Dépendance et résultats limites, quelques applications en finance et assurance

Arthur Charpentier

Arthur Charpentier

Dependence structures and limiting results, with applications in finance and insurance.

Promoters: Jan Beirlant (KUL) & Michel Denuit (UCL)

Katholieke Universiteit Leuven, June 2006.

Introduction

Importance of tail dependence in risk management.

In insurance, the cost of a claim is the sum of

- the loss amount (paid to the insured), X_1
- the allocated expenses (lawyers, expertise...), X_2

Consider the following excess-of-loss reinsurance treaty, with payoff

$$g(x_1,x_2)=\begin{cases} 0, \text{ if } x_1 \leq d, \\ x_1-d+\dfrac{x_1-d}{x_1}x_2, \text{ if } x_1>d. \end{cases}$$

The pure premium is then $\mathbb{E}(g(X_1,X_2))$.

In life insurance, the payoff for a joint life insurance is

$$g(x_1, x_2) = \sum_{k=1}^{\infty} v^k C_k \mathbf{1}(x_1 > k \text{ and } x_2 > k)$$

(capital is due as long as the spouses are both still alive) and the pure premium is then $\mathbb{E}(g(T_x, T_y))$, where T_x and T_y denote the survival life lengths, of the man at age x and his wife y.

In finance, the payoff of quanto derivatives is $g(x_1, x_2) = x_2(x_1 - K)_+$ where X_2 is the exchange rate, and X_1 some overseas asset.

Modeling dependence with copulas

Definition 1. *A d-dimensional copula is a d-dimensional cumulative distribution function restricted to $[0,1]^d$ with standard uniform margins, $d = 2, 3, \ldots$.*

Definition 2. *Given $F_1, ..., F_d$ some univariate distribution functions, the class of d-dimensional distribution functions F with marginal distributions $F_1, ..., F_d$ respectively, is called a Fréchet class, denoted $\mathcal{F}(F_1, ..., F_d)$.*

Definition 3. *If $\boldsymbol{U} = (U_1, ..., U_n)$ has cdf C then the cdf of $\mathbf{1} - \boldsymbol{U} = (1 - U_1, ..., 1 - U_n)$ is also a copula, called survival copula of C, and denoted C^*.*

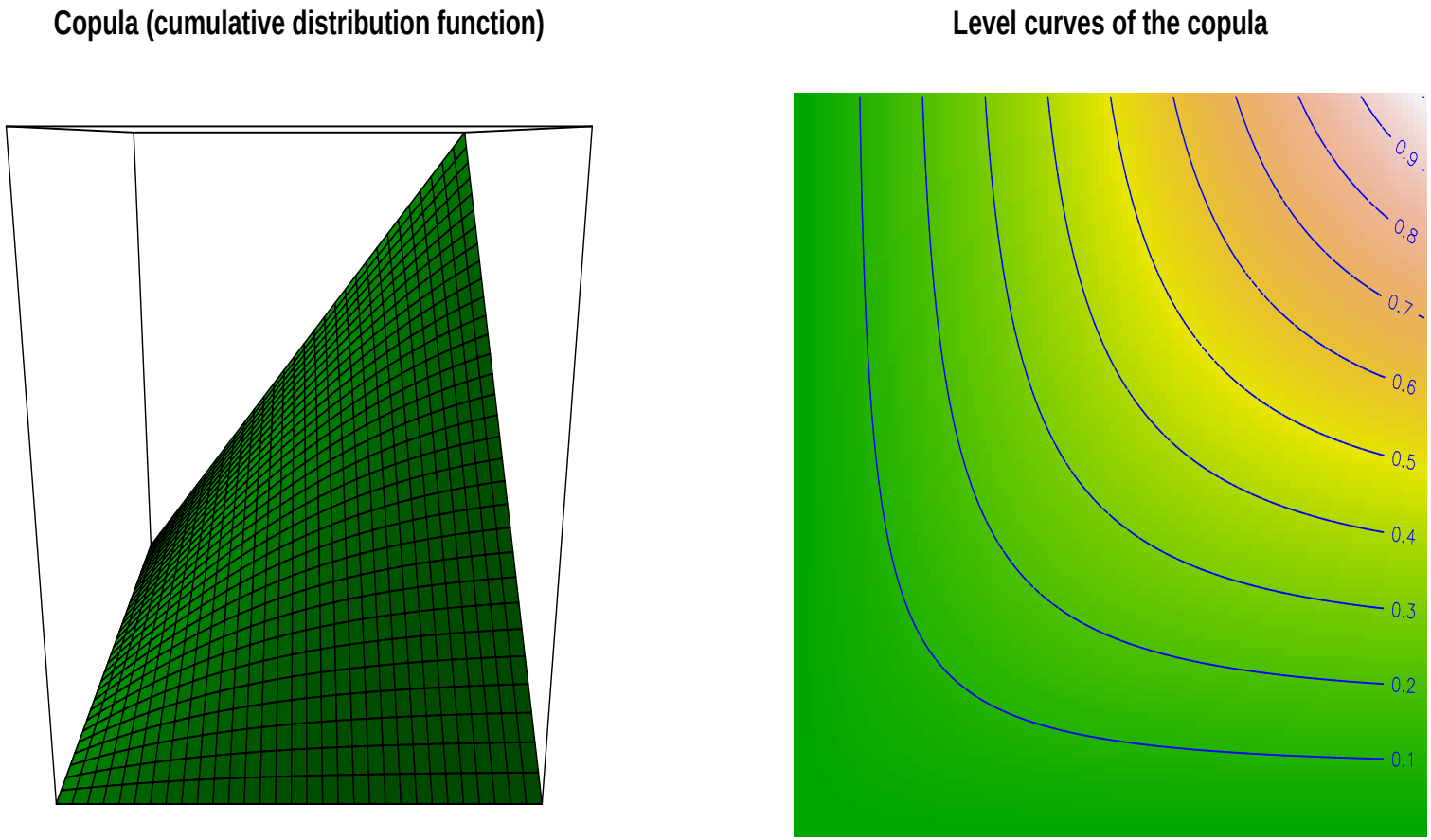


Figure 1: Copula, as a cumulative distribution function $C(u, v)$.

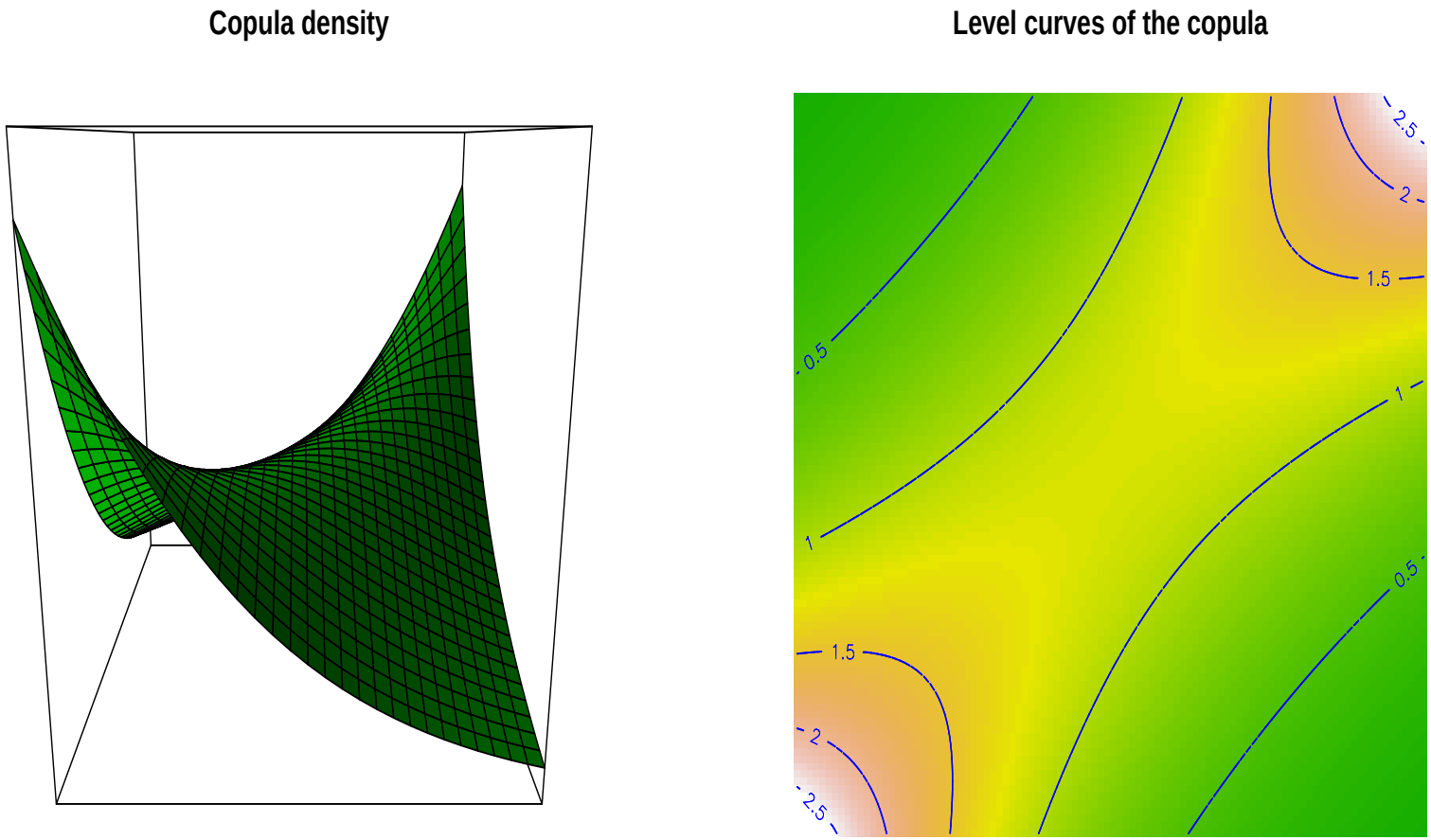


Figure 2: Density of a copula, $c(u,v) = \partial^2 C(u,v)/\partial u \partial v$.

Theorem 4. *Let C be a d-dimensional copula and $F_1, ..., F_d$ be univariate distribution functions. Then, for $\boldsymbol{x} = (x_1, ..., x_d) \in \mathbb{R}^d$,*

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_d(x_d)) \tag{1}$$

defines a distribution function with marginal distribution functions $F_1, ..., F_d$.Conversely, for a d-dimensional distribution function F with marginal distributions $F_1, ..., F_d$ there is a copula C satisfying Equation (1). This copula is not necessarily unique, but it is if $F_1, ..., F_d$ are continuous, given by

$$C(u_1, ..., u_d) = F(F_1^{\leftarrow}(u_1), ..., F_n^{\leftarrow}(x_n)), \tag{2}$$

for any $\boldsymbol{u} = (u_1, , ..., u_d) \in [0, 1]^d$, where $F_1^{\leftarrow}, ..., F_d^{\leftarrow}$ denote the generalized left continuous inverses of the F_i's, i.e.

$$F_i^{\leftarrow}(t) = \inf\{x \in \mathbb{R}, F_i(x) \geq t\} \text{ for all } 0 \leq t \leq 1.$$

Given, $\boldsymbol{X}$ with continuous marginals, *the* copula of $\boldsymbol{X}$ is the distribution of

$$\boldsymbol{U} = (U_1, \ldots, U_d) = (F_1(X_1), \ldots, F_d(X_d)).$$

Hence, it is the distribution of the ranks.

The survival copula of $\boldsymbol{X}$ is the distribution of $\mathbf{1} - \boldsymbol{U}$.

Hence,

$$\mathbb{P}(\boldsymbol{X} \leq \boldsymbol{x}) = C(\mathbb{P}(X_1 \leq x_1), \ldots, \mathbb{P}(X_d \leq x_d)),$$

$$\mathbb{P}(\boldsymbol{X} > \boldsymbol{x}) = C^*(\mathbb{P}(X_1 > x_1), \ldots, \mathbb{P}(X_d > x_d)),$$

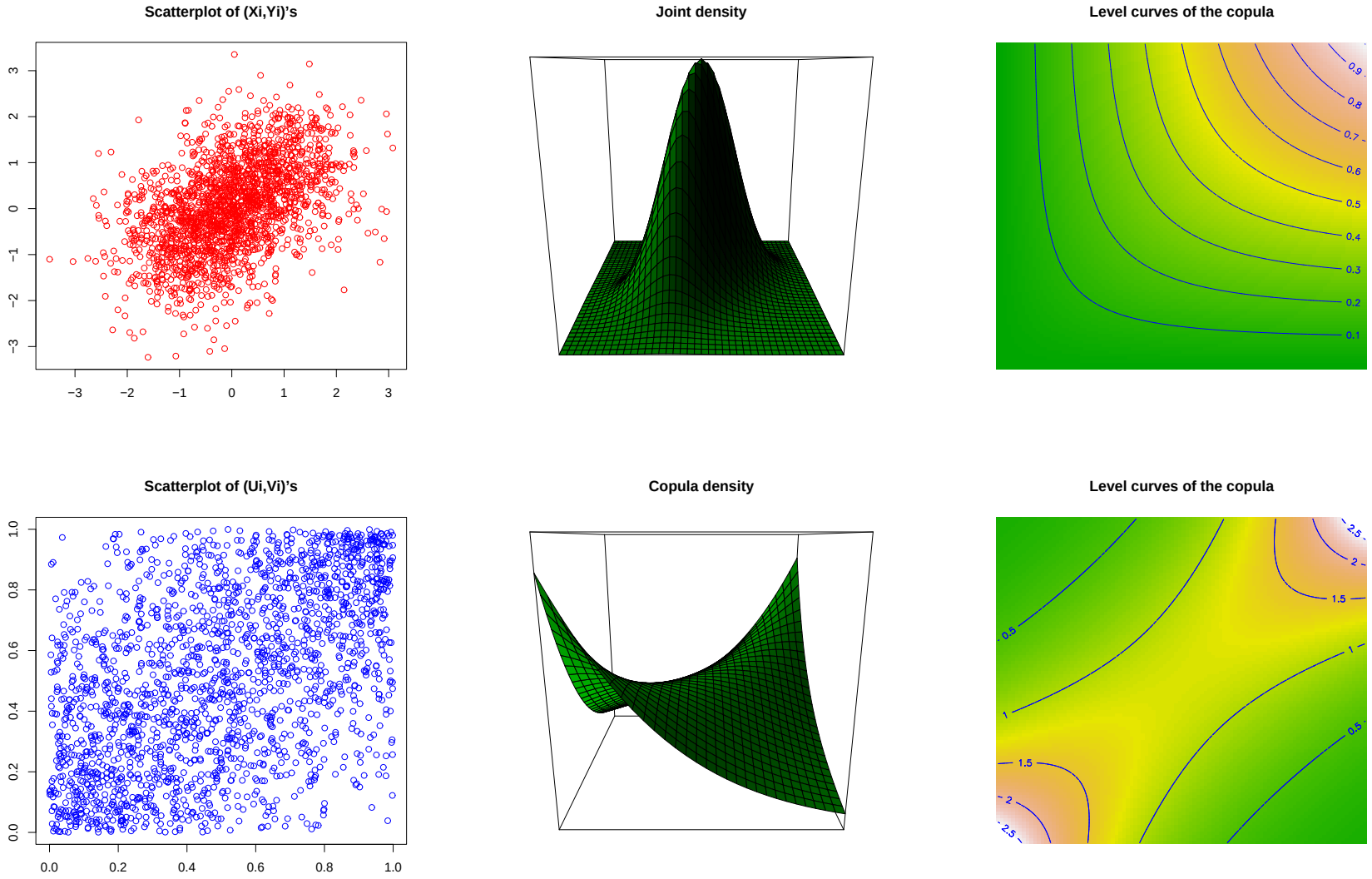


Figure 3: Scatterplot and densities of (X, Y) and (U, V).

In dimension 2, consider the following family of copulae

Definition 5. *Let ψ denote a convex decreasing function $(0,1] \to [0,\infty]$ such that $\psi(1)=0$. Define the inverse (or quasi-inverse if $\psi(0)<\infty$) as*

$$\psi^{\leftarrow}(t) = \begin{cases} \psi^{-1}(t) & \text{for } 0 \leq t \leq \psi(0) \\ 0 & \text{for } \psi(0) < t < \infty. \end{cases}$$

Then

$$C(u_1,u_2) = \psi^{\leftarrow}(\psi(u_1)+\psi(u_2)), u_1, u_2 \in [0,1],$$

is a copula, called an Archimedean copula, with generator ψ.

In higher dimension, one should add more conditions. Function f is d-completely monotonic if it is continuous and has derivatives which alternate in sign, i.e. for all $k = 0, 1, ..., d$, $(-1)^k d^k f(t)/dt^k \geq 0$ for all t.

Definition 6. *Assume further that $\psi^{\leftarrow}$ is d-completely monotonic, then*

$$C(u_1, ..., u_n) = \psi^{\leftarrow}(\psi(u_1) + ... + \psi(u_d)), u_1, ..., u_n \in [0, 1],$$

is a copula, called an Archimedean copula, with generator ψ.

Archimedean copulas can also be characterized through Kendall's cdf, K,

$$K(t) = \mathbb{P}(C(U_1, ..., U_d) \leq t), t \in [0, 1].$$

where $\boldsymbol{U} = (U_1, ..., U_d)$ has cdf C. Note that $K(t) = t - \lambda(t)$ where $\lambda(t) = \psi(t)/\psi'(t)$. And conversely ψ is

$$\psi(u) = \psi(u_0) \exp\left(\int_{u_0}^{u} \frac{1}{\lambda(t)} dt\right) \text{ pour } 0 < u_0 < 1.$$

The frailty representation: if the X_i's are $\mathcal{E}(\lambda_i\Theta)$ distributed, and that, given Θ the X_i's are independent, then

$$\begin{aligned}
\overline{F}(\boldsymbol{x}) &= \mathbb{P}(\boldsymbol{X} > \boldsymbol{x}) = \mathbb{E}(\mathbb{P}(\boldsymbol{X} > \boldsymbol{x}|\Theta)) \\
&= \mathbb{E}(\mathbb{P}(X_1 > x_1|\Theta) \cdot \ldots \cdot \mathbb{P}(X_d > x_d|\Theta)) \\
&= \mathbb{E}(\exp(-\Theta \cdot (\log \mathbb{P}(X_1 > x_1))) \cdot \ldots \cdot \exp(-\Theta \cdot (\log \mathbb{P}(X_d > x_d)))) \\
&= \phi(-\log \mathbb{P}(X_1 > x_1) - \ldots - \log \mathbb{P}(X_d > x_d)) \\
&= \phi(\phi^{\leftarrow}(\overline{F}_1(x_1)), \ldots, \phi^{\leftarrow}(\overline{F}_d(x_d))),
\end{aligned}$$

where $\phi(t) = \mathbb{E}(e^{-t\Theta})$ is the Laplace transform of Θ.

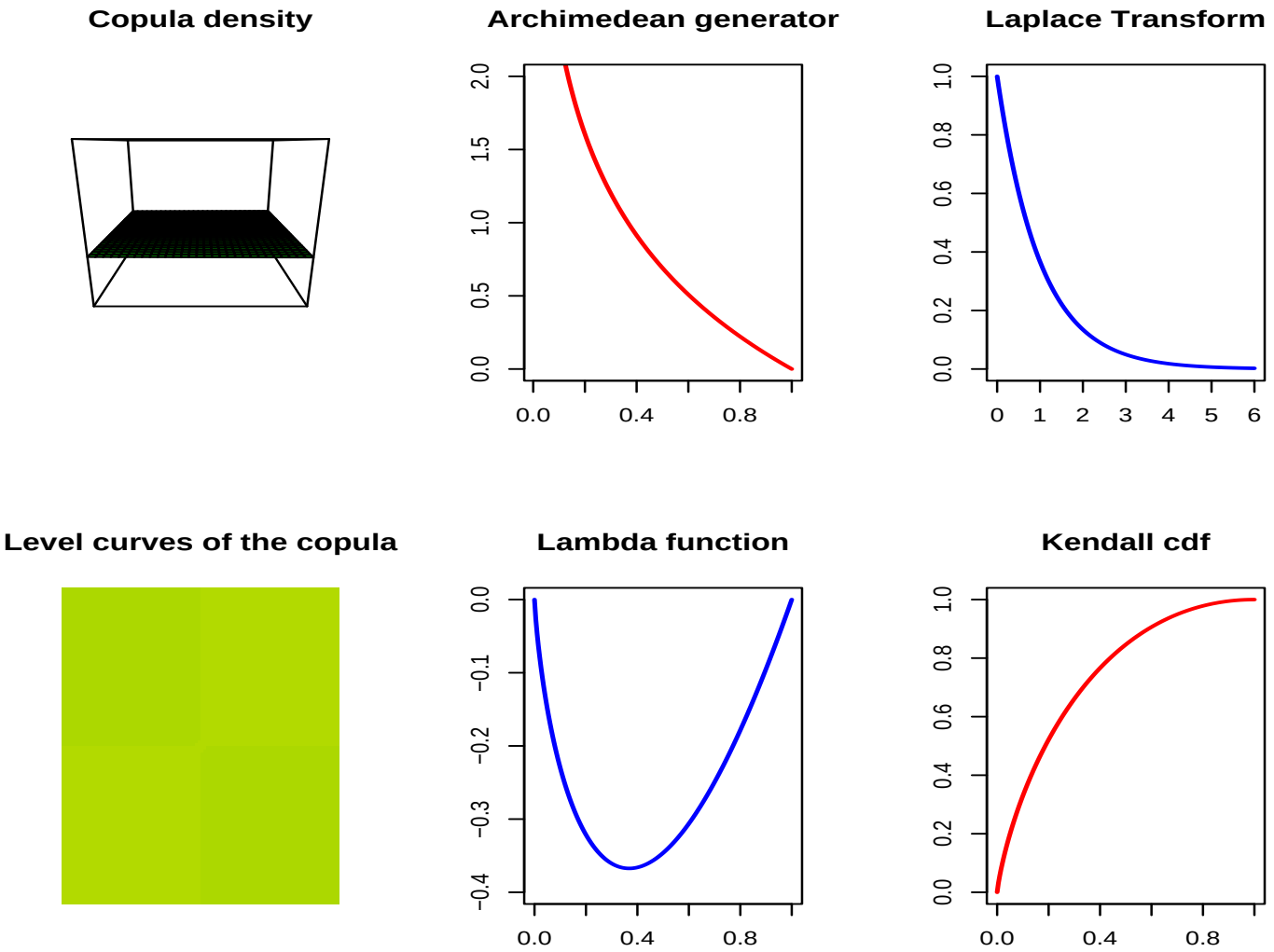


Figure 4: (Independent) Archimedean copula ($C = C^{\perp}, \psi(t) = -\log t$).

If Θ is Gamma distributed, one gets Clayton's copula (Figure 5), with parameter $\alpha \in [0, \infty)$ has generator

$$\psi(x; \alpha) = \frac{x^{-\alpha} - 1}{\alpha}$$

if $0 < \alpha < \infty$, with the limiting case $\psi(x; 0) = -\log(x)$, for any $0 < x \leq 1$.

The associated copula is

$$C(u_1, ..., u_d; \alpha) = (u_1^{-\alpha} + ... + u_d^{-\alpha} - (d-1))^{-1/\alpha}$$

if $0 < \alpha < \infty$, with the limiting case $C(\boldsymbol{u}; 0) = C^{\perp}(\boldsymbol{u})$, for any $\boldsymbol{u} \in (0, 1]^d$.

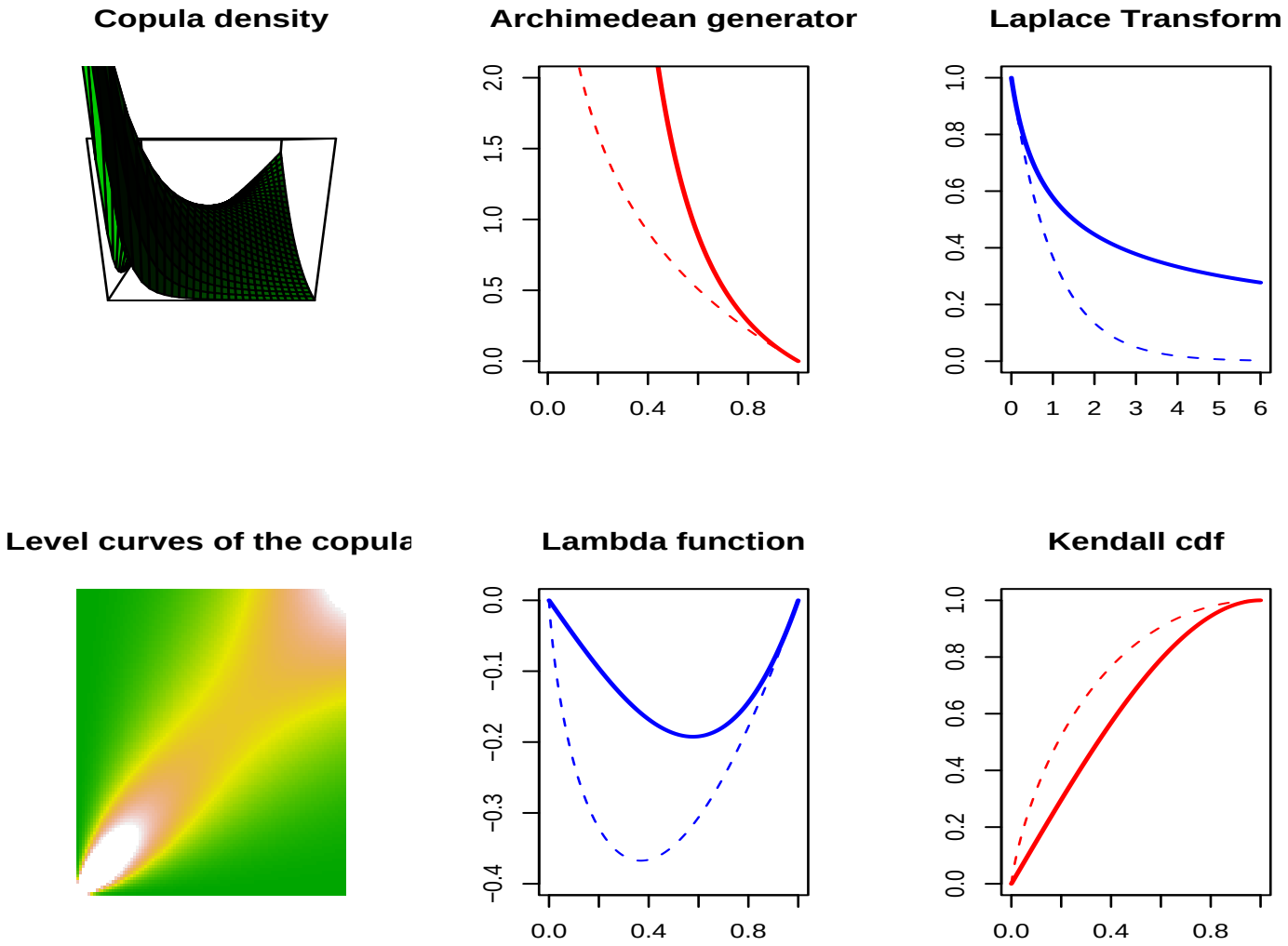


Figure 5: Clayton's copula.

If Θ is positive, stable, $\phi(t) = \exp(-t^{-1/\alpha})$, one gets Gumbel's copula (Figure 6), with parameter $\alpha \in [1, \infty)$ has generator

$$\psi(x; \alpha) = (-\log x)^{\alpha}$$

if $1 \leq \alpha < \infty$, with the limiting case $\psi(x; 0) = -\log(x)$, for any $0 < x \leq 1$.

The associated copula is

$$C(u_1, ..., u_d; \alpha) = -\frac{1}{\alpha} \log \left(1 + \frac{(e^{-\alpha u_1} - 1) ... (e^{-\alpha u_d} - 1)}{e^{-\alpha} - 1} \right),$$

if $1 \leq \alpha < \infty$, for any $\boldsymbol{u} \in (0, 1]^d$.

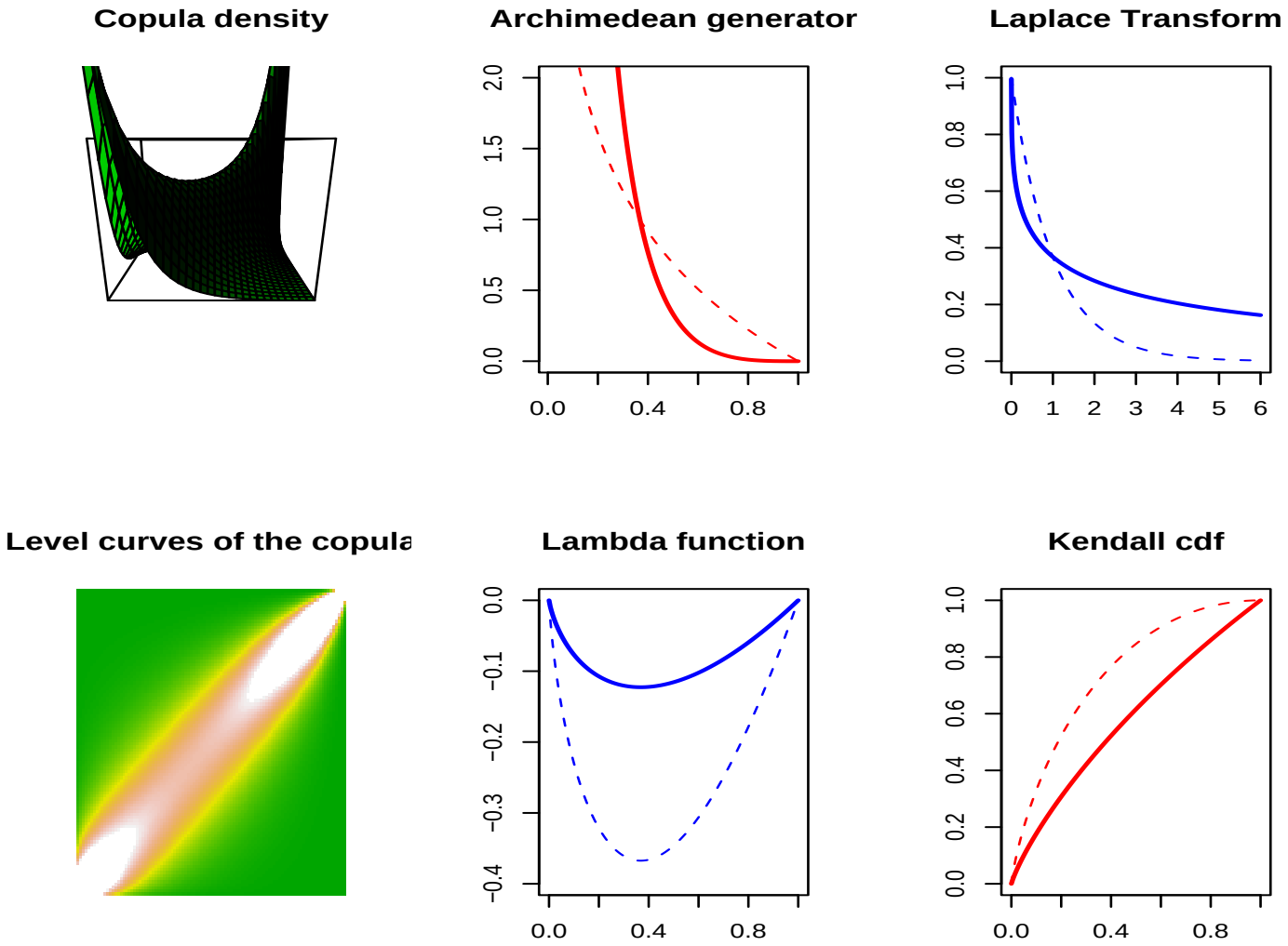


Figure 6: Gumbel's copula.

Frank's copula (Figure 7), with parameter $\alpha \in \mathbb{R}$ has generator

$$\psi(x;\alpha) = -\log \frac{e^{-\alpha x} - 1}{e^{-\alpha} - 1}$$

if $\alpha \neq 0$, for any $0 < x \leq 1$.

The associated copula is

$$C(u_1, ..., u_d; \alpha) = \exp\left(-\left[(-\log u_1)^\alpha + ... + (-\log u_d)^\alpha\right]^{1/\alpha}\right),$$

if $\alpha \neq 0$, with the limiting case $C(\boldsymbol{u}; 0) = C^\perp(\boldsymbol{u})$, for any $\boldsymbol{u} \in (0,1]^d$.

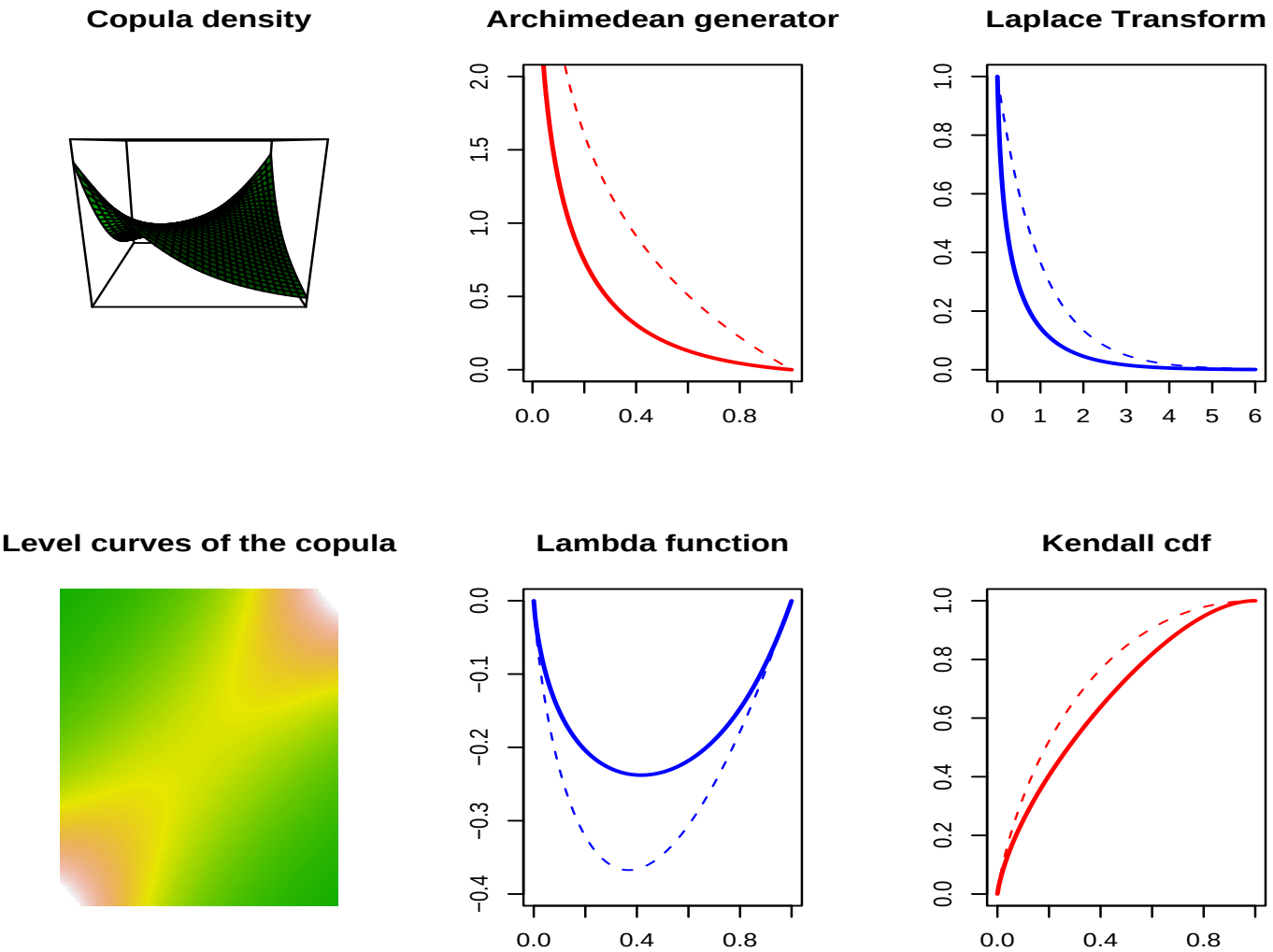


Figure 7: Frank's copula.

Agenda of the talk

- Presentation of conditional copulae,
- Some limiting results, with applications in credit risk,
- Conditional copulae of Archimedean copulae,
- A multivariate version of Pickands-Balkema-de Haan's theorem,
- Quantifying tail dependence using conditional concordance measures,
- Nonparametric estimation of copula densities,

Conditional copulae

Let $\boldsymbol{U} = (U_1, ..., U_n)$ be a random vector with uniform margins, and distribution funcion C. Let $C_{\boldsymbol{r}}$ denote the copula of random vector

$$(U_1, ..., U_n) | U_1 \leq r_1, ..., U_d \leq r_d, \tag{3}$$

where $r_1, ..., r_d \in (0, 1]$.

If $F_{i|\boldsymbol{r}}(\cdot)$ denotes the (marginal) distribution function of U_i given $\{U_1 \leq r_1, ..., U_i \leq r_i, ..., U_d \leq r_d\} = \{\boldsymbol{U} \leq \boldsymbol{r}\}$,

$$F_{i|\boldsymbol{r}}(x_i) = \frac{C(r_1, ..., r_{i-1}, x_i, r_{i+1}, ..., r_d)}{C(r_1, ..., r_{i-1}, r_i, r_{i+1}, ..., r_d)},$$

and therefore, the conditional copula is

$$C_{\boldsymbol{r}}(\boldsymbol{u}) = \frac{C(F^{\leftarrow}_{1|\boldsymbol{r}}(u_1), ..., F^{\leftarrow}_{d|\boldsymbol{r}}(u_d))}{C(r_1, ..., r_d)}. \tag{4}$$

Invariance properties

Theorem 7. *The application* $\boldsymbol{r} \to C_{\boldsymbol{r}}$ *is continuous (w.r.t. the* $\|\cdot\|_\infty$ *norm).*

Definition 8. *A copula is said to be invariant by truncature if* $C_{\boldsymbol{r}} = C$ *for all* $\boldsymbol{r} \in (0,1]^d$.

Theorem 9. *The only absolutely continuous copulae by truncature are Clayton's copulae, including the limiting case of independence.*

The comonotonic copula ($\alpha \to \infty$) is also invariant by truncature.

A weaker condition of invariance can be considered. Set

$$\mathcal{D} = \{(r_1(t), \ldots, r_d(t)), t \geq 0\},$$

where $r_i : [0, \infty) \to (0, 1]$ with $r_i(t) \to 0$ as $t \to \infty$, and $r_i(0) = 1$.

Definition 10. *A copula is said to be invariant by truncature under direction $\mathcal{D}$ if $C_{\boldsymbol{r}(t)} = C$ for all $t \geq 0$.*

Marshall & Olkin's copula is invariant when $\mathcal{D}$ is the discontinuity curve,

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

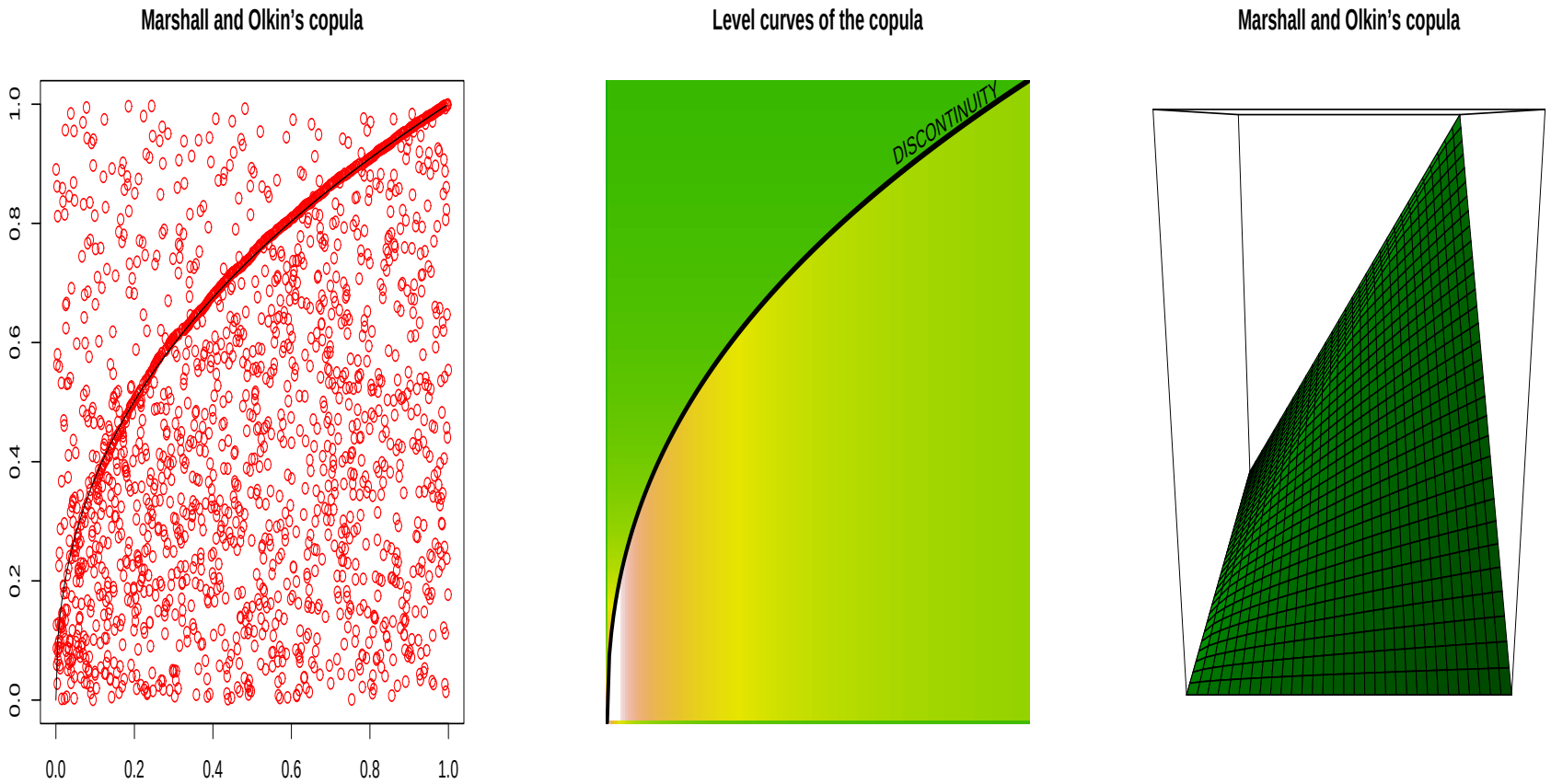


Figure 8: Marshall and Olkin's copula.

Regular variation and limiting results

Definition 11. *A real valued function r is regularly varying with index $\theta \in \mathbb{R}$ if $r(tx)/r(t) \to x^\theta$ as $t \to \infty$, and will be denoted $r \in \mathcal{R}_\theta$.*

Proposition 12. *A copula C_0 is limiting under direction $\mathcal{D}$ where the r_i's are $\mathcal{R}_{-\alpha_i}$ for some $\alpha_i > 0$, if and only if C_0 is an invariant copula under direction $\mathcal{D} = \{(t^{-\alpha_1}, \ldots, t^{-\alpha_d}), t \geq 0\}$.*

Extension of Juri & Wüthrich (2004) for nonsymmetric copulas: more general notion of regular variation in dimension 2.

de Haan, Omey & Resnick (1984): $f : [0, \infty)^2 \to [0, \infty)$ is regularly varying under direction (r, s) if there exists λ such that

$$\lim_{t \to \infty} \frac{f(r(t)x, s(t)y)}{f(r(t), s(t))} = \lambda(x, y) \text{ for all } x, y \in [0, \infty).$$

Proposition 13. *Assume that r and s are $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ respectively, with $\alpha, \beta > 0$, and that f is regularly varying under direction (r, s). Then λ satisfies functional equation*

$$\lambda(t^\alpha x, t^\beta y) = t^\theta \lambda(x, y),$$

for some $\theta > 0$, for any $x, y, t > 0$, and the general solution is

$$\lambda(x,y) = \begin{cases} x^{\theta/\alpha}\kappa(yx^{-\beta/\alpha}) & \text{if } x \neq 0 \\ cy^{\theta/\beta} & \text{if } x = 0 \text{ et } y \neq 0 \\ 0 & \text{if } x=0 \text{ et } y = 0, \end{cases}$$

where c is a positive constant, and $\kappa : [0, \infty) \to [0, \infty)$.

Let α, β, θ be positive constant, and P, Q cumulative distribution functions on $[0,1]$. Consider repartition function H defined on $[0,1]^2$ as

$$H(x,y) = x^{\theta/\alpha} h(yx^{-\beta/\alpha}), \text{ where } h(t) = \begin{cases} Q(t) & \text{if } t \in [0,1] \\ t^{\theta/\beta} P(t^{-\alpha/\beta}) & \text{if } t \in (1,\infty) \end{cases}$$

Let $\Gamma_{P,Q,\alpha,\beta,\theta}$ the associated copula, defined as

$$\Gamma_{P,Q,\alpha,\beta,\theta}(u,v) = \begin{cases} Q^{\leftarrow}(v)^{\theta/\beta} P(P^{\leftarrow}(u) Q^{\leftarrow}(v)^{-\alpha/\beta}), \text{ if } P^{\leftarrow}(u)^{\beta} \leq Q^{\leftarrow}(v)^{\alpha} \\ P^{\leftarrow}(u)^{\theta/\alpha} Q(P^{\leftarrow}(u)^{-\beta/\alpha} Q^{\leftarrow}(v)), \text{ if } P^{\leftarrow}(u)^{\beta} > Q^{\leftarrow}(v)^{\alpha} \end{cases}$$

Theorem 14. *If survival distributions $1 - F_X$ and $1 - F_Y$ are regularly varying with indices $\alpha, \beta \geq 0$ respectively, so that C^* is regularly varying at $(0,0)$ under direction $(1 - F_X(\cdot), 1 - F_Y(\cdot))$ then, there exists $\theta > 0$, P and Q two cdfs on $[0,1]$ such that the limiting copula is $\Gamma_{P,Q,\alpha,\beta,\theta}$.*

Applications in credit risk models

Default time defined using a cox process with intensity $(\lambda_s)_{s\geq 0}$ as

$$\tau = \inf\left\{t>0, \int_0^t \lambda_s ds \geq Z\right\},$$

where $Z \sim \mathcal{E}(1)$, see Lando (1998). Equivalently, $\tau = \inf\{t>0, \gamma_t \leq U\}$ where $\gamma_t = \int_0^t \lambda_s ds$ is called countdown process.

Then $\mathbb{P}(\tau_1 \leq t_1, \tau_2 \leq t_2 | \mathcal{G}_\infty) = C^*(\gamma_1(t_1), \gamma_2(t_2))$, where C is the cdf of (U_1, U_2).

For stress-scenarios, one needs conditional distributions such as

$$\mathbb{P}(\tau_1 \leq t_1, \tau_2 \leq t_2 | \mathcal{G}_\infty \wedge \{t_1 \leq T, t_2 \leq T\})$$

the associated copula is $(C_{1-\boldsymbol{\gamma}(T)})^*$.

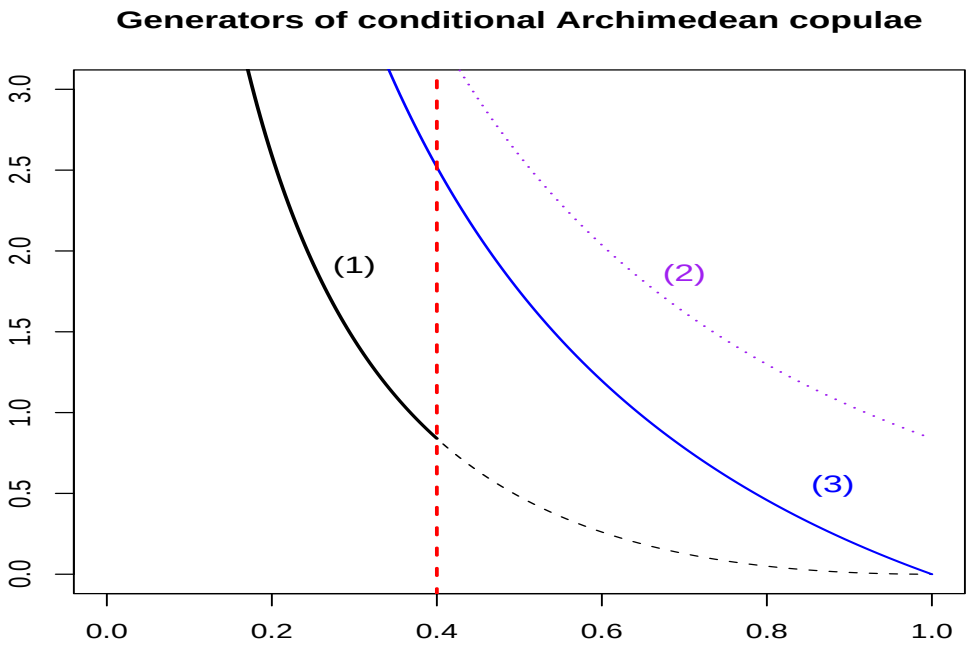
There are many perspectives (and work in progress) to go further

- use this conditional copula in the pricing of joint life-insurance and multiproduct reinsurance,
- more theoretical results on (conditional) dependence orderings (Colangelo, Shaked & Scarsini (2005)),
- investigate the link between the conditional copula and multivariate survival analysis (Bassan & Spizzichino (2005)).

Conditional dependence for Archimedean copulae

Proposition 15. *The class of Archimedean copulae is stable by truncature.*

More precisely, if $\boldsymbol{U}$ has cdf C, with generator ψ, $\boldsymbol{U}$ given $\{\boldsymbol{U} \leq \boldsymbol{r}\}$, for any $\boldsymbol{r} \in (0,1]^d$, will also have an Archimedean generator, with generator $\psi_{\boldsymbol{r}}(t) = \psi(tc) - \psi(c)$ where $c = C(r_1, ..., r_d)$.



On sequences of Archimedean copulae

Extension of results du to Genest & Rivest (1986),

Proposition The five following statements are equivalent,

(i) $\lim_{n\to\infty} C_n(u,v) = C(u,v)$ for all $(u,v) \in [0,1]^2$,

(ii) $\lim_{n\to\infty} \psi_n(x)/\psi_n'(y) = \psi(x)/\psi'(y)$ for all $x \in (0,1]$ and $y \in (0,1)$ such that ψ' such that is continuous in y,

(iii) $\lim_{n\to\infty} \lambda_n(x) = \lambda(x)$ for all $x \in (0,1)$ such that λ is continuous in x,

(iv) there exists positive constants κ_n such that $\lim_{n\to\infty} \kappa_n\psi_n(x) = \psi(x)$ for all $x \in [0,1]$,

(v) $\lim_{n\to\infty} K_n(x) = K(x)$ for all $x \in (0,1)$ such that K is continuous in x.

Proposition The four following statements are equivalent

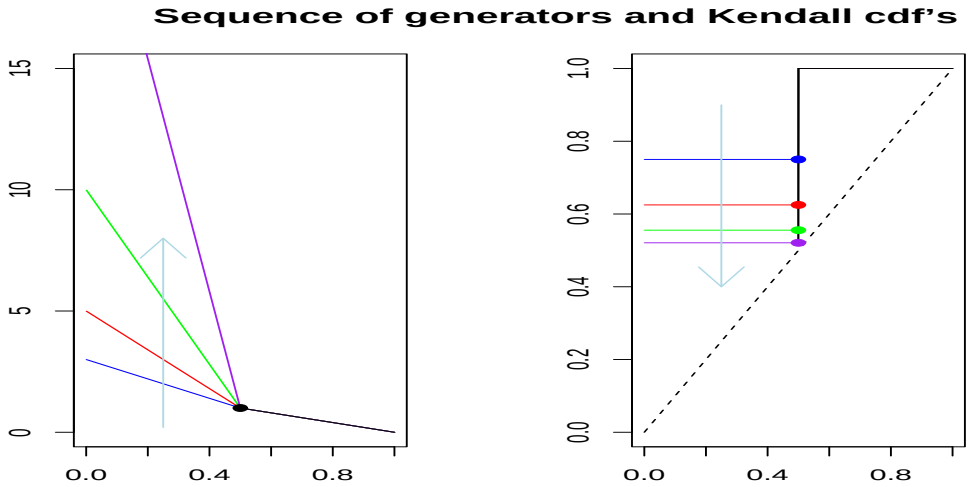
(i) $\lim_{n\to\infty} C_n(u,v) = C^+(u,v) = \min(u,v)$ for all $(u,v) \in [0,1]^2$,

(ii) $\lim_{n\to\infty} \lambda_n(x) = 0$ for all $x \in (0,1)$,

(iii) $\lim_{n\to\infty} \psi_n(y)/\psi_n(x) = 0$ for all $0 \leq x < y \leq 1$,

(iv) $\lim_{n\to\infty} K_n(x) = x$ for all $x \in (0,1)$.

Note that one can get non Archimedean limits,



Archimedean copulae in (lower) tails

Proposition 16. *Let C be an Archimedean copula with generator ψ, and $0 \leq \alpha \leq \infty$. If $C(\cdot,\cdot;\alpha)$ denote Clayton's copula with parameter α.*

(i) $\lim_{u\to 0} C_u(x,y) = C(x,y;\alpha)$ *for all* $(x,y) \in [0,1]^2$;

(ii) $-\psi' \in \mathcal{R}_{-\alpha-1}$.

(iii) $\psi \in \mathcal{R}_{-\alpha}$.

(iv) $\lim_{u\to 0} u\psi'(u)/\psi(u) = -\alpha$.

If $\alpha = 0$ (tail independence),

$$(i) \iff (ii) \Longrightarrow (iii) \iff (iv),$$

and if $\alpha \in (0,\infty]$ (tail dependence),

$$(i) \iff (ii) \iff (iii) \iff (iv).$$

Proposition 17. *There exists Archimedean copulae, with generators having continuous derivatives, slowly varying such that the conditional copula does not convergence to the independence.*

Generator ψ integration of a function piecewise linear, with knots $1/2^k$,

If $-\psi' \in \mathcal{R}_{-1}$, then $\psi \in \Pi_g$ (de Haan class), and not $\psi \notin \mathcal{R}_0$.

This generator is slowly varying, with the limiting copula is not $C^\perp$.

Lower tails study, in higher dimension

Tail dependence in lower tails

Proposition 18. *Let C be an Archimedean copula with generator ψ. If $\psi \in \mathcal{R}_{-\theta}$ with $\theta \in [-\infty, 0)$, if J is a nonempty set of $\{1, \ldots, d\}$ and if $0 < y_j < \infty$ for all $j \in J$, then, for all $(x_1, \ldots, x_d) \in (0, \infty)^d$,*

$$\lim_{s \to 0} \mathbb{P}(\forall i = 1, \ldots, d : U_i \leq s x_i \mid \forall j \in J : U_j \leq s y_j) \tag{5}$$

$$= \left[\frac{\sum_{j \in J^c} x_j^{-\theta} + \sum_{j \in J} \{\min(x_j, y_j)\}^{-\theta}}{\sum_{j \in J} y_j^{-\theta}} \right]^{-1/\theta}.$$

Note that the associated copula is Clayton's.

Tail **in**dependence in lower tails

Lemma 19. *Let C denote an Archimedean copula with generator ψ. If ψ is slowly varying at 0, then $C(s,\ldots,s)=o(s)$, and therefore* $\log\{\psi(s)\}/\log(s)\to 0$ *as* $s\to 0$.

Lemma 20. *If further*

$$\lim_{t\to\infty}\frac{D(\log\psi^{\leftarrow})(dt)}{D(\log\psi^{\leftarrow})(t)}=\frac{1}{d\eta},$$

then $s\mapsto C(s,\ldots,s)$ is regularly varying at 0, with index $1/\eta$.

Proposition 21. *Let C denote an Archimedean copula with generator ψ. If $\phi=-1/D(\log\psi^{\leftarrow})$ is regularly varying with index $-\infty<\tau\leq 1$ and if $\phi(t)=o(t)$ as $t\to\infty$, then, for all nonempty subset J of $\{1,\ldots,d\}$ and for all $\boldsymbol{x}\in(0,\infty)^d$ and $(u_j)_{j\in J}\in(0,1]^{|J|}$,*

$$\mathbb{P}(\forall j\in J:U_j\leq su_jx_j;\forall j\in J^c:U_j\leq\psi^{\leftarrow}\{x_j^{-1}\phi(\psi(s))\}\mid\forall j\in J:U_j\leq sx_j)$$
$$\to\prod_{j\in J}u_j^{|J|^{-\tau}}\prod_{j\in J^c}\exp\left(-|J|^{-\tau}x_j^{-1}\right),\ \text{as } s\to 0.$$

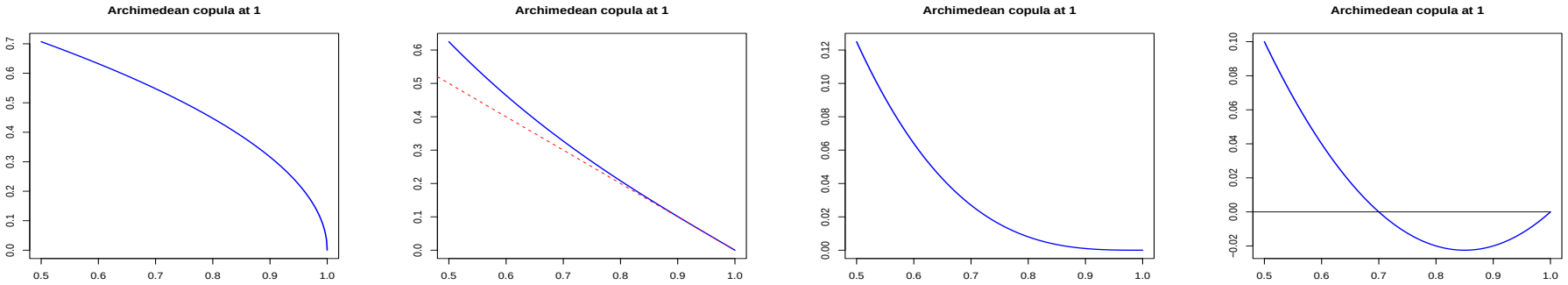
Upper tails study, in high dimension

Modeling tail independence is difficult (Ledford & Tawn (1997), or Draisma, Drees, Ferreira & de Haan (2004)).

Recall that $\psi(1)=0$, and therefore, using Taylor's expansion yields

$$\psi(1-s)=-sD\psi(1)+o(s) \text{ as } s\to 0.$$

And moreover, since ψ is convex, if $\psi(1-\cdot)$ is regularly varying with index θ, then necessarily $\theta\in[1,\infty)$. If if $(-D)\psi(1)>0$, then $\theta=1$ (but the converse is not true).



Proposition 22. *Let $\boldsymbol{U}$ be a random vector with cdf C, Archimedean with generator ψ. If $s \mapsto \psi(1-s)$ is regularly varying with index $\theta \in (1,\infty]$ at 0, then for all nonempty set J of $\{1,\dots,d\}$, for all $(x_1,\dots,x_d) \in (0,\infty)^d$ and all $(y_j)_{j\in J} \in (0,\infty)^{|J|}$,*

$$\mathbb{P}(\forall j = 1,\dots,d : U_j \geq 1 - sx_j \mid \forall j \in J : U_j \geq 1 - sy_j) \tag{6}$$
$$\to \quad \frac{r_d(z_1,\dots,z_d)}{r_{|J|}\left((y_j)_{j\in J}\right)} \text{ as } s \to 0,$$

where $z_j = min(x_j, y_j)$ when $j \in J$, and $z_j = x_j$ when $j \in J^c$, setting finally

$$r_k(u_1,\dots,u_k) = \begin{cases} \displaystyle\sum_{I\subset\{1,\dots,k\}:|I|\geq 1} (-1)^{|I|-1}\left(\sum_{i\in I} u_i^\theta\right)^{1/\theta} & \text{si } 1 < \theta < \infty, \\ \min(u_1,\dots,u_d) & \text{si } \theta = \infty, \end{cases}$$

for all integer k and all $(u_1,\dots,u_k) \in (0,\infty)^k$. (Note that this copula is max-stable: it is Gumbel's copula (the logistic copula)).

Proposition 23. *Let ψ be a d-dimensional generator, such that $\psi^{\leftarrow}$ is d times continuously differentiable, and $\boldsymbol{U}$ is a random vector with cdf C, the Archimedean copula induced by ψ. If $(-D)^d\psi^{\leftarrow}(0) < \infty$ then $(-D)\psi(1) > 0$. Let J be a nonempty set of $\{1,\dots,d\}$ such that J^c is nonempty. For all $\boldsymbol{v} \in (0,1]^d$,*

$$\mathbb{P}(\forall j \in J : U_j \geq 1 - sv_jx_j; \forall j \in J^c : U_j \leq v_j \mid \forall j \in J : U_j \geq 1 - sx_j)$$
$$\to \quad \frac{(-D)^{|J|}\psi^{\leftarrow}\left(\sum_{j\in J^c}\psi(v_j)\right)}{(-D)^{|J|}\psi^{\leftarrow}(0)} \prod_{j\in J} v_j \ \text{ as } s \downarrow 0.$$

Note that the conditional copula of $(U_j)_{j\in J^c}$ given $U_j \geq 1 - sx_j$ for all $j \in J$ tends to an Archimedean copula, with generator

$$\psi_{|J|}(\cdot) = \left(\frac{(-D)^{|J|}\psi^{\leftarrow}(\,\cdot\,)}{(-D)^{|J|}\psi^{\leftarrow}(0)}\right)^{\leftarrow}. \tag{7}$$

Proposition 24. *Let ψ be a d-dimensional generator, such that $\psi^{\leftarrow}$ is d times continuously differentiable. Set $f(s)=\psi(1-s)$. If $s^{-1}f(s)\to 0$ as $s\to 0$ and if $s\mapsto \mathcal{L}(s)=s(\mathrm{d}/\mathrm{d}s)\{s^{-1}f(s)\}$ is positive and slowly varying at 0, then function $g(s)=sf'(s)/f(s)-1$ is also positive and slowly varying, with $g(s)\to 0$ as $s\to 0$. If J is a set of $\{1,\ldots,d\}$ then, for all $\boldsymbol{x}\in(0,\infty)^d$ and $(y_j)_{j\in J}\in(0,\infty)^{|J|}$,*

$$\mathbb{P}(\forall i=1,\ldots,d: U_i\geq 1-sx_i \mid \forall j\in J: U_j\geq 1-sy_j)=\frac{r(z_1,\ldots,z_d)}{r((y_j)_{j\in J})}$$

as $s\downarrow 0$, where $z_j=\min(x_j,y_j)$ for all $j\in J$ and $z_j=x_j$ for $j\in J^c$, and where

$$\begin{aligned} r(x_1,\ldots,x_d) &= \sum_{I\subset\{1,\ldots,d\}:|I|\geq 1}(-1)^{|I|}\left(\textstyle\sum_I x_i\right)\log\left(\textstyle\sum_I x_i\right) \\ &= (d-2)!\int_0^{x_1}\cdots\int_0^{x_d}\left(\sum_{i=1}^d t_i\right)^{-(d-1)}dt_1\cdots dt_d. \end{aligned}$$

Multivariate extremes in $\mathbb{R}^2$

The standard approach in multivariate extremes is based on limiting distribution of maximum componentwise.

In Tiago de Olivera (1958), Geoffroy (1958), Sibuya (1961) or Resnick (1987) is proposed an extension of Fisher-Tippett's theorem, i.e. a nondegenerate distribution for a standardized version of $(X_{n:n}, Y_{n:n})$, the maximum componentwise, as $n \to \infty$ where $(X_1, Y_1), \ldots, (X_n, Y_n), ..$ is an iid sample.

In the univariate case, Pickands-Balkema-de Haan's theorem obtained limiting (nondegenerate) distribution for $X - u$ given $X > u$, when $u \to \infty$.

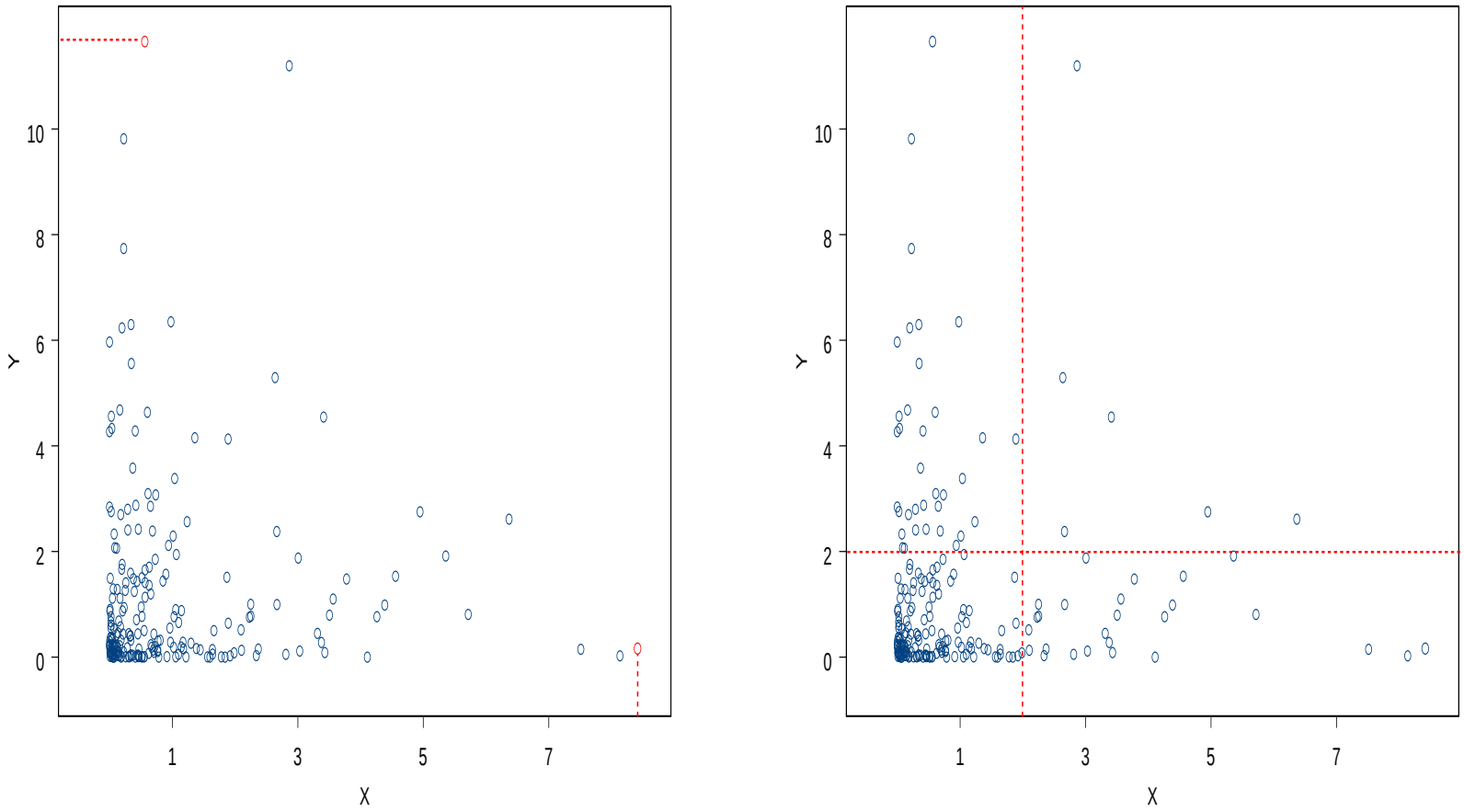


Figure 9: Defining multivariate extremes.

Pickands-Balkema-de Haan in $\mathbb{R}^2$

Two extensions will be considered,

$$(X, Y) \text{ given } \{X > F_X^{\leftarrow}(p) \text{ and } Y > F_Y^{\leftarrow}(p)\} \text{ as } p \to 1,$$

a quantile approach, and a level approach,

$$(X, Y) \text{ given } \{X > z \text{ and } Y > z\} \text{ as } z \to \infty.$$

Assume that X and Y are in the Fréchet domain of attraction, with parameters α and β, i.e. there exists $a(\cdot)$ and $b(\cdot)$ such that

$$\lim_{u\to\infty} 1 - \frac{1 - F_X(u + xa(u))}{1 - F_X(u)} = \lim_{u\to\infty} \mathbb{P}(X \leq u + a(u) \,|X > u) = G_\alpha(x),$$

$$\lim_{v\to\infty} 1 - \frac{1 - F_Y(v + yb(v))}{1 - F_Y(v)} = \lim_{v\to\infty} \mathbb{P}(Y \leq v + b(v) \,|Y > v) = G_\beta(y),$$

$$\text{where } G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp(-x) & \xi = 0, \end{cases}.$$

Pickands-Balkema-de Haan in $\mathbb{R}^2$ for quantiles

Theorem 25. *Assume that the survival copula of (X, Y), C^*, satisfies*

$$\lim_{u\to 0} \frac{C^*(xu, u)}{C^*(u, u)} = h(x) \text{ for all } x \geq 0, \tag{8}$$

for some continuous function $h : \mathbb{R}^+ \to \mathbb{R}^+$ with $h(x) > 0$ as $x > 0$. Then $h(0) = 0$, $h(1) = 1$, and there exists $\theta \in \mathbb{R}$ such that

$$h(x) = x^\theta h\left(\frac{1}{x}\right) \text{ pour tout } x > 0,$$

then, if $q_X(p) = F_X^{\leftarrow}(1-p)$ and $q_Y(p) = F_Y^{\leftarrow}(1-p)$,

$$\lim_{p\to 0} \mathbb{P}\left(\frac{X - F_X^{\leftarrow}(1-p)}{a\left(q_X(p)\right)} > x, \frac{Y - q_Y(p)}{b\left(q_Y(p)\right)} > y \middle| X > q_X(p), Y > q_Y(p) \right)$$

$$= (1+y)^{-\gamma} h\left(\frac{(1+x)^{-1/\alpha}}{(1+y)^{-1/\beta}} \right), \text{ where } \gamma = \frac{\theta}{\beta}.$$

Since the convergence is uniform, further

$$\lim_{p\to 0}\sup_{x,y}\left|\mathbb{P}\left(X-q_X(p)>x,Y-q_Y(p)>y\right|X>q_X(p),Y>q_Y(p)\right)$$
$$-(1+y)^{-\gamma}h\left(\frac{(1+x)^{-1/\alpha}}{(1+y)^{-1/\beta}}\right)|=0.$$

The associate copula is the survival copula of $C_h(x,y)=H(h^{\leftarrow}(x),h^{\leftarrow}(y))$, where $H(x,y)=y^{\theta}h(x/y)$ (cf Juri & Wüthrich (2004)).

Pickands-Balkema-de Haan in $\mathbb{R}^2$ for levels

Theorem 26. *Assume that the survival copula of (X, Y), C^*, is regularly varying under direction $(1 - F_X(\cdot), 1 - F_Y(\cdot))$, such that there exists λ satisfying*

$$\lim_{z\to\infty} \frac{C^*((1 - F_X(z))x, (1 - F_Y(z))y)}{C^*(1 - F_X(z), 1 - F_Y(z))} = \lambda(x, y).$$

From Proposition 13, there exists $\gamma > 0$ and function $h : [0, \infty) \to [0, \infty)$ such that $\lambda(x, y) = x^{\gamma/\alpha} h(yx^{-\beta/\alpha})$ if $x \neq 0$ and $\lambda(0, y) = cy^{\gamma/\beta}$ where c is a positive constant. Further,

$$\lim_{z\to\infty} \mathbb{P}\left(\frac{X - z}{a(z)} > x, \frac{Y - z}{b(z)} > y \middle| X > z, Y > z \right) = (1+y)^{-\gamma} h\left(\frac{(1 + x)^{-1/\alpha}}{(1 + y)^{-1/\alpha}} \right).$$

Since the convergence is uniform, we get

$$\lim_{z\to\infty} \sup_{x,y} |\mathbb{P}(X - z > x, Y - z | X > z, Y > z) = (1+y)^{-\gamma} h\left(\frac{(1 + x)^{-1/\alpha}}{(1 + y)^{-1/\alpha}} \right)| = 0.$$

There are many perspectives (and work in progress) to go further

- try to get a better understanding of the links between standard multivariate extremes and the one based on conditional copulae,
- statistical inference for multivariate extremes (e.g. estimate function h),

Concordance measures

Copula based dependence measure (scale free).

- Spearman's rho is defined as $\rho(X,Y) = 12\int_0^1\int_0^1 uvdC(u,v) - 3,$
- Kendall'tau is defined as $\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1.$

Conditional concordance measures

Definition 27. *Given X and Y two random variables, define upper tail rank correlations and Kendall's tau in upper tails respectively as*

$$\rho(u) = \rho((X,Y)|X > F_X^{\leftarrow}(u) \text{ and } Y > F_Y^{\leftarrow}(u)), u \in [0,1),$$

$$\tau(u) = \tau((X,Y)|X > F_X^{\leftarrow}(u) \text{ and } Y > F_Y^{\leftarrow}(u)), u \in [0,1).$$

On statistical inference for $\rho(u)$ and $\tau(u)$

Given a n-sample $(X_1, Y_1), ...(X_n, Y_n)$ consider natural estimates $\widehat{\rho}(u)$ and $\widehat{\tau}(u)$ based on the pseudo sample $(X_1, Y_1), ...(X_N, Y_N)$ such that $X_i > \widehat{F}_X^{\leftarrow}(u)$ and $Y_i > \widehat{F}_Y^{\leftarrow}(u)$.

Note that N is a random variable, $\mathcal{B}(m, \pi)$ where $m = [(n+1)(1-u)]$ and $\pi = C^*(1-u, 1-u)/(1-u)$.

Under the assumption of independence, given N, $\widehat{\rho}(u)$ and $\widehat{\tau}(u)$ are asymptotically normally distributed. Hence, as $n \to \infty$, the limiting distributions of $\widehat{\rho}(u)$ and $\widehat{\tau}(u)$ are mixture of Gaussian distributions.

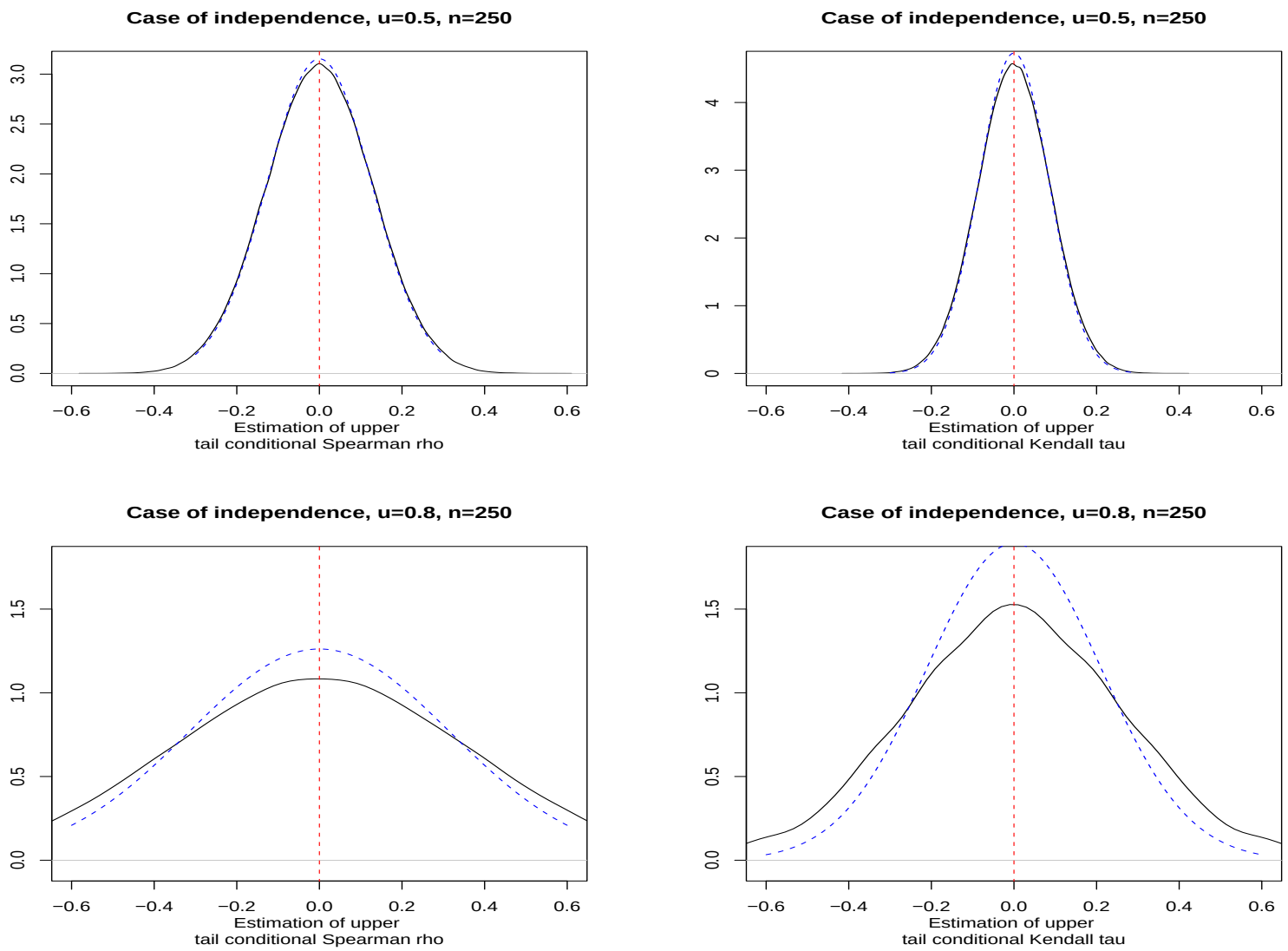


Figure 10: Distribution of $\widehat{\rho}_n(u)$ and $\widehat{\tau}_n(u)$.

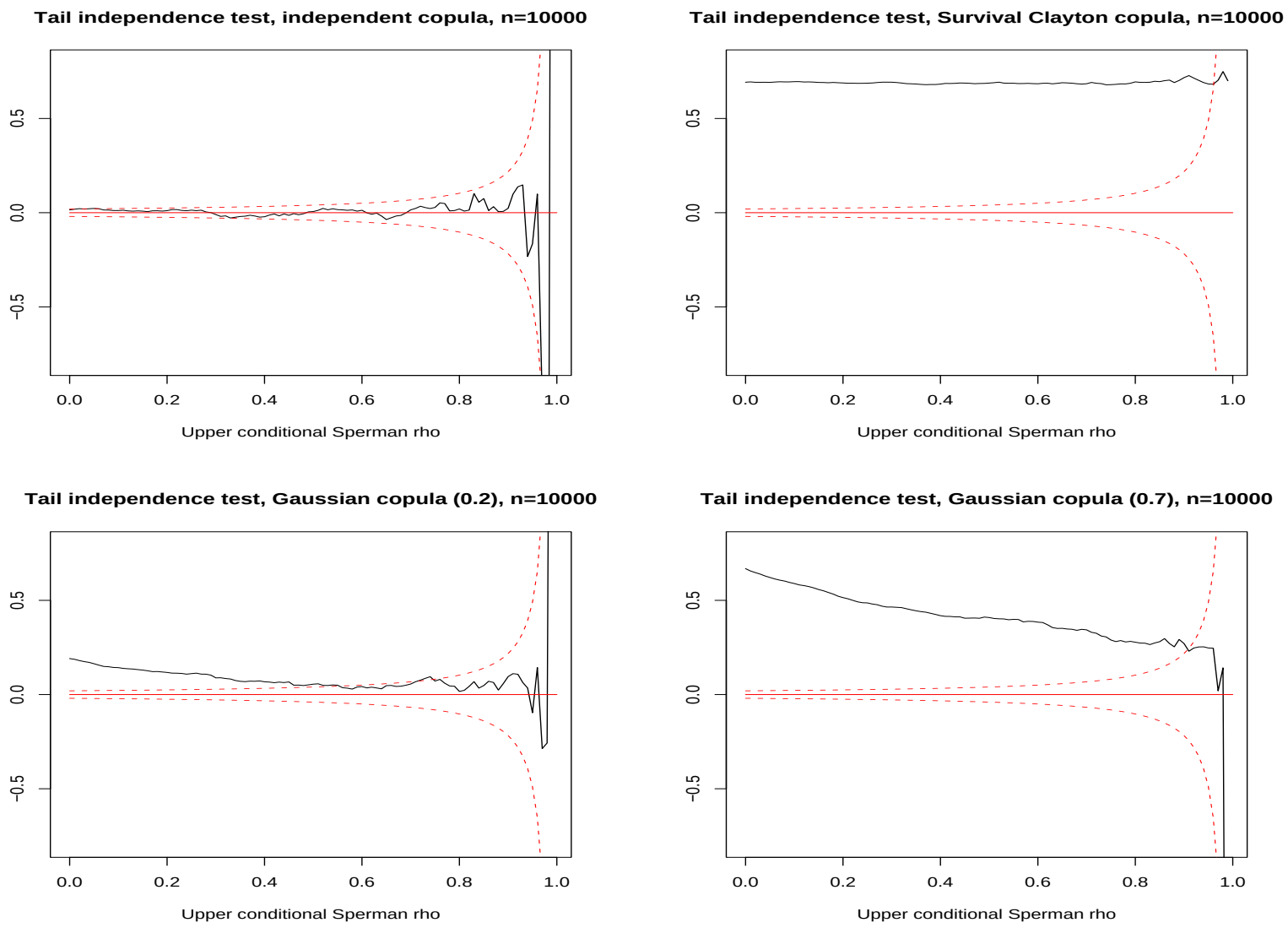


Figure 11: Evolution of $\widehat{\rho}_n(u)$, $n = 10,000$.

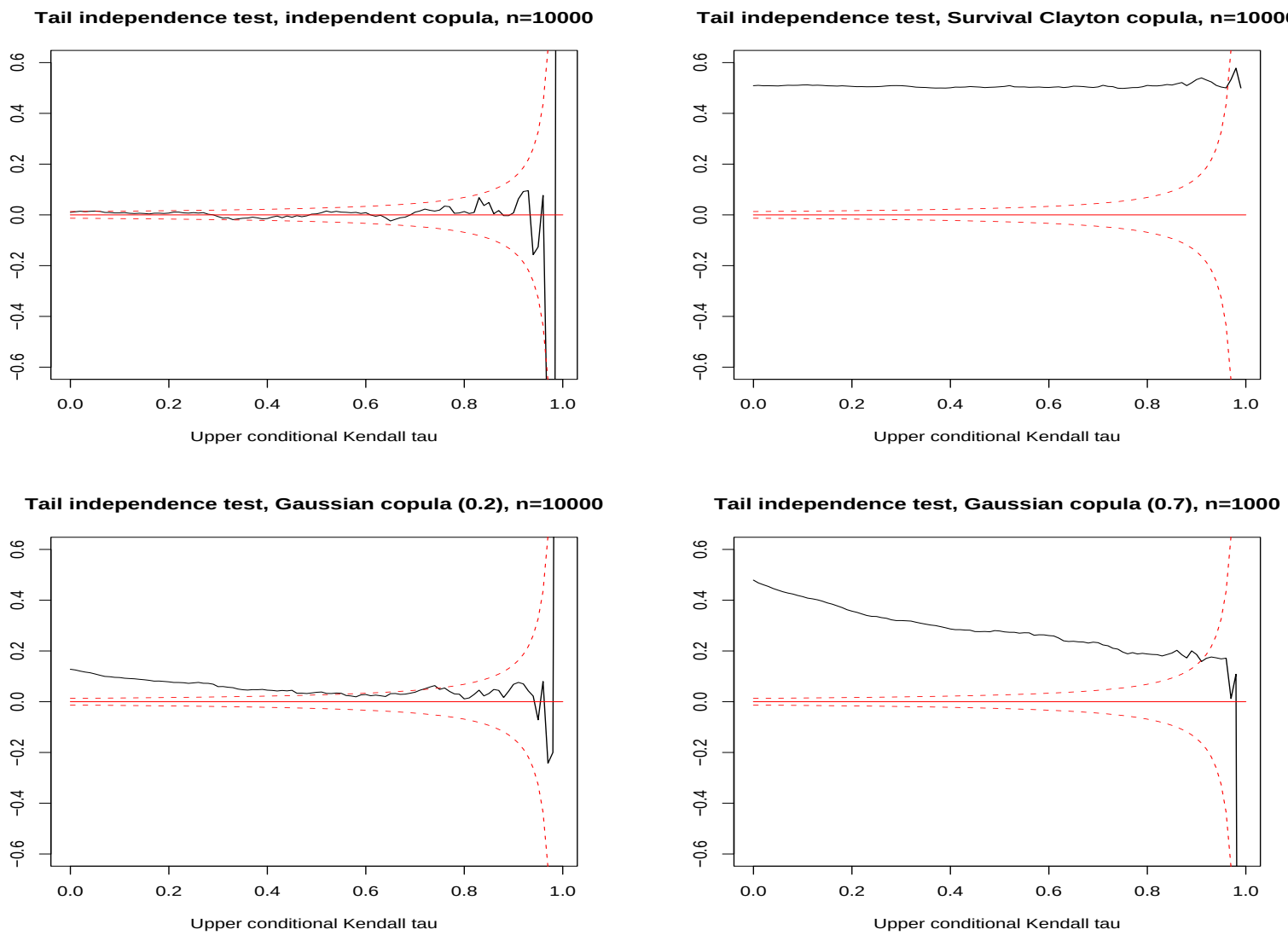


Figure 12: Evolution $\widehat{\tau}_n(u)$, $n = 10,000$.

There are perspectives to go further, since conditional correlations is a hot topic

- use of this tool in copula selection,
- statistical properties of the estimated correlation in the non-independent case.

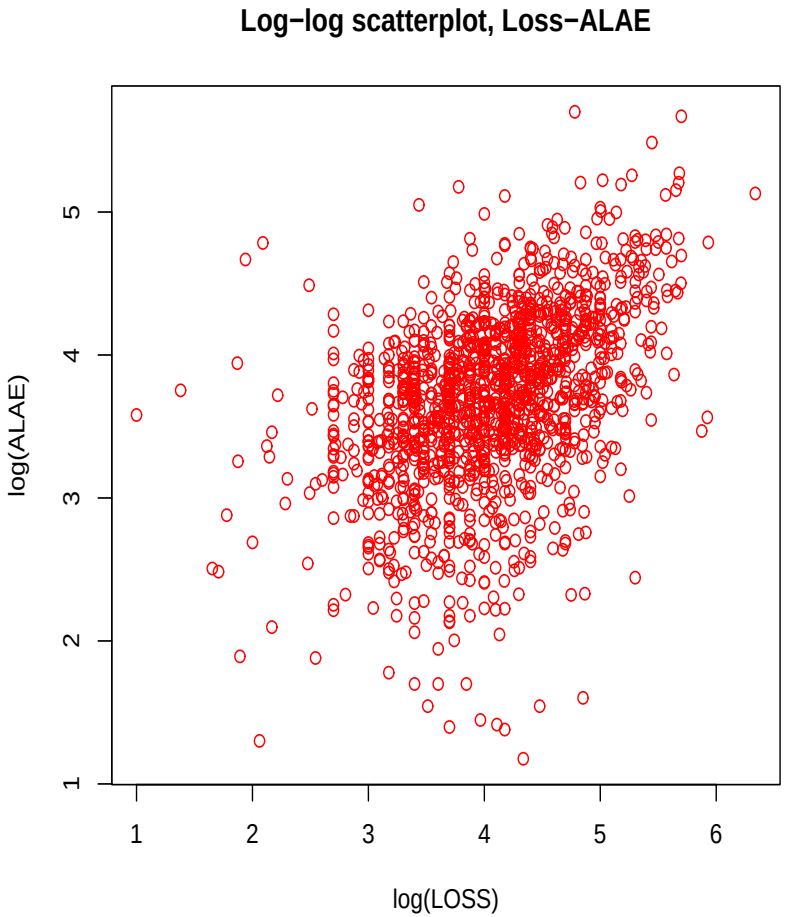
Estimation of a copula density, motivation

Example 28. *consider the following dataset, were the X_i's are loss amount (paid to the insured) and the Y_i's are allocated expenses. Denote by R_i and S_i the respective ranks of X_i and Y_i. Set $U_i = R_i/n = \hat{F}_X(X_i)$ and $V_i = S_i/n = \hat{F}_Y(Y_i)$.*

Figure 13 shows the log-log scatterplot $(\log X_i, \log Y_i)$, and the associate copula based scatterplot (U_i, V_i).

Figure 14 is simply an histogram of the (U_i, V_i), which is a nonparametric estimation of the copula density.

Note that the histogram suggests strong dependence in upper tails (the interesting part in an insurance/reinsurance context).



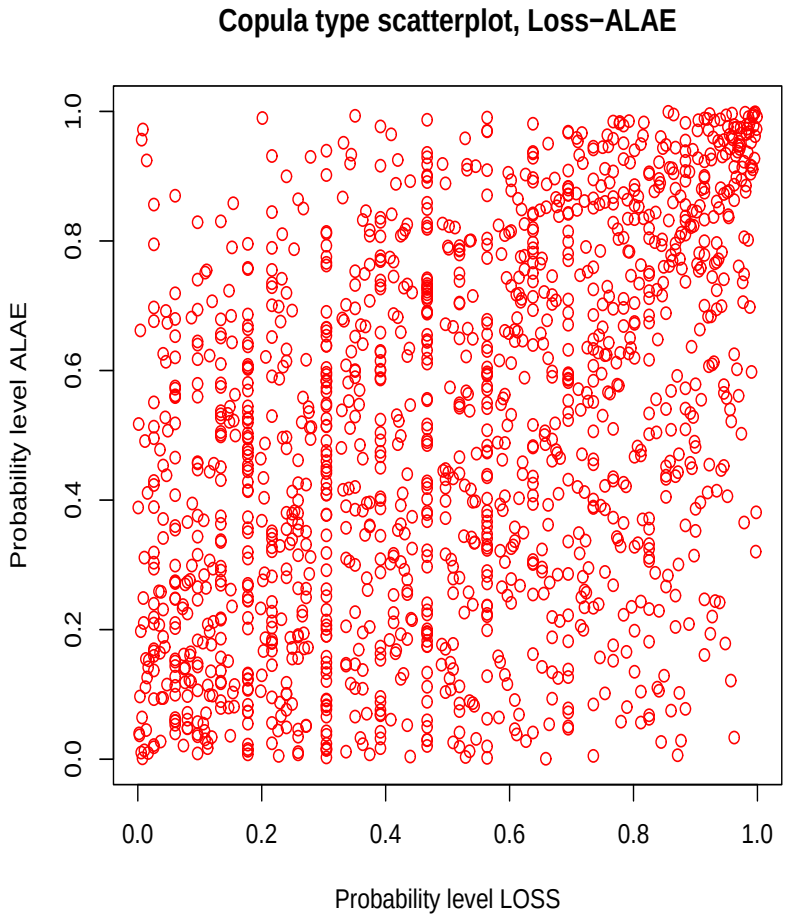


Figure 13: Loss-ALAE, scatterplots (log-log and copula type).

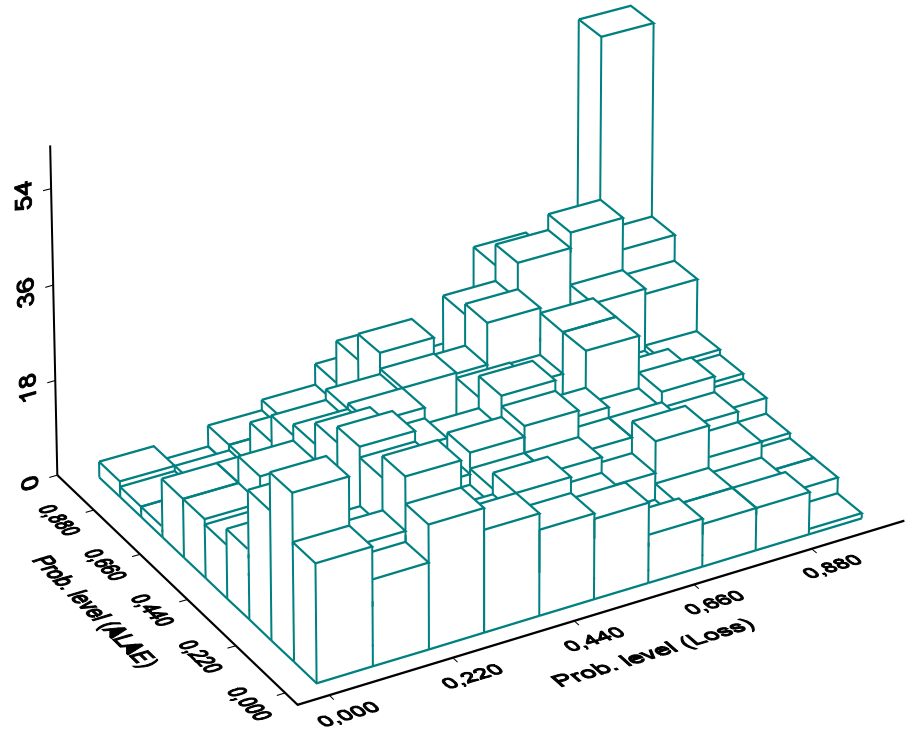


Figure 14: Loss-ALAE, histogram of copula type transformation.

Why nonparametrics, instead of parametrics ?

In parametric estimation, assume the the copula density c_θ belongs to some given family $\mathcal{C} = \{c_\theta, \theta \in \Theta\}$. The tail behavior will crucially depend on the tail behavior of the copulas in $\mathcal{C}$

Example 29. *Table below shows the probability that both X and Y exceed high thresholds ($X > F_X^{-1}(p)$ and $Y > F_Y^{-1}(p)$), for standard copula families, where parameter θ is such that $\tau = 0.5$.*

p	Clayton	Frank	Gaussian	Gumbel	Clayton*	max/min
0.9	1.935%	2.737%	4.737%	4.826%	5.668%	2.9
0.95	0.510%	0.784%	1.991%	2.300%	2.786%	5.4
0.99	0.021%	0.035%	0.2733%	0.442%	0.551%	25.8
0.999	0.000%	0.000%	0.016%	0.043%	0.054%	261.8

Basic notions on nonparametric estimation of densities

The basic idea to get an estimator of the density at some point x is to count how many observation are in the neighborhood of x (e.g. in $[x-h, x+h)$ for some $h>0$).

Therefore, consider the "*moving histogram*" or "*naive estimator*" as suggested by Rosenblatt (1956),

$$\widehat{f}(x) = \frac{1}{2nh}\sum_{i=1}^{n}\mathbb{I}(X_i \in [x-h, x+h)) = \widehat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$$

when other definitions of the neighborhood of x are considered, where K is a kernel function (e.g. $K(\omega) = \mathbb{I}(|\omega| \leq 1)/2$).

If X is a positive random variable, and K is symmetric,

$$\mathbb{E}(\widehat{f}(0,h)) = \frac{1}{2}f(0) + O(h)$$

Copula density estimation: the boundary problem

Let $(U_1, V_1), ..., (U_n, V_n)$ denote a sample with support $[0,1]^2$, and with density $c(u,v)$, which is assumed to be twice continuously differentiable on $(0,1)^2$.

If K denotes a symmetric kernel, with support $[-1,+1]$, then for all $(u,v) \in [0,1] \times [0,1]$, in any corners (e.g. $(0,0)$)

$$\mathbb{E}(\widehat{c}(0,0,h)) = \frac{1}{4} \cdot c(u,v) - \frac{1}{2}[c_1(0,0) + c_2(0,0)] \int_0^1 \omega K(\omega) d\omega \cdot h + o(h).$$

on the interior of the borders (e.g. $u = 0$ and $v \in (0,1)$),

$$\mathbb{E}(\widehat{c}(0,v,h)) = \frac{1}{2} \cdot c(u,v) - [c_1(0,v)] \int_0^1 \omega K(\omega) d\omega \cdot h + o(h).$$

and in the interior ($(u,v) \in (0,1) \times (0,1)$),

$$\mathbb{E}(\widehat{c}(u,v,h)) = c(u,v) + \frac{1}{2}[c_{1,1}(u,v) + c_{2,2}(u,v)] \int_{-1}^1 \omega^2 K(\omega) d\omega \cdot h^2 + o(h^2).$$

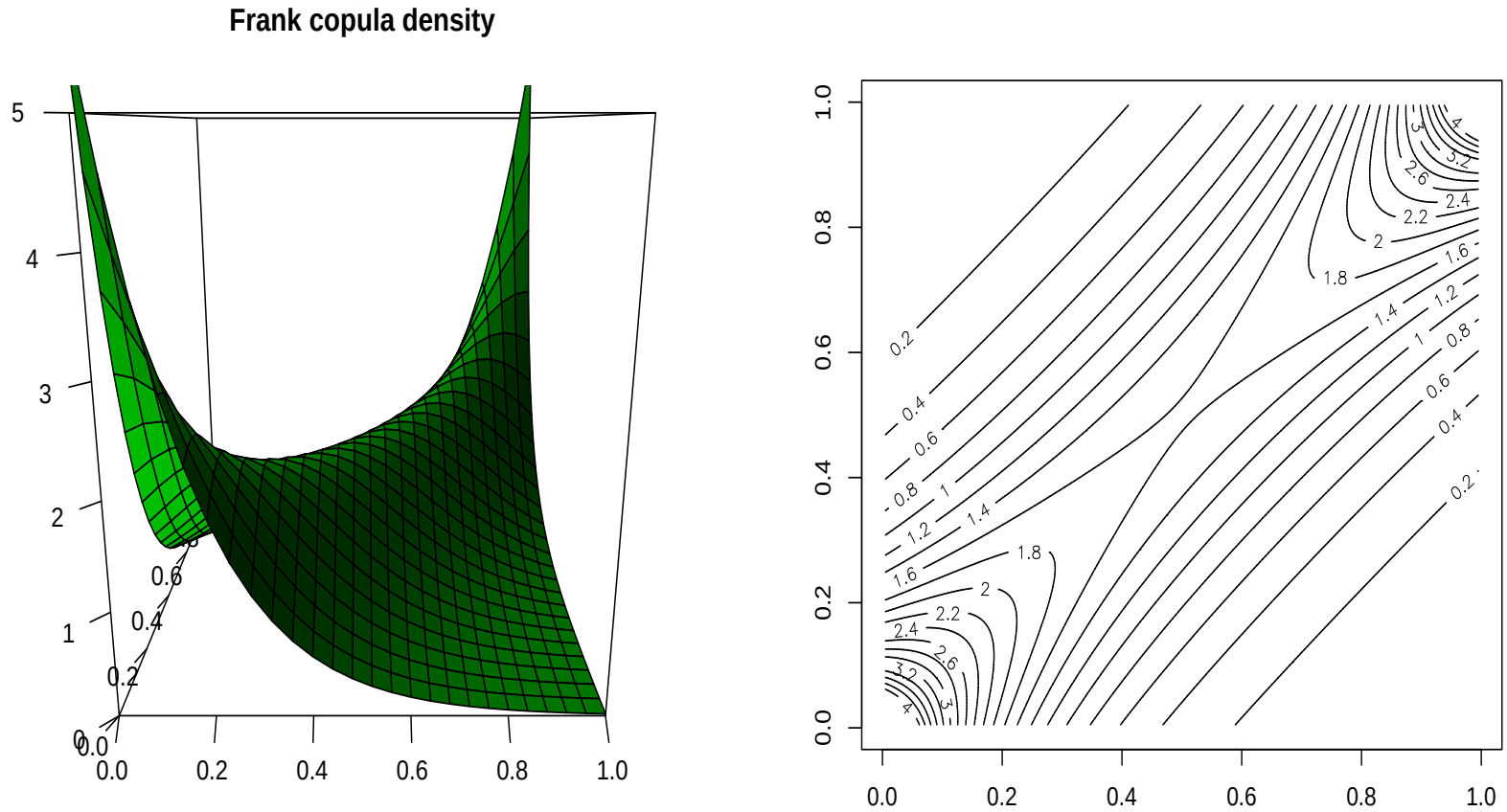


Figure 15: Theoretical density of Frank copula.

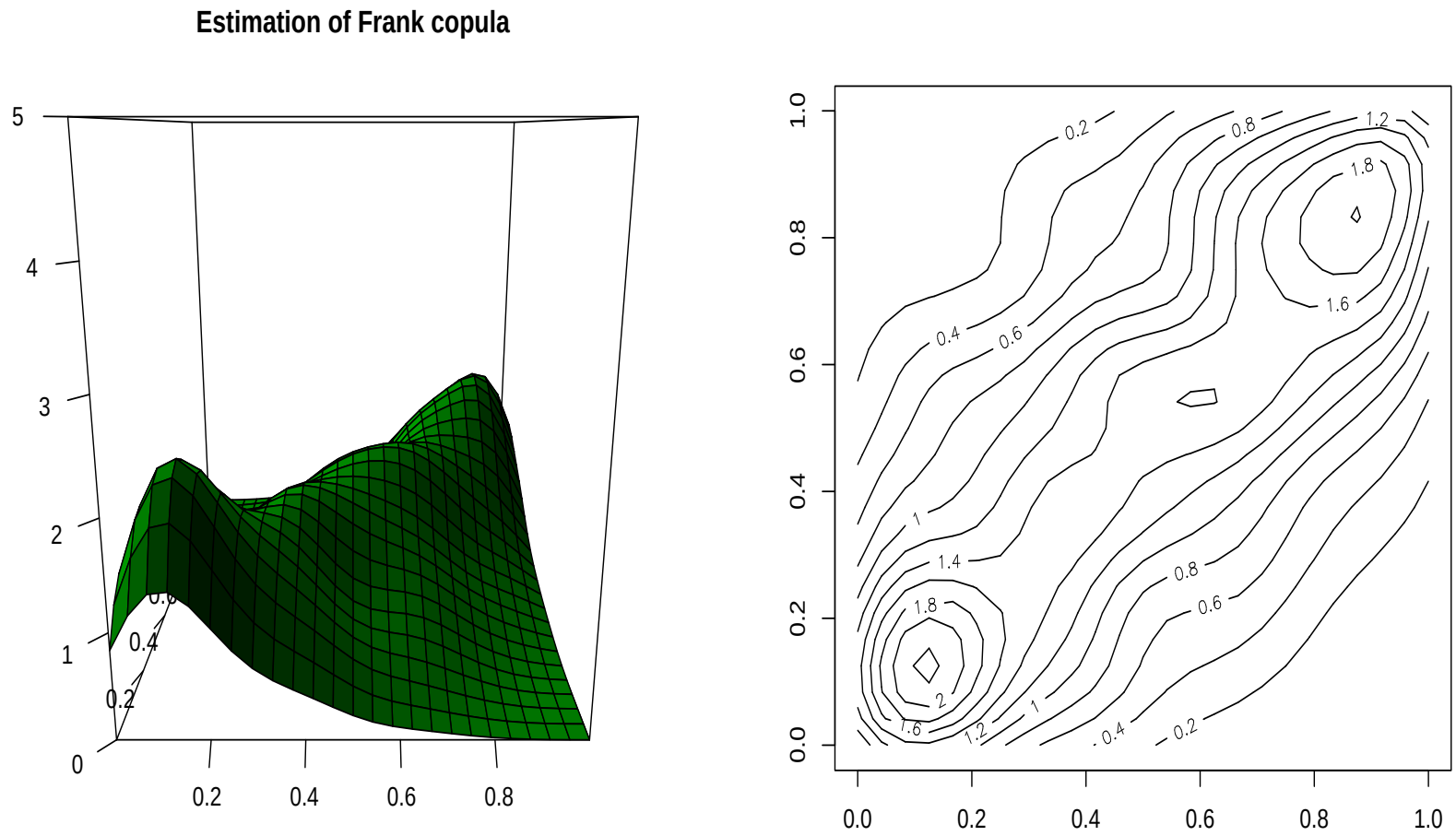


Figure 16: Estimation with standard Gaussian (independent) kernels.

How to get a proper estimation on the border

Several techniques have been introduce to get a better estimation on the border, in the univariate case

- boundary kernel (Müller (1991))
- mirror image modification (Deheuvels & Hominal (1989), Schuster (1985))
- transformed kernel (Devroye & Györfi (1981), Wand, Marron & Ruppert (1991))

In the particular case of densities on $[0, 1]$,

- Beta kernel (Brown & Chen (1999), Chen (1999, 2000)),

Transformed kernel technique

Set $(X_i, Y_i) = (G^{-1}(U_i), G^{-1}(V_i))$'s, where G is a strictly increasing distribution function, with a differentiable density, and consider a kernel estimate of the density. Since density f is continuous, twice differentiable, and bounded above, for all $(x, y) \in \mathbb{R}^2$, consider

$$\widehat{f}(x,y) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) K\left(\frac{y - Y_i}{h}\right).$$

Since $f(x,y) = g(x)g(y)c[G(x), G(y)]$ can be inverted in

$$c(u,v) = \frac{f(G^{-1}(u), G^{-1}(v))}{g(G^{-1}(u))g(G^{-1}(v))}, (u,v) \in [0,1] \times [0,1], \tag{9}$$

one gets, substituting $\widehat{f}$ in (9)

$$\widehat{c}(u,v) = \frac{1}{nh \cdot g(G^{-1}(u)) \cdot g(G^{-1}(v))} \sum_{i=1}^{n} K \quad \frac{G^{-1}(u) - G^{-1}(U_i)}{h}, \frac{G^{-1}(v) - G^{-1}(V_i)}{h} \quad , \tag{10}$$

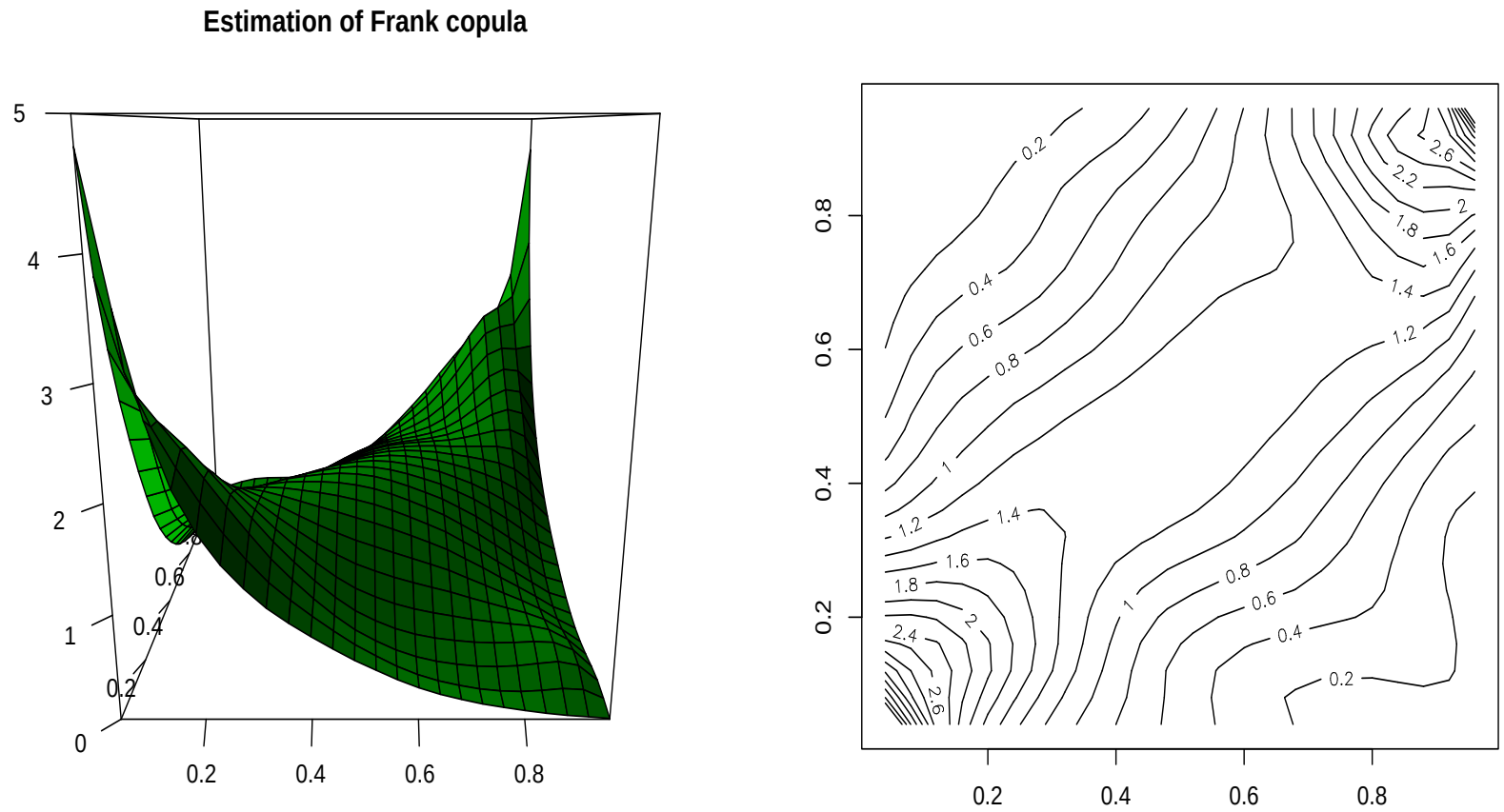


Figure 17: Estimated density with Gaussian normalization.

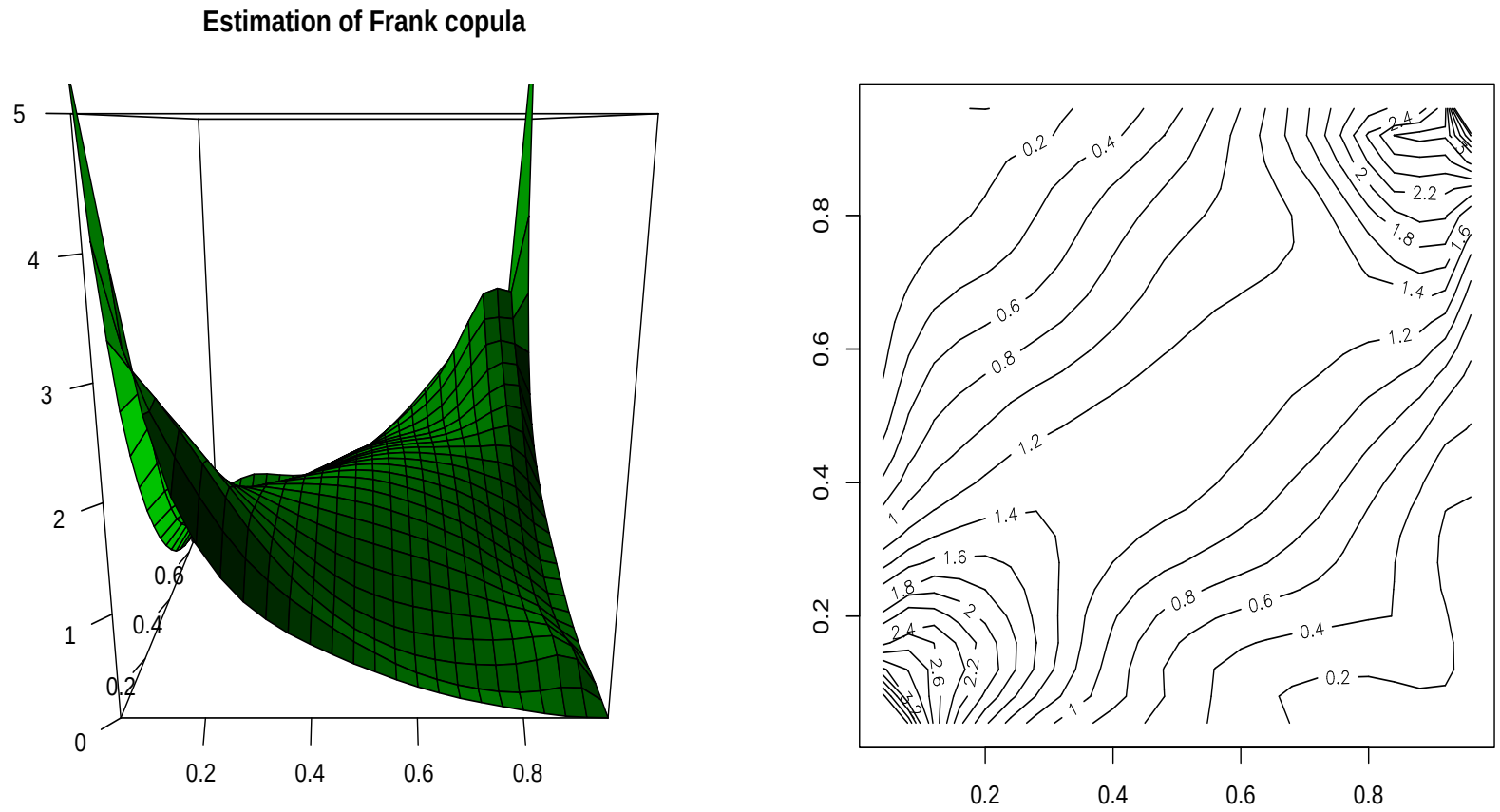


Figure 18: Estimated density with t normalization, $df = 5$.

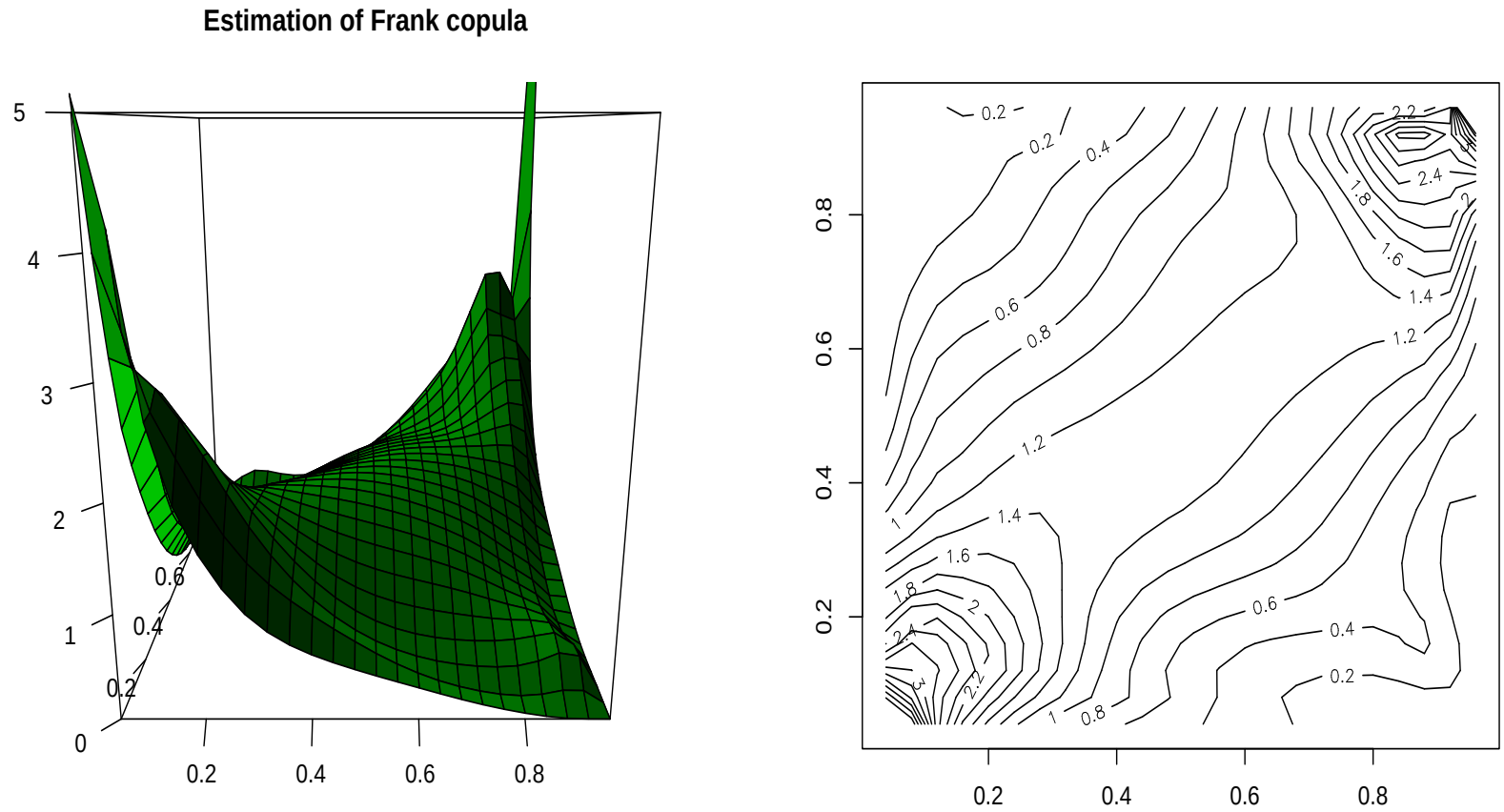


Figure 19: Estimated density with t normalization, $df = 3$.

Bivariate Beta kernels

The Beta-kernel based estimator of the copula density at point (u, v), is obtained using product beta kernels, which yields

$$\widehat{c}(u,v) = \frac{1}{n}\sum_{i=1}^{n} K\left(X_i, \frac{u}{b}+1, \frac{1-u}{b}+1\right) \cdot K\left(Y_i, \frac{v}{b}+1, \frac{1-v}{b}+1\right),$$

where $K(\cdot, \alpha, \beta)$ denotes the density of the Beta distribution with parameters α and β,

$$K(x, \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(1-x)^{\beta-1}x^{\alpha-1}\mathbf{1}_{x\in[0,1]}.$$

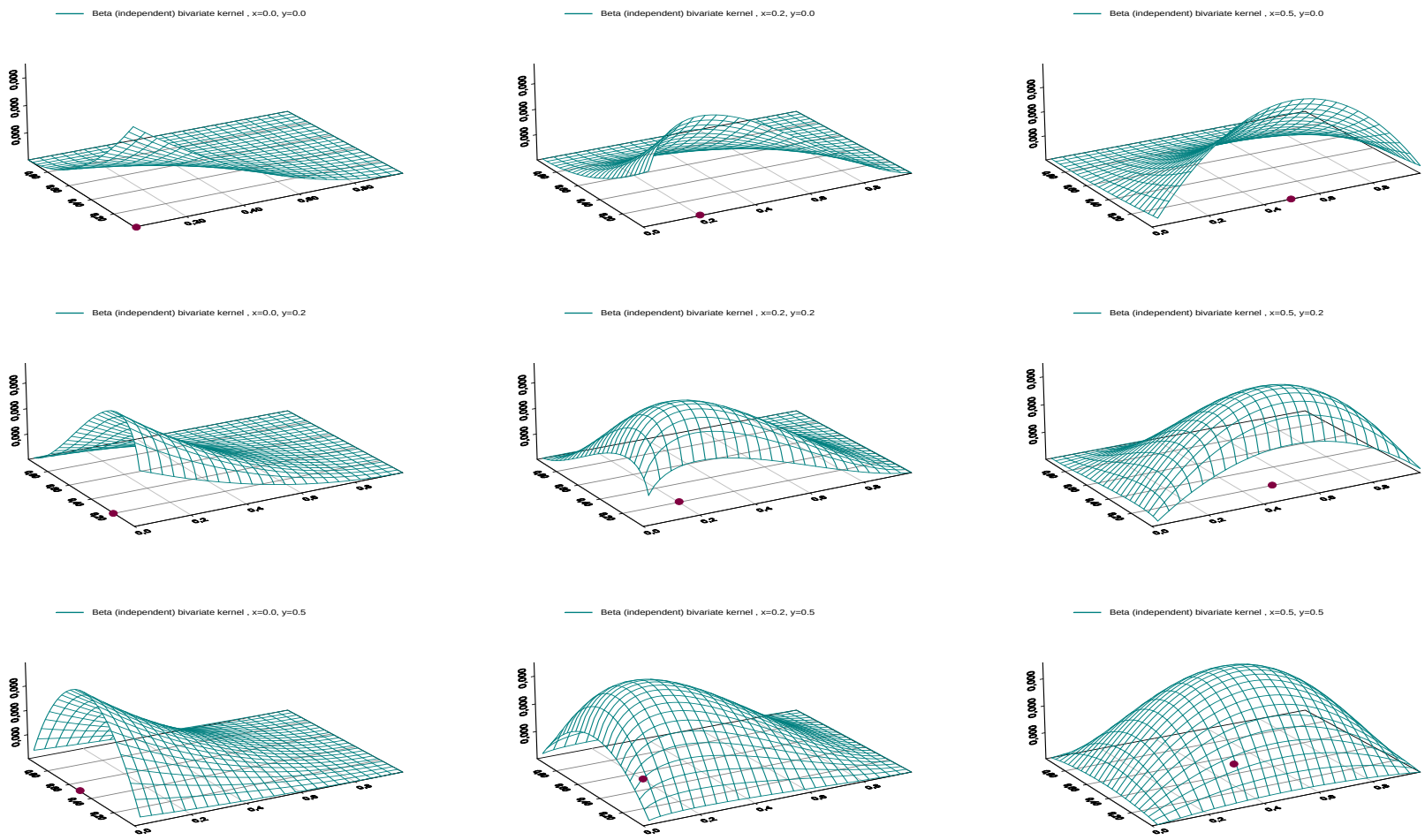


Figure 20: $K(\cdot, x/b+1, (1-x)/b+1) \times K(\cdot, y/b+1, (1-y)/b+1)$.

Assume that the copula density c is twice differentiable on $[0,1]\times[0,1]$. Let $(u,v)\in[0,1]\times[0,1]$. The bias of $\widehat{c}(u,v)$ is of order b, i.e.

$$\mathbb{E}(\widehat{c}(u,v)) = c(u,v) + \mathcal{Q}(u,v)\cdot b + o(b),$$

where the bias $\mathcal{Q}(u,v)$ is

$$(1-2u)c_1(u,v) + (1-2v)c_2(u,v) + \frac{1}{2}\left[u(1-u)c_{1,1}(u,v) + v(1-v)c_{2,2}(u,v)\right].$$

The bias here is $O(b)$ (everywhere) while it is $O(h^2)$ using standard kernels.

Assume that the copula density c is twice differentiable on $[0,1]\times[0,1]$. Let $(u,v)\in[0,1]\times[0,1]$. The variance of $\widehat{c}(u,v)$ is in corners, e.g. $(0,0)$,

$$Var(\widehat{c}(0,0)) = \frac{1}{nb^2}[c(0,0)+o(n^{-1})],$$

in the interior of borders, e.g. $u=0$ and $v\in(0,1)$

$$Var(\widehat{c}(0,v)) = \frac{1}{2nb^{3/2}\sqrt{\pi v(1-v)}}[c(0,v)+o(n^{-1})],$$

and in the interior,$(u,v)\in(0,1)\times(0,1)$

$$Var(\widehat{c}(u,v)) = \frac{1}{4nb\pi\sqrt{v(1-v)u(1-u)}}[c(u,v)+o(n^{-1})].$$

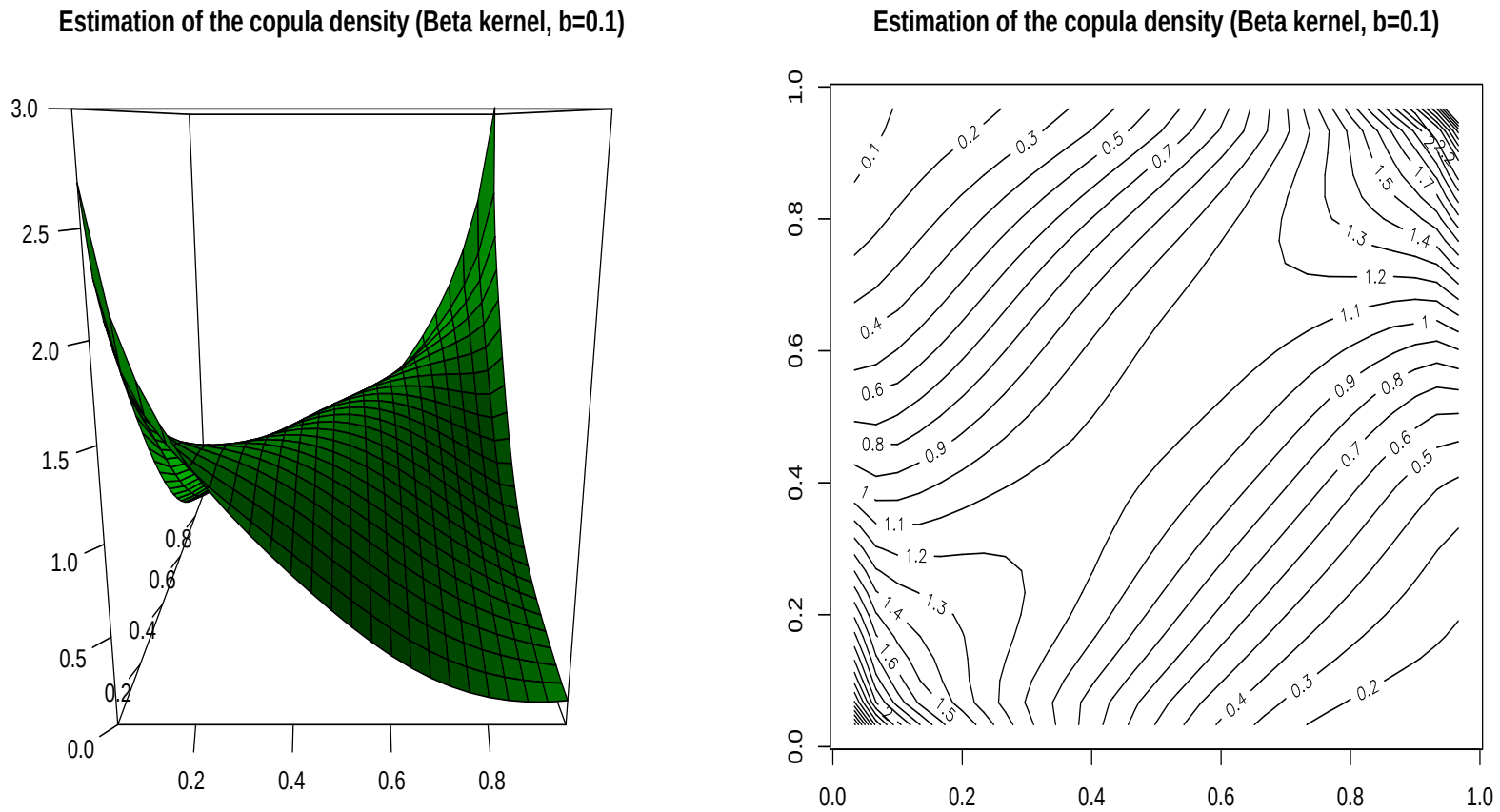


Figure 21: Estimated density of Frank copula, Beta kernels, $b = 0.1$

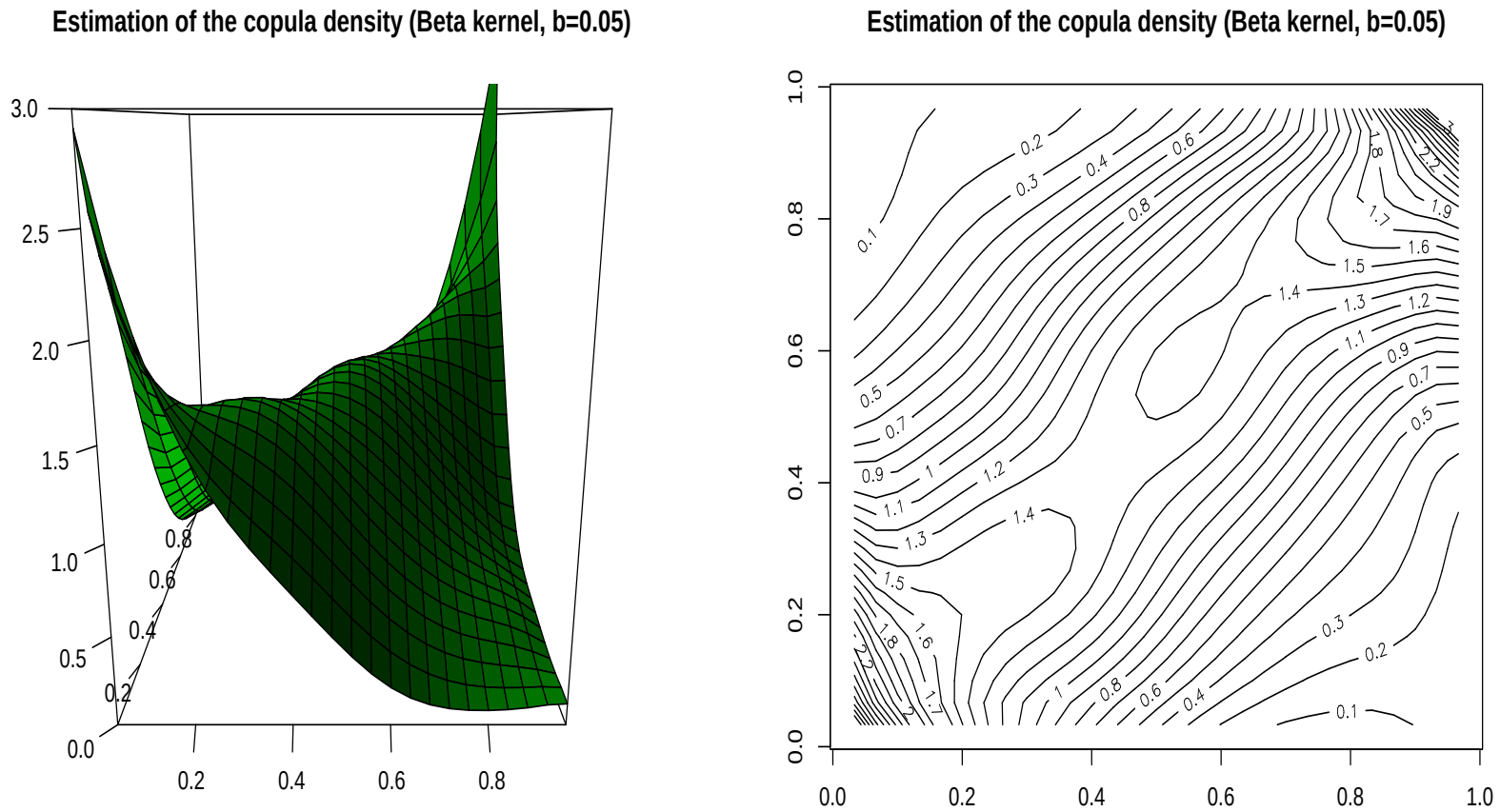


Figure 22: Estimated density of Frank copula, Beta kernels, $b = 0.05$

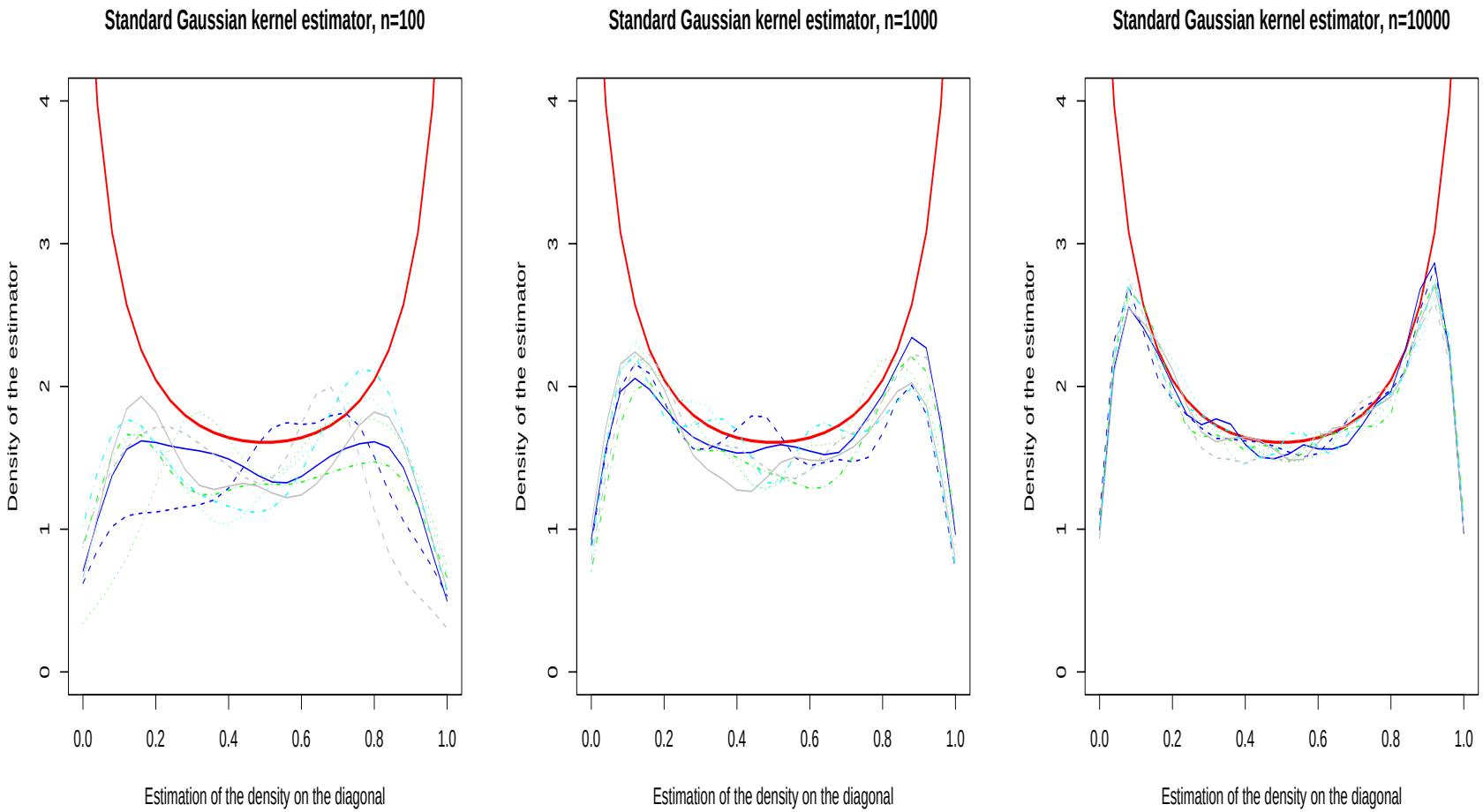


Figure 23: Density estimation on the diagonal, standard kernel.

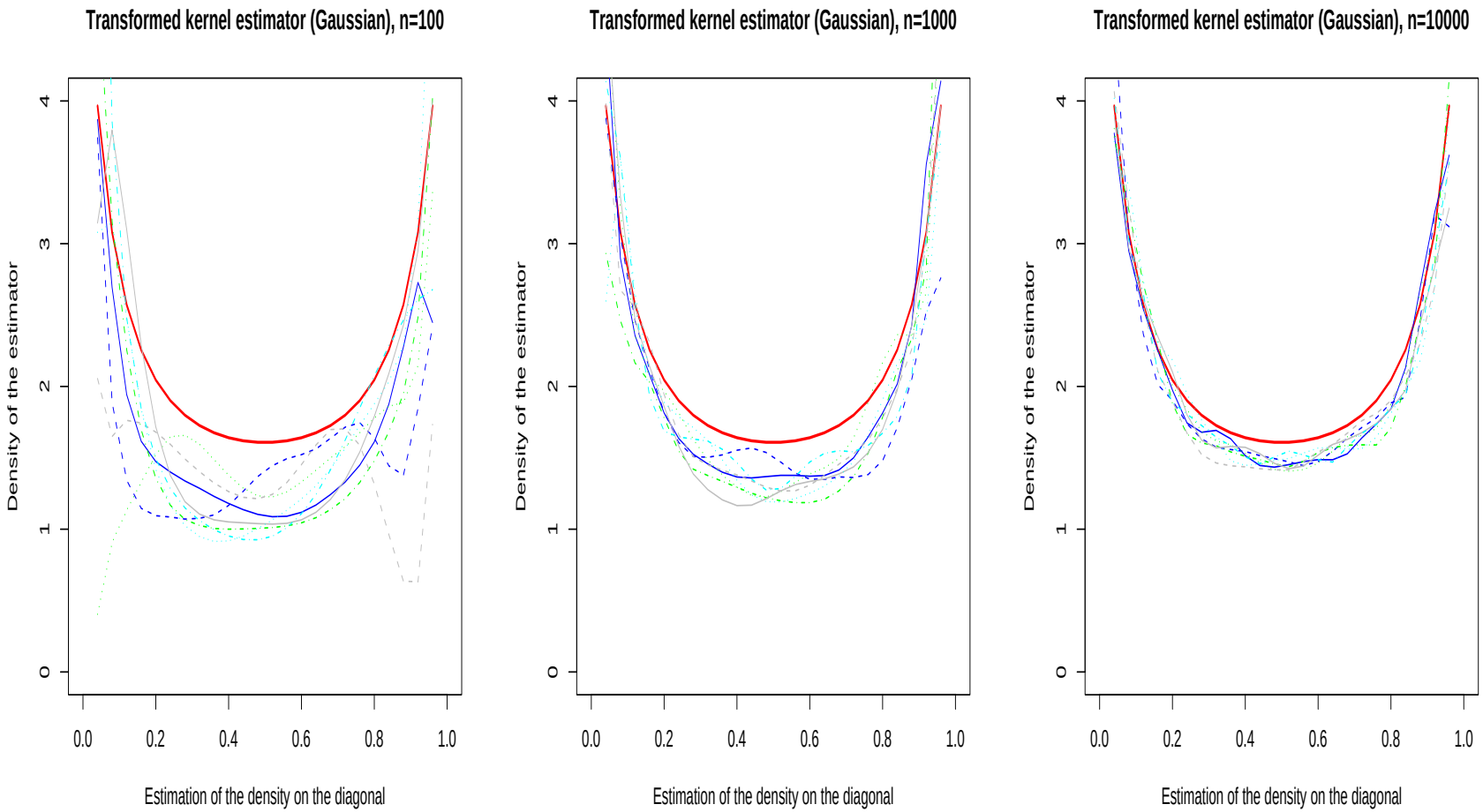


Figure 24: Density estimation on the diagonal, transformed kernel.

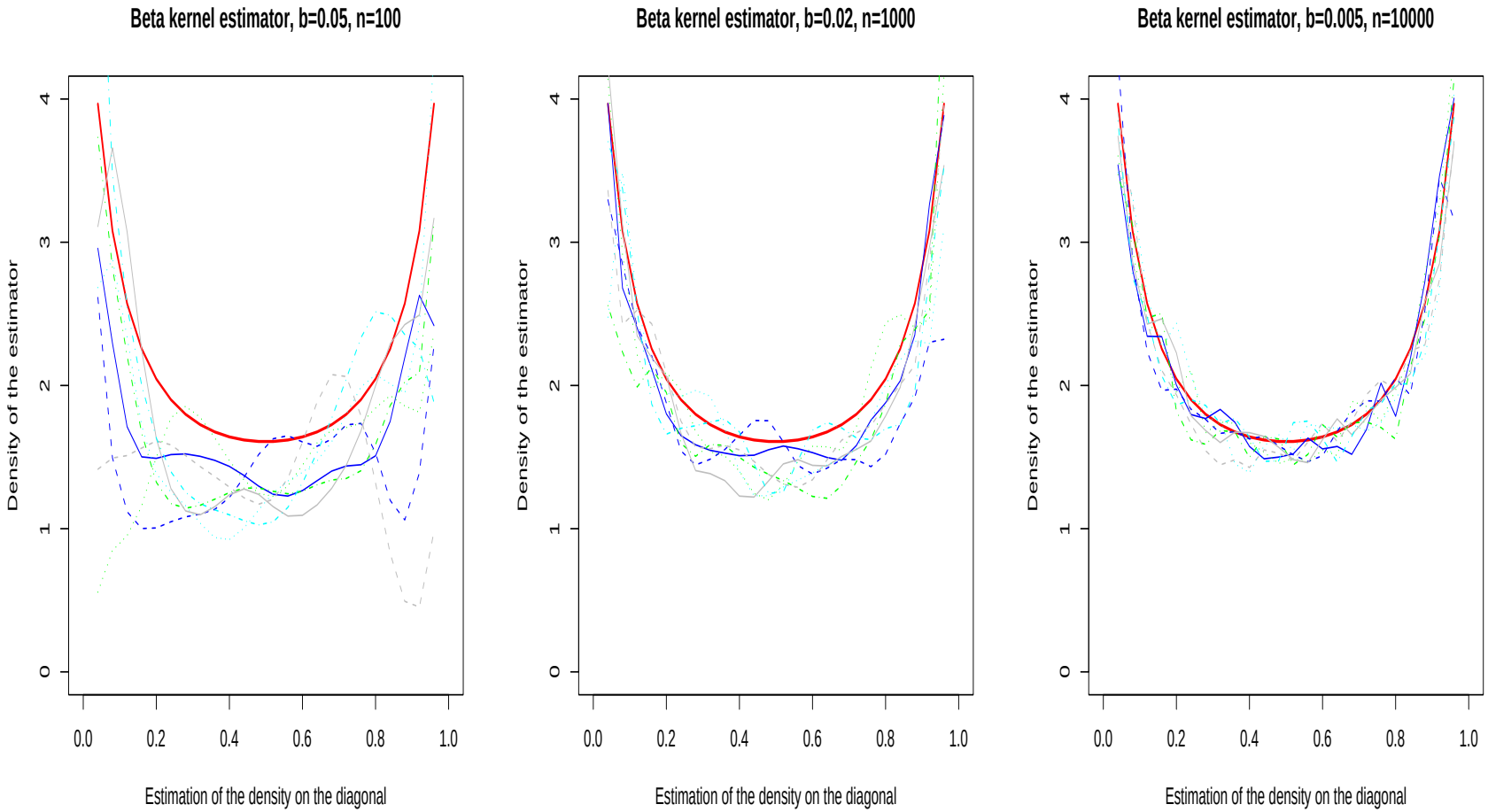


Figure 25: Density estimation on the diagonal, Beta kernel.

Copula density estimation

Gijbels & Mielniczuk (1990): given an i.i.d. sample, a natural estimate for the normed density is obtained using the transformed sample $(\widehat{F}_X(X_1), \widehat{F}_Y(Y_1)), ..., (\widehat{F}_X(X_n), \widehat{F}_Y(Y_n))$, where $\widehat{F}_X$ and $\widehat{F}_Y$ are the empirical distribution function of the marginal distribution. The copula density can be constructed as some density estimate based on this sample (Behnen, Husková & Neuhaus (1985) investigated the kernel method).

The natural kernel type estimator $\widehat{c}$ of c is

$$c(u,v) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left[\frac{u - \widehat{F}_X(X_i)}{h}, \frac{v - \widehat{F}_Y(Y_i)}{h}\right], (u,v) \in [0,1].$$

"*this estimator is not consistent in the points on the boundary of the unit square*."

Copula density estimation and pseudo-observations

Example: when dealing with copulas, ranks U_i, V_i yield pseudo-observations.

$$(U_i, V_i) = H(X_i, Y_i) = (F_X(X_i), F_Y(Y_i))$$

$$(\widehat{U}_i, \widehat{V}_i) = \widehat{H}_n(X_i, Y_i) = (\widehat{F}_X(X_i), \widehat{F}_Y(Y_i))$$

(see Genest & Rivest (1993)).

More formally, let $\boldsymbol{X}_1, ..., \boldsymbol{X}_n$ denote a series of observations of $\boldsymbol{X}$ $(\in \mathcal{X})$, stationary and ergodic.

Let $H : \mathcal{X} \to \mathbb{R}^d$ and set $\boldsymbol{\varepsilon}_i = H(\boldsymbol{X}_i)$ (non-observable).

If H is estimated by $\widehat{H}_n$ then $\widehat{\boldsymbol{\varepsilon}}_i = \widehat{H}_n(\boldsymbol{X}_i)$ are called pseudo-observations.

Let $\widehat{K}_n$ denote the empirical distribution function of those pseudo-observations,

$$\widehat{K}_n(\boldsymbol{t}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(\widehat{\boldsymbol{\varepsilon}}_i \leq \boldsymbol{t}) \text{ where } \boldsymbol{t} \in \mathbb{R}^d.$$

Further, if K denotes the distribution function of $\boldsymbol{\varepsilon} = H(\boldsymbol{X})$, then define the

empirical process based on pseudo-observations,

$$\mathbb{K}_n(t) = \sqrt{n}\left[\widehat{K}_n(\boldsymbol{t}) - K(\boldsymbol{t})\right]$$

As proved in Ghoudi & Rémillard (1998, 2004), this empirical process converges weakly.

Figure **??** shows scatterplots when margins are known (i.e. $(F_X(X_i), F_Y(Y_i))$'s), and when margins are estimated (i.e. $(\hat{F}_X(X_i), \hat{F}_Y(Y_i)$'s). Note that the pseudo sample is more "*uniform*", in the sense of a lower discrepancy (as in Quasi Monte Carlo techniques, see e.g. Niederreiter, H. (1992)).

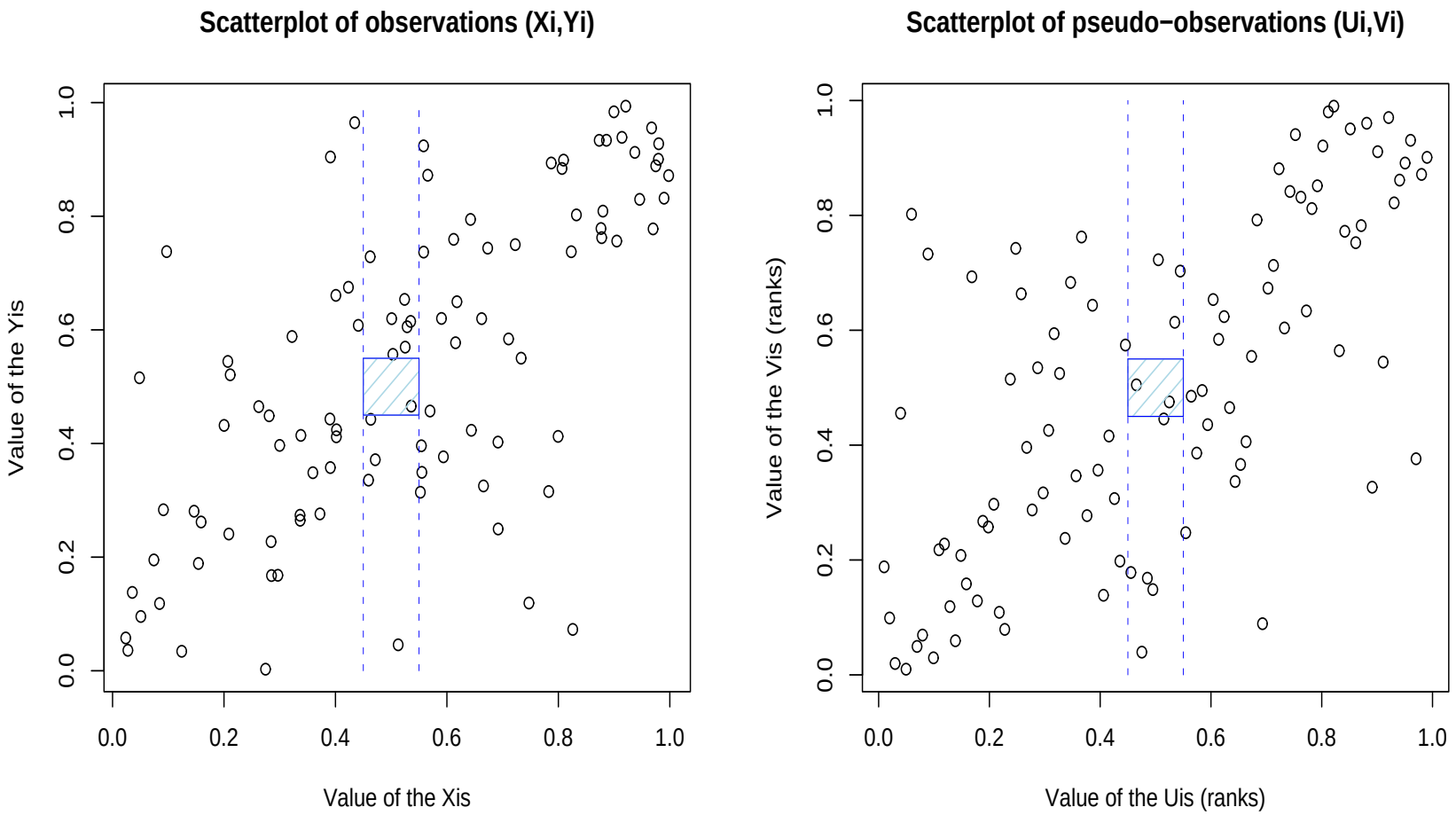


Figure 26: 500 simulated observations (X_i, Y_i) and the associate pseudo-sample $(\hat{F}_X(X_i), \hat{F}_Y(Y_i))$.

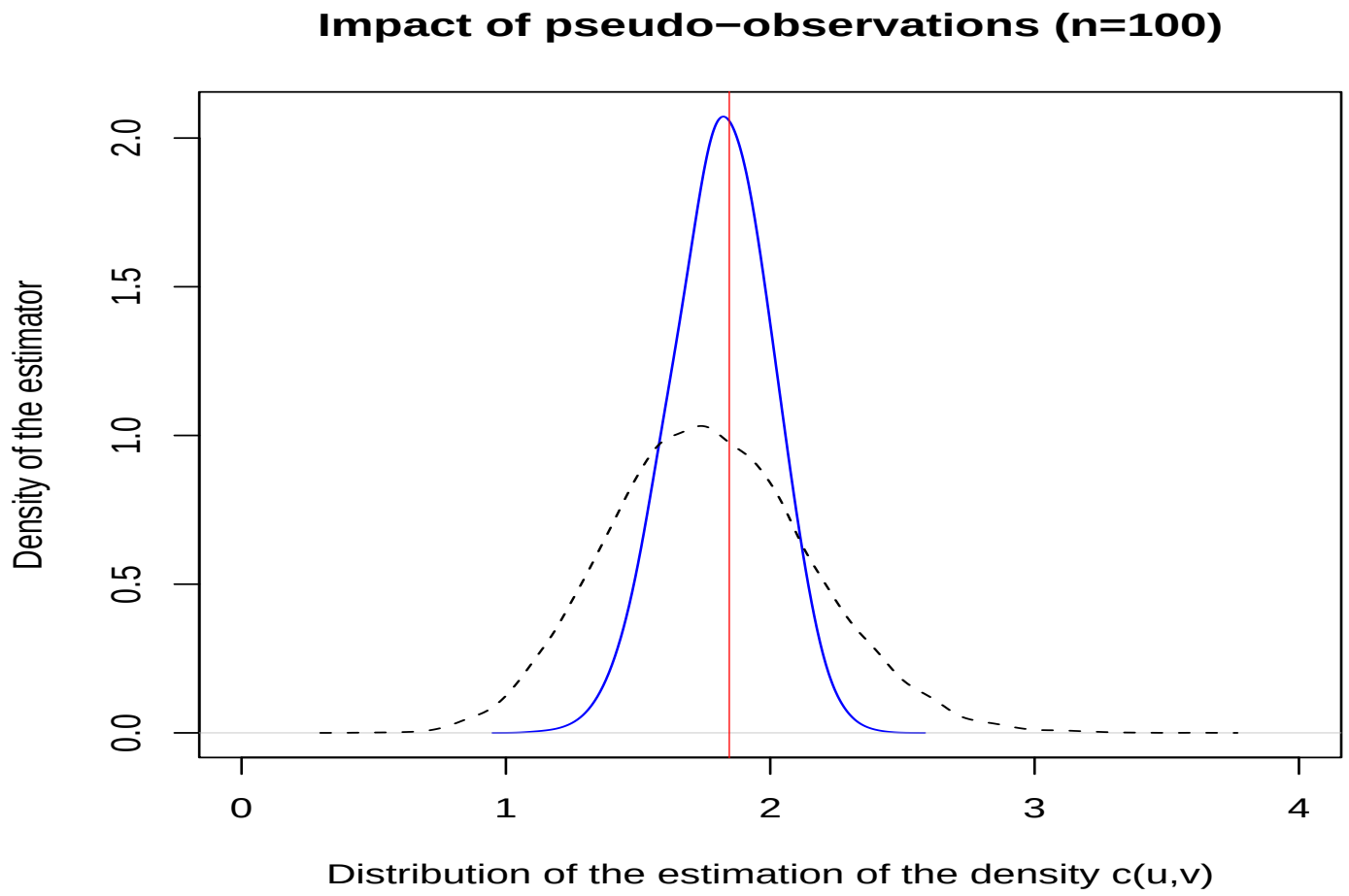


Figure 27: The impact of estimating from pseudo-observations.

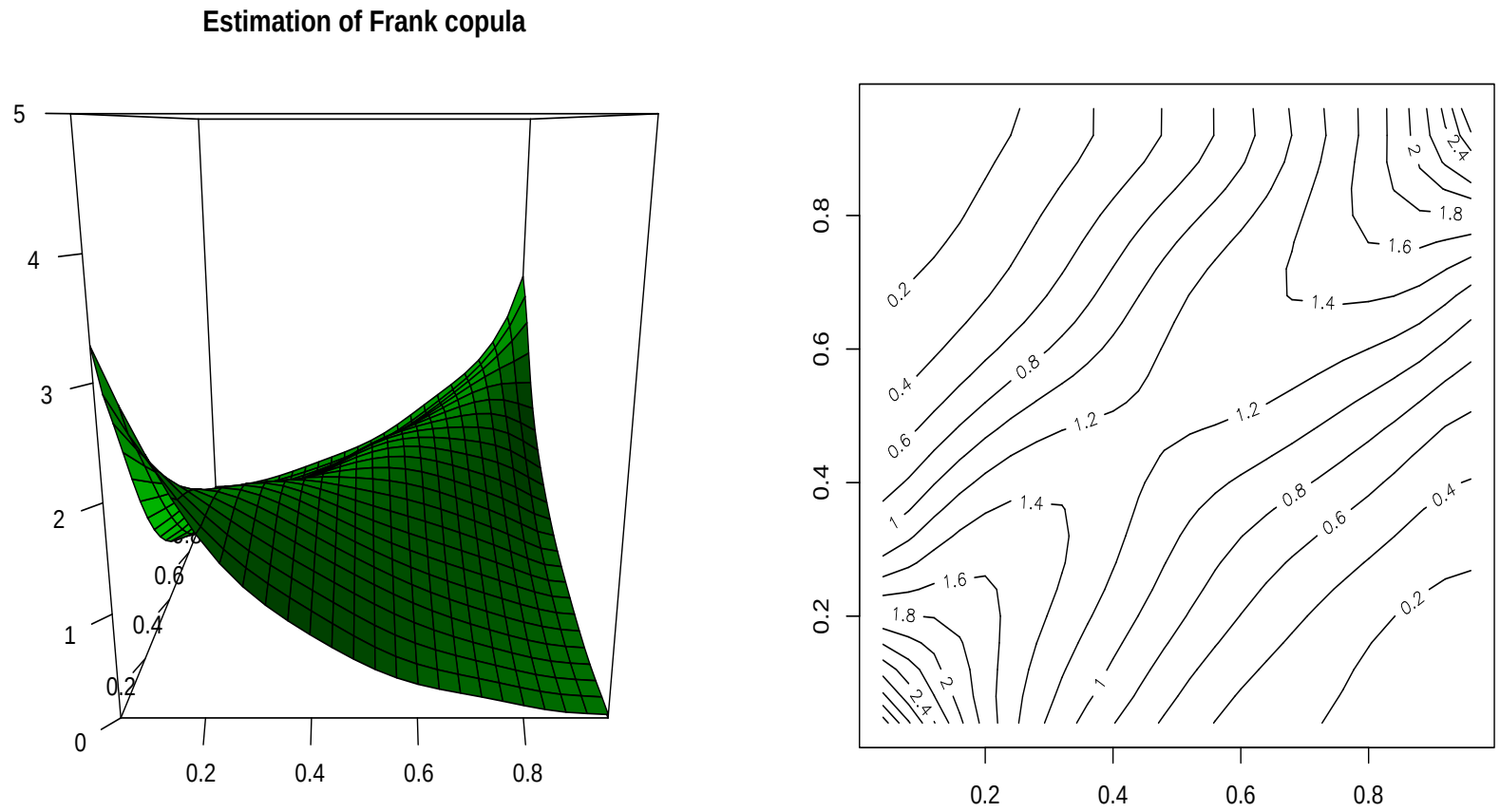


Figure 28: $n = 1000$, $b = 0.050$ and no-censoring.

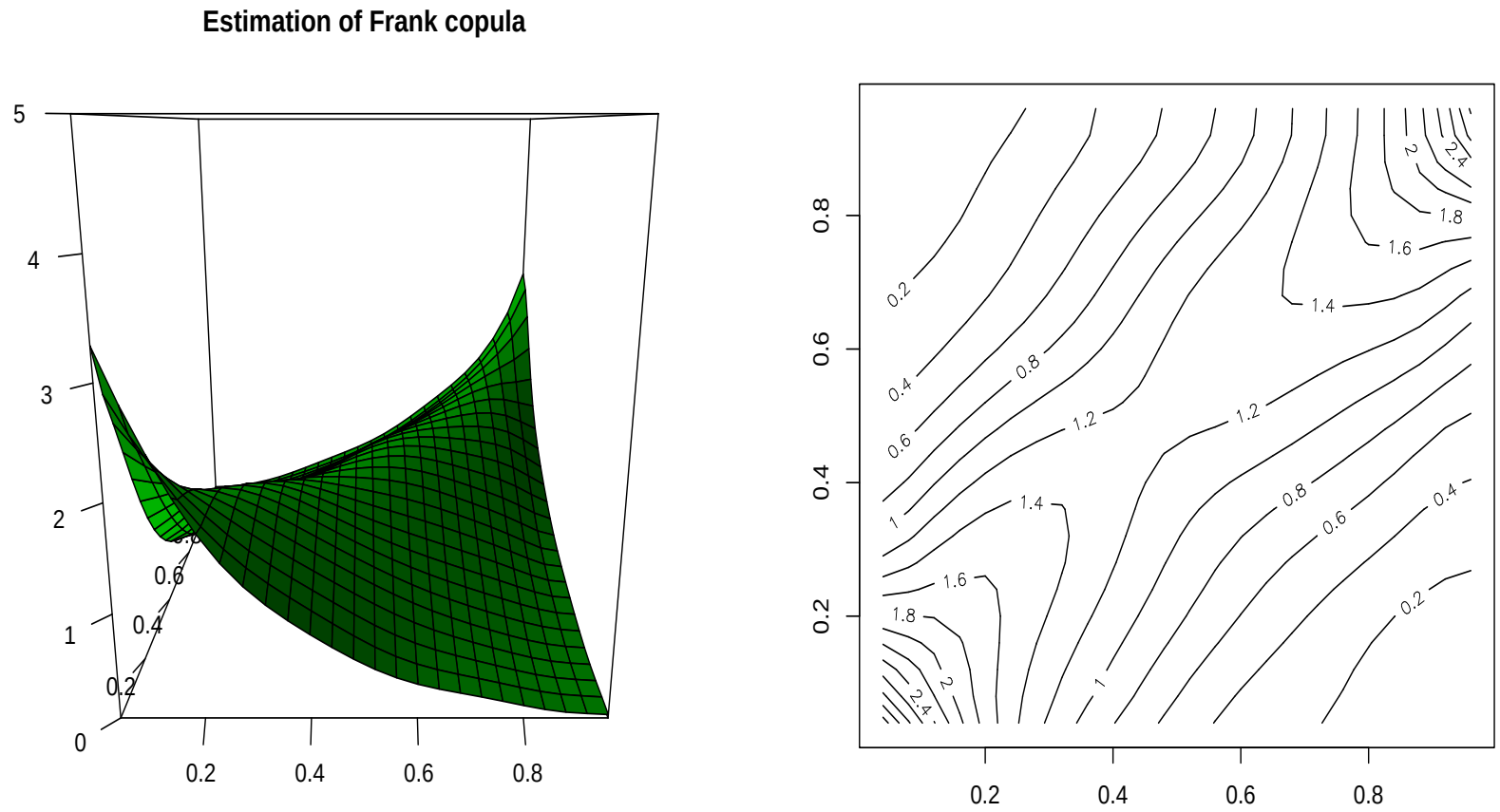


Figure 29: $n = 1000$, $b = 0.050$ and 1.7% censored data.

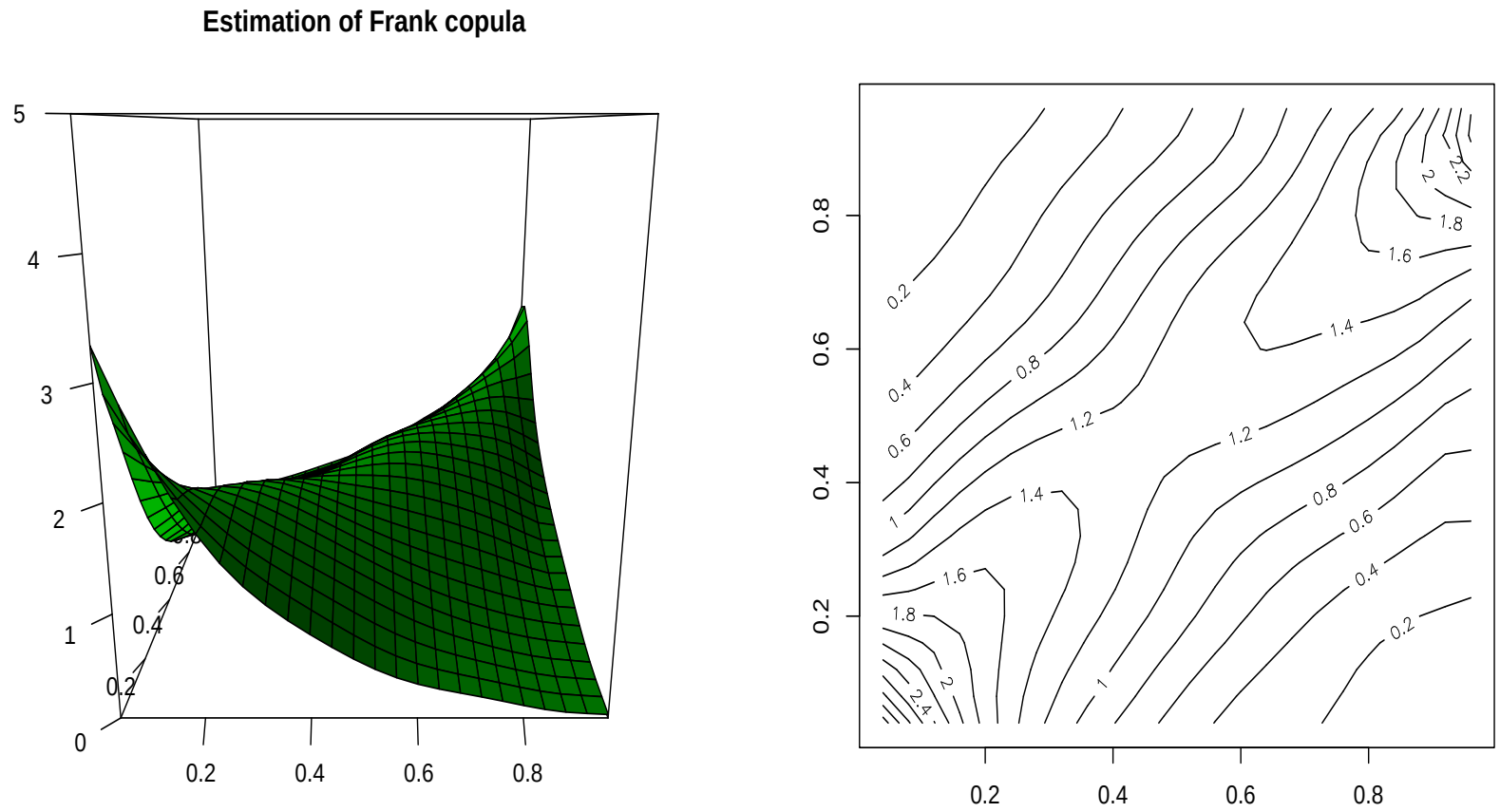


Figure 30: $n = 1000$, $b = 0.050$ and 7.8% censored data.

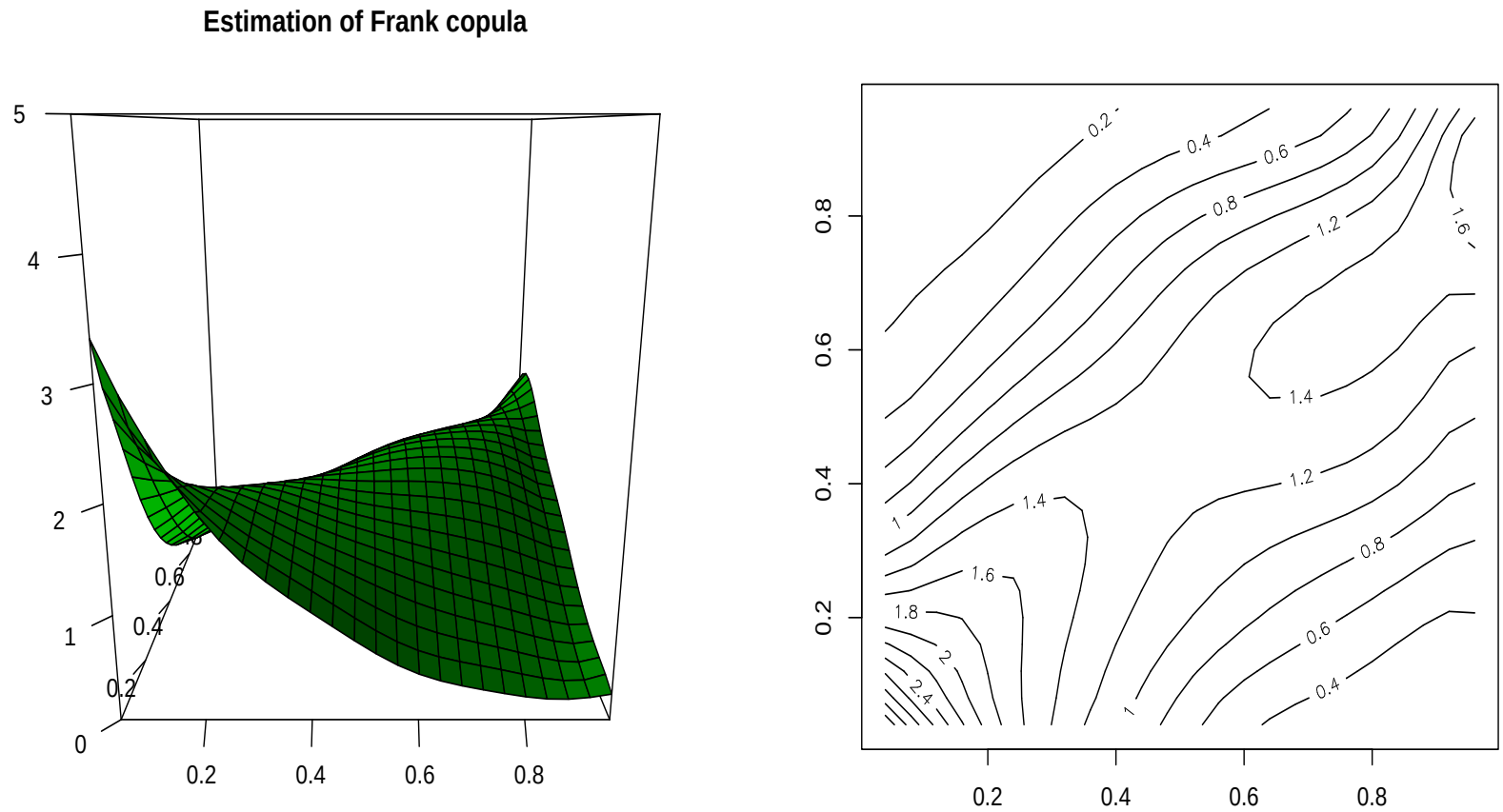


Figure 31: $n = 1000$, $b = 0.050$ and 24% censored data.

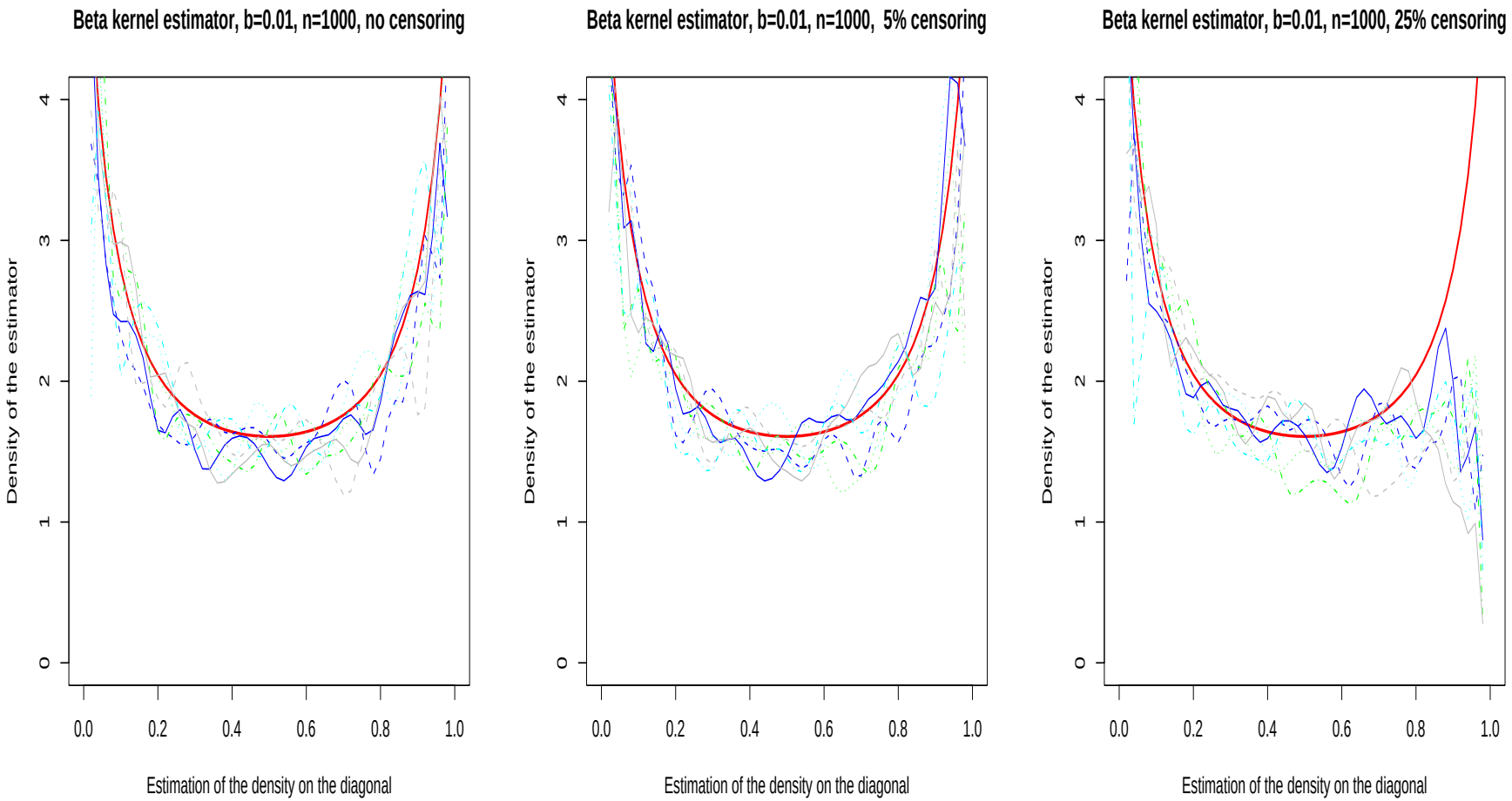


Figure 32: Estimation of the density on the diagonal.

Stock sampling can be observe as the bias increases as u goes to 1, and as the proportion of censored data increases.

Idea in Efron & Tibshirani (1993): bootstrap techniques can be used to estimate the bias. Consider B bootstrap samples drawn independently from the original censored data, with replacement, $X_{i,b}^*, Y_{i,b}^*, \delta_{i,b}^*$, $i = 1, ..., n$ and $b = 1, ..., B$. The bias of the density estimator at point (u, v) can be estimated from

$$\widehat{bias}\,(u,v) = \frac{1}{B}\sum_{i=1}^{B} \widehat{c}\,(u,v)^{*,b} - \widehat{c}\,(u,v)$$

where $\widehat{c}\,(u)^{*,b}$ is the density obtained from the bth bootstrap sample.

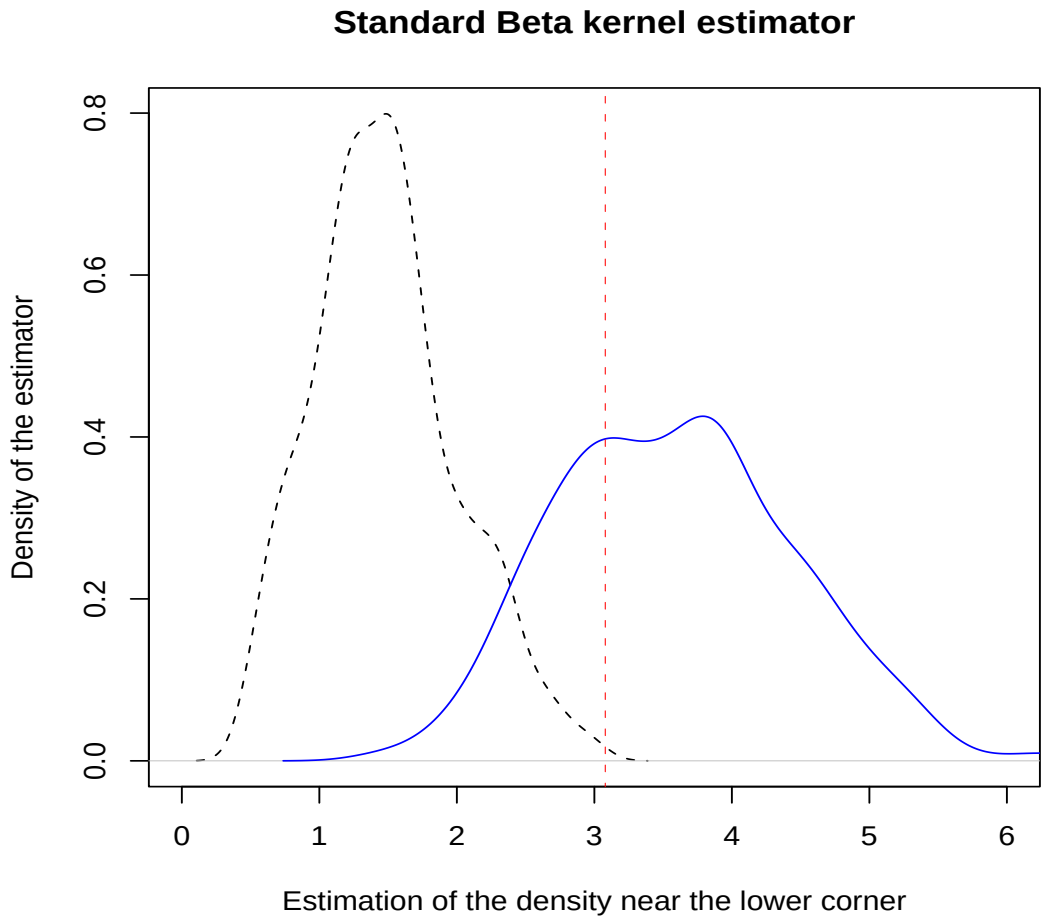


Figure 33: $\widehat{c}(12/13, 12/13)$ with 25% censored data.

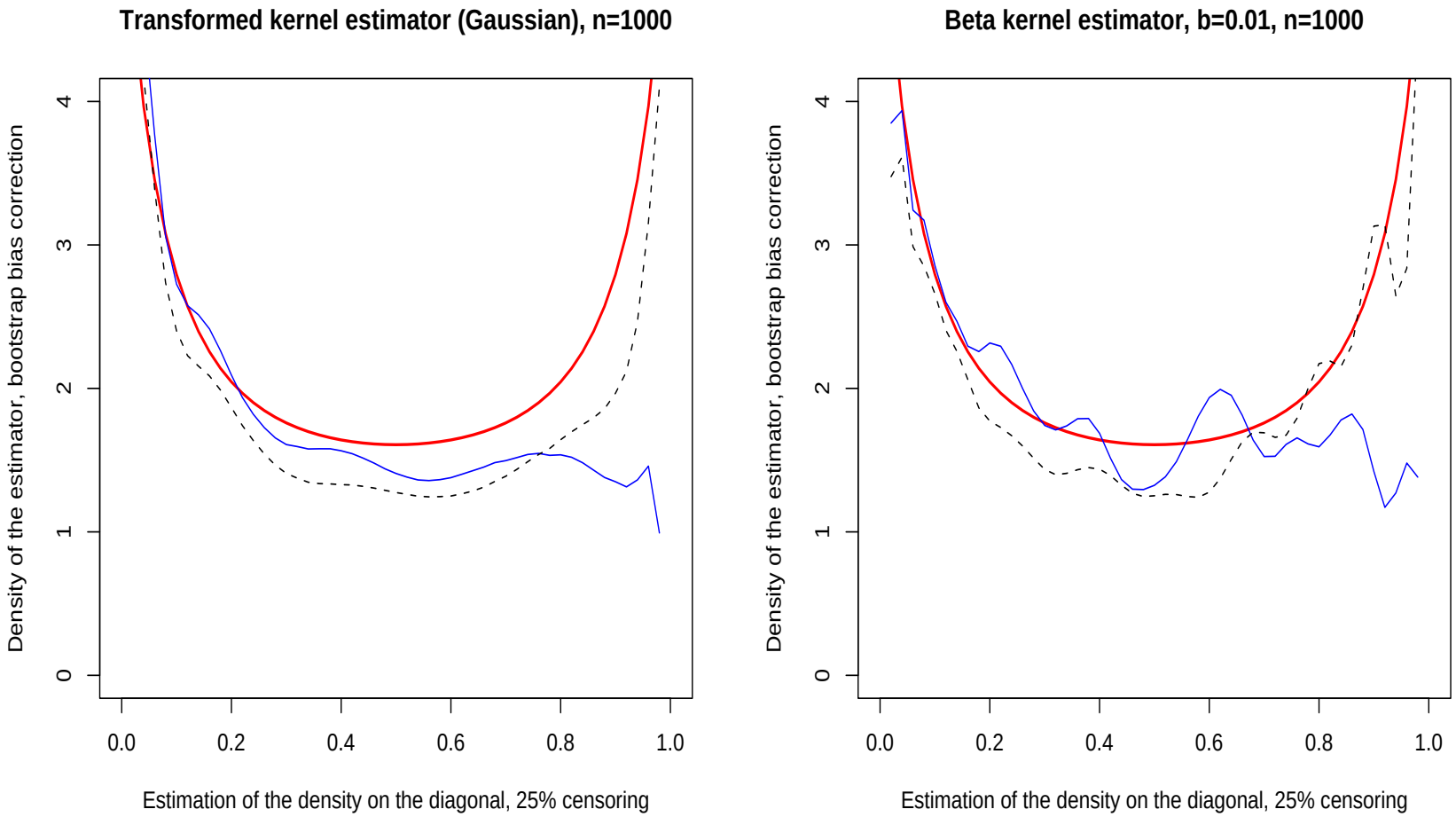


Figure 34: Censoring bias correction using bootstrap techniques.

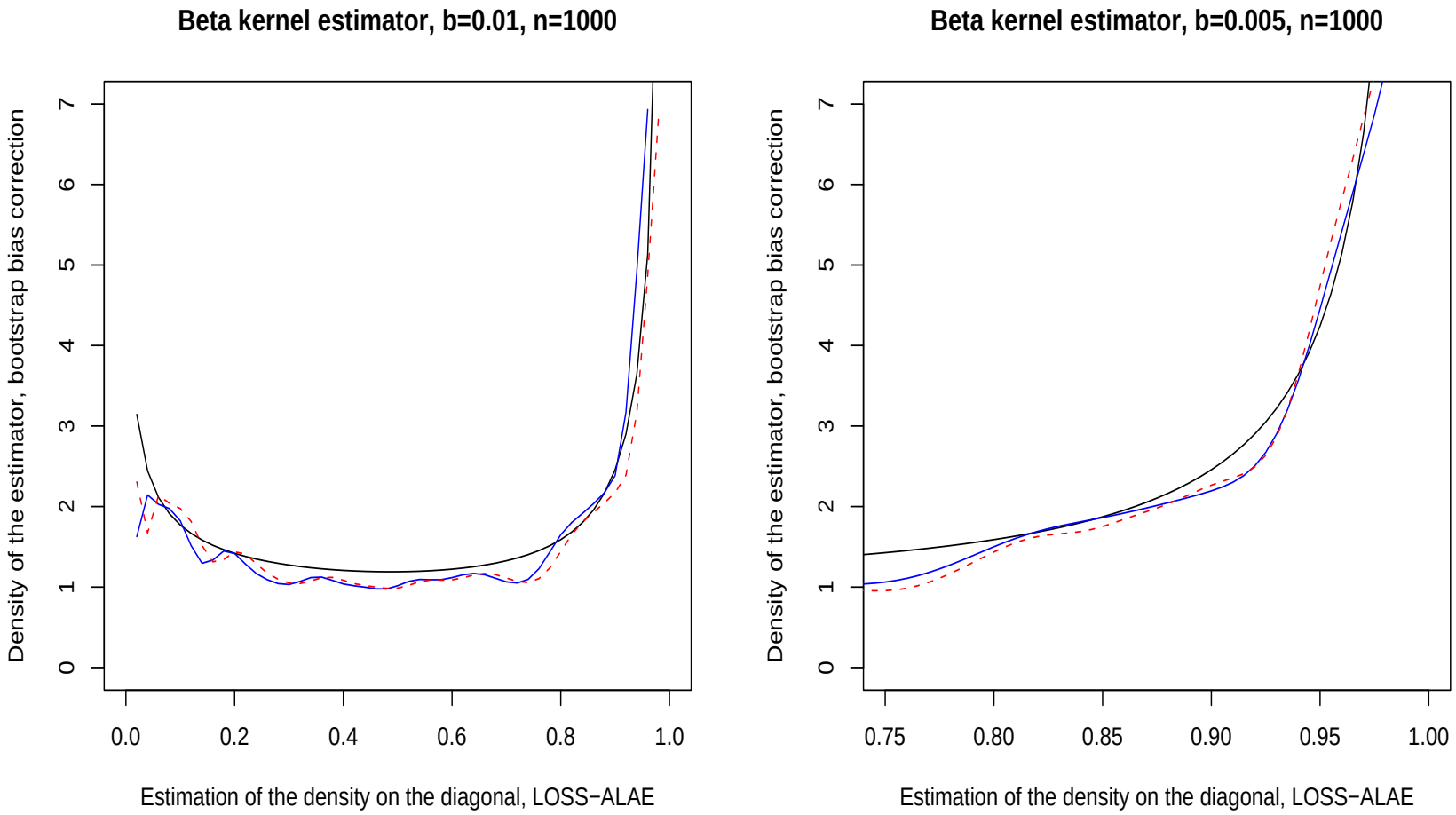


Figure 35: Loss-ALAE dataset, compared with Gumbel copula.

There are perspectives to go further,

- better knowledge of properties of the transformed kernel technique (impact of heavy tails of the transformation ?)
- impact of the biais correction on some risk measure (e.g. exceeding probabilities)

Structure of the PhD thesis

The thesis is based on several papers,

Charpentier, A. and Juri, A. (2006). Limiting dependence structure for tail events, with applications to credit derivatives. *Journal of Applied Probability*, **43**.

Charpentier, A. and Segers, J. (2006a). Convergence of Archimedean Copulas. submitted to *Probability and Statistical Letters*.

Charpentier, A. and Segers, J. (2006b). Lower tail dependence for Archimedean copulas: characterizations and pitfalls. submitted to *Insurance Mathematics and Economics*.

Charpentier, A. and Segers, J. (2006c). Upper tail dependence for Archimedean copulas. in progress.

Charpentier, A. and Segers, J. (2006d). Tail dependence copulas and extreme value theory. in progress.

Charpentier, A. (2003). Tail distribution and dependence measures. *Proceedings of the 34th ASTIN Conference*.

Charpentier, A. (2004). Extremes and dependence: a copula based approach. *Proceedings of the 3rd Conference in Actuarial Science & Finance in Samos*.

Charpentier, A., Fermanian, J.D. and Scaillet, 0. (2006). Beta kernel estimation of copula density for censored data. in progress.

Charpentier, A., Fermanian, J.D. and Scaillet, 0. (2005). Estimating copula densities. *to Appear in* Copula Methods in Derivatives and Risk Management: from Credit Risk to Market Risk. Risk Book.