Copula Construction Methods

Daniel Morales Fernando Corrêa

IME-USP

July 7, 2017

Outline

Preamble (Why copulas?)

- Copula is a d.f. in the unit square I^2 with uniform margins which, alongside two univariate d.f.'s F and G, defines a unique joint d.f. H (Sklar's Theorem).

$$H(x, y) = C(F(x), G(y))$$

- Why (and how) should we use Copula theory to analyse dependence structures?
- One possible dependence analysis: $Y = f(X) + \epsilon$ (regression: conditional inference).
- Copula dependence analysis: might be local, invariant over monotone scale transformations, might not use marginal.
- Objetive of this talk: How do we find Copulas that suits our needs?

Geometric Methods

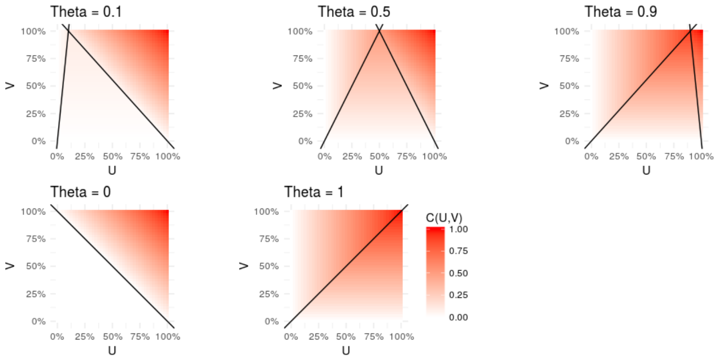
Geometric methods are ways of finding copulas satisfying some geometric properties. Some ideas are:

- prescribed support;
- prescribed vertical or horizontal sections;
- prescribed diagonal sections.

Geometric Method - Prescribed support

- Let θ be a number between 0 and 1. Consider a random point (X, Y) unif. choosen from the segment joining $(0, 0)$ to $(\theta, 1)$ with mass θ and unif. choosen from the segment joining $(\theta, 1)$ to $(0, 1)$ with mass $1 - \theta$.
- Which Copula arises from this example?

Geometric Method - Prescribed support

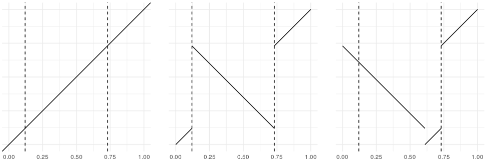


$$C_\theta(u,v) = \begin{cases} u,\ 0 \leq u \leq \theta v \leq \theta, \\ \theta v,\ 0 \leq \theta v \leq u \leq 1-(1-\theta)v, \\ u+v-1,\ \theta \leq 1-(1-\theta)v \leq u \leq 1 \end{cases}$$

Geometric Method - Prescribed support generalization - Shuffles of M

A similar idea may be generalized using this scheme (Mikusinski et al. 1992):

> *The mass distribution for a shuffle of $C_1(u, v)$ (also known as M or Frechét-Hoeffding upper bound) can be obtained by*
>
> 1. *Placing the mass for M on I^2;*
> 2. *Cutting I^2 vertically into a finite number of strips;*
> 3. *Shuffling the strips with perhaps some of them flipped around their vertical axes of symmetry;*
> 4. *Reassembling them to form the square again. [...].*



Geometric Method - Prescribed support generalization - Shuffles

Properties:

1. Shuffles of M are dense in the set of Copulas with the sup norm;
2. Shuffles of M provide a way of approximating copulas with fixed; prescribed points $C(a, b) = \theta$
3. Provides guidance on shuffling other copulas.

Geometric Method - Copulas with polynomial sections

- How many Copulas are, say, linear on u?
- One such Copula must satisfy

$$\begin{aligned} C(u,v) &= a(v)u + b(v) \\ C(0,v) &= b(v) = 0 \\ C(1,v) &= a(v) \propto v \\ C(1,1) = a(1) = 1 \iff a(v) &= v \end{aligned}$$

- So, the only Copula with linear section is $C(u,v) = uv$, the independence Copula.

Geometric Method - Copulas with polynomial sections

- How many Copulas are, say, quadratic on u?
- One such Copula must satisfy

$$C(u,v) = uv + \psi(v)u(1-u)$$
$$\psi(0) = 0,\ \psi(1) = 1$$

Example (Farlie-Gumbel-Morgenstern family)

If we ask for the C defined above to be quadratic on both u and v, then $\psi(v) = \theta v(1-v)$ for some $\theta \in [0,1]$ and then

$$C_\theta(u,v) = uv + \theta uv(1-u)(1-v)$$

Algebraic Methods

Algebraic methods are ways of finding copulas satisfying specific functional relations. Those main ideas stands above others:

1. Fixed odds ratio;
2. Fixed "dependence" for survival odds ratio.

Algebraic Methods - Plackett's Copula

- In categorical data analysis, the odds ratio is a measure of association between two discrete random variables.
- In continuous distributions, the generalization it not simple: the odd's ratio probably will depend on x and y. Is there a Copula such that the odd's ratio is fixed, in say, a number $\theta \in [0, 1]$, for every x and y?

$$\theta = \frac{\mathbb{P}(X \leq x, Y \leq y)\mathbb{P}(X > x, Y > y)}{\mathbb{P}(X \leq x, Y > y)\mathbb{P}(X > x, Y \leq y)} =$$

$$\frac{H(x, y)(1 - F(x) - G(x) + H(x, y))}{(F(x) - H(x, y))(G(y) - H(x, y))}$$

Algebraic Methods - Planckett's Copula

By Sklar's theorem, θ must have a unique representation in terms of the copula C:

$$\theta = \frac{C(u,v)(1-u-v+C(u,v))}{(u-C(u,v))(v-C(u,v))}$$

Solving for $C(u,v)$ and using copula properties, one finds that the only copula that satisfies the odds ratio property is

$$C(u,v) = \frac{1+(\theta-1)(u+v)-\sqrt{(1+(\theta-1)(u+v))^2-4uv\theta(\theta-1)}}{2(\theta-1)}.$$

Algebraic Methods - Ali-Mikhail-Haq distributions

In survival analysis is common to talk about the hazard ratio (or "odds of survival") defined as $\mathbb{P}[X > x]/\mathbb{P}[X \leq x]$. Analogaly, one might define a "bivariate odds of survival" as

$$\frac{\mathbb{P}[X > x, Y > y]}{\mathbb{P}[X \leq x, Y \leq y]} = \frac{1 - H(x,y)}{H(x,y)}.$$

Under independence, this quantity becomes:

$$\frac{\mathbb{P}[X > x, Y > y]}{\mathbb{P}[X \leq x, Y \leq y]} = \frac{1 - H(x,y)}{H(x,y)} =$$
$$\frac{1 - F(x)}{F(x)} + \frac{1 - G(y)}{G(y)} + \frac{1 - F(x)}{F(x)}\frac{1 - G(y)}{G(y)}.$$

Algebraic Methods - Ali-Mikhail-Haq distributions

To explore some forms of dependence, Ali-Mikhail-Haq suggests analysing copulas with the following "bivariate odds of survival"

$$\frac{1 - F(x)}{F(x)} + \frac{1 - G(y)}{G(y)} + (1 - \theta)\frac{1 - F(x)}{F(x)}\frac{1 - G(y)}{G(y)}.$$

It turns out that the copulas that satisfies this relation are given by

$$C_\theta(u, v) = \frac{uv}{1 - \theta(1 - u)(1 - v)}, \;\; \theta \in [-1, 1].$$

The Inversion Method

Having a bivariate distribution function H with continuous margins F and G, we derive the copula

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)).$$

Equally, for survival functions we have

$$\hat{C}(u, v) = \bar{H}(\bar{F}^{-1}(u), \bar{G}^{-1}(v)).$$

Example (Gaussian Copula)

Let $\Phi_\Sigma(\cdot)$ denote the distribution function of a p-variate normal random vector with zero means and variance-covariance matrix Σ and let $\Phi^{-1}(\cdot)$ denote the inverse of the distribution function of a standard normal random variable. The Gaussian copula with variance-covariance matrix Σ is defined by

$$C(u_1, \cdots, u_n) = \Phi_\Sigma(\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_n))$$

The Inversion Method

Example (Marshall-Olkin Bivariate Exponential Distribution)

$$Z_1 \sim \mathsf{Exp}(\lambda_1), Z_2 \sim \mathsf{Exp}(\lambda_2), Z_{12} \sim \mathsf{Exp}(\lambda_{12}) \text{ i.i.d.}$$

$$X = \min(Z_1, Z_{12}),\ Y = \min(Z_2, Z_{12})$$

$$\bar{F}(x) = \exp(-(\lambda_1 + \lambda_{12})x),\ \bar{G}(y) = \exp(-(\lambda_2 + \lambda_{12})y)$$

$$\begin{aligned} \bar{H}(x, y) &= P[Z_1 > x]P[Z_2 > y]P[Z_{12} > \max(x, y)] \\ &= \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)] \end{aligned}$$

$$\hat{C}(u, v) = uv \min(u^{-\alpha}, v^{-\beta}) = \min(u^{1-\alpha} v, uv^{1-\beta})$$

Cohen's Method

- Choose an arbitrary real integrable function f on the unit square. Then, define
$$V = \int_0^1 \int_0^1 f(u,v)\,du\,dv,$$
$$f_1(u) = \int_0^1 f(u,v)\,dv \text{ and } f_2(v) = \int_0^1 f(u,v)\,du.$$
- It follows that $f^1(u,v) = f(u,v) - f_1(u) - f_2(v) + V$ is such that
$$\int_0^1 \int_0^1 f^1(u,v)\,du\,dv = \int_0^1 f^1(u,v)\,dv = \int_0^1 f^1(u,v)\,du = 0.$$
- Finally, $C(u,v) = 1 + f^1(u,v)$ is a copula

Copulas Defined from a Distortion Function

- Pick a distortion function: an increasing function $\phi : [0,1] \to [0,1]$ with $\phi(0) = 0$, $\phi(1) = 1$.
- Starting with $H(x,y) = C(F(x), G(y))$, we define another distribution function

$$H^*(x,y) = \phi[H(x,y)],$$

with marginals $F^*(x) = \phi[F(x)]$ and $G^*(y) = \phi[G(y)]$.

- Then we create the associated copula via inversion method

$$C^*(u,v) = \phi[C(\phi^{-1}(u), \phi^{-1}(v))],$$

which does not depend on H, F or G.

Copulas Defined from a Distortion Function

Example (Frank's Copula)

Let

$$\phi(t) = \frac{1 - e^{\alpha t}}{1 - e^{-\alpha}},\ \alpha > 0$$

and $C(u,v) = \Pi(u,v) = uv$. Then the distorced copula is

$$C^*(u,v) = \log_\alpha \left(1 + \frac{(\alpha^u - 1)(\alpha^v - 1)}{\alpha - 1}\right)$$

Copulas with specified properties

- Any convex linear combination of copulas is a copula, i. e., $\sum_{i=1}^{n} \alpha_i C_i$ is a copula for $\alpha_i > 0$, $\sum_{i=1}^{n} \alpha_i = 1$, $\alpha_i \in \mathbb{R}$, C_i copula and $i = 1, \cdots, n$.

Definition (Comprehensive Copulas)

A family of copulas that includes $W(u,v) = \max(u+v-1, 0)$, $\Pi(u,v) = uv$ and $M = \min(u,v)$ is said to be comprehensive.

Example (Fréchet Copula)

The two-parameter comprehensive copula given below is due to Fréchet (1958):

$$C_{\alpha,\beta} = \alpha M(u,v) + \beta W(u,v) + (1-\alpha-\beta)\Pi(u,v)$$

Copulas with specified properties

Example (Mardia - 1970)

The Fréchet copula with $\alpha = \dfrac{\theta^2(1+\theta)}{2}$ and $\beta = \dfrac{\theta^2(1-\theta)}{2}$ gives us

$$C_\theta(u,v) = \frac{\theta^2(1+\theta)}{2} M(u,v) + \frac{\theta^2(1-\theta)}{2} W(u,v) + (1-\theta^2)\Pi(u,v)$$

Theorem

Let X and Y be continuous random variables with copula C_{XY}. C_{XY} is invariant under strictly increasing transformations of X and Y.

Copulas with specified properties

Definition (Homogeneous Copula)

A copula C is homogeneous of degree k if for some real number k and all u, v, λ in $[0,1]$,

$$C(\lambda u, \lambda v) = \lambda^k C(u,v)$$

Example (Cuadras-Augé Family)

For $\theta \in [0,1]$, we show a family of homogeneous copulas

$$\begin{aligned} C_\theta(\lambda u, \lambda v) &= [M(\lambda u, \lambda v)]^\theta [\Pi(\lambda u, \lambda v)]^{1-\theta} \\ &= \lambda^\theta [M(u,v)]^\theta (\lambda^2)^{1-\theta} [\Pi(u,v)]^{1-\theta} \\ &= \lambda^{2-\theta} C_\theta(u,v) \end{aligned}$$

Copulas with specified properties

Theorem

Suppose C is homogeneous of degree k. Then

- $1 \leq k \leq 2$
- C is a member of the Cuadras-Augé family with $\theta = 2 - k$

Definition (Harmonic Copulas)

Let C be a copula with continuous second-order partial derivatives on $[0,1]^2$. Then C is harmonic if satisfies Laplace's Equation in $[0,1]^2$:

$$\nabla^2 C(u,v) = \frac{\partial^2}{\partial u^2} C(u,v) + \frac{\partial^2}{\partial v^2} C(u,v) = 0$$

Theorem

$\Pi(u,v)$ is the only harmonic copula.

Conclusion

- Geometric methods may be used to approximate any copula
- Algebraic methods translates some well known dependence measures to the copula world
- Some undesirable properties of a Copula C might be transformed using a suitable distortion function
- Some methods, as the looking for harmonic copulas, do not result in interesting ways of constructing copulas
- Some Copulas might be represented and integrals of function on I^2

References

Roger B. Nelsen
An Introduction to Copulas
Springer, 2006

N. Balakrishnan, Chin-Diew Lai
Continuous Bivariate Distributions
Springer, 2009

S. Nadarajah, E. Afuecheta, S. Chan
A Compendium of Copulas
University of Manchester, 2016