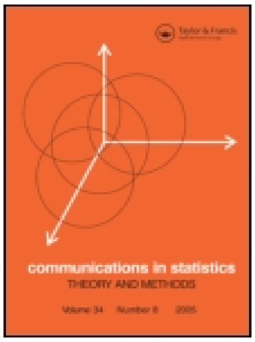

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

A recipe for bivariate copulas

Eric Key

To cite this article: Eric Key (2016) A recipe for bivariate copulas, Communications in Statistics - Theory and Methods, 45:21, 6416-6420, DOI: 10.1080/03610926.2014.915040

To link to this article: https://doi.org/10.1080/03610926.2014.915040

Published online: 11 Aug 2016.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2016, VOL. 45, NO. 21, 6416–6420
http://dx.doi.org/10.1080/03610926.2014.915040

Taylor & Francis
Taylor & Francis Group

A recipe for bivariate copulas

Eric Key

Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI, USA

ABSTRACT
We give conditions on $a \geq -1$, $b \in (-\infty, \infty)$, and f and g so that $C_{a,b}(x, y) = xy(1 + af(x)g(y))^b$ is a bivariate copula. Many well-known copulas are of this form, including the Ali–Mikhail–Haq Family, Huang–Kotz Family, Bairamov–Kotz Family, and Bekrizadeh–Parham–Zadkarmi Family. One result is that we produce an algorithm for producing such copulas. Another is a one-parameter family of copulas whose measures of concordance range from 0 to 1.

ARTICLE HISTORY
Received 18 July 2013
Accepted 3 April 2014

KEYWORDS
Copulas

MATHEMATICS SUBJECT CLASSIFICATION
62H20; 62F40; 62E10

1. Introduction

Briefly, a bivariate copula (copula for short) is the joint distribution function of a pair of random variables U and V that are each uniformly distributed on $[0, 1]$. Precisely, $C : [0, 1]^2 \to [0, 1]$ is a copula if and only if

- $C(1, u) = C(u, 1) = u$;
- $C(0, u) = C(u, 0) = 0$;
- $C(x, y) - C(x, b) - C(a, y) + C(a, b) \geq 0$ for $[a, x] \times [b, y] \in [0, 1]^2$.

If we have $C : [0, 1]^2 \to [0, 1]$ such that $C(1, u) = C(u, 1) = u$, $C(0, u) = C(u, 0) = 0$, and we can find $c : [0, 1]^2 \to [0, \infty)$ such that

$$C(x, y) = \int\int_{[0,x]\times[0,y]} c(u, v)\, du\, dv,$$

then C is a copula, and c is called a *copula density* for C.

Copulas are used to model multivariate data when one wants to separate modeling the dependence of the components of each sample point from modeling the marginal distributions of the components. For example, if we have a random sample (X_k, Y_k), $k = 1, 2, \ldots N$, from (X, Y) with distribution function $F(x, y)$ and marginal distribution functions F_X and F_Y, and we assume that F_X and F_Y are continuous, then we know that $F_X(X)$ and $F_Y(Y)$ are uniformly distributed on $(0, 1)$. Therefore,

$$\Pr(F_X(X) \leq x, F_Y(Y) \leq y) =: C(x, y)$$

defines a distribution function C supported on $[0, 1]^2$ whose marginals are uniform on $[0, 1]$. C is called a (bivariate) copula, and we can express F as

$$F(x, y) = C(F_X(x), F_Y(y)).$$

CONTACT Eric Key ericskey@uwm.edu Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA.

Presuming that we are satisfied with estimates of F_X and F_Y obtained from the data, we are left with the problem of choosing an appropriate copula. One way to do this is to choose a measure of concordance, such as Kendall's tau, compute its sample value, and then choose a copula with the same population value of that measure of dependence.

Some examples of copulas are

Farlie–Gumbel–Morgenstern Family: $C_{a,b}(x, y) = xy(1 + a(1 - x)(1 - y))$,
Ali–Mikhail–Haq Family: $C_{a,b}(x, y) = xy(1 + a(1 - x)(1 - y))^{-1}$,
Huang–Kotz Family: $C_{a,b,p}(x, y) = xy((1 + a(1 - x)^p(1 - y)^p)$,
Bekrizadeh–Parham–Zadkarmi Family: $C_{a,b,p}(x, y) = xy(1 + a(1 - x^p)(1 - y^p))^b$,

each for appropriate choices of a, b, and p.

Each of these copulas is of the form

$$F_{U,V}(x, y) = xy(1 + af(x)g(y))^b,$$

where f and g are continuous and strictly decreasing on $[0, 1]$ with $f(0) = g(0) = 1$ and $f(1) = g(1) = 0$.

The literature on copulas and copula methods is extensive. For a short introduction and overview of copulas, see Nelsen (2003).

2. One theorem, many well-known copulas!

Consider the class, $\mathcal{C}$, of decreasing continuous one-to-one functions f mapping $[0, 1]$ onto $[0, 1]$ with the additional properties

- f is continuously differentiable on $(0, 1)$,
- $xf'(x) \to 0$ as $x \searrow 0$,
- $xf'(x)$ converges as $x \nearrow 1$.

For example, if $p \geq 1$ and $q > 0$, then $f(u) = (1 - u^q)^p$ is in $\mathcal{C}$.

Theorem 1. *Suppose* $a \in [-1, \infty)$, $b \in (-\infty, \infty)$, $F > 0$, $G > 0$, $abFG \geq -1$, $abF \leq 1$, *and* $abG \leq 1$. *Suppose* $\{f, g\} \subset \mathcal{C}$ *and*

$$f(x) - \frac{xf'(x)}{F} \in [0, 1], \tag{1}$$

$$g(y) - \frac{yg'(y)}{G} \in [0, 1]. \tag{2}$$

For $(x, y) \in (0, 1)^2$ *define*

$$C(x, y) = xy(1 + af(x)g(y))^b, \tag{3}$$

and extend C *to* $[0, 1]^2$ *by continuity. Then* $C : [0, 1]^2 \to [0, 1]$ *is a bivariate copula.*

Proof. Since $f(1) = g(1) = 0$, we have $C(z, 1) = C(1, z) = z$. If $a > -1$ or $b > 0$ then $C(z, 0) = C(0, z) = 0$ for $z \in [0, 1]$. If $a = -1$ and $b < 0$, we need to show that $C(x, y) \to 0$ as $(x, y) \to (0^+, 0^+)$. Multiplying (1) by Fx^{-F-1} and integrating from z to 1 show that $f(z) \leq 1 - z^F$. In the same way, $g(z) \leq 1 - z^G$, so for (x, y) sufficiently close to $(0, 0)$ we have

$$xy(1 - f(x)g(y))^b \leq \frac{xy}{(x^F + y^G - x^F y^G)^{|b|}}.$$

The right-hand side converges to 0 as $(x, y) \to (0^+, 0^+)$, since we are assuming $1 \geq abF = |b|F$ and $1 \geq abG = |b|G$, $C(x, y) \to 0$ as $(x, y) \to (0^+, 0^+)$. $\square$

Next, on $(0, 1)^2$ we have

$$(1 + af(x)g(y))^{2-b} \frac{\partial^2 C}{\partial x \partial y} = H(x, y),$$

where

$$H(x, y) = \left(1 + af(x)g(y) + abxf'(x)g(y)\right)\left(1 + af(x)g(y) + abyg'(y)f(x)\right) + abxf'(x)yg'(y).$$

It remains to show that H is non negative.

Suppose that $ab > 0$, $abF \leq 1$, and $abG \leq 1$. Since $0 \geq xf'(x) \geq F(f(x) - 1)$,

$$\begin{aligned} 1 + af(x)g(y) + abxf'(x)g(y) &\geq 1 - f(x)g(y) + abF(f(x) - 1)g(y) \\ &\geq 1 - f(x)g(y) + (f(x) - 1)g(y) \\ &\geq 1 - g(y) \\ &\geq 0. \end{aligned}$$

Similarly, $1 + af(x)g(y) + abyg'(y)f(x) \geq 0$, so

$$H(x, y) \geq (1 + af(x)g(y) + abxf'(x)g(y))(1 + af(x)g(y) + abyg'(y)f(x)) \geq 0.$$

On the other hand, suppose that $ab < 0$ and $abFG \geq -1$. Then since
$0 \leq -xf'(x) \leq F(1 - f(x))$ and $0 \leq -yg'(y) \leq G(1 - g(y))$,

$$\begin{aligned} H(x, y) &\geq (1 - f(x)g(y))^2 + abxf'(x)yg'(y) \\ &= (1 - f(x)g(y))^2 + abFG(1 - f(x))(1 - g(y)) \\ &\geq (1 - f(x))(1 - g(y)) - (1 - f(x))(1 - g(y)) \\ &= 0. \end{aligned}$$

Corollary 1. *Suppose $a \in [-1, \infty)$, $b \in (-\infty, \infty)$, $p_i \geq 1$ and $q_i \geq 0$, $abq_i \leq 1$, $i \in \{1, 2\}$, and $abq_1q_2 \geq -1$. Then*

$$C(x, y) = xy\left(1 + a(1 - x^{q_1})^{p_1}(1 - y^{q_2})^{p_2}\right)^b$$

defines a bivariate copula.

Remark 1. We have now, in one stroke, the following copula families:

Farlie–Gumbel–Morgenstern Family: $C(x, y) = xy(1 + a(1 - x)(1 - y))$, $a \in [-1, 1]$;
Ali–Mikhail–Haq Family: $C(x, y) = xy(1 + a(1 - x)(1 - y))^{-1}$, $a \in [-1, 1]$;
Huang–Kotz Family: $C(x, y) = xy((1 + a(1 - x)^p(1 - y)^p)$, $a \in [-1, 1]$, $p \geq 1$;
Bairamov–Kotz Family: $C(x, y) = xy(1 + a(1 - x^q)^p(1 - y^q)^p)$, $a \in [-1, 1]$, $p \geq 1, q \geq 0$;
Bekrizadeh–Parham–Zadkarmi Family: $C(x, y) = xy(1 + a(1 - x^p)(1 - y^p))^b$, $b \in \{0, 1, 2, \ldots\}$, $abp \leq 1$, $abp^2 \geq -1$,

as well as a new four-parameter family $C(x, y) = (1 + a(1 - x^q)^p(1 - y^q)^p)^b$, $abq \leq 1$, $abq^2 \geq -1$, $a \geq -1$, $q \geq 0$, and $p \geq 1$, which we propose to call the **Fractional Bairamov–Kotz (FBK) Family**. Note also that we are weakening the requirements on b in the Bekrizadeh–Parham–Zadkarmi Family and that our copulas need not be symmetric in x and y.

Proof. If $abq_1q_2 = 0$ then $C(x, y) = xy$ which is the uniform distribution function on $[0, 1]^2$. Otherwise, suppose $p \geq 1$ and $q > 0$. Put $h(u) = (1 - u^q)^p$. Then setting $v = 1 - u^q \in [0, 1]$

$$\begin{aligned} h(u) - \frac{uh'(u)}{q} &= (1 - u^q)^p + p(1 - u^q)^{p-1}u^q \\ &= pv^{p-1} - (p - 1)v^p \\ &\in [0, 1] \end{aligned}$$

so the corollary follows by setting $F = q_1$ and $G = q_2$ in the Theorem. □

Remark 2. It is worth noting that the family

$$C_q(x, y) = \frac{xy}{(1 - (1 - x^q)(1 - y^q))^{1/q}}, \quad q > 0,$$

has the property that

$$\begin{aligned} \lim_{q \to 0^+} C_q(x, y) &= xy, \\ \lim_{q \to +\infty} C_q(x, y) &= \min(x, y), \end{aligned}$$

and is continuous in q. Since $C(x, y) = xy$ is the independent copula and $C(x, y) = \min(x, y)$ is the Fréchet–Hoeffding upper bound copula, for any measure of concordance, the theoretical range of this family's concordance values is $(0, 1)$; see Scarsini (1984) for more details.

Remark 3. If we set $a = 1/(nF)$ and $b = n$ then

$$C(x, y) = \lim_{n \to \infty} xy \left(1 + \frac{1}{nF} f(x) f(y)\right)^n = xy \exp\left(\frac{f(x) f(y)}{F}\right)$$

is a copula as is

$$C(x, y) = xy \exp\left(-\frac{f(x)g(y)}{FG}\right)$$

obtained by setting $a = -1/(nFG)$ and $b = n$ and letting $n \to \infty$.

3. A copula recipe

Theorem 1 provides us a recipe for constructing f and g, and thereby, a recipe for constructing copulas. Suppose we are given $\phi : [0, 1] \to [0, 1]$, where ϕ is continuous and $\phi(0) = 1$. The theorem tells us that f and g can be obtained by solving differential equations of the form

$$h(u) - \frac{uh'(u)}{H} = \phi(u), \quad h(1) = 0$$

on $(0, 1]$. A solution of such a differential equation is

$$h(u) = Hu^H \int_u^1 \frac{\phi(v)}{v^{H+1}}\, dv = H \int_1^{1/u} \frac{\phi(uz)}{z^{H+1}}\, dz.$$

Recall that we want $h \in \mathcal{C}$. We extend h to $[0, 1]$ by continuity and see that $h(0) = 1$. If we assume that $\phi(v)$ is non increasing then h is strictly decreasing. For example, if $\phi(v) = 0$ for all $v \in [0, 1]$, then $h(u) = 1 - u^H$. It is difficult to see what the most general condition on ϕ might be. For example, suppose N is a positive integer and we define

$$h(x) = 1 - x(1 + \mathrm{sinc}(N\pi x)),$$

where $\mathrm{sinc}(x) = \sin(x)/x$; we have $h(0) = 0$, $h(1) = 1$, $h'(x) = -(1 + \cos(N\pi x))$, and

$$\phi(x) := h(x) - \frac{xh'(x)}{H} = 1 - x\left(1 + \mathrm{sinc}(N\pi x)) - \frac{1 + \cos(N\pi x)}{H}\right).$$

To keep $\phi(x) \geq 0$, we need

$$H \geq \frac{1 + \cos(z)}{1 + \mathrm{sinc}(z)}.$$

Since the right-hand side is maximized as a function of $z \geq 0$ for some $z \in (\pi, 2\pi)$ and is the ratio of an increasing function to a decreasing function on this interval, we see that a lower bound for H is

$$H \geq \frac{1 + \cos(23\pi/12)}{1 + \mathrm{sinc}(22\pi/12)} \approx 2.153,$$

independent of N. Smaller values of H are possible, and numerical experiments indicate that H may be as low as 2.09 independent of N.

References

Bairamov, I., Kotz, S. (2002). Dependence structure and symmetry of Huang–Kotz FGM distributions and their extensions. *Metrika*. 56(1):55–72.

Bekrizadeh, H., Parham, G.A., Zadkarmi, M.R. (2012). The new generalization of Farlie–Gumbel–Morgenstern copulas. *Appl. Math. Sci.* 6(71):3527–3533.

Huang, J.S., Kotz, S. (1999). Modifications of the Farlie–Gumbel–Morgenstern distributions. A tough hill to climb. *Metrika* . 49(2):135–145.

Nelsen, R.B. (2003). *Properties and Applications of Copulas: A Brief Survey*. matematica.campusnet.unito.it/didattica/att/d11d.7507.file.pdf.

Scarsini, M. (1984). On measures of concordance. *Stochastica*. 8:201–218.