Kopule ako nástroj modelovania štruktúry stochastickej závislosti náhodných vektorov

Radko Mesiar
STU Bratislava, Faculty of Civil Engineering

ROBUST 2012, Nemčičky, 12. 9. 2012

Contents

Introduction

Abe Sklar 1957

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

Sklar theorem

$$Z = (X_1, \ldots, X_n)$$

$$F_Z(x_1, \ldots, x_n) = C\,(F_1(x_1), \ldots, F_n(x_n))$$

$$X_i \sim \text{ uniform on }]0, 1[$$

$$C \equiv F_Z \,|\, [0, 1]^n$$

Introduction

Abe Sklar 1957

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

Sklar theorem

$$Z = (X_1, \ldots, X_n)$$

$$F_Z(x_1, \ldots, x_n) = C\left(F_1(x_1), \ldots, F_n(x_n)\right)$$

$$X_i \sim \text{ uniform on }]0, 1[$$

$$C \equiv F_Z \mid [0, 1]^n$$

Introduction

Abe Sklar 1957

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

Sklar theorem

$$Z = (X_1, \ldots, X_n)$$

$$F_Z(x_1, \ldots, x_n) = C\,(F_1(x_1), \ldots, F_n(x_n))$$

$$X_i \sim \text{ uniform on }]0, 1[$$

$$C \equiv F_Z \,|\, [0, 1]^n$$

Axiomatic approach

i) $C(x_1, \ldots, x_n) = 0$ if some $x_i = 0$
C is **grounded**

ii) $C(x_1, \ldots, x_n) = x_i$ if $\forall j \neq i, x_j = 1$
C has **neutral element 1**

iii) $\forall u_1 \leq v_1, \ldots, u_n \leq v_n$
$\sum(-1)^{|I|} C(t_1^I, \ldots, t_n^I) \geq 0$
$I \subseteq \{1, \ldots, n\} \quad t_i^I = \begin{cases} u_i & \text{if } i \in I, \\ v_i & \text{if } i \notin I \end{cases}$
C is n–**increasing**

Axiomatic approach

i) $C(x_1, \ldots, x_n) = 0$ if some $x_i = 0$
C is **grounded**

ii) $C(x_1, \ldots, x_n) = x_i$ if $\forall j \neq i, x_j = 1$
C has **neutral element 1**

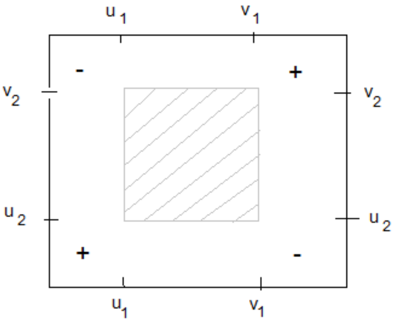
iii) $\forall u_1 \leq v_1, \ldots, u_n \leq v_n$
$\sum (-1)^{|I|} C(t_1^I, \ldots, t_n^I) \geq 0$
$I \subseteq \{1, \ldots, n\} \quad t_i^I = \begin{cases} u_i & \text{if } i \in I, \\ v_i & \text{if } i \notin I \end{cases}$
C is n–**increasing**

Axiomatic approach

i) $C(x_1, \ldots, x_n) = 0$ if some $x_i = 0$
C is **grounded**

ii) $C(x_1, \ldots, x_n) = x_i$ if $\forall j \neq i, x_j = 1$
C has **neutral element 1**

iii) $\forall u_1 \leq v_1, \ldots, u_n \leq v_n$
$\sum (-1)^{|I|} C(t_1^I, \ldots, t_n^I) \geq 0$
$I \subseteq \{1, \ldots, n\} \quad t_i^I = \begin{cases} u_i & \text{if } i \in I, \\ v_i & \text{if } i \notin I \end{cases}$
C is n–**increasing**



Each C is **1–Lipschitz** (wrt. L_1–norm)

$$C(v_1, v_2) - C(v_1, u_2) - C(v_2, u_1) + C(u_1, u_2) \geq 0$$

2–increasing, supermodular

$$W \leq C \leq M$$

$$M(u_1, \ldots, u_n) = min\ \{u_1, \ldots, u_n\}$$

$$W(u_1, \ldots, u_n) = max\ \left\{0, \sum_{i=1}^{n} u_i - n + 1\right\}$$

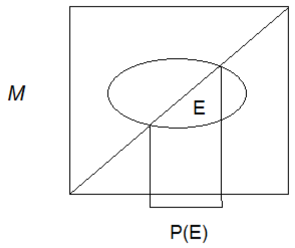
$$\Pi(u_1, \ldots, u_n) = \prod_{i=1}^{n} u_i$$

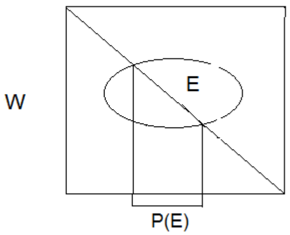
$\mathcal{C}_n$ class of all n–ary copulas is convex and compact

Example

Copulas are in 1–1 correspondence with probability measures on $\mathcal{B}([0,1]^n)$ with uniform 1–dimensional marginals

$$\Pi \sim \text{Lebesgue}$$





Statistical interpretation

$\Pi \sim$ independence

$M \sim$ comonotone dependence,

$$X_i = f_i(X_1), \quad f_i \nearrow$$

$W \sim$ counter monotone dependence,

$$X_2 = g(X_1), \quad g \searrow$$

Statistical interpretation

$\Pi \sim$ independence

$M \sim$ comonotone dependence,

$$X_i = f_i(X_1), \quad f_i \nearrow$$

$W \sim$ counter monotone dependence,

$$X_2 = g(X_1), \quad g \searrow$$

Statistical interpretation

$\Pi \sim$ independence

$M \sim$ comonotone dependence,

$$X_i = f_i(X_1), \quad f_i \nearrow$$

$W \sim$ counter monotone dependence,

$$X_2 = g(X_1), \quad g \searrow$$

Gaussian and t–copulas $\sim$ eliptic copulas, no closed form

$$C(u_1, \dots, u_n) = F_Z\left(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)\right)$$

X_i continuous, $F_i(X_i)$ uniform on $]0,1[$

in general, margins need not be Gaussian (Student)!

Gaussian and t–copulas $\sim$ eliptic copulas, no closed form

$$C(u_1,\ldots,u_n) = F_Z\left(F_1^{-1}(u_1),\ldots,F_n^{-1}(u_n)\right)$$

X_i continuous, $F_i(X_i)$ uniform on $]0,1[$

in general, margins need not be Gaussian (Student)!

Gaussian and t–copulas $\sim$ eliptic copulas, no closed form

$$C(u_1, \ldots, u_n) = F_Z\left(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\right)$$

X_i continuous, $F_i(X_i)$ uniform on $]0, 1[$

in general, margins need not be Gaussian (Student)!

flipping

$$C^{-}(u,v) = u - C(u, 1-v)$$

$$(X,Y) \rightarrow (X,-Y)$$

$$C_{-}(u,v) = v - C(1-u, v)$$

$$(X,Y) \rightarrow (-X,Y)$$

survival

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

$$(X,Y) \rightarrow (-X,-Y)$$

flipping

$$C^{-}(u,v) = u - C(u, 1-v)$$

$$(X,Y) \to (X,-Y)$$

$$C_{-}(u,v) = v - C(1-u, v)$$

$$(X,Y) \to (-X,Y)$$

survival

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

$$(X,Y) \to (-X,-Y)$$

flipping

$$C^{-}(u,v) = u - C(u, 1-v)$$

$$(X,Y) \rightarrow (X,-Y)$$

$$C_{-}(u,v) = v - C(1-u, v)$$

$$(X,Y) \rightarrow (-X,Y)$$

survival

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

$$(X,Y) \rightarrow (-X,-Y)$$

Ordinal sums

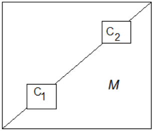


Figure: *M*–ordinal

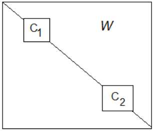


Figure: *W*–ordinal

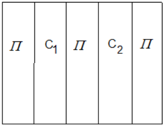


Figure: Π–ordinal

generators $f : [0,1] \to [0,\infty]$
strictly decreasing, continuous,

$$f(1) = 0$$

$$C_f(u_1, \ldots, u_n) = f^{-1}\left(min\left\{f(0), \sum_{i=1}^{n} f(u_i)\right\}\right)$$

$$g : [-\infty, 0] \to [0,1], \quad g(x) = f^{-1}(min\{f(0), -x\})$$

$$g' \geq 0, \ldots, g^{(n-2)} \geq 0, \quad g^{(n-2)} \text{ convex}$$

$$n = 2 \equiv f \text{ convex}$$

generators $f : [0,1] \to [0,\infty]$
strictly decreasing, continuous,

$$f(1) = 0$$

$$C_f(u_1, \ldots, u_n) = f^{-1}\left(min\left\{f(0), \sum_{i=1}^{n} f(u_i)\right\}\right)$$

$$g : [-\infty, 0] \to [0,1], \quad g(x) = f^{-1}(min\{f(0), -x\})$$

$$g' \geq 0, \ldots, g^{(n-2)} \geq 0, \quad g^{(n-2)} \text{ convex}$$

$$n = 2 \;\equiv\; f \text{ convex}$$

generators $f : [0,1] \to [0,\infty]$
strictly decreasing, continuous,

$$f(1) = 0$$

$$C_f(u_1,\ldots,u_n) = f^{-1}\left(min\left\{f(0), \sum_{i=1}^{n} f(u_i)\right\}\right)$$

$$g : [-\infty,0] \to [0,1], \quad g(x) = f^{-1}\left(min\left\{f(0), -x\right\}\right)$$

$$g' \geq 0,\ldots,g^{(n-2)} \geq 0, \quad g^{(n-2)} \text{ convex}$$

$$n = 2 \; \equiv \; f \text{ convex}$$

Archimedean copulas

associative and $C_f(u, \ldots, u) < u$ for $u \in]0, 1[$

$$f_{\Pi}(x) = -\log x \quad \forall n$$

$$f_W(x) = 1 - x \quad \text{only } n = 2$$

n fixed, weakest $C_{f[n]}$

Clayton copula

$$f^{[n]}(x) = 1 - x^{\frac{1}{n-1}}$$

Archimedean copulas

associative and $C_f(u, \ldots, u) < u$ for $u \in]0, 1[$

$$f_{\Pi}(x) = -log\, x \quad \forall n$$

$$f_W(x) = 1 - x \quad \text{only } n = 2$$

n fixed, weakest $C_{f[n]}$

Clayton copula

$$f^{[n]}(x) = 1 - x^{\frac{1}{n-1}}$$

Archimedean copulas

$$\text{associative and } C_f(u,\ldots,u) < u \text{ for } u \in]0,1[$$

$$f_\Pi(x) = -log\, x \quad \forall n$$

$$f_W(x) = 1 - x \quad \text{only } n = 2$$

n fixed, weakest $C_{f^{[n]}}$

Clayton copula

$$f^{[n]}(x) = 1 - x^{\frac{1}{n-1}}$$

Clayton

$$f_\lambda^C(x) = \frac{x^{-\lambda} - 1}{\lambda}, \quad \lambda \in [-1, \infty[, \quad \lambda \neq 0$$

$$f_0^\lambda = f_\Pi, \quad \lambda \geq 0 \forall n$$

$\lambda = 1$ ALI–MIKHAIL–HAQ copula

Hamacher product

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

Gumbel

$$\lambda \in [1, \infty[\quad f_\lambda^G(x) = (-log\, x)^\lambda, \quad \forall n$$

Clayton

$$f_\lambda^C(x) = \frac{x^{-\lambda} - 1}{\lambda}, \quad \lambda \in [-1, \infty[, \quad \lambda \neq 0$$

$$f_0^\lambda = f_\Pi, \quad \lambda \geq 0 \forall n$$

$\lambda = 1$ ALI–MIKHAIL–HAQ copula

Hamacher product

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

Gumbel

$$\lambda \in [1, \infty[\quad f_\lambda^G(x) = (-log\, x)^\lambda, \quad \forall n$$

Clayton

$$f_\lambda^C(x) = \frac{x^{-\lambda} - 1}{\lambda}, \quad \lambda \in [-1, \infty[, \quad \lambda \neq 0$$

$$f_0^\lambda = f_\Pi, \quad \lambda \geq 0 \forall n$$

$\lambda = 1$ ALI–MIKHAIL–HAQ copula

Hamacher product

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

Gumbel

$$\lambda \in [1, \infty[\quad f_\lambda^G(x) = (-log\, x)^\lambda, \quad \forall n$$

$n = 2$

$$C_{(f)}(u,v) = u\, f^{-1}\left(min\left\{ \frac{f(v)}{u},\ f(0) \right\} \right)$$

Univariate Conditioning Stable

$$C_{(f_1^C)}(u,v) = \frac{u^2\, v}{1 - u + uv}$$

$$C_{(f_\Pi)}(u,v) = u\, v^{-\frac{1}{u}}$$

$$C_{(f_W)} = W$$

$n = 2$

$$C_{(f)}(u,v) = u\, f^{-1}\left(min\left\{\frac{f(v)}{u},\, f(0)\right\}\right)$$

Univariate Conditioning Stable

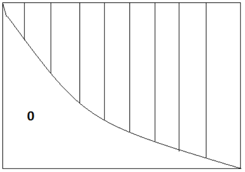
$$C_{(f_1^C)}(u,v) = \frac{u^2\, v}{1-u+uv}$$

$$C_{(f_\Pi)}(u,v) = u\, v^{-\frac{1}{u}}$$

$$C_{(f_W)} = W$$

Associative copula $\equiv$ M–ordinal sums of Archimedean

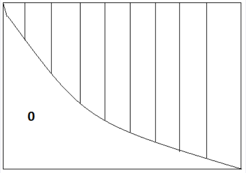
0 - y - semilinear copulas



$$v = 1 - d(u)$$
$$C_d(u,v) = u \cdot \max\left\{0, 1 + \frac{v-1}{d(u)}\right\}$$
$$\text{distortion } d : [0,1] \to [0,1]$$
$$d(x) \nearrow, \quad \frac{x}{d(x)} \nearrow, \quad d(1) = 1$$

$$C_{id} = W, \quad C_1 = \Pi$$

0 - y - semilinear copulas



$$v = 1 - d(u)$$
$$C_d(u,v) = u \cdot \max\left\{0, 1 + \frac{v-1}{d(u)}\right\}$$
$$\text{distortion } d : [0,1] \to [0,1]$$
$$d(x) \nearrow, \quad \frac{x}{d(x)} \nearrow, \quad d(1) = 1$$

$$C_{id} = W, \quad C_1 = \Pi$$

DUCS copulas
Distorted **U**nivariate **C**onditioning **S**table

$$C_{f,d}(u,v) = u \cdot f^{-1}\left(\min\left\{f(0), \frac{f(v)}{d(u)}\right\}\right) \quad u > 0$$

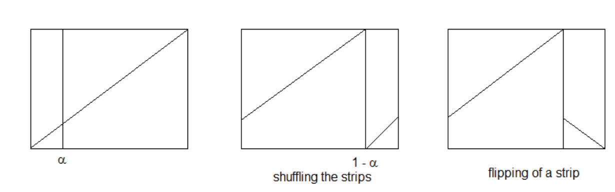
$$d(x) = x^p, \quad p \in]0,1]$$

$$f(x) = \frac{1}{x} - 1$$

$$C_{f,d}(u,v) = \frac{u^{p+1}v}{1 + u^p v - v}$$

Special singular copulas

Shuffles of M



Each copula can be seen as a limit of shuffles of M

Extreme values (EV) copulas

$$(X_1, Y_1) \sim C$$

$$\vdots$$

$$(X_n, Y_n) \sim C$$

$$\left(\bigvee X_i, \bigvee Y_i\right) \sim C_n$$

$$C_n(u,v) = \left(C(u^{\frac{1}{n}}, v^{\frac{1}{n}})\right)^n \xrightarrow[n\to\infty]{} C^*(u,v)$$

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

Extreme values (EV) copulas

$$(X_1, Y_1) \sim C$$

$$\vdots$$

$$(X_n, Y_n) \sim C$$

$$\left(\bigvee X_i, \bigvee Y_i\right) \sim C_n$$

$$C_n(u, v) = \left(C(u^{\frac{1}{n}}, v^{\frac{1}{n}})\right)^n \xrightarrow[n\to\infty]{} C^*(u, v)$$

$$C^*\left(u^\lambda, v^\lambda\right) = \left(C^*(u, v)\right)^\lambda$$

Extreme values (EV) copulas

$$C^{*}(u,v)=(u\,v)^{A\left(\frac{\log\ u}{\log\ uv}\right)}=$$

$$=g^{-1}\left(\min\left\{g(0),(g(u)+g(v))\right)\cdot A\left(\frac{g(u)}{g(u)+g(v)}\right\}\right)$$

$$g(x)=-\log x$$

$$A:[0,1]\to[0,1]\quad \text{convex}$$

$$t\vee(1-t)\leq A(t)\leq 1$$

Archimax

$$C_{f,A}(u,v)=f^{-1}\left(\min\left(f(0),(f(u)+f(v))\right)\cdot A\left(\frac{f(u)}{f(u)+f(v)}\right)\right)$$

Extreme values (EV) copulas

$$C^* (u,v) = (u\,v)^{A\left(\frac{\log\ u}{\log\ uv}\right)} =$$

$$= g^{-1}\left(\min\left\{g(0),(g(u)+g(v))\right) \cdot A\left(\frac{g(u)}{g(u)+g(v)}\right\}\right)$$

$$g(x) = -\log x$$

$$A : [0,1] \to [0,1] \quad \text{convex}$$

$$t \vee (1-t) \leq A(t) \leq 1$$

Archimax

$$C_{f,A}(u,v) = f^{-1}\left(\min\left(f(0),(f(u)+f(v))\right) \cdot A\left(\frac{f(u)}{f(u)+f(v)}\right)\right)$$

Dependence (association) measures

Kendall's tau

$$t = \frac{c-d}{\binom{n}{2}}$$

c number of concordant pairs
d number of discordant pairs
IID $(X_1, Y_1), \ (X_2, Y_2) \sim (X, Y)$

$$\tau_{X,Y} = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0)$$

$$\tau_C = 4 \int_0^1 \int_0^1 C(u,v) \, dC(u,v) - 1$$

C Archimedean, ad. gen. f:

$$\tau_C = 1 + 4 \int_0^1 \frac{f(t)}{f'(t)} dt$$

Dependence (association) measures

Kendall's tau

$$t = \frac{c - d}{\binom{n}{2}}$$

c number of concordant pairs
d number of discordant pairs
IID $(X_1, Y_1),\ (X_2, Y_2) \sim (X, Y)$

$$\tau_{X,Y} = P\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - P\left((X_1 - X_2)(Y_1 - Y_2) < 0\right)$$

$$\tau_C = 4 \int_0^1 \int_0^1 C(u, v)\, dC(u, v) - 1$$

C Archimedean, ad. gen. f:

$$\tau_C = 1 + 4 \int_0^1 \frac{f(t)}{f'(t)} dt$$

Dependence (association) measures

Kendall's tau

$$t = \frac{c - d}{\binom{n}{2}}$$

c number of concordant pairs
d number of discordant pairs
IID $(X_1, Y_1), \ (X_2, Y_2) \sim (X, Y)$

$$\tau_{X,Y} = P\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - P\left((X_1 - X_2)(Y_1 - Y_2) < 0\right)$$

$$\tau_C = 4 \int_0^1 \int_0^1 C(u, v)\, dC(u, v) - 1$$

C Archimedean, ad. gen. f:

$$\tau_C = 1 + 4 \int_0^1 \frac{f(t)}{f'(t)} dt$$

Spearman's rho

IID $(X_1, Y_1),\ (X_2, Y_2),\ (X_3, Y_3) \sim (X, Y)$

$$\rho_{X,Y} = 3\ \left(P\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) - P\left((X_1 - X_2)(Y_1 - Y_3) < 0\right)\right)$$

$$\rho_C = 12 \int_0^1 \int_0^1 C(u, v)\ du\ dv - 3$$

rank correlation
$\equiv$ Pearson's rho for $U = F_X(X),\ V = F_Y(Y)$
(if $X,\ Y$ are continuous r. v.)

Spearman's rho

IID $(X_1, Y_1),\ (X_2, Y_2),\ (X_3, Y_3) \sim (X, Y)$

$$\rho_{X,Y} = 3\ \left(P\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) - P\left((X_1 - X_2)(Y_1 - Y_3) < 0\right)\right)$$

$$\rho_C = 12 \int_0^1 \int_0^1 C(u, v)\, du\, dv - 3$$

rank correlation
$\equiv$ Pearson's rho for $U = F_X(X),\ V = F_Y(Y)$
(if $X,\ Y$ are continuous r. v.)

if C is EV-copula with dependence function A

$$\tau_C = \int_0^1 \frac{t\,(1-t)}{A(t)} dA'(t)$$

$$\rho_C = \int_0^1 \frac{1}{(1+A(t))^2} dt - 3$$

general relationship

$$-1 \leq 3\tau - 2\rho \leq 1$$

$$\tau = 0 : \quad -\frac{1}{2} \leq \rho \leq \frac{1}{2}$$

$$\rho = 0 : \quad -\frac{1}{3} \leq \tau \leq \frac{1}{3}$$

$$\frac{1+\rho}{2} \geq \frac{(1+\tau)^2}{2}, \quad \frac{1-\rho}{2} \geq \frac{(1-\tau)^2}{2}$$

best bounds:

$$\tau \geq 0 : \quad \frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2}$$

$$\tau \leq 0 : \quad \frac{\tau^2 + 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2}$$

general relationship

$$-1 \leq 3\tau - 2\rho \leq 1$$

$$\tau = 0: \quad -\frac{1}{2} \leq \rho \leq \frac{1}{2}$$

$$\rho = 0: \quad -\frac{1}{3} \leq \tau \leq \frac{1}{3}$$

$$\frac{1+\rho}{2} \geq \frac{(1+\tau)^2}{2}, \quad \frac{1-\rho}{2} \geq \frac{(1-\tau)^2}{2}$$

best bounds:

$$\tau \geq 0: \quad \frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2}$$

$$\tau \leq 0: \quad \frac{\tau^2 + 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2}$$

general relationship

$$-1 \leq 3\tau - 2\rho \leq 1$$

$$\tau = 0: \quad -\frac{1}{2} \leq \rho \leq \frac{1}{2}$$

$$\rho = 0: \quad -\frac{1}{3} \leq \tau \leq \frac{1}{3}$$

$$\frac{1+\rho}{2} \geq \frac{(1+\tau)^2}{2}, \quad \frac{1-\rho}{2} \geq \frac{(1-\tau)^2}{2}$$

best bounds:

$$\tau \geq 0: \quad \frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2}$$

$$\tau \leq 0: \quad \frac{\tau^2 + 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2}$$

GINI

$$\gamma_C = 4\left[\int_0^1 C(u, 1-u)\,du - \int_0^1 (u - C(u,u))\,du\right]$$

BLOMQVIST

$$\beta_C = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1$$

for any dependence measure

$$M \longrightarrow 1$$

$$\Pi \longrightarrow 0$$

$$W \longrightarrow -1$$

GINI

$$\gamma_C = 4\left[\int_0^1 C(u, 1-u)\, du - \int_0^1 (u - C(u,u))\, du\right]$$

BLOMQVIST

$$\beta_C = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1$$

for any dependence measure

$$M \longrightarrow 1$$

$$\Pi \longrightarrow 0$$

$$W \longrightarrow -1$$

GINI

$$\gamma_C = 4\left[\int_0^1 C(u, 1-u)\, du - \int_0^1 (u - C(u,u))\, du\right]$$

BLOMQVIST

$$\beta_C = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1$$

for any dependence measure

$$M \longrightarrow 1$$

$$\Pi \longrightarrow 0$$

$$W \longrightarrow -1$$

Thanks for your attention