^k \phi^{(k)}(t) \geq 0$, for $k \in \{0, 1, ..., n-2\}, t \in (0, \infty)$

Exemplary families of Archimedean Copulas

Families of Archimedean copulas include:

- Clayton, with $\phi(t) = (1+t)^{-\frac{1}{\theta}}$, with $\theta \in (0, \infty)$
- Frank, with $\phi(t) = -\log\left(1 - (1 - \exp(-\theta))\exp(-t)\right)/\theta$, with $\theta \in (0, \infty)$
- Gumbel-Hougaard, with $\phi(t) = \exp\left(-t^{\frac{1}{\theta}}\right)$, with $\theta \in [1, \infty)$
- Ali-Mikhail-Haq, with $\phi(t) = (1-\theta)/\exp(t) - \theta$, with $\theta \in [0, 1)$
- ...

Section 5

Appendix

A.1 Proof of transformation property

Let $X = F_X^{-1}(U)$ with $U \sim \mathcal{U}(0,1)$

$$\begin{aligned} P(X \leq x) &= P(F_X^{-1}(U) \leq x) && (1) \\ &= P(U \leq F_X(x)) && (2) \\ &= F_X(x) && (3) \end{aligned}$$

Which is the cdf of a uniform distribution.

A.2 Proof of grade property

Let $X \sim F$ and F be a continuous cdf, then:

$$\begin{align} F_U(u) &= P(U \leq u) \tag{4} \\ &= P(F_X(X) \leq u) \tag{5} \\ &= P(X \leq F_X^{-1}(u)) \tag{6} \\ &= F_X(F_X^{-1}(u)) = u \tag{7} \end{align}$$

Which is the cdf of a uniform distribution.

A.3 n-increasing function

A function $C : [0,1]^n \to [0,1]$ is n-increasing if for any hyper-rectangle

$$R = \prod_{i=1}^{n} [x_i, y_i] \subseteq [0,1]^n$$

the C-measure / C-volume is non-negative:

$$V_C(R) \geq 0$$

A.4 Fréchet-Hoeffding bounds

For a bivariate copula, the following bounds are called the Fréchet-Hoeffding bounds:

$$max(u_1 + u_2 - 1, 0) \leq C(u_1, u_2) \leq min(u_1, u_2) \tag{8}$$

The upper bound is always (not just the bivariate case) a copula itself. The lower boundary only in this case of two variables.

A.5 Origin of the term Archimedean

The Archimedean axiom for the positive real numbers: If a, b are positive real numbers, then there exists an integer n, such that $na > b$.
Since an Archimedean copula behaves like a binary operation on the unit interval (assigns each pair $(u_1, u_2) \in [0,1]^2$ a value $C(u)$), one can prove that a version of the axiom holds for the Abelian semi-group $(\mathbf{I}, C)$. The enjoyability of the proof is rather limited, therefore we will ignore the details here.

Properties of Archimedean Copulas

The class of Archimedean copulas is an associative class of copulas, i.e.:

$$C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3))$$

Properties of Archimedean Copulas

The class of Archimedean copulas is a permutation-symmetric class of copulas, i.e.:

$$C(u_1, u_2, u_3) = C(u_3, u_2, u_1)$$

Properties of Archimedean Copulas

Generators can be easily extended, since ϕ being a generator is equivalent to $a\phi$ being a generator, for $a > 0$.