Methods of constructing copulas. Inversion Method and Algebraic Method

Duc T. Nguyen

Ulm University
Advisor: **Prof. Evgeny Spodarev**

February 5, 2021

References

1. **An introduction to copulas.**
   Roger B. Nelsen - (2006) Springer(Springer series in statistics).

2. **Elements of Copula Modeling with R-Springer.**
   Marius Hofert, Ivan Kojadinovic, Martin Machler, Jun Yan - (2018) Springer.

Outline

- First part: Survival copula.
- Second part: Continuity and Singurlarity.
- Third part: Inversion Method.
  - The Marshall-Olkin Bivariate Exponential Distribution.
  - The Circular Uniform Distribution.
- Fourth part: Survival copula.
  - Plackett Distributions.
  - Ali-Mikhail-Haq Distributions.

PART 1: Survival Copula.

Definition

- The probability of an individual living X or "surviving" beyond time x, which is called **the survival function**

$$\bar{F}(x) = \mathbb{P}[X > x] = 1 - F(x).$$

- For a pair (X, Y) of random variables with joint distribution function H, **the joint survival function**

$$\begin{aligned}\bar{H}(x, y) &= \mathbb{P}[X > x, Y > y] \\ &= 1 - F(x) - G(y) + H(x, y) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + H(x, y)\end{aligned}$$

Survival Copula

- Assume that $\boldsymbol{C_{XY}}$ is the copula of two random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$. From Sklar's Theorem, one has

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

- Rewrite the survival joint distribution function

$$\bar{H}(x, y) = \bar{F}(x) + \bar{G}(y) - 1 + C(1 - \bar{F}(x), 1 - \bar{G}(y))$$

- If we define a function $\hat{C} : \mathbb{I}^2 \to \mathbb{I}$ given by

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

we have

$$\bar{H}(x, y) = \hat{C}(\bar{F}(x), \bar{G}(y)).$$

$\hat{C}$ is a copula

- We check the following conditions
  - $Dom\hat{C} = \mathbb{I}^2$.
  - For any rectangle $K = [x_1, x_2] \times [y_1, y_2]$ lying in $\mathbb{I}^2$ then $K' = [1 - x_2, 1 - x_1] \times [1 - y_2, 1 - y_1]$ also is a rectangle within $\mathbb{I}^2$ and

$$\begin{aligned} V_{\hat{C}}(K) &= \hat{C}(x_2, y_2) + \hat{C}(x_1, y_1) - \hat{C}(x_1, y_2) - \hat{C}(x_2, y_1) \\ &= C(1 - x_2, 1 - y_2) - C(1 - x_2, 1 - y_1) \\ &\qquad - C(1 - x_1, 1 - y_2) + C(1 - x_1, 1 - y_1) \end{aligned}$$

  - $C(u, 0) = 0 = C(0, v)$ for any $(u, v) \in \mathbb{I}^2$.
  - $C(u, 1) = u$ and $C(1, v) = v$ for any $(u, v) \in \mathbb{I}^2$.

Example

- We have the following relation

$$\bar{H}(x,y) = \hat{C}(\bar{F}(x), \bar{G}(y)).$$

and

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

- One can find the survival copula of X and Y if
  - The copula C_{XY} is known.
  - Using Sklar's Theorem when the survival joint distribution function is known.

Example

- *(Gumbel's bivariate exponential distribution)*. Let H_θ be the joint distribution function of X and Y given by

$$H(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)} & \text{if } x \geq 0, y \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

where $\theta \in [0,1]$. Then the copula C_{XY} is

$$C_\theta(u,v) = u + v - 1 + (1-u)(1-v)e^{-\theta \ln(1-u)\ln(1-v)}$$

hence

$$\hat{C}_\theta(u,v) = uve^{-\theta \ln u \ln v}$$

Example

- Let X and Y be random variables whose joint survival function is given by

$$\bar{H}_\theta(x,y) = \begin{cases} (1+x+y)^{-\theta} & \text{if } x \geq 0, y \geq 0 \\ (1+x)^{-\theta} & \text{if } x \geq 0, y < 0 \\ (1+y)^{-\theta} & \text{if } x < 0, y \geq 0 \\ 1 & \text{if } x < 0, y < 0. \end{cases}$$

and the marginal survival function $\hat{F}$ and $\hat{G}$ are

$$\hat{F}(x) = \begin{cases} (1+x)^{-\theta} & x \geq 0 \\ 1, & x < 0 \end{cases} \text{ and } \hat{G}(y) = \begin{cases} (1+y)^{-\theta} & y \geq 0 \\ 1, & y < 0 \end{cases}$$

From the Corollary of Sklar's theorem, one has

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

PART 2: Continuity and Singularity.

Singular copula and Absolutely continuous copula

- For each copula C induces a probability measure on $\mathbb{I}^2$ by the following equation

$$V_C([0,u] \times [0,v]) = C(u,v).$$

- For any copula C, let

$$C(u,v) = A_C(u,v) + S_C(u,v)$$

where

$$A_C(u,v) = \int_0^u \int_0^v \frac{\delta^2}{\delta s \delta t} C(s,t) dt ds$$

and

$$S_C(u,v) = C(u,v) - A_C(u,v)$$

Singular copula and Absolutely continuous copula

- If $C \equiv A_C$, then C is absolutely continuous and it has a joint density function given by $\delta^2 C(u,v)/\delta u \delta v$.
- If $\delta^2 C(u,v)/\delta u \delta v = 0$ almost everywhere in $\mathbb{I}^2$, then C is singular.
- From the relation between copula and its components, one has

$$A_C(1,1) + S_C(1,1) = C(1,1) = 1$$

- The Fréchet - Hoeffding upper bound $M(u,v)$ is singular and the product copula is absolutely continuous.

PART 3: The Inversion Method.

Sklar's Theorem

- Let H be a bivariate function with continuous margins F and G. From Sklar's theorem,there exists a copula such that

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

and

$$\hat{C}(u, v) = \bar{H}(barF^{-1}(u), barG^{-1}(v))$$

where $\bar{F}^{-1}(t) = F^{-1}(1 - t)$.

Outline

Problem: Consider a system with two components whose life times are X and Y respectively, compute the probabilities that one or both components may fail.

- It is necessary to find the survival function

$$\bar{H}(x, y) = \mathbb{P}[X > x, Y > y]$$

- Assume that the "shock" of these components follow the Poisson processes with certain parameters:
  - Component 1 fail: λ_1
  - Component 2 fail: λ_2
  - Both components fail: λ_{12}

Problem Setting

- Let Z_1, Z_2 and Z_{12} be the times of the shocks and they are independent exponential random variables, then

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

- For all $x, y \geq 0$,

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

- The marginal distribution functions are

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

Survival copula

- Since $\max(x,y) = x + y - \min(x,y)$, one has

$$\bar{H}(x,y) = \bar{F}(x)\bar{G}(y)\min(\exp\{\lambda_{12}x\}, \exp\{\lambda_{12}y\})$$

- Denote by

$$\alpha = \lambda_{12}/(\lambda_1 + \lambda_{12}), \quad \beta = \lambda_{12}/(\lambda_2 + \lambda_{12})$$

and

$$u = \bar{F}(x), \quad v = \bar{G}(y)$$

then

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

- Denote by

$$C_{\alpha,\beta}(u,v) := \hat{C}(u,v).$$

- Note that for any $\alpha, \beta \in [0,1]$,

$$C_{\alpha,0} = \prod = C_{0,\beta} \quad \text{and} \quad C_{1,1} = M$$

Survival copula

- Recall the copula $C_{\alpha,\beta}$

$$C_{\alpha,\beta}(u,v) = \begin{cases} u^{1-\alpha}v & u^\alpha \geq v^\beta \\ uv^{1-\beta}, & u^\alpha \leq v^\beta \end{cases}$$

- For $u^\alpha \neq v^\beta$, the absolutely continuous component is

$$A_{\alpha,\beta}(u,v) = C_{\alpha,\beta}(u,v) - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}[\min(u^\alpha, v^\beta)]^{(\alpha+\beta-\alpha\beta)/(\alpha\beta)}$$

- The singular component is

$$S_{\alpha,\beta}(u,v) = \int_0^{\min(u^\alpha,v^\beta)} t^{\frac{1}{\alpha}+\frac{1}{\beta}-2}dt$$

Singular Component

- Recall that

$$A_{\alpha,\beta}(1,1) + S_{\alpha,\beta}(1,1) = 1$$

Hence,

$$S_{\alpha,\beta}(1,1) = \frac{\alpha\beta}{\alpha + \beta - \alpha\beta}$$

- If $U, V \sim U[0,1]$ with $C_{UV} = C_{\alpha,\beta}$, then

$$\mathbb{P}[U^\alpha = V^\beta] = \frac{\alpha\beta}{\alpha + \beta - \alpha\beta}$$

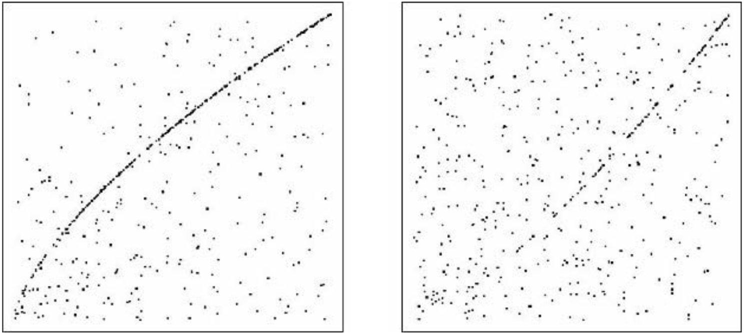

Figure: Marshall - Olkin copulas, $(\alpha, \beta) = (1/2, 3/4), (1/3, 1/4)$

The Circular Uniform Distribution

- Let (X, Y) denote the coordinates of the uniformly chosen point on the unit circle. In polar coordinate system, its coordinate is $(1, \theta)$, then

$$X = \cos(\theta) \quad \text{and} \quad Y = \sin(\theta)$$

or

$$(X, Y) \in [0, 1]^2$$

- To find a copula C_{XY}, one needs to find its joint distribution function $H(x, y)$ for $(x, y) \in [0, 1]^2$.

Joint Distribution function

- The joint distribution function

$$H(x, y) = \mathbb{P}[\cos\theta < x, \sin\theta < y]$$

- Since (X, Y) is uniformly chosen on the circle, hence the probability of this point belongs to an arc is proportional to the length of the arc.

Intuition about JDF

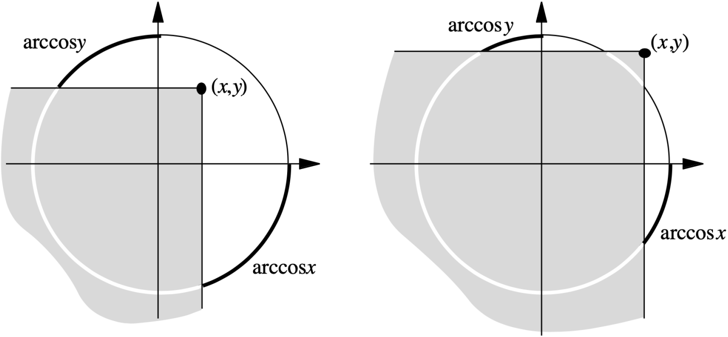

Figure: Joint distribution function

Joint distribution function

$$H(x,y)=\begin{cases}\frac{3}{4}-\frac{\arccos x+\arccos y}{2\pi}, & x^2+y^2\le 1\\ 1-\frac{\arccos x+\arccos y}{\pi}, & x^2+y^2>1; x,y\ge 0\\ 1-\frac{\arccos x}{\pi}, & x^2+y^2>1; x<0, y\ge 0\\ 1-\frac{\arccos y}{\pi}, & x^2+y^2>1; y<0, x\ge 0\\ 0 & x^2+y^2>1; x,y<0\end{cases}$$

- Its marginal distribution functions are

$$F(x)=1-\frac{\arccos x}{\pi} \quad \text{and} \quad G(y)=1-\frac{\arccos y}{\pi}.$$

- Its quasi functions are

$$F^{-1}(u)=\cos(\pi(u-1)) \quad \text{and} \quad G^{-1}(v)=\cos(\pi(v-1))$$

The copula C_{XY}

- The inverse image of the circle $x^2 + y^1 = 1$ via transformations $x = F^{-1}(u), y = G^{-1}(v)$ is

$$|u - 1/2| + |v - 1/2| = 1/2.$$

- From the marginal distribution functions and the support of the copula C_{XY}, one has the formula of the copula

$$C(u,v) = \begin{cases} M(u,v), & |u-v| > 1/2 \\ W(u,v), & |u-v+1| > 1/2 \\ \frac{u+v}{2} - \frac{1}{4}, & otherwise \end{cases}$$

- Note that $\delta^2 C/\delta u \delta v = 0$ almost everywhere in $\mathbb{I}^2$, hence C_{XY} is singular.

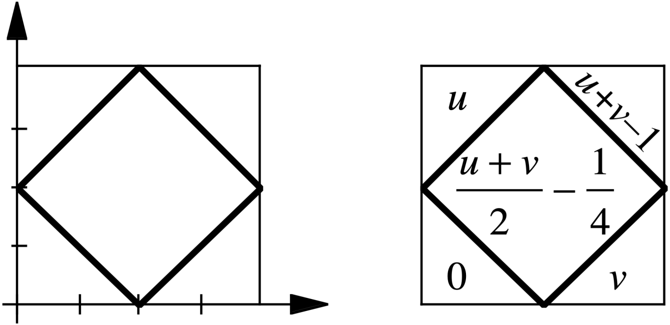

Figure: The copula C_{XY}

Example

- Now one has the copula C_{XY}, how to construct a bivariate distribution function with other margins?
- Assume that our desired margin is standard Cauchy margin, which means

$$F(x) = \frac{1}{2} + \frac{\arctan x}{\pi} \quad \text{and} \quad G(y) = \frac{1}{2} + \frac{\arctan y}{\pi}$$

- The procedure to get a bivariate distribution function is
  1. Find the quasi functions from the marginal distribution functions
  2. Find the image of the circle through these quasi functions
  3. Replace the formula of quasi function to the formula of the copula

PART 4: The Algebraic Method.

Problem setting

- Let X and Y be continuous random variables with a joint distribution function $H(x, y)$.
- Considering the following quantity

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

where

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

Question

Question Which joint distribution function H that θ is a constant?

Constructing a copula

- From Sklar's theorem, assume that C_{XY} is the copula of two random variables X and Y, then

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

- For different values of θ, one can find different forms of C_θ

1. If $\theta = 1$,

$$C_\theta(u,v) = uv$$

2. If $\theta \neq 1$,

$$C_\theta(u,v) = \frac{[1+(\theta-1)(u+v)]}{2(\theta-1)} \pm \frac{\sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}{2(\theta-1)}$$

The signs

- In case of $\theta \neq 1$, one can find 2 different copulas with different signs in this expression.

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

and

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

Margins

- One has

$$C_\theta(u,0) = \frac{[1+(\theta-1)u] \pm [1+(\theta-1)u]}{2(\theta-1)}$$
$$C_\theta(u,1) = \frac{[\theta+(\theta-1)u] \pm [\theta+(\theta-1)u]}{2(\theta-1)}$$

- One can show that $\delta^2 C(u,v)/\delta u \delta v \geq 0$ and

$$C_\theta(u,v) = \int_0^u \int_0^v \frac{\delta^2 C(u,v)}{\delta u \delta v} dudv$$

hence $C_\theta(u,v)$ is an absolutely continuous copula.

If $\theta \to 0$ and $\theta \to 1$

- When $\theta = 0$, then

$$C_0(u, v) = \max(0, u + v - 1)$$

- When $\theta = 1$, then

$$C_1(u, v) = uv$$

These copulas are the Fréchet - Hoeffding bounds for an arbitrary copula.

Estimate θ from a data set

- Recall the definition of θ

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

- Our goal is to choose the optimal point (x_0, y_0) such that

$$\frac{H(x_0,y_0)[1 - F(x_0) - G(y_0) + H(x_0,y_0)]}{[F(x_0) - H(x_0,y_0)][G(y_0) - H(x_0,y_0)]} = constant$$

The above quantity can be compute with respect to (x_0, y_0) by analyzing the frequencies of points belonging to quadrant from the point (x_0, y_0) .

Estimation

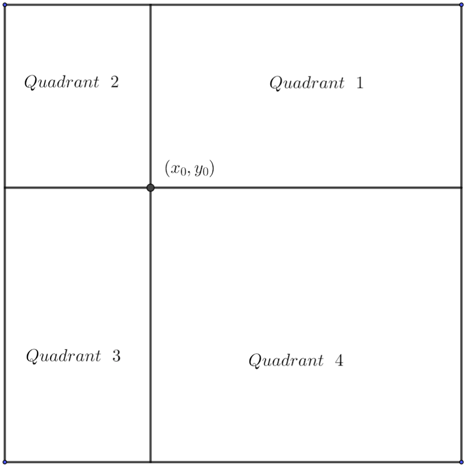

Figure: Quadrants

- One is interested in the ratio $\mathbb{P}[X > x]/\mathbb{P}[X \leq x]$, where X is a random variable representing the lifetime of an object.

- Analogously, one is interested in the ratio $(1 - H(x, y))/H(x, y)$

Two typical examples

- Let X, Y be two random variables with joint distribution function

$$H(x,y) = (1 + e^{-x} + e^{-y})^{-1}$$

then

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

- Let X, Y be two independent random variables, then

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

Question

Question: Which bivariate function $H(x, y)$ that satisfies the following equation

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

where $\theta \in (0, 1)$.

- Denote $C_\theta(u, v)$ by the copula of X and Y, then

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

then

$$C_\theta(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)}$$

One can check C_θ is a copula and it is absolutely continuous.

Transformation Method

- Let $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ be IID random variables with common joint distribution function H and marginal distribution function F and G.
- Denote $C_{(n)}, H_{(n)}$ by the copula and joint distribution function of $X_{(n)}$ and $Y_{(n)}$ where

$$X_{(n)} = \max\{Y_i\}, \qquad Y_{(n)} = \max\{Y_i\}$$

- Denote $F_{(n)}$ and $G_{(n)}$ by the margins of $X_{(n)}$ and $Y_{(n)}$. then

$$F_{(n)}(x) = [F(x)]^n, \qquad \text{and} \qquad G_{(n)}(y) = [G(y)]^n$$

Find $C_{(n)}$

- If C is the copula of $(X_i, Y_i), i = 1, \ldots, n$, then

$$H_{(n)}(x,y) = H^n(x,y) = C^n\left([F_{(n)}(x)]^{1/n}, [G_{(n)}(y)]^{1/n}\right)$$

or

$$C_{(n)}(u,v) = C^n(u^{1/n}, v^{1/n})$$

where $(u,v) \in \mathbb{I}^2$.

Image of a copula through a concave function

Theorem

Let $\gamma : [0, 1] \to [0, 1]$ be continuous and strictly increasing with $\gamma(0) = 0, \gamma(1)1 = 1$, and let γ^{-1} denote the inverse of γ. For an arbitrary copula C any $(u, v) \in \mathbb{I}^2$, define the function C_γ by

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

Then C_γ is a copula if and only if γ is concave(or γ^{-1} is convex).

γ^{-1} is convex $\Rightarrow C_\gamma$ is a copula

- Since $\gamma(0) = 0 = \gamma^{-1}(0)$ and $\gamma(1) = 1 = \gamma^{-1}(1)$, then for $(u, v) \in \mathbb{I}^2$,

$$C_\gamma(u, 0) = \gamma^{-1}(C(\gamma(u), 0)) = 0 = \gamma^{-1}(C(0, \gamma(v))) = C_\gamma(0, v)$$

$$C_\gamma(u, 1) = \gamma^{-1}(C(\gamma(u), 1)) = u$$

$$C_\gamma(1, v) = \gamma^{-1}(C(1, \gamma(v))) = v$$

- Let $K = [u_1, u_2] \times [v_1, v_2]$ be an arbitrary rectangle within in $\mathbb{I}^2$, denote by

$$a = C(\gamma(u_1), \gamma(v_1)), b = C(\gamma(u_1), \gamma(v_2)),$$

$$c = C(\gamma(u_2), \gamma(v_1)), d = C(\gamma(u_2, v_2), \gamma())$$

then,

$$\gamma^{-1}(b) - \gamma^{-1}(a) \leq \gamma^{-1}(d) - \gamma^{-1}(c)$$

or

$$V_{C_\gamma}(K) \geq 0$$

C_γ is a copula $\Rightarrow \gamma^{-1}$ is convex

For any a, d in $[0, 1]$ such that $a \leq d$, let

$$u_1 = v_1 = \gamma^{-1}((a+1)/2), \qquad u_2 = v_2 = \gamma^{-1}((d+1)/2)$$

Since C_γ is a copula, then the C_γ - measure of the rectangle $K = [u_1, v_1] \times [u_2, v_2]$ is non-negative. Hence

$$\gamma^{-1}(a) - 2\gamma^{-1}((a+d)/2) + \gamma^{-1}(d) \geq 0$$

which implies γ^{-1} is convex.