Conference 3: Multivariate distorted models and applications to quantile regression

Jorge Navarro[1]
Universidad de Murcia, Murcia, Spain.
E-mail: jorgenav@um.es.

[1]Supported by Ministerio de Ciencia e Innovación of Spain under grant PID2019-108079GB-C22/AEI/10.13039/501100011033.

References

The conference is based mainly on the following reference:

- Navarro J., Calì C., Longobardi, M., Durante F. (2021). Distortion Representations of Multivariate Distributions. *Submitted*.

Outline

Notation

- $(X_1,\dots,X_n)$ random vector over $(\Omega,\mathcal{S},\Pr)$.

Notation

- $(X_1, \dots, X_n)$ random vector over $(\Omega, \mathcal{S}, \Pr)$.
- Joint distribution function (DF)

$$\mathbf{F}(x_1, \dots, x_n) = \Pr(X_1 \leq x_1, \dots, X_n \leq x_n).$$

Notation

- $(X_1,\dots,X_n)$ random vector over $(\Omega,\mathcal{S},\Pr)$.
- Joint distribution function (DF)

$$\mathbf{F}(x_1,\dots,x_n)=\Pr(X_1\le x_1,\dots,X_n\le x_n).$$

- Copula representation

$$\mathbf{F}(x_1,\dots,x_n)=C(F_1(x_1),\dots,F_n(x_n)),$$

where $F_1,\dots,F_n$ are the mariginals.

Notation

- $(X_1,\dots,X_n)$ random vector over $(\Omega,\mathcal{S},\Pr)$.
- Joint distribution function (DF)

$$\mathbf{F}(x_1,\dots,x_n)=\Pr(X_1\le x_1,\dots,X_n\le x_n).$$

- Copula representation

$$\mathbf{F}(x_1,\dots,x_n)=C(F_1(x_1),\dots,F_n(x_n)),$$

where $F_1,\dots,F_n$ are the mariginals.
- A similar representation holds for the joint survival function

$$\overline{\mathbf{F}}(x_1,\dots,x_n)=\Pr(X_1> x_1,\dots,X_n> x_n).$$

Preceding multivariate distortions

- Several multivariate distortions have been proposed in the literature with the purpose of changing (shift) the distribution function of a given random vector $(X_1, \dots, X_n)$.

Preceding multivariate distortions

- Several multivariate distortions have been proposed in the literature with the purpose of changing (shift) the distribution function of a given random vector $(X_1, \dots, X_n)$.
- The distortion of the first kind proposed in Valdez and Xiao (2011) maintains the copula and distorts the marginals.

Preceding multivariate distortions

- Several multivariate distortions have been proposed in the literature with the purpose of changing (shift) the distribution function of a given random vector $(X_1, \ldots, X_n)$.
- The distortion of the first kind proposed in Valdez and Xiao (2011) maintains the copula and distorts the marginals.
- The distortion of the third kind proposed there maintains the marginals and replaces the copula by a distorted copula

Preceding multivariate distortions

- Other authors propose alternative representations (to the classical copula one) for a multivariate distribution $\mathbf{F}$.

Preceding multivariate distortions

- Other authors propose alternative representations (to the classical copula one) for a multivariate distribution **F**.
- Klüppelberg and Resnick (2008) proposed to use the Pareto-copula C_P to represent **F**.

Preceding multivariate distortions

- Other authors propose alternative representations (to the classical copula one) for a multivariate distribution **F**.
- Klüppelberg and Resnick (2008) proposed to use the Pareto-copula C_P to represent **F**.
- If G is a continuous univariate distribution function and C is a copula, we can define

$$C_G(v_1,\dots,v_n) = C(G(v_1),\dots,G(v_n)),\ (v_1,\dots,v_n)\in\mathbb{R}^n.$$

Preceding multivariate distortions

- Other authors propose alternative representations (to the classical copula one) for a multivariate distribution **F**.
- Klüppelberg and Resnick (2008) proposed to use the Pareto-copula C_P to represent **F**.
- If G is a continuous univariate distribution function and C is a copula, we can define

$$C_G(v_1,\dots,v_n)=C(G(v_1),\dots,G(v_n)),\ (v_1,\dots,v_n)\in\mathbb{R}^n.$$

- C_G is a copula only when $G(x)=x$ for $x\in[0,1]$.

Preceding multivariate distortions

- Other authors propose alternative representations (to the classical copula one) for a multivariate distribution $\mathbf{F}$.
- Klüppelberg and Resnick (2008) proposed to use the Pareto-copula C_P to represent $\mathbf{F}$.
- If G is a continuous univariate distribution function and C is a copula, we can define

$$C_G(v_1,\dots,v_n)=C(G(v_1),\dots,G(v_n)),\ (v_1,\dots,v_n)\in\mathbb{R}^n.$$

- C_G is a copula only when $G(x)=x$ for $x\in[0,1]$.
- C_G can be used to obtain the following representation

$$\mathbf{F}(x_1,\dots,x_n)=C_G(G^{-1}(F_1(x_1)),\dots,G^{-1}(F_n(x_n))).\quad (1.1)$$

Preceding multivariate distortions

- Other authors propose alternative representations (to the classical copula one) for a multivariate distribution $\mathbf{F}$.
- Klüppelberg and Resnick (2008) proposed to use the Pareto-copula C_P to represent $\mathbf{F}$.
- If G is a continuous univariate distribution function and C is a copula, we can define

$$C_G(v_1,\dots,v_n)=C(G(v_1),\dots,G(v_n)),\ (v_1,\dots,v_n)\in\mathbb{R}^n.$$

- C_G is a copula only when $G(x)=x$ for $x\in[0,1]$.
- C_G can be used to obtain the following representation

$$\mathbf{F}(x_1,\dots,x_n)=C_G(G^{-1}(F_1(x_1)),\dots,G^{-1}(F_n(x_n))). \quad (1.1)$$

- If G is the standard Pareto, then C_G is the Pareto copula proposed in Klüppelberg and Resnick (2008).

Definition

Definition (Navarro et al. (2021))

A multivariate distribution function **F** is said to be a *multivariate distorted distribution* (MDD) of the univariate distribution functions $G_1, \ldots, G_n$ if there exists a *distortion* function D such that

$$\mathbf{F}(x_1, \ldots, x_n) = D(G_1(x_1), \ldots, G_n(x_n)), \ \forall (x_1, \ldots, x_n) \in \mathbb{R}^n. \quad (1.2)$$

We write $\mathbf{F} \equiv MDD(G_1, \ldots, G_n)$, when **F** is a MDD of $G_1, \ldots, G_n$.

Definition

Definition (Navarro et al. (2021))

A continuous function $D : [0,1]^n \to [0,1]$ is called *(n–dimensional) distortion function* (shortly written as $D \in \mathcal{D}_n$) if:

(i) $D(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_n) = 0$ for all $u_1, \dots, u_n \in [0,1]$.

(ii) $D(1, \dots, 1) = 1$.

(iii) D is n–increasing, i.e. for all $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ with $x_i \le y_i$, it holds $\triangle_{\mathbf{x}}^{\mathbf{y}} D \ge 0$, where

$$\triangle_{(x_1,\dots,x_n)}^{(y_1,\dots,y_n)} D := \sum_{z_i \in \{x_i, y_i\}} (-1)^{\mathbf{1}(z_1,\dots,z_n)} D(z_1, \dots, z_n),$$

with $\mathbf{1}(z_1, \dots, z_n) = \sum_{i=1}^n 1(z_i = x_i)$ and $1(A) = 1$ (respectively, 0) if A is true (respectively, false).

Examples

- All the copula representations are also MDD representation.

Examples

- All the copula representations are also MDD representation.
- The copula distortion of the first kind proposed in Valdez and Xiao (2011) is

$$\mathbf{F}_{d_1,\ldots,d_n}(x_1,\ldots,x_n) := C(d_1(F_1(x_1)),\ldots,d_n(F_n(x_n))),$$

for given univariate distortion functions $d_1,\ldots,d_n$.

Examples

- All the copula representations are also MDD representation.
- The copula distortion of the first kind proposed in Valdez and Xiao (2011) is

$$\mathbf{F}_{d_1,\dots,d_n}(x_1,\dots,x_n) := C(d_1(F_1(x_1)),\dots,d_n(F_n(x_n))),$$

  for given univariate distortion functions $d_1,\dots,d_n$.
- It is a MDD with

$$D(u_1,\dots,u_n) := C(d_1(u_1),\dots,d_n(u_n))$$

  for all $(u_1,\dots,u_n) \in [0,1]^n$.

Main properties

- According to Sklar's theorem, any multivariate distribution function can be expressed in terms of its univariate marginal distributions via a copula representation.

Main properties

- According to Sklar's theorem, any multivariate distribution function can be expressed in terms of its univariate marginal distributions via a copula representation.
- If the marginals are continuous then this representation (copula) is unique.

Main properties

- According to Sklar's theorem, any multivariate distribution function can be expressed in terms of its univariate marginal distributions via a copula representation.
- If the marginals are continuous then this representation (copula) is unique.
- In the following result, we state a similar Sklar-type theorem for MDD under mild conditions.

Sklar-type theorem

Proposition

Let $(X_1,\dots,X_n)$ be a random vector with joint continuous distribution function $\mathbf{F}$. Let $G_1,\dots,G_n$ be arbitrary continuous distribution functions and let us assume that G_i is strictly increasing in the support of X_i for $i=1,\dots,n$. Then there exists a unique distortion $D\in\mathcal{D}_n$ such that

$$\mathbf{F}(x_1,\dots,x_n)=D(G_1(x_1),\dots,G_n(x_n))$$

holds for all $(x_1,\dots,x_n)\in\mathbb{R}^n$.

Sklar-type theorem

- From the proof, it follows that D is the multivariate distribution function of $(G_1(X_1), \ldots, G_n(X_n))$.

Sklar-type theorem

- From the proof, it follows that D is the multivariate distribution function of $(G_1(X_1), \ldots, G_n(X_n))$.
- $V_i = G_i(X_i)$ is a componentwise increasing transformation of X_i for $i = 1, \ldots, n$.

Sklar-type theorem

- From the proof, it follows that D is the multivariate distribution function of $(G_1(X_1), \dots, G_n(X_n))$.
- $V_i = G_i(X_i)$ is a componentwise increasing transformation of X_i for $i = 1, \dots, n$.
- Thus, for any measure of concordance κ (as Kendall's tau or Spearman's rho), $\kappa(V_1, \dots, V_n) = \kappa(X_1, \dots, X_n)$.

Sklar-type theorem

- From the proof, it follows that D is the multivariate distribution function of $(G_1(X_1), \dots, G_n(X_n))$.
- $V_i = G_i(X_i)$ is a componentwise increasing transformation of X_i for $i = 1, \dots, n$.
- Thus, for any measure of concordance κ (as Kendall's tau or Spearman's rho), $\kappa(V_1, \dots, V_n) = \kappa(X_1, \dots, X_n)$.
- In essence, D contains all the information about the (rank-invariant) dependence structure of $\mathbf{F}$.

Sklar-type theorem

- From the proof, it follows that D is the multivariate distribution function of $(G_1(X_1), \dots, G_n(X_n))$.
- $V_i = G_i(X_i)$ is a componentwise increasing transformation of X_i for $i = 1, \dots, n$.
- Thus, for any measure of concordance κ (as Kendall's tau or Spearman's rho), $\kappa(V_1, \dots, V_n) = \kappa(X_1, \dots, X_n)$.
- In essence, D contains all the information about the (rank-invariant) dependence structure of $\mathbf{F}$.
- Actually $\mathbf{F}$ and D share the same copula C.

Construction of new multivariate models

The converse of the preceding proposition can be stated as follows.

Proposition

If $D \in \mathcal{D}_n$, then the function defined by the right-hand side of (1.2) *is a multivariate distribution function for all univariate distribution functions $G_1, \ldots, G_n$.*

Relationship with C

Proposition

Let $(X_1,\dots,X_n)$ be a random vector with joint continuous distribution function $\mathbf{F}$*. Let $G_1,\dots,G_n$ be arbitrary continuous distribution functions. Suppose that $\mathbf{F}\equiv MDD(G_1,\dots,G_n)$ with distortion D. Then,*

$$D(u_1,\dots,u_n)=C(F_1(G_1^{-1}(u_1)),\dots,F_n(G_n^{-1}(u_n)))$$

for all $(u_1,\dots,u_n)\in[0,1]^n$, where G_i^{-1} is the quasi-inverse of G_i and F_i is the ith marginal of $\mathbf{F}$ *for $i=1,\dots,n$.*

Joint survival function.

Proposition

Let $(X_1, \ldots, X_n)$ be a random vector with distribution function $\mathbf{F}$. If (1.2) holds for $G_1, \ldots, G_n$ and $D \in \mathcal{D}_n$, then the joint survival function of $(X_1, \ldots, X_n)$ can be written as

$$\overline{\mathbf{F}}(x_1, \ldots, x_n) = \hat{D}(\bar{G}_1(x_1), \ldots, \bar{G}_n(x_n)) \quad (1.3)$$

for all $x_1, \ldots, x_n$, where $\bar{G}_i = 1 - G_i$ is the survival function associated to G_i for $i = 1, \ldots, n$ and $\hat{D} \in \mathcal{D}_n$.

Marginal distributions

- A relevant property of the MDD representation $\mathbf{F} \equiv MDD(G_1, \ldots, G_n)$ is that all the multivariate marginal distributions of $\mathbf{F}$ are also MDD from $G_1, \ldots, G_n$.

Marginal distributions

- A relevant property of the MDD representation $\mathbf{F} \equiv MDD(G_1, \dots, G_n)$ is that all the multivariate marginal distributions of $\mathbf{F}$ are also MDD from $G_1, \dots, G_n$.
- Let $F_{1,\dots,m}$ be the distribution function of $(X_1, \dots, X_m)$.

Marginal distributions

- A relevant property of the MDD representation $\mathbf{F} \equiv MDD(G_1, \ldots, G_n)$ is that all the multivariate marginal distributions of $\mathbf{F}$ are also MDD from $G_1, \ldots, G_n$.
- Let $F_{1,\ldots,m}$ be the distribution function of $(X_1, \ldots, X_m)$.

- Proposition

If $\mathbf{F} \equiv MDD(G_1, \ldots, G_n)$ *and* $1 \leq m \leq n$, *then*

$$F_{1,\ldots,m}(x_1, \ldots, x_m) = D_{1,\ldots,m}(G_1(x_1), \ldots, G_m(x_m)) \qquad (1.4)$$

for all $(x_1, \ldots, x_m) \in \mathbb{R}^m$, *where*

$$D_{1,\ldots,m}(u_1, \ldots, u_m) := D(u_1, \ldots, u_m, 1, \ldots, 1)$$

for all $(u_1, \ldots, u_m) \in [0, 1]^m$ *and* $D_{1,\ldots,m} \in \mathcal{D}_m$.

Univariate marginal distributions.

- In particular, the ith marginal distribution function of X_i can be written as

$$F_i(x_i) = D(1,\dots,1,G_i(x_i),1,\dots,1) = D_i(G_i(x_i)) \qquad (1.5)$$

for all $x_i \in \mathbb{R}$, where

$$D_i(u) := D(1,\dots,1,u,1,\dots,1)$$

and the value u is placed at the ith position.

Univariate marginal distributions.

- In particular, the ith marginal distribution function of X_i can be written as

$$F_i(x_i) = D(1, \dots, 1, G_i(x_i), 1, \dots, 1) = D_i(G_i(x_i)) \qquad (1.5)$$

for all $x_i \in \mathbb{R}$, where

$$D_i(u) := D(1, \dots, 1, u, 1, \dots, 1)$$

and the value u is placed at the ith position.

- Clearly, we have $G_i = F_i$ for a fixed $i \in \{1, \dots, n\}$ when $D_i(u) = u$ for all $u \in [0, 1]$.

Probability density function

Let us assume that **F** is absolutely continuous with joint probability density function (PDF) **f**, where

$$\mathbf{f}(x_1,\ldots,x_n)=\partial_{1,\ldots,n}\mathbf{F}(x_1,\ldots,x_n)\ (a.e.).$$

Proposition

If $\mathbf{F}\equiv MDD(G_1,\ldots,G_n)$ *for absolutely continuous distribution functions* $G_1,\ldots,G_n$ *with PDFs* $g_1,\ldots,g_n$*, respectively, and a distortion function* D *that admits continuous mixed derivatives of order* n*, then*

$$\mathbf{f}(x_1,\ldots,x_n)=g_1(x_1)\ldots g_n(x_n)\ \partial_{1,\ldots,n}D(G_1(x_1),\ldots,G_n(x_n)). \tag{1.6}$$

Conditional distributions

- All the conditional distributions of $\mathbf{F} \equiv MDD(G_1, \ldots, G_n)$ have MDD representations.

Conditional distributions

- All the conditional distributions of $\mathbf{F} \equiv MDD(G_1, \dots, G_n)$ have MDD representations.
- We just consider the DF $F_{2|1}$ of $(X_2|X_1 = x_1)$.

Conditional distributions

- All the conditional distributions of $\mathbf{F} \equiv MDD(G_1, \ldots, G_n)$ have MDD representations.
- We just consider the DF $F_{2|1}$ of $(X_2|X_1 = x_1)$.

- Proposition

Let (X_1, X_2) with $\mathbf{F} \equiv MDD(G_1, G_2)$ for a distortion function $D \in \mathcal{D}_2$ that admits continuous mixed derivatives of order 2, then

$$F_{2|1}(x_2|x_1) = D_{2|1}(G_2(x_2)|G_1(x_1)) \tag{1.7}$$

whenever $\lim_{v \to 0^+} \partial_1 D(G_1(x_1), v) = 0$, *where*

$$D_{2|1}(v|G_1(x_1)) = \frac{\partial_1 D(G_1(x_1), v)}{\partial_1 D(G_1(x_1), 1)}$$

for $0 < v < 1$ *and* x_1 *such that* $\partial_1 D(G_1(x_1), 1) > 0$.

Theoretical Quantile Regression

- The regression curve to predict X_2 from X_1 is

$$m_{2|1}(x_1) = E(X_2|X_1 = x_1)$$

Theoretical Quantile Regression

- The regression curve to predict X_2 from X_1 is

$$m_{2|1}(x_1) = E(X_2|X_1 = x_1)$$

- If $d_{2|1}(v|u) := D'_{2|1}(v|u)$, then

$$m_{2|1}(x_1) = \int_{-\infty}^{+\infty} x_2 g_2(x_2) d_{2|1}(G_2(x_2)|G_1(x_1)) dx_2.$$

Theoretical Quantile Regression

- The regression curve to predict X_2 from X_1 is

$$m_{2|1}(x_1) = E(X_2|X_1 = x_1)$$

- If $d_{2|1}(v|u) := D'_{2|1}(v|u)$, then

$$m_{2|1}(x_1) = \int_{-\infty}^{+\infty} x_2 g_2(x_2) d_{2|1}(G_2(x_2)|G_1(x_1))dx_2.$$

- Another option is the *conditional median regression curve*

$$\widetilde{m}_{2|1}(x_1) := F_{2|1}^{-1}(0.5|x_1)$$

(see Koenker (2005) or Nelsen (2006), p. 217).

Theoretical Quantile Regression

- The regression curve to predict X_2 from X_1 is

$$m_{2|1}(x_1) = E(X_2|X_1 = x_1)$$

- If $d_{2|1}(v|u) := D'_{2|1}(v|u)$, then

$$m_{2|1}(x_1) = \int_{-\infty}^{+\infty} x_2 g_2(x_2) d_{2|1}(G_2(x_2)|G_1(x_1)) dx_2.$$

- Another option is the *conditional median regression curve*

$$\widetilde{m}_{2|1}(x_1) := F_{2|1}^{-1}(0.5|x_1)$$

(see Koenker (2005) or Nelsen (2006), p. 217).

- This quantile function $F_{2|1}^{-1}$ can be computed from (1.7) as

$$F_{2|1}^{-1}(q|x_1) = G_2^{-1}(D_{2|1}^{-1}(q|G_1(x_1))),\ 0 < q < 1.$$

Confidence bands

- Moreover, we can obtain α-confidence bands in a similar way (see Koenker (2005)) with

$$\left[F_{2|1}^{-1}(\beta_1|x_1), F_{2|1}^{-1}(\beta_1|x_1)\right]$$

taking $0 \leq \beta_1 < \beta_2 \leq 1$ such that $\beta_2 - \beta_1 = \alpha$.

Confidence bands

- Moreover, we can obtain α-confidence bands in a similar way (see Koenker (2005)) with

$$\left[F_{2|1}^{-1}(\beta_1|x_1), F_{2|1}^{-1}(\beta_1|x_1)\right]$$

taking $0 \leq \beta_1 < \beta_2 \leq 1$ such that $\beta_2 - \beta_1 = \alpha$.

- For example, the centered 50% and 90% quantile-confidence bands for $(X_2|X_1 = x_1)$ are determined, respectively, by

$$\left[F_{2|1}^{-1}(0.25|x_1), F_{2|1}^{-1}(0.75|x_1)\right]$$

and

$$\left[F_{2|1}^{-1}(0.05|x_1), F_{2|1}^{-1}(0.95|x_1)\right].$$

Generating data from (X_1, X_2)

- This procedure can also be used the generate data from (X_1, X_2).

Generating data from (X_1, X_2)

- This procedure can also be used the generate data from (X_1, X_2).
- First we use the inverse transform method to generate a data from X_1 as
$$x_1 := F_1^{-1}(u),$$
where u is a random number in $(0, 1)$.

Generating data from (X_1, X_2)

- This procedure can also be used the generate data from (X_1, X_2).
- First we use the inverse transform method to generate a data from X_1 as

$$x_1 := F_1^{-1}(u),$$

where u is a random number in $(0, 1)$.
- Then we generate the associated data for X_2 as

$$x_2 := F_{2|1}^{-1}(v|x_1),$$

where v is another (independent) random number in $(0, 1)$.

Generating data from (X_1, X_2)

- This procedure can also be used the generate data from (X_1, X_2).
- First we use the inverse transform method to generate a data from X_1 as
$$x_1 := F_1^{-1}(u),$$
where u is a random number in $(0, 1)$.
- Then we generate the associated data for X_2 as
$$x_2 := F_{2|1}^{-1}(v|x_1),$$
where v is another (independent) random number in $(0, 1)$.
- By repeating n times this procedure we get a sample of size n from (X_1, X_2).

An example with a Clayton copula

- Let us assume that the DF of (X, Y) is

$$C(x,y) = \frac{xy}{x+y-xy}, \; x,y \in (0,1).$$

An example with a Clayton copula

- Let us assume that the DF of (X, Y) is

$$C(x,y) = \frac{xy}{x+y-xy}, \quad x,y \in (0,1).$$

- The marginals are

$$F_1(x) = F_2(x) = C(x,1) = x, \quad x \in (0,1).$$

An example with a Clayton copula

- Let us assume that the DF of (X, Y) is

$$C(x,y) = \frac{xy}{x+y-xy}, \ x, y \in (0,1).$$

- The marginals are

$$F_1(x) = F_2(x) = C(x,1) = x, \ x \in (0,1).$$

- The conditional distribution is

$$F_{2|1}(y|x) = \frac{\partial_1 C(x,y)}{\partial_1 C(x,1)} = \frac{y^2}{(x+y-xy)^2}, \ x, y \in (0,1).$$

An example with a Clayton copula

- Let us assume that the DF of (X, Y) is

$$C(x,y) = \frac{xy}{x+y-xy},\ \ x,y \in (0,1).$$

- The marginals are

$$F_1(x) = F_2(x) = C(x,1) = x,\ \ x \in (0,1).$$

- The conditional distribution is

$$F_{2|1}(y|x) = \frac{\partial_1 C(x,y)}{\partial_1 C(x,1)} = \frac{y^2}{(x+y-xy)^2},\ \ x,y \in (0,1).$$

- Then the quantile regression curve is

$$\widetilde{m}_{2|1}(x) = F_{2|1}^{-1}(0.5|x).$$

An example with a Clayton copula

- To get the inverse of $F_{2|1}(y|x)$ for $0 < q < 1$ we solve

$$\frac{y^2}{(x+y-xy)^2} = q$$

obtaining

$$x = (q^{-1/2} + x - 1)y.$$

An example with a Clayton copula

- To get the inverse of $F_{2|1}(y|x)$ for $0 < q < 1$ we solve

$$\frac{y^2}{(x+y-xy)^2} = q$$

obtaining

$$x = (q^{-1/2} + x - 1)y.$$

- Hence

$$F_{2|1}^{-1}(q|x) = \frac{x}{q^{-1/2} + x - 1}$$

and

$$\widetilde{m}_{2|1}(x) = \frac{x}{0.5^{-1/2} + x - 1}, x \in (0,1).$$

An example with a Clayton copula

- To get the inverse of $F_{2|1}(y|x)$ for $0<q<1$ we solve

$$\frac{y^2}{(x+y-xy)^2}=q$$

obtaining

$$x=(q^{-1/2}+x-1)y.$$

- Hence

$$F^{-1}_{2|1}(q|x)=\frac{x}{q^{-1/2}+x-1}$$

and

$$\widetilde{m}_{2|1}(x)=\frac{x}{0.5^{-1/2}+x-1},x\in(0,1).$$

- The confidence bands are obtained in a similar way.

Quantile regression curve and confidence bands

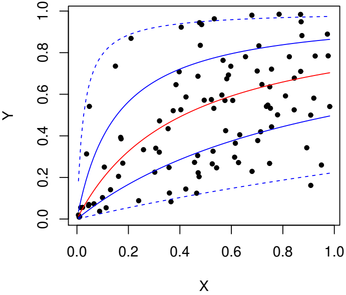

Figure: Quantile regression curve (red) and confidence bands (50% continuous blue, 90% dashed blue) for a Clayton copula jointly with 100 data from this model.

```
FI<-function(q,x) x/(q^(-1/2)+x-1)
m<-function(x) FI(0.5,x)
n<-100
x<-1:n
y<-1:n
set.seed(201)
for (i in 1:n){
  x[i]<-runif(1)
  y[i]<-FI(runif(1),x[i])
}
plot(x,y,xlab='X',ylab='Y',pch=20)
curve(m(x),add=T,col='red')
curve(FI(0.25,x),add=T,col='blue')
curve(FI(0.75,x),add=T,col='blue')
curve(FI(0.05,x),add=T,col='blue',lty=2)
curve(FI(0.95,x),add=T,col='blue',lty=2)
```

Predictions

- The first data is $(X_1, X_2) = (0.6125842, 0.2972452)$.

Predictions

- The first data is $(X_1, X_2) = (0.6125842, 0.2972452)$.
- If we want to predict X_2 from $X_1 = x$ we get

$$\widetilde{m}_{2|1}(0.6125842) = \frac{0.6125842}{0.5^{-1/2} + 0.6125842 - 1} = 0.5965967.$$

Predictions

- The first data is $(X_1, X_2) = (0.6125842, 0.2972452)$.
- If we want to predict X_2 from $X_1 = x$ we get

$$\widetilde{m}_{2|1}(0.6125842) = \frac{0.6125842}{0.5^{-1/2} + 0.6125842 - 1} = 0.5965967.$$

- The 90% confidence interval is $[0.1499697, 0.9593175]$.

Predictions

- The first data is $(X_1, X_2) = (0.6125842, 0.2972452)$.
- If we want to predict X_2 from $X_1 = x$ we get

$$\widetilde{m}_{2|1}(0.6125842) = \frac{0.6125842}{0.5^{-1/2} + 0.6125842 - 1} = 0.5965967.$$

- The 90% confidence interval is $[0.1499697, 0.9593175]$.
- The prediction is not good since the dispersion of this conditional variable is big.

Predictions

- The first data is $(X_1, X_2) = (0.6125842, 0.2972452)$.
- If we want to predict X_2 from $X_1 = x$ we get

$$\widetilde{m}_{2|1}(0.6125842) = \frac{0.6125842}{0.5^{-1/2} + 0.6125842 - 1} = 0.5965967.$$

- The 90% confidence interval is $[0.1499697, 0.9593175]$.
- The prediction is not good since the dispersion of this conditional variable is big.
- The 50% confidence band contains 59 and the 90%, 94.

Parametric estimation

- If we just assume a copula family with a parameter, we can use the sample to estimate this parameter.

Parametric estimation

- If we just assume a copula family with a parameter, we can use the sample to estimate this parameter.
- For example, for the Clayton family of copulas C_θ with a dependence parameter θ, we can use the Kendall tau to estimate θ (see Nelsen (2006)).

Parametric estimation

- If we just assume a copula family with a parameter, we can use the sample to estimate this parameter.
- For example, for the Clayton family of copulas C_θ with a dependence parameter θ, we can use the Kendall tau to estimate θ (see Nelsen (2006)).
- Then we use the estimation $\hat{\theta}$ and $C_{\hat{\theta}}$ to compute the quantile regression curves.

Non-parametric estimation

- Another option is to choose a parametric form for $\widetilde{m}$ and try to estimate it from the data.

Non-parametric estimation

- Another option is to choose a parametric form for $\widetilde{m}$ and try to estimate it from the data.
- For example, in the linear quantile regression (LQR) we assume that $\widetilde{m}_\theta(x) = \theta_0 + \theta_1 x$ with $\theta = (\theta_0, \theta_1)$.

Non-parametric estimation

- Another option is to choose a parametric form for $\widetilde{m}$ and try to estimate it from the data.
- For example, in the linear quantile regression (LQR) we assume that $\widetilde{m}_\theta(x) = \theta_0 + \theta_1 x$ with $\theta = (\theta_0, \theta_1)$.
- Then we will try to minimize

$$J^*(\theta) = \sum_{i=1}^{n} |m(X_i) - Y_i| = \sum_{i=1}^{n} |\theta_0 + \theta_1 X_i - Y_i|.$$

- This procedure was proposed by Koenker and Bassett (1978).

Non-parametric estimation

- Another option is to choose a parametric form for $\widetilde{m}$ and try to estimate it from the data.
- For example, in the linear quantile regression (LQR) we assume that $\widetilde{m}_\theta(x) = \theta_0 + \theta_1 x$ with $\theta = (\theta_0, \theta_1)$.
- Then we will try to minimize

$$J^*(\theta) = \sum_{i=1}^{n} |m(X_i) - Y_i| = \sum_{i=1}^{n} |\theta_0 + \theta_1 X_i - Y_i|.$$

- This procedure was proposed by Koenker and Bassett (1978).
- For the q-quantile line $m_q(x) = a_q + b_q x$ we minimize

$$J_q(a,b) = q \sum_{i:Y_i > a+bX_i} (Y_i - a - bX_i) + (1-q) \sum_{i:Y_i < a+bX_i} (a + bX_i - Y_i).$$

Non-parametric estimation

- Another option is to choose a parametric form for $\widetilde{m}$ and try to estimate it from the data.
- For example, in the linear quantile regression (LQR) we assume that $\widetilde{m}_\theta(x) = \theta_0 + \theta_1 x$ with $\theta = (\theta_0, \theta_1)$.
- Then we will try to minimize

$$J^*(\theta) = \sum_{i=1}^{n} |m(X_i) - Y_i| = \sum_{i=1}^{n} |\theta_0 + \theta_1 X_i - Y_i|.$$

- This procedure was proposed by Koenker and Bassett (1978).
- For the q-quantile line $m_q(x) = a_q + b_q x$ we minimize

$$J_q(a,b) = q \sum_{i:Y_i > a+bX_i} (Y_i - a - bX_i) + (1-q) \sum_{i:Y_i < a+bX_i} (a + bX_i - Y_i).$$

- The solutions can be obtained with the R package `quantreg`.

```
install.packages('quantreg')
rq(y~x)
plot(x,y,xlab='X',ylab='Y',pch=20)
abline(rq(y~x),col='red')
abline(rq(y~x,0.25),col='blue')
abline(rq(y~x,0.75),col='blue')
abline(rq(y~x,0.05),col='green')
abline(rq(y~x,0.95),col='green')
d<-data.frame(y,x,x^2,x^3)
rq(d)
```

Non-parametric estimation

- The quantile regression line is

$$\widehat{m}(x) = 0.09029761 + 0.71142861x$$

Non-parametric estimation

- The quantile regression line is

$$\widehat{m}(x) = 0.09029761 + 0.71142861x$$

- The prediction for X_2 for the first data with this QR line is

$$\widehat{m}(0.6125842) = 0.09029761 + 0.71142861 \cdot 0.6125842 = 0.5261075.$$

Non-parametric estimation

- The quantile regression line is

$$\widehat{m}(x) = 0.09029761 + 0.71142861x$$

- The prediction for X_2 for the first data with this QR line is

$$\widehat{m}(0.6125842) = 0.09029761 + 0.71142861 \cdot 0.6125842 = 0.5261075.$$

- The real data was $X_2 = 0.2972452$ and the prediction with the exact QR curve was $\widetilde{m}_{2|1}(0.6125842) = 0.5965967$.

Non-parametric estimation

- The quantile regression line is

$$\widehat{m}(x) = 0.09029761 + 0.71142861x$$

- The prediction for X_2 for the first data with this QR line is

$$\widehat{m}(0.6125842) = 0.09029761 + 0.71142861 \cdot 0.6125842 = 0.5261075.$$

- The real data was $X_2 = 0.2972452$ and the prediction with the exact QR curve was $\widetilde{m}_{2|1}(0.6125842) = 0.5965967$.
- With a polynomial of degree 3 we get $\widehat{X}_2 = 0.5727091$ with

$$\widehat{m}_3(x) = 0.00045114 + 1.54985173x - 1.01832625x^2 + 0.02166253x^3.$$

Quantile regression curve and confidence bands

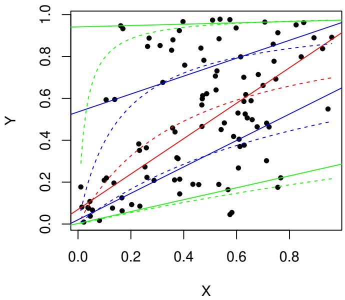

Figure: Estimated Quantile Regression line (red) and confidence bands (50% continuous blue, 90% continuous green) for the 100 data from a Clayton model. The dashed lines are the exact curves.

Residual lifetimes

- $X_1, \dots, X_n$ represent the lifetimes of n components.

Residual lifetimes

- $X_1, \ldots, X_n$ represent the lifetimes of n components.
- $(X_i - t | X_i > t)$ denotes the univariate residual lifetimes at time $t > 0$ with

$$\bar{F}_{i,t}(x) := \Pr(X_i - t > x | X_i > t) = \frac{\bar{F}_i(t + x)}{\bar{F}_i(t)}$$

for every $x \geq 0$, whenever $\bar{F}_i(t) > 0$.

Residual lifetimes

- $X_1, \ldots, X_n$ represent the lifetimes of n components.
- $(X_i - t | X_i > t)$ denotes the univariate residual lifetimes at time $t > 0$ with

$$\bar{F}_{i,t}(x) := \Pr(X_i - t > x | X_i > t) = \frac{\bar{F}_i(t + x)}{\bar{F}_i(t)}$$

  for every $x \geq 0$, whenever $\bar{F}_i(t) > 0$.
- The mean residual lifetime is $m_i(t) = E(X_i - t | X_i > t)$.

Residual lifetimes

- $X_1, \ldots, X_n$ represent the lifetimes of n components.
- $(X_i - t | X_i > t)$ denotes the univariate residual lifetimes at time $t > 0$ with

$$\bar{F}_{i,t}(x) := \Pr(X_i - t > x | X_i > t) = \frac{\bar{F}_i(t+x)}{\bar{F}_i(t)}$$

  for every $x \geq 0$, whenever $\bar{F}_i(t) > 0$.
- The mean residual lifetime is $m_i(t) = E(X_i - t | X_i > t)$.
- From $\mathbf{X} = (X_1, \ldots, X_n)$, we can consider

$$\mathbf{X}_t = (X_1 - t, \ldots, X_n - t | X_1 > t, \ldots, X_n > t)$$

  whose survival function for $x_1, \ldots, x_n \geq$ is

$$\bar{F}_t(x_1, \ldots, x_n) := \Pr(X_1 > x_1 + t, \ldots, X_n > x_n + t | X_1 > t, \ldots, X_n > t).$$

Proposition

If $\bar{F}(t,\dots,t)>0$ *for some* $t\geq 0$*, then*

$$\bar{F}_t(x_1,\dots,x_n)=\widehat{D}_t(\bar{F}_{1,t}(x_1),\dots,\bar{F}_{n,t}(x_n)) \tag{2.1}$$

for all $x_1,\dots,x_n\geq t$ *and distortion function*

$$\widehat{D}_t(u_1,\dots,u_n):=\frac{\widehat{C}(\bar{F}_1(t)u_1,\dots,\bar{F}_n(t)u_n)}{\widehat{C}(\bar{F}_1(t),\dots,\bar{F}_n(t))},\ \ u_1,\dots,u_n\in[0,1], \tag{2.2}$$

which depends on $\bar{F}_1(t),\dots,\bar{F}_n(t)$*.*

Residual lifetimes

- Note that $\bar{F}_{i,t}$ is not the ith marginal survival function of the random vector $\mathbf{X}_t$.

Residual lifetimes

- Note that $\bar{F}_{i,t}$ is not the ith marginal survival function of the random vector $\mathbf{X}_t$.
- The ith marginal survival function is

$$\bar{H}_{i,t}(x) = \Pr(X_i - t > x | X_1 > t, \ldots, X_n > t).$$

Residual lifetimes

- Note that $\bar{F}_{i,t}$ is not the ith marginal survival function of the random vector $\mathbf{X}_t$.
- The ith marginal survival function is

$$\bar{H}_{i,t}(x) = \Pr(X_i - t > x | X_1 > t, \ldots, X_n > t).$$

- Hence (2.1) is not a copula representation and $\widehat{D}_t$ is not always a copula.

Residual lifetimes

- Note that $\bar{F}_{i,t}$ is not the ith marginal survival function of the random vector $\mathbf{X}_t$.
- The ith marginal survival function is

$$\bar{H}_{i,t}(x) = \Pr(X_i - t > x | X_1 > t, \dots, X_n > t).$$

- Hence (2.1) is not a copula representation and $\widehat{D}_t$ is not always a copula.
- If $X_1, \dots, X_n$ are exponential, then $\bar{F}_{i,t} = \bar{F}_i \neq \bar{H}_{i,t}$.

Ordered paired data

- Let us assume that X and Y have a common absolutely continuous distribution function. Then

$$F_{X,Y}(x,y) = C(F(x), F(y)).$$

Ordered paired data

- Let us assume that X and Y have a common absolutely continuous distribution function. Then

$$F_{X,Y}(x,y) = C(F(x), F(y)).$$

- We may also assume that C is permutation symmetric, i.e., (X, Y) is exchangeable (EXC).

Ordered paired data

- Let us assume that X and Y have a common absolutely continuous distribution function. Then

$$F_{X,Y}(x,y) = C(F(x), F(y)).$$

- We may also assume that C is permutation symmetric, i.e., (X, Y) is exchangeable (EXC).
- We assume that $L = \min(X, Y)$ is known and that we want to predict $U = \max(X, Y)$.

Ordered paired data

- Let us assume that X and Y have a common absolutely continuous distribution function. Then

$$F_{X,Y}(x,y) = C(F(x), F(y)).$$

- We may also assume that C is permutation symmetric, i.e., (X, Y) is exchangeable (EXC).
- We assume that $L = \min(X, Y)$ is known and that we want to predict $U = \max(X, Y)$.
- To this purpose we need the conditional distribution function

$$G_{2|1}(x|t) := \Pr(U \leq x | L = t), \ x \geq t.$$

Ordered paired data

- Let us assume that X and Y have a common absolutely continuous distribution function. Then

$$F_{X,Y}(x,y) = C(F(x), F(y)).$$

- We may also assume that C is permutation symmetric, i.e., (X, Y) is exchangeable (EXC).
- We assume that $L = \min(X, Y)$ is known and that we want to predict $U = \max(X, Y)$.
- To this purpose we need the conditional distribution function

$$G_{2|1}(x|t) := \Pr(U \leq x | L = t), \ x \geq t.$$

- It can be used to compute the median regression curve and the confidence bands.

Ordered paired data

- We assume that we have a training sample $(X_1, Y_1), \dots, (X_m, Y_m)$ from (X, Y).

Ordered paired data

- We assume that we have a training sample $(X_1, Y_1), \dots, (X_m, Y_m)$ from (X, Y).
- However, for other individuals, we may just know $L = \min(X, Y)$ and we want to estimate $U = \max(X, Y)$.

Ordered paired data

- We assume that we have a training sample $(X_1, Y_1), \ldots, (X_m, Y_m)$ from (X, Y).
- However, for other individuals, we may just know $L = \min(X, Y)$ and we want to estimate $U = \max(X, Y)$.
- Note that both F and C can be estimated from the training sample by using parametric models or empirical or kernel type estimators.

Ordered paired data

- We assume that we have a training sample $(X_1, Y_1), \dots, (X_m, Y_m)$ from (X, Y).
- However, for other individuals, we may just know $L = \min(X, Y)$ and we want to estimate $U = \max(X, Y)$.
- Note that both F and C can be estimated from the training sample by using parametric models or empirical or kernel type estimators.
- We want to obtain a MDD representation for the random vector (L, U) in terms of F and C.

Ordered paired data

- Its joint distribution function $\mathbf{G}(x,y)=\Pr(L\le x,U\le y)$ is

$$\mathbf{G}(x,y)=\Pr(U\le y)=\Pr(X\le y,Y\le y)=C(F(y),F(y))$$

for $y\le x$, while for $x<y$ it is

$$\mathbf{G}(x,y)=\Pr(L\le x,U\le y)=\Pr((\{X\le x\}\cup\{Y\le x\})\cap\{X\le y\}\cap\{Y\le y\})$$

Ordered paired data

- Its joint distribution function $\mathbf{G}(x,y)=\Pr(L\le x, U\le y)$ is

$$\mathbf{G}(x,y)=\Pr(U\le y)=\Pr(X\le y, Y\le y)=C(F(y),F(y))$$

for $y\le x$, while for $x<y$ it is

$$\mathbf{G}(x,y)=\Pr(L\le x, U\le y)=\Pr((\{X\le x\}\cup\{Y\le x\})\cap\{X\le y\}\cap\{Y\le y\})$$

- Hence, by using the inclusion-exclusion formula, we get

$$\begin{aligned}\mathbf{G}(x,y)&=\Pr(X\le x, Y\le y)+\Pr(X\le y, Y\le x)-\Pr(X\le x, Y\le x)\\&=C(F(x),F(y))+C(F(y),F(x))-C(F(x),F(x))\end{aligned}$$

for $x<y$.

Ordered paired data

- Therefore, $\mathbf{G} \equiv MDD(F, F)$, i.e.

$$\mathbf{G}(x, y) = D(F(x), F(y)) \tag{2.3}$$

with the following distortion function

$$D(u, v) = \begin{cases} C(v, v) & \text{for} \quad v \leq u; \\ C(u, v) + C(v, u) - C(u, u) & \text{for} \quad u < v. \end{cases} \tag{2.4}$$

Ordered paired data

- Therefore, $\mathbf{G} \equiv MDD(F,F)$, i.e.

$$\mathbf{G}(x,y) = D(F(x),F(y)) \quad (2.3)$$

with the following distortion function

$$D(u,v) = \begin{cases} C(v,v) & \text{for} \quad v \le u; \\ C(u,v) + C(v,u) - C(u,u) & \text{for} \quad u < v. \end{cases} \quad (2.4)$$

- Then the marginal distributions of (L,U) can be written as

$$G_1(x) := \Pr(L \le x) = D(F(x),1) = D_1(F(x)),$$

$$G_2(y) := \Pr(U \le y) = D(1,F(y)) = D_2(F(y)),$$

where $D_1(u) = D(u,1) = 2u - C(u,u)$ and
$D_2(v) = D(1,v) = C(v,v)$ for all $u,v \in [0,1]$.

Ordered paired data

- For example, if X and Y are independent, then $D_1(u) = D(u, 1) = 2u - u^2 \neq u$ and $D_2(u) = D(1, u) = u^2 \neq u$ for all $u \in (0, 1)$.

Ordered paired data

- For example, if X and Y are independent, then $D_1(u) = D(u,1) = 2u - u^2 \neq u$ and $D_2(u) = D(1,u) = u^2 \neq u$ for all $u \in (0,1)$.
- Note that

$$D(u,v) = \begin{cases} v^2 & \text{for} \quad v \leq u; \\ 2uv - u^2 & \text{for} \quad u < v. \end{cases} \tag{2.5}$$

is not a copula and that the marginals G_1 and G_2 of $\mathbf{G}$ do not appear in (2.3).

Ordered paired data

- From (1.7) and (2.3), the distribution function of $(U|L=x)$ is

$$G_{2|1}(y|x) = D_{2|1}(F(y)|F(x)) \tag{2.6}$$

for $y \geq x$, where

$$D_{2|1}(v|F(x)) := \frac{\partial_1 D(F(x), v)}{\partial_1 D(F(x), 1)},$$

$$\partial_1 D(u,v) = \partial_1 C(u,v) + \partial_2 C(v,u) - \partial_1 C(u,u) - \partial_2 C(u,u),$$

and $v > u$.

Ordered paired data

- From (1.7) and (2.3), the distribution function of $(U|L=x)$ is
$$G_{2|1}(y|x) = D_{2|1}(F(y)|F(x)) \tag{2.6}$$
for $y \geq x$, where
$$D_{2|1}(v|F(x)) := \frac{\partial_1 D(F(x), v)}{\partial_1 D(F(x), 1)},$$
$$\partial_1 D(u,v) = \partial_1 C(u,v) + \partial_2 C(v,u) - \partial_1 C(u,u) - \partial_2 C(u,u),$$
and $v > u$.
- In the EXC case, we have
$$\partial_1 D(u,v) = 2\partial_1 C(u,v) - 2\partial_1 C(u,u), \ \ u \leq v \leq 1.$$

Ordered paired data

- From (1.7) and (2.3), the distribution function of $(U|L = x)$ is

$$G_{2|1}(y|x) = D_{2|1}(F(y)|F(x)) \tag{2.6}$$

for $y \geq x$, where

$$D_{2|1}(v|F(x)) := \frac{\partial_1 D(F(x), v)}{\partial_1 D(F(x), 1)},$$

$$\partial_1 D(u, v) = \partial_1 C(u, v) + \partial_2 C(v, u) - \partial_1 C(u, u) - \partial_2 C(u, u),$$

and $v > u$.

- In the EXC case, we have

$$\partial_1 D(u, v) = 2\partial_1 C(u, v) - 2\partial_1 C(u, u), \; u \leq v \leq 1.$$

- In the IID case, we get $\partial_1 D(u, v) = 2(v - u)$ for $u \leq v \leq 1$ and $D_{2|1}(v|F(x)) = (v - F(x))/\bar{F}(x)$ for $F(x) \leq v \leq 1$.

Coherent systems

Theorem

If T_1 and T_2 are two coherent systems with ID$\sim F$ components $(X_1,\dots,X_n)$, then its joint distribution is MDD(F,F).

Coherent systems

Theorem

If T_1 and T_2 are two coherent systems with ID$\sim F$ components $(X_1,\dots,X_n)$, then its joint distribution is MDD(F,F).

- In particular, it can be applied to the k-out-of-n systems (order statistics).

Coherent systems

Theorem

If T_1 and T_2 are two coherent systems with ID$\sim F$ components $(X_1, \dots, X_n)$, then its joint distribution is MDD(F,F).

- In particular, it can be applied to the k-out-of-n systems (order statistics).
- In a more particular case, for $X_{1:2}$ and $X_{2:2}$ we obtain the distortion D of the preceding subsection.

Coherent systems

Theorem

If T_1 and T_2 are two coherent systems with ID$\sim F$ components $(X_1, \dots, X_n)$, then its joint distribution is MDD(F,F).

- In particular, it can be applied to the k-out-of-n systems (order statistics).
- In a more particular case, for $X_{1:2}$ and $X_{2:2}$ we obtain the distortion D of the preceding subsection.
- Other examples: Sequential order statistics, record values, ...

Exact QR curves for paired ordered data. IID case.

- Let (X_i, Y_i) be a sample from (X, Y) where X, Y are IID$\sim F$.

Exact QR curves for paired ordered data. IID case.

- Let (X_i, Y_i) be a sample from (X, Y) where X, Y are IID$\sim F$.
- Let $L_i = \min(X_i, Y_i)$ and $U_i = \max(X_i, Y_i)$.

Exact QR curves for paired ordered data. IID case.

- Let (X_i, Y_i) be a sample from (X, Y) where X, Y are IID$\sim F$.
- Let $L_i = \min(X_i, Y_i)$ and $U_i = \max(X_i, Y_i)$.
- Note that L_i and U_i are dependent.

Exact QR curves for paired ordered data. IID case.

- Let (X_i, Y_i) be a sample from (X, Y) where X, Y are IID$\sim F$.
- Let $L_i = \min(X_i, Y_i)$ and $U_i = \max(X_i, Y_i)$.
- Note that L_i and U_i are dependent.
- From (2.6), the distribution function of $(U|L = x)$ is

$$G_{2|1}(y|x) = D_{2|1}(F(y)|F(x)) \tag{3.1}$$

for $y \geq x$, where

$$D_{2|1}(v|F(x)) = \frac{v - F(x)}{\bar{F}(x)}$$

for $F(x) \leq v \leq 1$.

Paired ordered data. IID case.

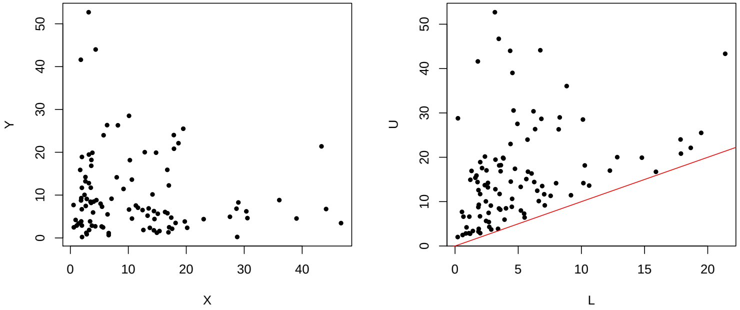

Figure: Independent data from two exponential distributions with mean $\mu = 10$ (left) and the associated paired ordered data (right).

```
# Code
n<-100
set.seed(202)
mu<-10
x<-rexp(n,1/mu)
y<-rexp(n,1/mu)
plot(x,y,pch=20)
L<-pmin(x,y)
U<-pmax(x,y)
plot(L,U,pch=20)
abline(0,1,col='red')
```

Exact QR curves for paired ordered data. IID case.

- The quantile function $F_{2|1}^{-1}$ can be computed as

$$F_{2|1}^{-1}(q|x) = F^{-1}(D_{2|1}^{-1}(q|F(x)))$$

for $0 < v < 1$, where $D_{2|1}^{-1}(q|F(x)) = F(x) + q\bar{F}(x)$ and $F^{-1}(y) = -\mu \log(1-y)$. Then

$$F_{2|1}^{-1}(q|x) = -\mu \log\left((1-q)e^{-x/\mu}\right) = x - \mu \log(1-q).$$

Exact QR curves for paired ordered data. IID case.

- The quantile function $F_{2|1}^{-1}$ can be computed as

$$F_{2|1}^{-1}(q|x) = F^{-1}(D_{2|1}^{-1}(q|F(x)))$$

for $0 < v < 1$, where $D_{2|1}^{-1}(q|F(x)) = F(x) + q\bar{F}(x)$ and $F^{-1}(y) = -\mu \log(1-y)$. Then

$$F_{2|1}^{-1}(q|x) = -\mu \log\left((1-q)e^{-x/\mu}\right) = x - \mu \log(1-q).$$

- Therefore, the exact QR curve is

$$m(x) = x - \mu \log(0.5).$$

Exact QR confidence bands for paired ordered data

- Analogously, the exact QR centered 90% confidence band is

$$[x - \mu \log(0.05), x - \mu \log(0.95)] .$$

Exact QR confidence bands for paired ordered data

- Analogously, the exact QR centered 90% confidence band is

$$[x-\mu\log(0.05), x-\mu\log(0.95)].$$

- The 50% centered confidence band is obtained in a similar way.

Exact QR confidence bands for paired ordered data

- Analogously, the exact QR centered 90% confidence band is

$$[x - \mu \log(0.05), x - \mu \log(0.95)] \, .$$

- The 50% centered confidence band is obtained in a similar way.
- The exact QR lower 90% confidence band is

$$[x, x - \mu \log(0.90)] \, .$$

QR for paired ordered data. IID case.

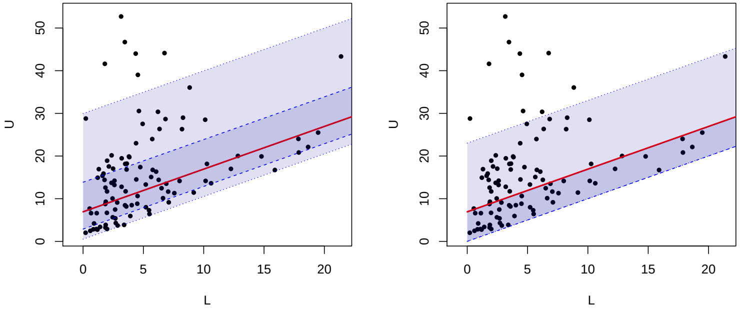

Figure: QR for the paired ordered data (L, U) associated to independent data (X, Y) from two exponential distributions with mean $\mu = 10$ jointly with 50% and 90% centered (left) or bottom (right) confidence bands.

Predictions

- The first ordered pair in our sample is $L_1 = 10.15771$ and $U_1 = 14.17195$.

Predictions

- The first ordered pair in our sample is $L_1 = 10.15771$ and $U_1 = 14.17195$.
- The prediction for U_1 from L_1 is

$$m(L_1) = m(10.15771) = 10.15771 - \mu \log(0.5) = 17.08918.$$

Predictions

- The first ordered pair in our sample is $L_1 = 10.15771$ and $U_1 = 14.17195$.
- The prediction for U_1 from L_1 is

$$m(L_1) = m(10.15771) = 10.15771 - \mu \log(0.5) = 17.08918.$$

- The centered 90% confidence interval for this prediction is $[10.67064, 40.11503]$.

Predictions

- The first ordered pair in our sample is $L_1 = 10.15771$ and $U_1 = 14.17195$.
- The prediction for U_1 from L_1 is

$$m(L_1) = m(10.15771) = 10.15771 - \mu \log(0.5) = 17.08918.$$

- The centered 90% confidence interval for this prediction is $[10.67064, 40.11503]$.
- The centered 50% confidence interval for this prediction is $[13.03453, 24.02065]$.

Dependent EXC data

- Let us consider now that (X, Y) are DID with a copula C and a common marginal distribution F.

Dependent EXC data

- Let us consider now that (X, Y) are DID with a copula C and a common marginal distribution F.
- We consider again the exponential model

$$\bar{F}(t) = \exp(-t/\mu), \ \ t \geq 0$$

and the Clayton EXC copula

$$C(u, v) = \frac{uv}{u + v - uv}, \ \ (u, v) \in [0, 1]^2. \qquad (3.2)$$

Dependent EXC data

- To get the QR curves we need the distribution $G_{2|1}(y|x)$ of $(U|L=x)$. From (2.6) we need

$$\partial_1 D(u,v) = 2\partial_1 C(u,v) - 2\partial_1 C(u,u) = \frac{2v^2}{(u+v-uv)^2} - \frac{2}{(2-u)^2}$$

and

$$\partial_1 D(u,1) = \frac{2}{(u+1-u)^2} - \frac{2}{(2-u)^2} = 2 - \frac{2}{(2-u)^2}.$$

Dependent EXC data

- To get the QR curves we need the distribution $G_{2|1}(y|x)$ of $(U|L = x)$. From (2.6) we need

$$\partial_1 D(u,v) = 2\partial_1 C(u,v) - 2\partial_1 C(u,u) = \frac{2v^2}{(u+v-uv)^2} - \frac{2}{(2-u)^2}$$

and

$$\partial_1 D(u,1) = \frac{2}{(u+1-u)^2} - \frac{2}{(2-u)^2} = 2 - \frac{2}{(2-u)^2}.$$

- Hence, for $v \geq u$, we get

$$D_{2|1}(v|u) = \frac{\frac{v^2}{(u+v-uv)^2} - \frac{1}{(2-u)^2}}{1 - \frac{1}{(2-u)^2}}$$

Dependent EXC data

- To get the QR curves we need the distribution $G_{2|1}(y|x)$ of $(U|L=x)$. From (2.6) we need

$$\partial_1 D(u,v) = 2\partial_1 C(u,v) - 2\partial_1 C(u,u) = \frac{2v^2}{(u+v-uv)^2} - \frac{2}{(2-u)^2}$$

and

$$\partial_1 D(u,1) = \frac{2}{(u+1-u)^2} - \frac{2}{(2-u)^2} = 2 - \frac{2}{(2-u)^2}.$$

- Hence, for $v \geq u$, we get

$$D_{2|1}(v|u) = \frac{\frac{v^2}{(u+v-uv)^2} - \frac{1}{(2-u)^2}}{1 - \frac{1}{(2-u)^2}}$$

- To compute the inverse, we need to solve in y the equation $G_{2|1}(y|x) = q$ for $q \in (0,1)$.

Dependent EXC data

- This leads to

$$\frac{F^2(y)}{(F(x)+F(y)-F(x)F(y))^2}=\frac{1-q+q(2-F(x))^2}{(2-F(x))^2}$$

or

$$\frac{(F(x)+F(y)-F(x)F(y))^2}{F^2(y)}=\frac{(2-F(x))^2}{1-q+q(2-F(x))^2}.$$

Dependent EXC data

- This leads to

$$\frac{F^2(y)}{(F(x)+F(y)-F(x)F(y))^2}=\frac{1-q+q(2-F(x))^2}{(2-F(x))^2}$$

or

$$\frac{(F(x)+F(y)-F(x)F(y))^2}{F^2(y)}=\frac{(2-F(x))^2}{1-q+q(2-F(x))^2}.$$

- Therefore

$$G_{2|1}^{-1}(q|x)=y=F^{-1}\left(\frac{F(x)}{F(x)-1+\frac{2-F(x)}{\sqrt{1-q+q(2-F(x))^2}}}\right).$$

Dependent EXC data

- Therefore, the exact median regression curve to predict U from $L = x$ is
$$m(x) = G_{2|1}^{-1}(0.5|x).$$

Dependent EXC data

- Therefore, the exact median regression curve to predict U from $L = x$ is
$$m(x) = G_{2|1}^{-1}(0.5|x).$$
- Analogously, the 90% centered confidence band is
$$\left[G_{2|1}^{-1}(0.05|x), G_{2|1}^{-1}(0.95|x)\right].$$

Dependent EXC data

- Therefore, the exact median regression curve to predict U from $L = x$ is
$$m(x) = G_{2|1}^{-1}(0.5|x).$$
- Analogously, the 90% centered confidence band is
$$\left[G_{2|1}^{-1}(0.05|x), G_{2|1}^{-1}(0.95|x)\right].$$
- The other confidence bands can be obtained in a similar way.

Dependent EXC data

- Therefore, the exact median regression curve to predict U from $L = x$ is
$$m(x) = G_{2|1}^{-1}(0.5|x).$$
- Analogously, the 90% centered confidence band is
$$\left[G_{2|1}^{-1}(0.05|x), G_{2|1}^{-1}(0.95|x)\right].$$
- The other confidence bands can be obtained in a similar way.
- For an exponential distribution with mean $\mu = 10$ we get the following curves.

QR for paired ordered data. ID case.

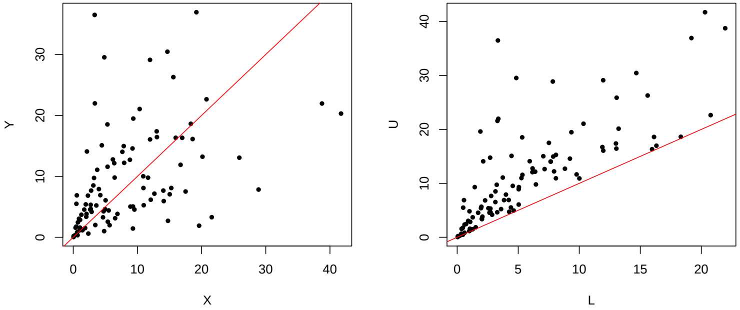

Figure: Paired ordered data (L, U) associated to dependent data (X, Y) from two exponential distributions and a Clayton copula.

QR for paired ordered data. ID case.

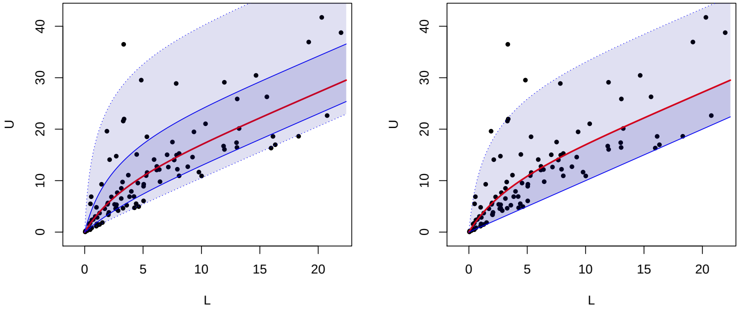

Figure: QR curves for paired ordered data (L, U) associated to dependent data (X, Y) from two exponential distributions with centered (left) and bottom (right) confidence bands.

Parametric QR curves

- Our model can contain some unknown parameters.

Parametric QR curves

- Our model can contain some unknown parameters.
- They can be both in F or in C.

Parametric QR curves

- Our model can contain some unknown parameters.
- They can be both in F or in C.
- In the last case we need the training sample (X_i, Y_i) from (X, Y) to estimate the copula parameter.

Parametric QR curves. IID case

- If (X, Y) are IID with $F(t) = 1 - \exp(-t/\mu)$, then μ can be estimated as

$$\hat{\mu} = \bar{X} = \frac{X_1 + \cdots + X_n}{n}$$

or as

$$\hat{\mu} = \bar{Y} = \frac{Y_1 + \cdots + Y_n}{n}.$$

Parametric QR curves. IID case

- If (X, Y) are IID with $F(t) = 1 - \exp(-t/\mu)$, then μ can be estimated as
$$\hat{\mu} = \bar{X} = \frac{X_1 + \cdots + X_n}{n}$$
or as
$$\hat{\mu} = \bar{Y} = \frac{Y_1 + \cdots + Y_n}{n}.$$
- As $L = \min(X, Y) \sim Exp(\mu/2)$ it can also be estimated as
$$\hat{\mu} = 2\bar{L} = 2\frac{L_1 + \cdots + L_n}{n}.$$

Parametric QR curves. IID case

- If (X, Y) are IID with $F(t) = 1 - \exp(-t/\mu)$, then μ can be estimated as
$$\hat{\mu} = \bar{X} = \frac{X_1 + \cdots + X_n}{n}$$
or as
$$\hat{\mu} = \bar{Y} = \frac{Y_1 + \cdots + Y_n}{n}.$$
- As $L = \min(X, Y) \sim Exp(\mu/2)$ it can also be estimated as
$$\hat{\mu} = 2\bar{L} = 2\frac{L_1 + \cdots + L_n}{n}.$$
- In our sample we get $\bar{X} = 11.3661$, $\bar{Y} = 9.956799$ and $2\bar{L} = 10.32929$.

Parametric QR for paired ordered data IID case

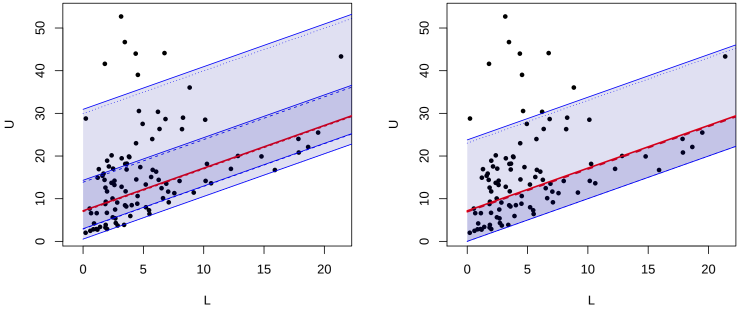

Figure: Parametric QR curves for (L, U) associated to IID data (X, Y) from an exponential distribution jointly with centered (left) and bottom (right) confidence bands. The dashed lines are the exact curves.

Parametric QR curves. Clayton copula

- If (X, Y) are ID with $F(t) = 1 - \exp(-t/\mu)$, with $\mu = 10$, then μ can be estimated as

$$\hat{\mu} = \frac{\bar{X} + \bar{Y}}{2} = \frac{8.298329 + 9.229868}{2} = 8.764098.$$

Parametric QR curves. Clayton copula

- If (X, Y) are ID with $F(t) = 1 - \exp(-t/\mu)$, with $\mu = 10$, then μ can be estimated as

$$\hat{\mu} = \frac{\bar{X} + \bar{Y}}{2} = \frac{8.298329 + 9.229868}{2} = 8.764098.$$

- If (X, Y) has a Clayton copula with an unknown parameter $\theta \geq 0$

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

Parametric QR curves. Clayton copula

- If (X,Y) are ID with $F(t)=1-\exp(-t/\mu)$, with $\mu=10$, then μ can be estimated as

$$\hat{\mu}=\frac{\bar{X}+\bar{Y}}{2}=\frac{8.298329+9.229868}{2}=8.764098.$$

- If (X,Y) has a Clayton copula with an unknown parameter $\theta\geq 0$

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

- Then its Kendall's tau coefficient is (Nelsen 2006, p. 163)

$$\tau=\frac{\theta}{\theta+2}.$$

Parametric QR curves. Clayton copula

- Then θ can be obtained from τ as

$$\theta = \frac{2\tau}{1-\tau}.$$

Parametric QR curves. Clayton copula

- Then θ can be obtained from τ as

$$\theta = \frac{2\tau}{1-\tau}.$$

- We can use `library('Kendall')` and `Kendall(X,Y)` to estimate τ from (X_i, Y_i).

Parametric QR curves. Clayton copula

- Then θ can be obtained from τ as

$$\theta = \frac{2\tau}{1-\tau}.$$

- We can use `library('Kendall')` and `Kendall(X,Y)` to estimate τ from (X_i, Y_i).
- In our sample from $\theta = 1$ we get $\hat{\tau} = 0.445$ and

$$\hat{\theta} = \frac{2\hat{\tau}}{1-\hat{\tau}} = \frac{2 \cdot 0.445}{1-0.445} = 1.603604.$$

- By replacing these estimations in F and C we obtain the following QR curves.

Parametric QR for paired ordered data ID case

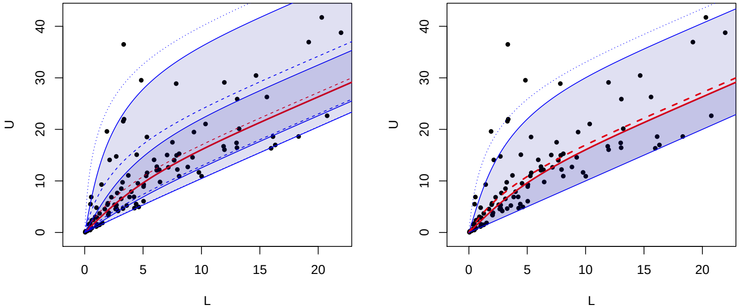

Figure: Parametric QR curves for (L, U) associated to data (X, Y) from an exponential distribution with unknown mean μ and a Clayton copula with unknown parameter θ. The dashed lines are the exact curves.

Non-parametric QR curves.

- If we do not have a parametric model, we can use the non-parametric models mentioned above.

Non-parametric QR curves.

- If we do not have a parametric model, we can use the non-parametric models mentioned above.
- We can use `library('quantreg')` and `rq(d)` where d<-data.frame(y,x,x^ 2,...) to estimate the exact curves.

Non-parametric QR curves. IID case.

- For the sample from two IID exponential distributions we obtain

Non-parametric QR curves. IID case.

- For the sample from two IID exponential distributions we obtain
- With `d<-data.frame(U,L)`:

$$\widehat{m}(x) = 8.9336392 + 0.8448752x.$$

Non-parametric QR curves. IID case.

- For the sample from two IID exponential distributions we obtain
- With `d<-data.frame(U,L)`:

$$\widehat{m}(x) = 8.9336392 + 0.8448752x.$$

- With `d<-data.frame(U,L,L^2)`:

$$\widehat{m}(x) = 6.90953311 + 1.43820751x - 0.03337719x^2.$$

Non-parametric QR curves. IID case.

- For the sample from two IID exponential distributions we obtain
- With `d<-data.frame(U,L)`:

$$\widehat{m}(x) = 8.9336392 + 0.8448752x.$$

- With `d<-data.frame(U,L,L^2)`:

$$\widehat{m}(x) = 6.90953311 + 1.43820751x - 0.03337719x^2.$$

- The exact curve (line) is

$$m(x) = -\mu \log(0.5) + x = 6.931472 + x.$$

Non-parametric QR for paired ordered data, IID case

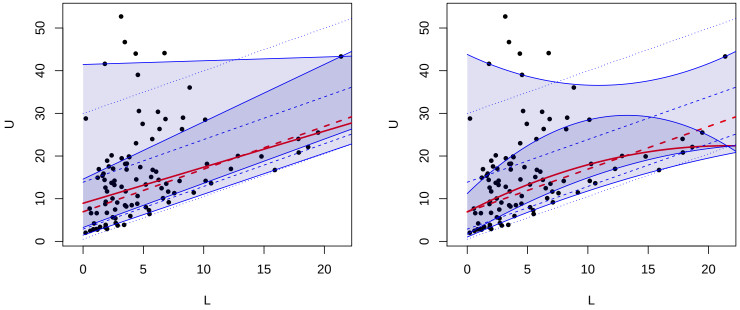

Figure: Non-parametric QR curves for paired ordered data (L, U) associated to IID data (X, Y) from an exponential distribution with $\mu = 10$ and $k = 1$ (left) or $k = 2$ (right).

Non-parametric QR curves. ID case.

- For the sample from the Clayton copula we obtain

Non-parametric QR curves. ID case.

- For the sample from the Clayton copula we obtain
- With d<-data.frame(U,L):

$$m(x) = 1.290607 + 1.601958x.$$

Non-parametric QR curves. ID case.

- For the sample from the Clayton copula we obtain
- With d<-data.frame(U,L):

$$m(x) = 1.290607 + 1.601958x.$$

- With d<-data.frame(U,L,Lˆ 2):

$$m(x) = 0.66257664 + 2.10657913x - 0.04794258x^2.$$

Non-parametric QR for paired ordered data, ID case

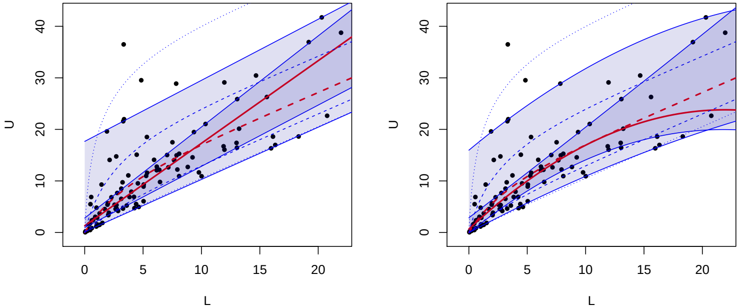

Figure: Non-parametric QR curves for (L, U) associated to data (X, Y) from an exponential distribution and a Clayton copula with $\theta = 1$ and $k = 1$ (left) or $k = 2$ (right). The dashed lines are the exact curves.

Conclusions

- It is not easy to manage (predict) paired ordered data.

Conclusions

- It is not easy to manage (predict) paired ordered data.
- Multivariate distortions and QR techniques may help in this difficult task.

Conclusions

- It is not easy to manage (predict) paired ordered data.
- Multivariate distortions and QR techniques may help in this difficult task.
- In general, the (correct) parametric models provide better predictions.

Conclusions

- It is not easy to manage (predict) paired ordered data.
- Multivariate distortions and QR techniques may help in this difficult task.
- In general, the (correct) parametric models provide better predictions.
- But it is difficult to know when we have a "correct" parametric model.

Conclusions

- It is not easy to manage (predict) paired ordered data.
- Multivariate distortions and QR techniques may help in this difficult task.
- In general, the (correct) parametric models provide better predictions.
- But it is difficult to know when we have a "correct" parametric model.
- Fit tests should be developed to check (confirm) these models.

Conclusions

- It is not easy to manage (predict) paired ordered data.
- Multivariate distortions and QR techniques may help in this difficult task.
- In general, the (correct) parametric models provide better predictions.
- But it is difficult to know when we have a "correct" parametric model.
- Fit tests should be developed to check (confirm) these models.
- This approach can be applied to other relevant models (order statistics, systems, records,...).

References

Klüppelberg C, Resnick SI (2008). The Pareto copula, aggregation of risks and the emperor's socks. *Journal of Applied Probability* 45, 67–84.

Koenker R (2005). Quantile Regression. Cambridge University Press.

Koenker R, Bassett Jr G (1978). Regression quantiles. Econometrica 46, 33–50.

Navarro, J. (2021). Prediction of record values by using quantile regression curves and distortion functions. *Submitted*.

Navarro J, Calì C, Longobardi M, Durante F (2021). Distortion Representations of Multivariate Distributions. *Submitted*.

Nelsen RB (2006). An introduction to copulas. Second edition. Springer, New York.

Valdez EA, Xiao Y (2011). On the distortion of a copula and its margins. *Scandinavian Actuarial Journal* 4, 292–317.

The slides and more references can be seen in my webpage:

https://webs.um.es/jorgenav/miwiki/doku.php

Exercises

1. Generate a sample from a copula and plot it jointly with the quantile regression curves.
2. Generate a sample from a copula and plot it jointly with the estimated quantile regression lines.
3. Obtain the multivariate distortion representation for the residual lifetimes of two components (X_1, X_2) with standard exponential distributions and a given survival copula. Try to obtain the copula representation.
4. Obtain the multivariate distortion representation for $(X_{1:3}, X_{2:3})$ with IID components with a standard exponential distribution. Try to obtain the copula representation.

5. Simulate a sample from $(X_{1:2}, X_{2:2})$ with IID components with a standard exponential distribution and compute the quantile regression curves to predict $X_{2:2}$ from $X_{1:2}$. What is the prediction from $X_{1:2} = 3$?

6. Simulate a sample from $(X_{1:2}, X_{2:2})$ with ID components with a standard exponential distribution and a copula C and compute the quantile regression curves to predict $X_{2:2}$ from $X_{1:2}$. What is the prediction from $X_{1:2} = 3$?

7. Simulate a sample from $(X_{1:3}, X_{2:3})$ with IID components with a standard exponential distribution and compute the quantile regression curves to predict $X_{2:3}$ from $X_{1:3}$. What is the prediction from $X_{1:3} = 3$?

- **That's all.**

- That's all.
- Thank you for your attention!!

- That's all.
- Thank you for your attention!!
- Questions?