Exercise Session #3

Camilla Calì

Università degli Studi di Napoli Federico II, Italy
camilla.calì@unina.it

Stochastic Models and Complex Systems
Summer School
June 10th-July 7th, 2021

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

Exercise 1. Generate a sample from a copula and plot it jointly with the quantile regression curves.

Exercise 1. Generate a sample from a copula and plot it jointly with the quantile regression curves.

- We choose as copula the Farlie-Gumbel-Morgenstern (FGM) copula with $-1 \leq \theta \leq 1$.

$$C(x, y) = xy[1 + \theta(1 - x)(1 - y)].$$

Exercise 1. Generate a sample from a copula and plot it jointly with the quantile regression curves.

- We choose as copula the Farlie-Gumbel-Morgenstern (FGM) copula with $-1 \leq \theta \leq 1$.

$$C(x, y) = xy[1 + \theta(1 - x)(1 - y)].$$

- We assume that the DF of (X, Y) is FGM copula with $\theta = -1$, so

$$C(x, y) = xy[1 - (1 - x)(1 - y)], \; x, y \in (0, 1).$$

Exercise 1. Generate a sample from a copula and plot it jointly with the quantile regression curves.

- We choose as copula the Farlie-Gumbel-Morgenstern (FGM) copula with $-1 \leq \theta \leq 1$.

$$C(x, y) = xy[1 + \theta(1 - x)(1 - y)].$$

- We assume that the DF of (X, Y) is FGM copula with $\theta = -1$, so

$$C(x, y) = xy[1 - (1 - x)(1 - y)], \; x, y \in (0, 1).$$

- The marginals are

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

Exercise 1. Generate a sample from a copula and plot it jointly with the quantile regression curves.

- We choose as copula the Farlie-Gumbel-Morgenstern (FGM) copula with $-1 \leq \theta \leq 1$.

$$C(x, y) = xy[1 + \theta(1 - x)(1 - y)].$$

- We assume that the DF of (X, Y) is FGM copula with $\theta = -1$, so

$$C(x, y) = xy[1 - (1 - x)(1 - y)], \; x, y \in (0, 1).$$

- The marginals are

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

- The conditional distribution is

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

- The quantile regression curve is

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

- The quantile regression curve is

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

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

$$2xy + xy^2 - x^2y^2 = q$$

- The quantile regression curve is

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

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

$$2xy + xy^2 - x^2y^2 = q$$

- For $x = 1/2$ we have $y = q$.

- The quantile regression curve is

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

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

$$2xy + xy^2 - x^2y^2 = q$$

- For $x = 1/2$ we have $y = q$.

- In the other cases

$$2xy + xy^2 - x^2y^2 = q$$
$$y^2(1 - 2x) + 2xy - q = 0$$
$$y_{1,2} = \frac{-x \pm (x^2 - 2xq + q)^{1/2}}{1 - 2x}$$

- For $0 < q < 1$ and $0 < x < 1$

$$x^2 - 2xq + q \geq 0.$$

- For $0 < q < 1$ and $0 < x < 1$

$$x^2 - 2xq + q \geq 0.$$

- The only possible solution is

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

- For $0 < q < 1$ and $0 < x < 1$

$$x^2 - 2xq + q \geq 0.$$

- The only possible solution is

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

- Hence

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

- For $0 < q < 1$ and $0 < x < 1$

$$x^2 - 2xq + q \geq 0.$$

- The only possible solution is

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

- Hence

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

- So the quantile regression curve is

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

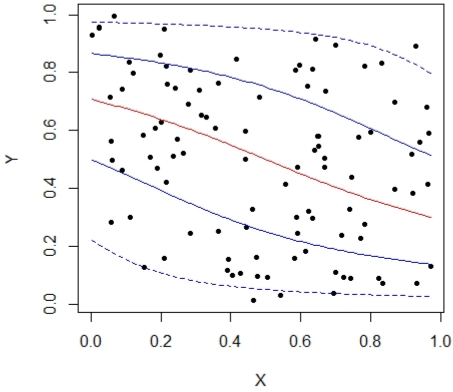

Figure: Quantile regression curve (red) and confidence bands (50% continuous blue, 90% dashed blue) for a FGM copula jointly with 100 data from this model.

- The first data are $(X_1, X_2) = (0.614549, 0.1837631)$.

- The first data are $(X_1, X_2) = (0.614549, 0.1837631)$.
- If we want to predict X_2 from $X_1 = 0.614549$, we have to calculate

$$\begin{aligned}\tilde{m}_{2|1}(0.614549) &= \frac{-0.614549 + (0.614549^2 - 0.614549 + 0.5)^{1/2}}{1 - 2 \cdot 0.614549} \\ &= 0.4434579\end{aligned}$$

- The first data are $(X_1, X_2) = (0.614549, 0.1837631)$.
- If we want to predict X_2 from $X_1 = 0.614549$, we have to calculate

$$\begin{aligned}\tilde{m}_{2|1}(0.614549) &= \frac{-0.614549 + (0.614549^2 - 0.614549 + 0.5)^{1/2}}{1 - 2 \cdot 0.614549} \\ &= 0.4434579\end{aligned}$$

- The 90% confidence interval is

$$[0.04099347, 0.9363451]$$

- The prediction is not good since the dispersion of this conditional variable is very big.

Exercise 2. Generate a sample from a copula and plot it jointly with the estimated quantile regression lines.

Exercise 2. Generate a sample from a copula and plot it jointly with the estimated quantile regression lines.

- The estimated quantile regression line is

$$\hat{m}(x) = 0.6351853 - 0.2629179x$$

Exercise 2. Generate a sample from a copula and plot it jointly with the estimated quantile regression lines.

- The estimated quantile regression line is

$$\hat{m}(x) = 0.6351853 - 0.2629179x$$

- The prediction for X_2 from $X_1 = 0.614549$ with this QR line is

$$\hat{m}(0.614549) = 0.6351853 - 0.2629179 \cdot 0.614549 = 0.4736094$$

Exercise 2. Generate a sample from a copula and plot it jointly with the estimated quantile regression lines.

- The estimated quantile regression line is

$$\hat{m}(x) = 0.6351853 - 0.2629179x$$

- The prediction for X_2 from $X_1 = 0.614549$ with this QR line is

$$\hat{m}(0.614549) = 0.6351853 - 0.2629179 \cdot 0.614549 = 0.4736094$$

- The estimated quantile regression line with a polynomial of degree 3 is

$$\hat{m}_3(x) = 0.9090624 - 2.0883789x + 2.5793994x^2 - 0.8383797x^3$$

- The real data was $X_2 = 0.1837631$ and the prediction with the exact QR curve was $\tilde{m}_{2|1}(0.614549) = 0.4434579$.

- The real data was $X_2 = 0.1837631$ and the prediction with the exact QR curve was $\tilde{m}_{2|1}(0.614549) = 0.4434579$.
- The prediction with a polynomial of degree 3 is

$$\hat{m}_3(0.614549) = 0.4052288.$$

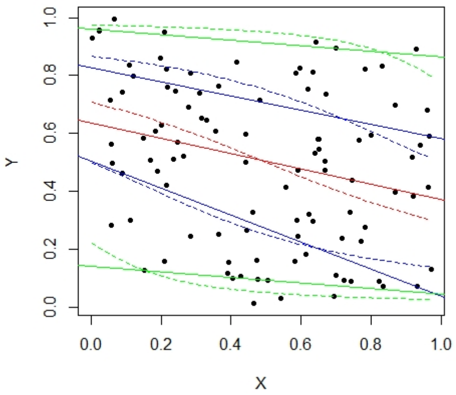

Figure: Estimated Quantile Regression line (red) and confidence bands (50% continuous blue, 90% continuous green) for the 100 data from a FGM model. The dashed lines are the exact curves.

Exercise 3. 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 $X_{1:2} = 3$?

- (X_1, X_2) are IID with a common standard exponential distribution function F.

Exercise 3. 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 $X_{1:2} = 3$?

- (X_1, X_2) are IID with a common standard exponential distribution function F.
- $X_{1:2} = min(X_1, X_2)$ and $X_{2:2} = max(X_1, X_2)$.

Exercise 3. 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 $X_{1:2} = 3$?

- (X_1, X_2) are IID with a common standard exponential distribution function F.
- $X_{1:2} = min(X_1, X_2)$ and $X_{2:2} = max(X_1, X_2)$.
- We want to estimate $X_{2:2}$ from $X_{1:2}$, that is we want to calculate the distribution function of $(X_{2:2}|X_{1:2} = x_1)$.

Exercise 3. 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 $X_{1:2} = 3$?

- (X_1, X_2) are IID with a common standard exponential distribution function F.
- $X_{1:2} = min(X_1, X_2)$ and $X_{2:2} = max(X_1, X_2)$.
- We want to estimate $X_{2:2}$ from $X_{1:2}$, that is we want to calculate the distribution function of $(X_{2:2}|X_{1:2} = x_1)$.
- The distribution function of $(X_{2:2}|X_{1:2} = x_1)$ is

$$G_{2|1}(x_2|x_1) = D_{2|1}(F(x_2)|F(x_1)), \text{ for } x_2 \geq x_1$$

where, in IID case,

$$D_{2|1}(v|F(x_1)) = \frac{v - F(x_1)}{\bar{F}(x_1)}, \text{ for } F(x_1) \leq v < 1.$$

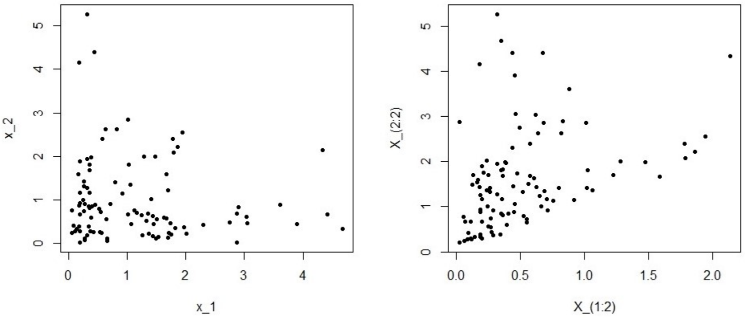

Figure: Independent data from two standard exponential distributions (left) and the associated paired ordered data (right).

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

$$F_{2|1}^{-1}(q|x_1) = F^{-1}(D_{2|1}^{-1}(q|F(x_1))).$$

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

$$F_{2|1}^{-1}(q|x_1) = F^{-1}(D_{2|1}^{-1}(q|F(x_1))).$$

- We have

$$D_{2|1}^{-1}(q|F(x_1)) = F(x_1) + q\bar{F}(x_1).$$

- F is a standard exponential distribution, so

$$F(x) = 1 - e^{-x} \quad \Rightarrow \quad F^{-1}(y) = -\log(1-y)$$

- The quantile function $F_{2|1}^{-1}$ can be computed as
$$F_{2|1}^{-1}(q|x_1) = F^{-1}(D_{2|1}^{-1}(q|F(x_1))).$$
- We have
$$D_{2|1}^{-1}(q|F(x_1)) = F(x_1) + q\bar{F}(x_1).$$
- F is a standard exponential distribution, so
$$F(x) = 1 - e^{-x} \quad \Rightarrow \quad F^{-1}(y) = -\log(1-y)$$
- Then
$$\begin{aligned} F^{-1}(q|x_1) &= F^{-1}(F(x_1) + q\bar{F}(x_1)) \\ &= F^{-1}(1 - e^{-x_1} + qe^{-x_1}) \\ &= -\log(e^{-x_1}(1-q)) \\ &= x_1 - \log(1-q) \end{aligned}$$

- So the exact QR curve is

$$m(x) = x - \log(1 - 0.5) = x - \log(0.5).$$

- So the exact QR curve is

$$m(x) = x - \log(1 - 0.5) = x - \log(0.5).$$

- The exact QR centered 50% confidence band is

$$[x - \log(1 - 0.25), x - \log(1 - 0.75)].$$

- So the exact QR curve is

$$m(x) = x - \log(1 - 0.5) = x - \log(0.5).$$

- The exact QR centered 50% confidence band is

$$[x - \log(1 - 0.25), x - \log(1 - 0.75)].$$

- The exact QR centered 90% confidence band is

$$[x - \log(1 - 0.05), x - \log(1 - 0.95)].$$

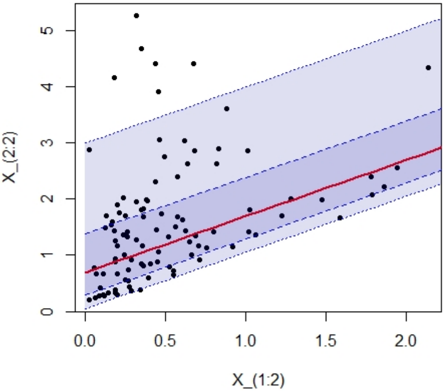

Figure: QR for the paired ordered data $(X_{1:2}, X_{2:2})$ associated to independent data (X_1, X_2) from two standard exponential distributions jointly with 50% and 90% centered confidence bands.

What is the prediction from $X_{1:2} = 3$?

What is the prediction from $X_{1:2} = 3$?

- The prediction for $X_{2:2}$ from $X_{1:2} = 3$ is

$$m(3) = 3 - \log(1 - 0.5) = 3.693147$$

What is the prediction from $X_{1:2} = 3$?

- The prediction for $X_{2:2}$ from $X_{1:2} = 3$ is
$$m(3) = 3 - \log(1 - 0.5) = 3.693147$$

- The centered 90% confidence interval for this prediction is
$$[3.051293, 5.995732]$$

What is the prediction from $X_{1:2} = 3$?

- The prediction for $X_{2:2}$ from $X_{1:2} = 3$ is

$$m(3) = 3 - \log(1 - 0.5) = 3.693147$$

- The centered 90% confidence interval for this prediction is

$$[3.051293, 5.995732]$$

- The centered 50% confidence interval for this prediction is

$$[3.287682, 4.386294]$$

Exercise 4. 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 $X_{1:2} = 3$?

- (X_1, X_2) are DID with a common standard exponential distribution function F and a Clayton copula C.

Exercise 4. 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 $X_{1:2} = 3$?

- (X_1, X_2) are DID with a common standard exponential distribution function F and a Clayton copula C.
- $X_{1:2} = min(X_1, X_2)$ and $X_{2:2} = max(X_1, X_2)$.

Exercise 4. 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 $X_{1:2} = 3$?

- (X_1, X_2) are DID with a common standard exponential distribution function F and a Clayton copula C.
- $X_{1:2} = min(X_1, X_2)$ and $X_{2:2} = max(X_1, X_2)$.
- We want to estimate $X_{2:2}$ from $X_{1:2}$, that is we want to calculate the distribution function of $(X_{2:2}|X_{1:2} = x_1)$.

Exercise 4. 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 $X_{1:2} = 3$?

- (X_1, X_2) are DID with a common standard exponential distribution function F and a Clayton copula C.
- $X_{1:2} = min(X_1, X_2)$ and $X_{2:2} = max(X_1, X_2)$.
- We want to estimate $X_{2:2}$ from $X_{1:2}$, that is we want to calculate the distribution function of $(X_{2:2} | X_{1:2} = x_1)$.
- The distribution function of $(X_{2:2} | X_{1:2} = x_1)$ is

$$G_{2|1}(x_2|x_1) = D_{2|1}(F(x_2)|F(x_1)), \text{ for } x_2 \geq x_1.$$

- During the lesson we have seen that

$$G_{2|1}^{-1}(q|x_1) = F^{-1}\left(\frac{F(x_1)}{F(x_1) - 1 + \frac{2-F(x_1)}{\sqrt{1-q+q(2-F(x_1))^2}}}\right)$$

where

$$F(x) = 1 - e^{-x} \quad \Rightarrow \quad F^{-1}(y) = -\log(1-y).$$

- During the lesson we have seen that

$$G_{2|1}^{-1}(q|x_1) = F^{-1}\left(\frac{F(x_1)}{F(x_1) - 1 + \frac{2-F(x_1)}{\sqrt{1-q+q(2-F(x_1))^2}}}\right)$$

where

$$F(x) = 1 - e^{-x} \quad \Rightarrow \quad F^{-1}(y) = -\log(1-y).$$

- So the exact median regression curve to predict $X_{2:2}$ from $X_{1:2} = x_1$ is

$$m(x_1) = G_{2|1}^{-1}(0.5|x_1)$$

- During the lesson we have seen that

$$G_{2|1}^{-1}(q|x_1) = F^{-1}\left(\frac{F(x_1)}{F(x_1) - 1 + \frac{2-F(x_1)}{\sqrt{1-q+q(2-F(x_1))^2}}} \right)$$

where

$$F(x) = 1 - e^{-x} \quad \Rightarrow \quad F^{-1}(y) = -\log(1-y).$$

- So the exact median regression curve to predict $X_{2:2}$ from $X_{1:2} = x_1$ is

$$m(x_1) = G_{2|1}^{-1}(0.5|x_1)$$

- The 50% and 90% centered confidence bands are, respectively

$$[G_{2|1}^{-1}(0.25|x_1), G_{2|1}^{-1}(0.75|x_1)] \quad \text{and} \quad [G_{2|1}^{-1}(0.05|x_1), G_{2|1}^{-1}(0.95|x_1)].$$

What is the prediction from $X_{1:2} = 3$?

What is the prediction from $X_{1:2} = 3$?

- The prediction for $X_{2:2}$ from $X_{1:2} = 3$ is

$$m(3) = G_{2|1}^{-1}(0.5|3) = 3.704109$$

What is the prediction from $X_{1:2} = 3$?

- The prediction for $X_{2:2}$ from $X_{1:2} = 3$ is

$$m(3) = G_{2|1}^{-1}(0.5|3) = 3.704109$$

- The centered 90% confidence interval for this prediction is

$$[3.052409, 6.016211]$$

What is the prediction from $X_{1:2} = 3$?

- The prediction for $X_{2:2}$ from $X_{1:2} = 3$ is
$$m(3) = G_{2|1}^{-1}(0.5|3) = 3.704109$$

- The centered 90% confidence interval for this prediction is
$$[3.052409, 6.016211]$$

- The centered 50% confidence interval for this prediction is
$$[3.293216, 4.402582]$$

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

- (X_1, X_2, X_3) are IID with a common standard exponential distribution function $F = 1 - e^{-x}$.

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

- (X_1, X_2, X_3) are IID with a common standard exponential distribution function $F = 1 - e^{-x}$.
- $X_{1:3} = min(X_1, X_2, X_3)$ and $X_{3:3} = max(X_1, X_2, X_3)$.

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

- (X_1, X_2, X_3) are IID with a common standard exponential distribution function $F = 1 - e^{-x}$.
- $X_{1:3} = min(X_1, X_2, X_3)$ and $X_{3:3} = max(X_1, X_2, X_3)$.
- We want to estimate $X_{3:3}$ from $X_{1:3}$, that is we want to calculate the distribution function of $(X_{3:3}|X_{1:3} = x_1)$.

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

- (X_1, X_2, X_3) are IID with a common standard exponential distribution function $F = 1 - e^{-x}$.
- $X_{1:3} = min(X_1, X_2, X_3)$ and $X_{3:3} = max(X_1, X_2, X_3)$.
- We want to estimate $X_{3:3}$ from $X_{1:3}$, that is we want to calculate the distribution function of $(X_{3:3}|X_{1:3} = x_1)$.
- The distribution function of $(X_{3:3}|X_{1:3} = x_1)$ is

$$G_{2|1}(x_2|x_1) = D_{2|1}(F(x_2)|F(x_1)), \text{ for } x_2 \geq x_1$$

- We have to calculate

$$D_{2|1}(v|u) = \frac{\partial_1 D(u,v)}{\partial_1 D(u,1)}, \text{ for } 0 \leq u \leq v \leq 1$$

- We have to calculate

$$D_{2|1}(v|u) = \frac{\partial_1 D(u,v)}{\partial_1 D(u,1)}, \text{ for } 0 \le u \le v \le 1$$

- In general case, for a parallel system with 3 components, $D(u,v)$ is equal to

$$C(u,v,v)+C(v,u,v)+C(v,v,u)-C(u,u,v)-C(u,v,u)-C(v,u,u)+C(u,u,u)$$

for $0 \le u \le v \le 1$.

- We have to calculate

$$D_{2|1}(v|u) = \frac{\partial_1 D(u,v)}{\partial_1 D(u,1)}, \text{ for } 0 \le u \le v \le 1$$

- In general case, for a parallel system with 3 components, $D(u,v)$ is equal to

$$C(u,v,v)+C(v,u,v)+C(v,v,u)-C(u,u,v)-C(u,v,u)-C(v,u,u)+C(u,u,u)$$

for $0 \le u \le v \le 1$.

- In IID case,

$$D(u,v) = 3uv^2 - 3u^2v + u^3, \text{ for } 0 \le u \le v \le 1.$$

- So

$$D_{2|1}(v|u) = \frac{\partial_1(3uv^2 - 3u^2v + u^3)}{\partial_1(3u - 3u^2 + u^3)} = \frac{3v^2 - 6uv + 3u^2}{3 - 6u + 3u^2},$$

for $0 \leq u \leq v \leq 1$.

- So

$$D_{2|1}(v|u) = \frac{\partial_1(3uv^2 - 3u^2v + u^3)}{\partial_1(3u - 3u^2 + u^3)} = \frac{3v^2 - 6uv + 3u^2}{3 - 6u + 3u^2},$$

for $0 \leq u \leq v \leq 1$.

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

$$F_{2|1}^{-1}(q|x_1) = F^{-1}(D_{2|1}^{-1}(q|F(x_1))).$$

- So

$$D_{2|1}(v|u) = \frac{\partial_1(3uv^2 - 3u^2v + u^3)}{\partial_1(3u - 3u^2 + u^3)} = \frac{3v^2 - 6uv + 3u^2}{3 - 6u + 3u^2},$$

for $0 \le u \le v \le 1$.

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

$$F_{2|1}^{-1}(q|x_1) = F^{-1}(D_{2|1}^{-1}(q|F(x_1))).$$

- To compute the inverse of $D_{2|1}$ we need to solve in v the equation

$$D_{2|1}(v|u) = \frac{3v^2 - 6uv + 3u^2}{3 - 6u + 3u^2} = q$$

- We have

$$v = q^{1/2}(1 - u) + u$$

- We have

$$v = q^{1/2}(1-u) + u$$

- So

$$D_{2|1}^{-1}(q|F(x_1)) = q^{1/2}(1-F(x_1)) + F(x_1)$$

- We have

$$v = q^{1/2}(1-u) + u$$

- So

$$D_{2|1}^{-1}(q|F(x_1)) = q^{1/2}(1-F(x_1)) + F(x_1)$$

- F is a standard exponential distribution, so

$$F(x) = 1 - e^{-x} \quad \Rightarrow \quad F^{-1}(y) = -\log(1-y)$$

- Then

$$
\begin{aligned}
F^{-1}(q|x_1) &= F^{-1}(q^{1/2}(1 - F(x_1)) + F(x_1)) \\
&= F^{-1}(q^{1/2}e^{-x_1} + 1 - e^{-x_1}) \\
&= -\log(e^{-x_1}(1 - q^{1/2})) \\
&= x - \log(1 - q^{1/2})
\end{aligned}
$$

- Then

$$\begin{aligned} F^{-1}(q|x_1) &= F^{-1}(q^{1/2}(1-F(x_1))+F(x_1)) \\ &= F^{-1}(q^{1/2}e^{-x_1}+1-e^{-x_1}) \\ &= -\log(e^{-x_1}(1-q^{1/2})) \\ &= x-\log(1-q^{1/2}) \end{aligned}$$

- So the exact QR curve is

$$m(x) = x - \log(1-0.5^{1/2}).$$

- The exact QR centered 50% confidence band is

$$[x-\log(1-0.25^{1/2}), x-\log(1-0.75^{1/2})].$$

- Then

$$
\begin{aligned}
F^{-1}(q|x_1) &= F^{-1}(q^{1/2}(1-F(x_1))+F(x_1)) \\
&= F^{-1}(q^{1/2}e^{-x_1}+1-e^{-x_1}) \\
&= -\log(e^{-x_1}(1-q^{1/2})) \\
&= x-\log(1-q^{1/2})
\end{aligned}
$$

- So the exact QR curve is

$$m(x) = x - \log(1-0.5^{1/2}).$$

- The exact QR centered 50% confidence band is

$$[x-\log(1-0.25^{1/2}), x-\log(1-0.75^{1/2})].$$

- The exact QR centered 90% confidence band is

$$[x-\log(1-0.05^{1/2}), x-\log(1-0.95^{1/2})].$$

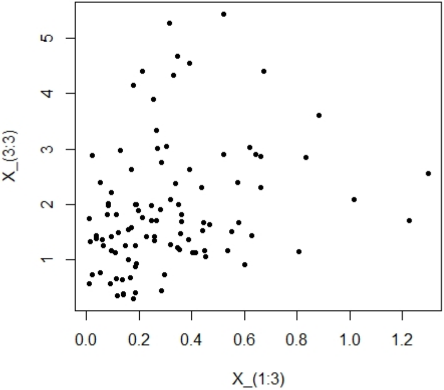

Figure: Paired ordered data of a sample from $(X_{1:3}, X_{3:3})$ with IID components with a standard exponential distribution

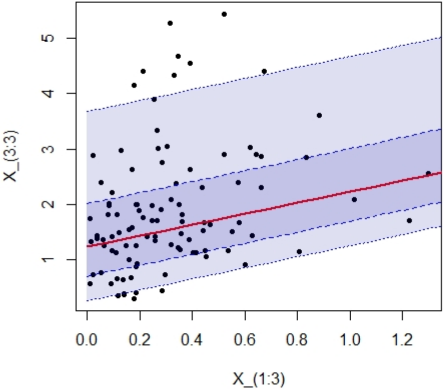

Figure: QR for the paired ordered data $(X_{1:3}, X_{3:3})$ associated to independent data (X_1, X_2, X_3) from a common standard exponential distribution jointly with 50% and 90% centered confidence bands.

What is the prediction from $X_{1:3} = 3$?

What is the prediction from $X_{1:3} = 3$?

- The prediction for $X_{3:3}$ from $X_{1:3} = 3$ is

$$m(3) = 3 - \log(1 - 0.5^{1/2}) = 4.227947$$

What is the prediction from $X_{1:3} = 3$?

- The prediction for $X_{3:3}$ from $X_{1:3} = 3$ is
$$m(3) = 3 - \log(1 - 0.5^{1/2}) = 4.227947$$

- The centered 90% confidence interval for this prediction is
$$[3.253096, 6.676138]$$

What is the prediction from $X_{1:3} = 3$?

- The prediction for $X_{3:3}$ from $X_{1:3} = 3$ is

$$m(3) = 3 - \log(1 - 0.5^{1/2}) = 4.227947$$

- The centered 90% confidence interval for this prediction is

$$[3.253096, 6.676138]$$

- The centered 50% confidence interval for this prediction is

$$[3.693147, 5.010105]$$

```
#EXERCISE 1
FI<-function(q,x) (-x+(x^2-2*x*q+q)^(1/2))/(1-2*x)
m<-function(x) FI(0.5,x) #Quantile regression curve is FI evaluated in q=0.5
n<-100 #Sample of size 100
x<-1:n #Generate 100 x
y<-1:n #Generate 100 y
set.seed(110) #Set the seed of random number generator:
#is useful for creating simulations or random objects that can be reproduced.
for (i in 1:n){
  x[i]<-runif(1) #Generate random variables, uniformly distributed
  y[i]<-FI(runif(1),x[i]) #Generate y from x
}
plot(x,y,xlab="X",ylab="Y",pch=20) #Plot sample
curve(m(x),add=T,col='red',ylim=c(0,1)) #Add quantile regression curve
curve(FI(0.25,x),add=T,col='blue') #Add 50% centered confidence band (low)
curve(FI(0.75,x),add=T,col='blue') #Add 50% centered confidence band (up)
curve(FI(0.05,x),add=T,col='blue',lty=2) #Add 90% centered confidence band (low)
curve(FI(0.95,x),add=T,col='blue',lty=2) #Add 90% centered confidence band (up)

#Prediction
x[1] #First data for x
y[1] #First data for y
m(x[1]) #Quantile regression curve evaluated in x[1]
FI(0.05,x[1])  #90% centered confidence band (low)
FI(0.95,x[1])  #90% centered confidence band (up)


#EXERCISE 2
#Non parametric estimation
install.packages('quantreg')
library('quantreg')

rq(y~x) #Quantile regression fit of y as a function of x

#Predictions
m2<-function(x) 0.6351853-0.2629179*x #Quantile regression line from coefficient
#estimated by rq(y~x)
```

```
m2(x[1]) #Quantile regression line evaluated in y[1] from x[1]
m(x[1]) #Quantile regression curve evaluated in y[1] from x[1]
y[1] #Real value y[1]
d<-data.frame(y,x,x^2,x^3)
rq(d) #Quantile regression fit of data frame d

m3<-function(x) 0.9090624-2.0883789*x+ 2.5793994*x^2 -0.8383797*x^3 #Quantile
#regression line from coefficient estimated by rq(d)

m3(x[1]) #Quantile regression line evaluated in y[1] from x[1]

plot(x,y,xlab='X',ylab='Y',pch=20) #Plot sample
abline(rq(y~x),col='red') #Add estimated quantile regression line (EQRL)
abline(rq(y~x,0.25),col='blue') #Add 50% centered confidence band (low) for EQRL
abline(rq(y~x,0.75),col='blue') #Add 50% centered confidence band (up) for EQRL
abline(rq(y~x,0.05),col='green') #Add 90% centered confidence band (low) for EQRL
abline(rq(y~x,0.95),col='green') #Add 90% centered confidence band (up) for EQRL
curve(m(x),add=T,col='red',lty=2) #Add quantile regression curve
curve(FI(0.25,x),add=T,col='blue',lty=2) #Add 50% centered confidence band (low)
curve(FI(0.75,x),add=T,col='blue',lty=2) #Add 50% centered confidence band (up)
curve(FI(0.05,x),add=T,col='green',lty=2) #Add 90% centered confidence band (low)
curve(FI(0.95,x),add=T,col='green',lty=2) #Add 90% centered confidence band (up)
```

```
#EXERCISE 3
#Paired ordered data IID case
# Independent Standard Exponentials
n<-100 #Sample of size 100
set.seed(202) #Set the seed of random number generator:
#is useful for creating simulations or random objects that can be reproduced.
mu<-1 #Standard exponential mu=1
#'rexp' do a random generation for the exponential distribution
x<-rexp(n,1/mu) #x is a sample from a standard exponential distribution
y<-rexp(n,1/mu) #y is a sample from a standard exponential distribution
plot(x,y,xlab='x_1',ylab='x_2',pch=20) #Plot sample

L<-pmin(x,y) #L is the minimum between x and y for each i=1,...,n
U<-pmax(x,y) #U is the maximum between x and y for each i=1,...,n
plot(L,U,xlab='X_(1:2)',ylab='X_(2:2)',pch=20) #Plot the paired ordered data from the sample


xmax=max(L) #xmax is the maximum of L




x=seq(0,xmax+1,by=0.1)
qr=x-log(1-0.5) #Exact quantile regression (QR) curve

qr.90=x-log(1-0.95) #Exact QR centered 90% confidence band (up)
qr.10=x-log(1-0.05) #Exact QR centered 90% confidence band (low)

qr.75=x-log(1-0.75) #Exact QR centered 50% confidence band (up)
qr.25=x-log(1-0.25) #Exact QR centered 50% confidence band (low)




lines(x,qr,col="red",lwd=2) #Add the exact quantile regression (QR) curve
lines(x,qr.75, lty="dashed", col="blue", lwd=1) #Add the QR centered 50% confidence band (up)
lines(x,qr.25, lty="dashed", col="blue", lwd=1) #Add the QR centered 50% confidence band (low)
polygon(c(x,rev(x)), c(qr.25,rev(qr.75)), col="#00009920", border=NA)
```

```
lines(x,qr.90, lty="dotted", col="blue", lwd=1) #Add the QR centered 90% confidence band (up)
lines(x,qr.10, lty="dotted", col="blue", lwd=1) #Add the QR centered 90% confidence band (low)
polygon(c(x,rev(x)), c(qr.10,rev(qr.90)), col="#00009920", border=NA)

#Predictions
L.ex=3 #What is the prediction from X_1:2 = 3?
m<-function(x) x-log(1-0.5) #Function for exact quantile regression (QR) curve in x
m(L.ex) #Exact quantile regression (QR) curve evaluated in 3

CI<-function(x,q) x-log(1-q) #Function for confidence band of (QR) curve in x and q
CI(L.ex,0.05)  #QR centered 90% confidence band (low) for x=3
CI(L.ex,0.95)  #QR centered 90% confidence band (up) for x=3
CI(L.ex,0.25)  #QR centered 50% confidence band (low) for x=3
CI(L.ex,0.75)  #QR centered 50% confidence band (up) for x=3

####
```

```
#EXERCISE 4
#QR curves Clayton
mu<-1
#'qexp(x,1/mu)' calculates the quantile function of x for the exponential distribution with mean mu
#'pexp(x,1/mu)' calculates the DF of x for the exponential distribution with mean mu
G=function(x,q){
  qexp(pexp(x,1/mu)/(pexp(x,1/mu)-1+(2-pexp(x,1/mu))*(1-q+q*(2-pexp(x,1/mu))^2 )^-0.5),1/mu)
} #Regression curve

L.ex=3 #What is the prediction from X_1:2 = 3?
m=G(L.ex,0.5) #Median regression curve evaluated in 3

# Confidence band
G(L.ex,0.05)  #QR centered 90% confidence band (low) for x=3
G(L.ex,0.95)  #QR centered 90% confidence band (up) for x=3
G(L.ex,0.25)  #QR centered 50% confidence band (low) for x=3
G(L.ex,0.75)  #QR centered 50% confidence band (up) for x=3
```

```
#EXERCISE 5
#Paired ordered data IID case-3 components
# Independent Standard Exponentials
n<-100 #Sample of size 100
set.seed(202)#Set the seed of random number generator:
#is useful for creating simulations or random objects that can be reproduced.
mu<-1 #Standard exponential mu=1
x<-rexp(n,1/mu) #x is a sample from a standard exponential distribution
y<-rexp(n,1/mu) #y is a sample from a standard exponential distribution
z<-rexp(n,1/mu) #z is a sample from a standard exponential distribution

L<-pmin(x,y,z) #L is the minimum between x, y and z for each i=1,...,n
U<-pmax(x,y,z) #U is the maximum between x, y and z for each i=1,...,n
plot(L,U,xlab='X_(1:3)',ylab='X_(3:3)',pch=20) #Plot the paired ordered data from the sample


xmax=max(L) #xmax is the maximum of L


x=seq(0,xmax+1,by=0.1)
qr=x-log(1-(0.5)^(0.5)) #Exact quantile regression (QR) curve

qr.90=x-log(1-(0.95)^(0.5)) #Exact QR centered 90% confidence band (up)
qr.10=x-log(1-(0.05)^(0.5)) #Exact QR centered 90% confidence band (low)

qr.75=x-log(1-(0.75)^(0.5)) #Exact QR centered 50% confidence band (up)
qr.25=x-log(1-(0.25)^(0.5)) #Exact QR centered 50% confidence band (low)


lines(x,qr,col="red",lwd=2) #Add the exact quantile regression (QR) curve
lines(x,qr.75, lty="dashed", col="blue", lwd=1) #Add the QR centered 50% confidence band (up)
lines(x,qr.25, lty="dashed", col="blue", lwd=1) #Add the QR centered 50% confidence band (low)
polygon(c(x,rev(x)), c(qr.25,rev(qr.75)), col="#00009920", border=NA)

lines(x,qr.90, lty="dotted", col="blue", lwd=1) #Add the QR centered 90% confidence band (up)
lines(x,qr.10, lty="dotted", col="blue", lwd=1) #Add the QR centered 90% confidence band (low)
polygon(c(x,rev(x)), c(qr.10,rev(qr.90)), col="#00009920", border=NA)
```

```
#Predictions
L.ex=3  #What is the prediction from X_1:3 = 3?
m<-function(x) x-log(1-(0.5)^(0.5)) #Function for exact quantile regression (QR) curve in x
m(L.ex) #Exact quantile regression (QR) curve evaluated in 3

CI<-function(x,q) x-log(1-(q)^(0.5)) #Function for confidence band of (QR) curve in x and q
CI(L.ex,0.05)  #QR centered 90% confidence band (low) for x=3
CI(L.ex,0.95)  #QR centered 90% confidence band (up) for x=3
CI(L.ex,0.25)  #QR centered 50% confidence band (low) for x=3
CI(L.ex,0.75)  #QR centered 50% confidence band (up) for x=3

####
```