Single-index copulae

Jean-David Fermanian (Ensae-Crest),
& Olivier Lopez (Univ. Paris 6).

Copulae & Extremes workshop, Luminy, February 2016

Agenda

- Conditional copulae
- Single-index copula models
- Consistency
- Asymptotic normality

Conditional copulae

In a lot of situations, we need to model the dependence among several r.v.s' (copulae), conditionally to some amount of information.

Conditional copulae

In a lot of situations, we need to model the dependence among several r.v.s' (copulae), conditionally to some amount of information. Typically:

- Time series $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$, $\boldsymbol{X}_t \in \mathbb{R}^d$, not necessarily stationary: (joint) law of $\boldsymbol{X}_t$ knowing $\boldsymbol{X}_{t-1}$?

Conditional copulae

In a lot of situations, we need to model the dependence among several r.v.s' (copulae), conditionally to some amount of information. Typically:

- Time series $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$, $\boldsymbol{X}_t \in \mathbb{R}^d$, not necessarily stationary: (joint) law of $\boldsymbol{X}_t$ knowing $\boldsymbol{X}_{t-1}$?
- Econometric models with exogenous variables: what is the (joint) law of $\boldsymbol{X} \in \mathbb{R}^d$ knowing $\boldsymbol{Z} \in \mathbb{R}^p$?

Conditional copulae

In a lot of situations, we need to model the dependence among several r.v.s' (copulae), conditionally to some amount of information. Typically:

- Time series $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$, $\boldsymbol{X}_t \in \mathbb{R}^d$, not necessarily stationary: (joint) law of $\boldsymbol{X}_t$ knowing $\boldsymbol{X}_{t-1}$?
- Econometric models with exogenous variables: what is the (joint) law of $\boldsymbol{X} \in \mathbb{R}^d$ knowing $\boldsymbol{Z} \in \mathbb{R}^p$?
- Pair-copula constructions (vines): a conditional bivariate distribution at every node of a tree-based structure.

Conditional copulae

In a lot of situations, we need to model the dependence among several r.v.s' (copulae), conditionally to some amount of information. Typically:

- Time series $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$, $\boldsymbol{X}_t \in \mathbb{R}^d$, not necessarily stationary: (joint) law of $\boldsymbol{X}_t$ knowing $\boldsymbol{X}_{t-1}$?
- Econometric models with exogenous variables: what is the (joint) law of $\boldsymbol{X} \in \mathbb{R}^d$ knowing $\boldsymbol{Z} \in \mathbb{R}^p$?
- Pair-copula constructions (vines): a conditional bivariate distribution at every node of a tree-based structure.

$\Rightarrow$ we need conditional copulae (Patton 2005)

Conditional copulae

Definition 1

A conditional copula associated to $(\boldsymbol{X}, \mathcal{A})$ is a $\mathcal{B}([0,1]^d) \otimes \mathcal{A}$ measurable function C such that, for any $x_1, \ldots, x_d \in \mathbb{R}$,

$$\mathbb{P}\left(\boldsymbol{X} \leq \boldsymbol{x} | \mathcal{A}\right) = C\left\{\mathbb{P}(X_1 \leq x_1 | \mathcal{A}), \ldots, \mathbb{P}(X_d \leq x_d | \mathcal{A}) | \mathcal{A}\right\}.$$

For us, $\mathcal{A} = \sigma(\boldsymbol{Z})$.

Conditional copulae

Definition 1

A conditional copula associated to $(\boldsymbol{X}, \mathcal{A})$ is a $\mathcal{B}([0,1]^d) \otimes \mathcal{A}$ measurable function C such that, for any $x_1, \ldots, x_d \in \mathbb{R}$,

$$\mathbb{P}\left(\boldsymbol{X} \leq \boldsymbol{x} | \mathcal{A}\right) = C\left\{\mathbb{P}(X_1 \leq x_1 | \mathcal{A}), \ldots, \mathbb{P}(X_d \leq x_d | \mathcal{A}) | \mathcal{A}\right\}.$$

For us, $\mathcal{A} = \sigma(\boldsymbol{Z})$.

Fermanian and Wegkamp (2012) have extended this concept when different conditioning subsets are introduced ("pseudo-copulae").

Conditional copulae

In practice, a lot of conditioning variables potentially, particularly in econometrics.

Conditional copulae

In practice, a lot of conditioning variables potentially, particularly in econometrics. Diverse solutions:

1. the fully nonparametric approach: empirical counterparts of all conditional distributions.

$$C(\boldsymbol{u}|\boldsymbol{Z}=\boldsymbol{z}) = \hat{F}\left(\hat{F}_1^{-1}(u_1|\boldsymbol{Z}=\boldsymbol{z}),\ldots,\hat{F}_d^{-1}(u_d|\boldsymbol{Z}=\boldsymbol{z})\,|\,\boldsymbol{Z}=\boldsymbol{z}\right),$$

$$\hat{F}(\boldsymbol{x}|\boldsymbol{Z}=\boldsymbol{z}) = \sum_{i=1}^{n} w_{i,n}(\boldsymbol{Z}_i,\boldsymbol{z})\mathbf{1}(\boldsymbol{X}_i \leq \boldsymbol{x}),$$

for some weights (Nadaraya-Watson, Gasser-Müller, Priestley-Chao...).

Conditional copulae

In practice, a lot of conditioning variables potentially, particularly in econometrics. Diverse solutions:

1. the fully nonparametric approach: empirical counterparts of all conditional distributions.

$$C(\boldsymbol{u}|\boldsymbol{Z}=\boldsymbol{z}) = \hat{F}\left(\hat{F}_1^{-1}(u_1|\boldsymbol{Z}=\boldsymbol{z}),\ldots,\hat{F}_d^{-1}(u_d|\boldsymbol{Z}=\boldsymbol{z})\,|\,\boldsymbol{Z}=\boldsymbol{z}\right),$$

$$\hat{F}(\boldsymbol{x}|\boldsymbol{Z}=\boldsymbol{z}) = \sum_{i=1}^{n} w_{i,n}(\boldsymbol{Z}_i,\boldsymbol{z})\mathbf{1}(\boldsymbol{X}_i \leq \boldsymbol{x}),$$

for some weights (Nadaraya-Watson, Gasser-Müller, Priestley-Chao...).

Ok..., but unfeasible when $\dim(\boldsymbol{Z}) > 3$.

See Fermanian and Wegkamp (2012), Gijbels et al. (2011).

Conditional copulae

2. the fully parametric approach: $C(\cdot|\boldsymbol{Z}) = C_{\theta(\boldsymbol{Z},\beta_0)}(\cdot)$, $\beta_0 \in \mathbb{R}^q$, where
   - C_θ belongs to a **known** parametric copula family $\mathcal{C}$, and

Conditional copulae

2. the fully parametric approach: $C(\cdot|\boldsymbol{Z}) = C_{\theta(\boldsymbol{Z},\beta_0)}(\cdot)$, $\beta_0 \in \mathbb{R}^q$, where

   - C_θ belongs to a **known** parametric copula family $\mathcal{C}$, and
   - $\theta(\cdot, \beta)$ is **known.**

Conditional copulae

2. the fully parametric approach: $C(\cdot|\boldsymbol{Z}) = C_{\theta(\boldsymbol{Z},\beta_0)}(\cdot)$, $\beta_0 \in \mathbb{R}^q$, where
   - C_θ belongs to a **known** parametric copula family $\mathcal{C}$, and
   - $\theta(\cdot,\beta)$ is **known.**

Ok...but a lot of assumptions, and difficult to specify the influence of a covariate on a model parameter, in general.

See Patton (2006), Rockinger and Jondeau (2006), etc.

3. An hybrid parametric-NP approach: only $\theta(\cdot)$ is **unknown**.
   - C_θ belongs to a **known** parametric family $\mathcal{C}$;

3. An hybrid parametric-NP approach: only $\theta(\cdot)$ is **unknown**.
   - C_θ belongs to a **known** parametric family $\mathcal{C}$;
   - $\theta(\cdot)$ is estimated nonparametrically, or through local likelihood techniques : for a given $\boldsymbol{z}$, a limited expansion of the type

$$\theta(\boldsymbol{Z}_i) = \theta(\boldsymbol{z}) + d\theta(\boldsymbol{z}).(\boldsymbol{Z}_i - \boldsymbol{z}) + \frac{1}{2}d^2\theta(\boldsymbol{z}).(\boldsymbol{Z}_i - \boldsymbol{z})^{(2)} + \ldots,$$

and MLE, but restricted to the observations s.t. $\boldsymbol{Z}_i$ is "close" to $\boldsymbol{z}$, providing $\hat{\theta}(\boldsymbol{z})$.

3. An hybrid parametric-NP approach: only $\theta(\cdot)$ is **unknown**.
   - C_θ belongs to a **known** parametric family $\mathcal{C}$;
   - $\theta(\cdot)$ is estimated nonparametrically, or through local likelihood techniques : for a given $\boldsymbol{z}$, a limited expansion of the type

$$\theta(\boldsymbol{Z}_i) = \theta(\boldsymbol{z}) + d\theta(\boldsymbol{z}).(\boldsymbol{Z}_i - \boldsymbol{z}) + \frac{1}{2}d^2\theta(\boldsymbol{z}).(\boldsymbol{Z}_i - \boldsymbol{z})^{(2)} + \ldots,$$

   and MLE, but restricted to the observations s.t. $\boldsymbol{Z}_i$ is "close" to $\boldsymbol{z}$, providing $\hat{\theta}(\boldsymbol{z})$.

See Acar et al. (2011), Hafner and Reznikova (2011), Abegaz et al. (2012), Craiu and Sabeti (2012).

3. An hybrid parametric-NP approach: only $\theta(\cdot)$ is **unknown**.
   - C_θ belongs to a **known** parametric family $\mathcal{C}$;
   - $\theta(\cdot)$ is estimated nonparametrically, or through local likelihood techniques : for a given $\boldsymbol{z}$, a limited expansion of the type

$$\theta(\boldsymbol{Z}_i) = \theta(\boldsymbol{z}) + d\theta(\boldsymbol{z}).(\boldsymbol{Z}_i - \boldsymbol{z}) + \frac{1}{2}d^2\theta(\boldsymbol{z}).(\boldsymbol{Z}_i - \boldsymbol{z})^{(2)} + \ldots,$$

     and MLE, but restricted to the observations s.t. $\boldsymbol{Z}_i$ is "close" to $\boldsymbol{z}$, providing $\hat{\theta}(\boldsymbol{z})$.

See Acar et al. (2011), Hafner and Reznikova (2011), Abegaz et al. (2012), Craiu and Sabeti (2012).

Pb: still the curse of dimensionality !

Conditional copulae

4. A "true" semi-parametric approach: a priori a good intermediate solution, to avoid the $\boldsymbol{Z}$-curse of dimensionality.

   This is the topic of this work, under the single-index framework.

4. A "true" semi-parametric approach: a priori a good intermediate solution, to avoid the $\boldsymbol{Z}$-curse of dimensionality.

   This is the topic of this work, under the single-index framework.

   An alternative approach: additive models, as in Craiu and Sabeti (2014), Vatter and Chavez-Demoulin (2015), Acar (2015).

Single-index copulae: some notations

- The vector $\boldsymbol{X} \in \mathbb{R}^d$ is the endogenous vector, and $\boldsymbol{Z}$ is the vector of covariates.
- $F(\cdot|\boldsymbol{z})$ is the law of $\boldsymbol{X}$ knowing $\boldsymbol{Z} = \boldsymbol{z}$
- $F_k(\cdot|\boldsymbol{z})$, $k = 1, \ldots, d$, is the (marginal) law of X_k knowing $\boldsymbol{Z} = \boldsymbol{z}$
- The unobserved random vector $\boldsymbol{U}_{\boldsymbol{z}} = (U_{1,\boldsymbol{z}}, \ldots, U_{d,\boldsymbol{z}})$, with $U_{k,\boldsymbol{z}} = F_k(X_k|\boldsymbol{z})$, $k = 1, \ldots, d$.
- By definition, the law of $\boldsymbol{U}_{\boldsymbol{Z}}$ conditionally to $\boldsymbol{Z} = \boldsymbol{z}$ is the conditional copula of $\boldsymbol{X}$ knowing $\boldsymbol{Z} = \boldsymbol{z}$, denoted by $C(\cdot|\boldsymbol{z})$.

Single-index copulae: the model

A conditional copula framework: For any $\boldsymbol{u} \in [0,1]^d$ and $\boldsymbol{z} \in \mathbb{R}^p$,

$$C(\boldsymbol{u}|\boldsymbol{z}) = C_{\theta(\boldsymbol{z})}(\boldsymbol{u}),$$

where $\theta : \mathbb{R}^p \to \mathbb{R}^q$ maps the vector of covariates to the (true) parameter of the conditional copula knowing $\boldsymbol{Z} = \boldsymbol{z}$, and $\mathcal{C} = \{C_\theta : \theta \in \Theta \subset \mathbb{R}^q\}$ is a **known** parametric family of copulae.

A conditional copula framework: For any $\boldsymbol{u} \in [0,1]^d$ and $\boldsymbol{z} \in \mathbb{R}^p$,

$$C(\boldsymbol{u}|\boldsymbol{z}) = C_{\theta(\boldsymbol{z})}(\boldsymbol{u}),$$

where $\theta : \mathbb{R}^p \to \mathbb{R}^q$ maps the vector of covariates to the (true) parameter of the conditional copula knowing $\boldsymbol{Z} = \boldsymbol{z}$, and $\mathcal{C} = \{C_\theta : \theta \in \Theta \subset \mathbb{R}^q\}$ is a **known** parametric family of copulae.

+ A single-index assumption: There exists an **unknown** function ψ s.t.

$$\theta(\boldsymbol{z}) = \psi(\beta_0, \beta_0' \boldsymbol{z}), \tag{1}$$

where the true parameter $\beta_0 \in \mathcal{B}$, a compact subset in $\mathbb{R}^m$, with $\beta_{0,1} = 1$.

Single-index copulae: the model

A conditional copula framework: For any $\boldsymbol{u} \in [0,1]^d$ and $\boldsymbol{z} \in \mathbb{R}^p$,

$$C(\boldsymbol{u}|\boldsymbol{z}) = C_{\theta(\boldsymbol{z})}(\boldsymbol{u}),$$

where $\theta : \mathbb{R}^p \to \mathbb{R}^q$ maps the vector of covariates to the (true) parameter of the conditional copula knowing $\boldsymbol{Z} = \boldsymbol{z}$, and $\mathcal{C} = \{C_\theta : \theta \in \Theta \subset \mathbb{R}^q\}$ is a **known** parametric family of copulae.

+ A single-index assumption: There exists an **unknown** function ψ s.t.

$$\theta(\boldsymbol{z}) = \psi(\beta_0, \beta_0' \boldsymbol{z}), \quad (1)$$

where the true parameter $\beta_0 \in \mathcal{B}$, a compact subset in $\mathbb{R}^m$, with $\beta_{0,1} = 1$.

Notation: $C(\cdot|\boldsymbol{z}) = C_\beta(\cdot|\beta' \boldsymbol{z})$.

In general, $C(\cdot|\boldsymbol{z})$ (the conditional copula of $\boldsymbol{X}$ knowing $\boldsymbol{Z}=\boldsymbol{z}$) is not equal to $\tilde{C}(\cdot|\beta_0'\boldsymbol{z})$, the conditional copula of $\boldsymbol{X}$ knowing $\beta_0'\boldsymbol{Z}=\beta_0'\boldsymbol{z}$.

In general, $C(\cdot|\boldsymbol{z})$ (the conditional copula of $\boldsymbol{X}$ knowing $\boldsymbol{Z} = \boldsymbol{z}$) is not equal to $\tilde{C}(\cdot|\beta_0'\boldsymbol{z})$, the conditional copula of $\boldsymbol{X}$ knowing $\beta_0'\boldsymbol{Z} = \beta_0'\boldsymbol{z}$.

In the former case, the margins are $F_k(\cdot|\boldsymbol{z})$, $k = 1, \ldots, d$, and in the latter case, they are $\tilde{F}_k(\cdot|\beta_0'\boldsymbol{z}) : x_k \mapsto P(X_k \leq x_k|\beta_0'\boldsymbol{z})$.

Single-index copulae: a warning

In general, $C(\cdot|\boldsymbol{z})$ (the conditional copula of $\boldsymbol{X}$ knowing $\boldsymbol{Z}=\boldsymbol{z}$) is not equal to $\tilde{C}(\cdot|\beta_0'\boldsymbol{z})$, the conditional copula of $\boldsymbol{X}$ knowing $\beta_0'\boldsymbol{Z}=\beta_0'\boldsymbol{z}$.

In the former case, the margins are $F_k(\cdot|\boldsymbol{z})$, $k=1,\ldots,d$, and in the latter case, they are $\tilde{F}_k(\cdot|\beta_0'\boldsymbol{z}) : x_k \mapsto P(X_k \leq x_k|\beta_0'\boldsymbol{z})$.

Denote $\tilde{\boldsymbol{U}}_\beta = (\tilde{F}_1(X_1|\beta'\boldsymbol{Z}),\ldots,\tilde{F}_d(X_d|\beta'\boldsymbol{Z}))$.

$\tilde{C}(\cdot|\beta'\boldsymbol{Z}=y)$ is the copula of $\tilde{\boldsymbol{U}}_\beta$ knowing $\beta'\boldsymbol{Z}=y$.

Single-index copulae: $\hat{\psi}(\beta, \beta' \boldsymbol{z})$

Estimation of $\psi(\cdot)$?

(A1) There exists a known functional Ψ s.t., for any $\beta \in \mathbb{R}^m$,

$$\psi(\beta, \beta' \boldsymbol{z}) = \Psi \left(C_\beta(\cdot | \beta' \boldsymbol{z}) \right). \tag{2}$$

Estimation of $\psi(\cdot)$?

(A1) There exists a known functional Ψ s.t., for any $\beta \in \mathbb{R}^m$,

$$\psi(\beta, \beta' \boldsymbol{z}) = \Psi\left(C_\beta(\cdot|\beta' \boldsymbol{z})\right). \tag{2}$$

(A2) There exists a known functional Ψ s.t., for any $\beta \in \mathbb{R}^m$,

$$\psi(\beta, \beta' \boldsymbol{z}) = \Psi\left(H_\beta(\cdot|\beta' \boldsymbol{z})\right), \tag{3}$$

where $H_\beta(\cdot|y)$ is the cdf of $(\boldsymbol{X}, \boldsymbol{Z})$ given $\beta' \boldsymbol{Z} = y$.

Estimation of $\psi(\cdot)$?

(A1) There exists a known functional Ψ s.t., for any $\beta \in \mathbb{R}^m$,

$$\psi(\beta, \beta' \boldsymbol{z}) = \Psi\left(C_\beta(\cdot|\beta' \boldsymbol{z})\right). \tag{2}$$

(A2) There exists a known functional Ψ s.t., for any $\beta \in \mathbb{R}^m$,

$$\psi(\beta, \beta' \boldsymbol{z}) = \Psi\left(H_\beta(\cdot|\beta' \boldsymbol{z})\right), \tag{3}$$

where $H_\beta(\cdot|y)$ is the cdf of $(\boldsymbol{X}, \boldsymbol{Z})$ given $\beta' \boldsymbol{Z} = y$.

$\Rightarrow$ empirical counterparts provide $\hat{\psi}(\beta, \beta' \boldsymbol{z})$.

Single-index copulae: $\hat{\psi}(\beta, \beta' \boldsymbol{z})$

Assumptions (2) and (3) are often moment-like conditions, as in GMM: there is a map $g : \mathbb{R}^{\bar{m}} \to \mathbb{R}^q$, $\bar{m} \geq m$, such that

$$\theta(\boldsymbol{z}) = g(m_1(\beta_0, \beta_0' \boldsymbol{z}), \ldots, m_{\bar{m}}(\beta_0, \beta_0' \boldsymbol{z})),$$

where $m_k(\beta, y) \in \mathbb{R}$, $k = 1, 2, \ldots$, are "moment" relations.

Single-index copulae: $\hat{\psi}(\beta, \beta' \boldsymbol{z})$

Assumptions (2) and (3) are often moment-like conditions, as in GMM: there is a map $g : \mathbb{R}^{\bar{m}} \to \mathbb{R}^q$, $\bar{m} \geq m$, such that

$$\theta(\boldsymbol{z}) = g(m_1(\beta_0, \beta_0' \boldsymbol{z}), \ldots, m_{\bar{m}}(\beta_0, \beta_0' \boldsymbol{z})),$$

where $m_k(\beta, y) \in \mathbb{R}$, $k = 1, 2, \ldots$, are "moment" relations.

In the case of (A1),

$$m_k(\beta, y) = \int \chi_k(\boldsymbol{u}, y) \, C_\beta(d\boldsymbol{u} | \beta' \boldsymbol{Z} = y),$$

for some known functions χ_k, $k = 1, \ldots, \bar{m}$.

Single-index copulae: $\hat{\psi}(\beta, \beta' z)$

Assumptions (2) and (3) are often moment-like conditions, as in GMM: there is a map $g : \mathbb{R}^{\bar{m}} \to \mathbb{R}^q$, $\bar{m} \geq m$, such that

$$\theta(\boldsymbol{z}) = g(m_1(\beta_0, \beta_0' \boldsymbol{z}), \ldots, m_{\bar{m}}(\beta_0, \beta_0' \boldsymbol{z})),$$

where $m_k(\beta, y) \in \mathbb{R}$, $k = 1, 2, \ldots$, are “moment” relations.

In the case of (A1),

$$m_k(\beta, y) = \int \chi_k(\boldsymbol{u}, y)\, C_\beta(d\boldsymbol{u} | \beta' \boldsymbol{Z} = y),$$

for some known functions χ_k, $k = 1, \ldots, \bar{m}$.

In the case of (A2),

$$m_k(\beta, y) = \int \chi_k(\boldsymbol{x}, \boldsymbol{z})\, H_\beta(d\boldsymbol{x}, d\boldsymbol{z} | \beta' \boldsymbol{Z} = y).$$

Example: Spearman's rho.

$m_k(\beta, \beta' \boldsymbol{z}) = \rho(\beta, \beta' \boldsymbol{z})$, a multivariate extension of the usual Spearman's rho, defined by

$$\rho(\beta, y) = \int \left(C_\beta(\boldsymbol{u} | \beta' \boldsymbol{Z} = y) - \prod_{j=1}^{d} u_j \right) d\boldsymbol{u}.$$

Through a d-dimensional integration by parts, check this moment is of the type (A1).

Example: Spearman's rho.

$m_k(\beta, \beta' \boldsymbol{z}) = \rho(\beta, \beta' \boldsymbol{z})$, a multivariate extension of the usual Spearman's rho, defined by

$$\rho(\beta, y) = \int \left(C_\beta(\boldsymbol{u} | \beta' \boldsymbol{Z} = y) - \prod_{j=1}^{d} u_j \right) d\boldsymbol{u}.$$

Through a d-dimensional integration by parts, check this moment is of the type (A1).

Other definitions of Spearman's rho are possible with an arbitrary dimension d: see Schmidt and Schmid (2007), for instance.

Example: Kendall's tau.

When $d = 2$, the Kendall's tau of $\boldsymbol{X}$ conditionally to $\boldsymbol{Z} = \boldsymbol{z}$ is

$$\tau_{\boldsymbol{z}} = 4 \int C(\boldsymbol{u}|\boldsymbol{z})C(d\boldsymbol{u}|\boldsymbol{z}) - 1 = 4 \int C_\beta(\boldsymbol{u}|\beta'\boldsymbol{z})C_\beta(d\boldsymbol{u}|\,\beta'\boldsymbol{z}) - 1. \tag{4}$$

Example: Kendall's tau.

When $d = 2$, the Kendall's tau of $\boldsymbol{X}$ conditionally to $\boldsymbol{Z} = \boldsymbol{z}$ is

$$\tau_{\boldsymbol{z}} = 4 \int C(\boldsymbol{u}|\boldsymbol{z}) C(d\boldsymbol{u}|\boldsymbol{z}) - 1 = 4 \int C_\beta(\boldsymbol{u}|\beta'\boldsymbol{z}) C_\beta(d\boldsymbol{u}|\,\beta'\boldsymbol{z}) - 1. \tag{4}$$

It will be denoted by $\tau(\beta, \beta'\boldsymbol{z})$.

Managing Kendall's tau, we work under Assumption (A1).

Single-index copulae: $\hat{\psi}(\beta, \beta' \boldsymbol{z})$ by Kendall's tau

$$\int C_\beta(\boldsymbol{u}|y)\, C_\beta(d\boldsymbol{u}|y) = \int \tilde{C}_\beta(\boldsymbol{u}|y)\, \tilde{C}_\beta(d\boldsymbol{u}|y), \text{ and}$$

$$\tau(\beta, \beta' \boldsymbol{Z} = y) = 4 \int \tilde{C}_\beta(\boldsymbol{u}|y)\, \tilde{C}_\beta(d\boldsymbol{u}|y) - 1. \tag{5}$$

$$\int C_\beta(\boldsymbol{u}|y)\, C_\beta(d\boldsymbol{u}|y) = \int \tilde{C}_\beta(\boldsymbol{u}|y)\, \tilde{C}_\beta(d\boldsymbol{u}|y), \text{ and}$$

$$\tau(\beta, \beta'\boldsymbol{Z} = y) = 4 \int \tilde{C}_\beta(\boldsymbol{u}|y)\, \tilde{C}_\beta(d\boldsymbol{u}|y) - 1. \qquad (5)$$

Moreover,

$$\tau(\beta, \beta'\boldsymbol{z}) = 4 \int H_\beta(\boldsymbol{x}, +\infty|\beta'\boldsymbol{z})\, H_\beta(d\boldsymbol{x}, +\infty|\beta'\boldsymbol{z}) - 1. \qquad (6)$$

$\Rightarrow$ Kendall's tau are of the two types (A1) and (A2) together.

The relations (5) and (6) are very useful for inference: the estimation of $H_\beta(\cdot|y)$ or $\tilde{C}_\beta(\cdot|y)$ is a lot less demanding than the non parametric estimation of $C_\beta(\cdot|\beta'\boldsymbol{z})$, that involves conditioning wrt $\boldsymbol{z} \in \mathbb{R}^p$ to manage its marginal laws.

The relations (5) and (6) are very useful for inference: the estimation of $H_\beta(\cdot|y)$ or $\tilde{C}_\beta(\cdot|y)$ is a lot less demanding than the non parametric estimation of $C_\beta(\cdot|\beta'\boldsymbol{z})$, that involves conditioning wrt $\boldsymbol{z} \in \mathbb{R}^p$ to manage its marginal laws.

In dimension d, many Kendall's tau can be built.

They may be associated to any couple of variables (X_i, X_j), $i, j = 1, \ldots, d$, $i \neq j$.

The relations (5) and (6) are very useful for inference: the estimation of $H_\beta(\cdot|y)$ or $\tilde{C}_\beta(\cdot|y)$ is a lot less demanding than the non parametric estimation of $C_\beta(\cdot|\beta'\boldsymbol{z})$, that involves conditioning wrt $\boldsymbol{z} \in \mathbb{R}^p$ to manage its marginal laws.

In dimension d, many Kendall's tau can be built.

They may be associated to any couple of variables (X_i, X_j), $i, j = 1, \ldots, d$, $i \neq j$.

Or they can be defined formally as in (4), with d'-dimension integrals, $d' \leq d$, focusing on sub-vectors of $\boldsymbol{X}$.

The relations (5) and (6) are very useful for inference: the estimation of $H_\beta(\cdot|y)$ or $\tilde{C}_\beta(\cdot|y)$ is a lot less demanding than the non parametric estimation of $C_\beta(\cdot|\beta'\boldsymbol{z})$, that involves conditioning wrt $\boldsymbol{z} \in \mathbb{R}^p$ to manage its marginal laws.

In dimension d, many Kendall's tau can be built.

They may be associated to any couple of variables (X_i, X_j), $i, j = 1, \ldots, d$, $i \neq j$.

Or they can be defined formally as in (4), with d'-dimension integrals, $d' \leq d$, focusing on sub-vectors of $\boldsymbol{X}$.

$\Rightarrow$ a lot of moments are available.

Inference: the criterion

An i.i.d. sample of observations $(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ in $\mathbb{R}^d \times \mathbb{R}^p$, that are drawn from the law of $(\boldsymbol{X}, \boldsymbol{Z})$.

We will rely on M-estimators of single-index models

An i.i.d. sample of observations $(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ in $\mathbb{R}^d \times \mathbb{R}^p$, that are drawn from the law of $(\boldsymbol{X}, \boldsymbol{Z})$.

We will rely on M-estimators of single-index models

If we were able to observe a sample of the random vector $\boldsymbol{U}_{\boldsymbol{Z}}$, i.e. $\boldsymbol{U}_i$, $i = 1, \ldots, n$, then our "naive" estimator of β_0 could be

$$\hat{\beta}_{naive} = \arg\max_{\beta \in \mathcal{B}} \sum_{i=1}^{n} \ln c_{\hat{\psi}(\beta, \beta' \boldsymbol{Z}_i)}(\boldsymbol{U}_i),$$

for some function $\hat{\psi}$ that estimates $\psi(\cdot, \cdot)$ consistently.

Inference: the criterion

Pb: we do not observe realizations of $\boldsymbol{U}$

$\Rightarrow$ replace the unknown vectors $\boldsymbol{U}_i$ by some estimates $\hat{\boldsymbol{U}}_i$, conditionally to $\boldsymbol{Z}_i$

We get a so-called pseudo-sample $\hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_n$.

Inference: the criterion

Pb: we do not observe realizations of $\boldsymbol{U}$

$\Rightarrow$ replace the unknown vectors $\boldsymbol{U}_i$ by some estimates $\hat{\boldsymbol{U}}_i$, conditionally to $\boldsymbol{Z}_i$

We get a so-called pseudo-sample $\hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_n$.

$$\hat{\beta} = \arg\max_{\beta \in \mathcal{B}} \sum_{i=1}^{n} \hat{\omega}_{i,n} \ln c_{\hat{\psi}(\beta, \beta' \boldsymbol{Z}_i)}(\hat{\boldsymbol{U}}_i), \qquad (7)$$

for some sequence of trimming functions $\hat{\omega}_{i,n}$.

Pb: we do not observe realizations of $\boldsymbol{U}$

$\Rightarrow$ replace the unknown vectors $\boldsymbol{U}_i$ by some estimates $\hat{\boldsymbol{U}}_i$, conditionally to $\boldsymbol{Z}_i$

We get a so-called pseudo-sample $\hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_n$.

$$\hat{\beta} = \arg\max_{\beta \in \mathcal{B}} \sum_{i=1}^{n} \hat{\omega}_{i,n} \ln c_{\hat{\psi}(\beta, \beta' \boldsymbol{Z}_i)}(\hat{\boldsymbol{U}}_i), \qquad (7)$$

for some sequence of trimming functions $\hat{\omega}_{i,n}$.

Such trimming functions allow to control some boundary effects and the uniform convergence of our kernel estimates.

We set a fixed trimming for $\mathcal{Z}$. This is permitted, because the law of the $\boldsymbol{U}$ knowing $\boldsymbol{Z} \in \mathcal{Z}$ depends on the true parameter β_0 only.

Inference: choice of $\hat{\boldsymbol{U}}$

Several possibilities:

1. parametric marginal conditional distributions: for every $k = 1, \ldots, d$ and $\boldsymbol{z}$, $F_k(\cdot|\boldsymbol{z})$ belongs to a parametric family $\mathcal{G}_k = \{G_{k,\theta_k},\ \theta_k \in \Theta_k\}$. And the true parameter $\theta_k(\boldsymbol{z})$ is estimated by $\hat{\theta}_k(\boldsymbol{z})$.

Inference: choice of $\hat{\boldsymbol{U}}$

Several possibilities:

1. parametric marginal conditional distributions: for every $k = 1, \ldots, d$ and $\boldsymbol{z}$, $F_k(\cdot|\boldsymbol{z})$ belongs to a parametric family $\mathcal{G}_k = \{G_{k,\theta_k},\ \theta_k \in \Theta_k\}$. And the true parameter $\theta_k(\boldsymbol{z})$ is estimated by $\hat{\theta}_k(\boldsymbol{z})$.
2. nonparametric estimates of conditional expectations:

$$\hat{F}(\boldsymbol{x}|\boldsymbol{z}) = \sum_{j=1}^{n} w_{j,n}(\boldsymbol{z})\mathbf{1}(\boldsymbol{X}_j \leq \boldsymbol{x}), \tag{8}$$

with weights

$$w_{j,n}(\boldsymbol{z}) = \boldsymbol{K}\left(\boldsymbol{Z}_j - \boldsymbol{z}, \boldsymbol{h}\right) / \sum_{l=1}^{n} \boldsymbol{K}\left(\boldsymbol{Z}_l - \boldsymbol{z}, \boldsymbol{h}\right), \tag{9}$$

$\boldsymbol{K}$ is a p-dimensional kernel functions and $\boldsymbol{h} := (h_1, \ldots, h_p)$ is a p-vector of bandwidths $h_k > 0$.

Inference: choice of $\hat{\boldsymbol{U}}$

2. For example,

$$\boldsymbol{K}\left(\boldsymbol{Z}_j - \boldsymbol{z}, \boldsymbol{h}\right) = \prod_{k=1}^{p} K_k\left(\frac{Z_{j,k} - z_k}{h_k}\right),$$

for some univariate kernel functions K_k.

Inference: choice of $\hat{\boldsymbol{U}}$

2. For example,

$$\boldsymbol{K}\left(\boldsymbol{Z}_j - \boldsymbol{z}, \boldsymbol{h}\right) = \prod_{k=1}^{p} K_k \left(\frac{Z_{j,k} - z_k}{h_k} \right),$$

   for some univariate kernel functions K_k.

   Nonparametric estimators of the cdf $F_k(\boldsymbol{x}|\boldsymbol{z})$ are obtained using $\hat{F}_k(x|\boldsymbol{z}) = \hat{F}(x, +\infty_{(-k)}|\boldsymbol{z})$.

   The marginal "unfeasible" observations $U_{i,k} = F_k(X_{i,k}|\boldsymbol{Z}_i)$ are estimated by $\hat{U}_{i,k} = \hat{F}_k(X_{i,k}|\boldsymbol{Z}_i)$.

2. For example,

$$\boldsymbol{K}\left(\boldsymbol{Z}_j - \boldsymbol{z}, \boldsymbol{h}\right) = \prod_{k=1}^{p} K_k\left(\frac{Z_{j,k} - z_k}{h_k}\right),$$

   for some univariate kernel functions K_k.

   Nonparametric estimators of the cdf $F_k(\boldsymbol{x}|\boldsymbol{z})$ are obtained using $\hat{F}_k(x|\boldsymbol{z}) = \hat{F}(x, +\infty_{(-k)}|\boldsymbol{z})$.

   The marginal "unfeasible" observations $U_{i,k} = F_k(X_{i,k}|\boldsymbol{Z}_i)$ are estimated by $\hat{U}_{i,k} = \hat{F}_k(X_{i,k}|\boldsymbol{Z}_i)$.

3. Others: marginal single-index distributions, additive models... to avoid the curse of dimensionality on margins.

Consistency

Assumption 1

Let us set $\mathcal{Z} := [-M, M]^p$ *and* $\mathcal{E}_n = [\nu_n, 1-\nu_n]^d$ *for some positive sequence* (ν_n), $\nu_n \in (0, 1/2)$, $\nu_n \to 0$.

The trimming functions are $\omega_n : [0,1]^d \times \mathbb{R}^p \to [0,1]$, $(\boldsymbol{u}, \boldsymbol{z}) \mapsto \mathbf{1}(\boldsymbol{u} \in \mathcal{E}_n, \boldsymbol{z} \in \mathcal{Z})$.

Notations: $\hat{\omega}_{i,n} = \omega_n(\hat{\boldsymbol{U}}_i, \boldsymbol{Z}_i)$, $\omega_{i,n} := \omega_n(\boldsymbol{U}_i, \boldsymbol{Z}_i)$ and $\omega_i = \omega_{i,\infty} = \mathbf{1}(\boldsymbol{Z}_i \in \mathcal{Z})$.

Assumption 2

The parameter β_0 *is identifiable, i.e. two different parameters induce two different laws of* $\boldsymbol{U}_{\boldsymbol{Z}}$*, knowing* $\boldsymbol{Z} \in \mathcal{Z}$*.*

For every $\boldsymbol{z} \in \mathcal{Z}$*, the function* $\mathcal{M}(\boldsymbol{z}) : \beta \mapsto E[\ln c_{\psi(\beta,\beta'\boldsymbol{z})}(\boldsymbol{U}_{\boldsymbol{z}})]$ *is uniquely maximized at* $\beta = \beta_0$*.*

There exists a function g *s.t., for every* $\boldsymbol{z} \in \mathcal{Z}$ *and some* $a > 1$*,*

$$\sup_{\beta \in \mathcal{B}} |\ln c_{\psi(\beta,\beta'\boldsymbol{z})}(\boldsymbol{U}_{\boldsymbol{z}})| \leq g(\boldsymbol{U}_{\boldsymbol{z}}, \boldsymbol{z}), \; E[g^a(\boldsymbol{U}_{\boldsymbol{Z}}, \boldsymbol{Z}).1(\boldsymbol{Z} \in \mathcal{Z})] < \infty. \tag{10}$$

The limiting objective function will be

$$M(\beta) := E\left[\ln c_{\psi(\beta,\beta'\boldsymbol{Z})}(\boldsymbol{U}) \,|\, \boldsymbol{Z} \in \mathcal{Z}\right].$$

Consistency

Assumption 3

$$\sup_{\boldsymbol{z}\in\mathcal{Z}} \sup_{\beta\in\mathcal{B}} \left| \hat{\psi}(\beta, \beta'\boldsymbol{z}) - \psi(\beta, \beta'\boldsymbol{z}) \right| = o_P(1). \tag{11}$$

Moreover, there exists a deterministic sequence (δ_n), $\delta_n = o(\nu_n)$, *s.t.*

$$\sup_i |\hat{\boldsymbol{U}}_i - \boldsymbol{U}_i| . \mathbf{1}(\boldsymbol{Z}_i \in \mathcal{Z}) = O_P(\delta_n). \tag{12}$$

Definition 2

- A function $f : (0,1) \to (0,\infty)$ is called u-shaped if it is symmetric about $1/2$ and decreasing on $(0,1/2]$.
- For $\beta \in (0,1)$ and a u-shaped function r, define

$$r_\beta(t) = \begin{cases} r(\beta u) & \text{if} \quad 0 < u \leq 1/2; \\ r(1-\beta(1-u)) & \text{if} \quad 1/2 < u \leq 1. \end{cases}$$

If, for every $\beta > 0$ in a neighborhood of 0, there exists a constant M_β, such that $r_\beta < M_\beta . r$ on $(0,1)$, then r is called a reproducing u-shaped function.

Definition 2

- A function $f : (0,1) \to (0,\infty)$ is called u-shaped if it is symmetric about $1/2$ and decreasing on $(0,1/2]$.
- For $\beta \in (0,1)$ and a u-shaped function r, define

$$r_\beta(t) = \begin{cases} r(\beta u) & \text{if} \quad 0 < u \leq 1/2; \\ r(1-\beta(1-u)) & \text{if} \quad 1/2 < u \leq 1. \end{cases}$$

  If, for every $\beta > 0$ in a neighborhood of 0, there exists a constant M_β, such that $r_\beta < M_\beta . r$ on $(0,1)$, then r is called a reproducing u-shaped function.
- We denote by $\mathcal{R}$ the set of univariate reproducing u-shaped functions. The set $\mathcal{R}_d$ is the set of functions $r : (0,1)^d \to \mathbb{R}^+$, $r(\boldsymbol{u}) = \prod_{k=1}^{d} r_k(u_k)$, and $r_k \in \mathcal{R}$ for every k.

Consistency

Definition 2

- A function $f : (0,1) \to (0,\infty)$ is called u-shaped if it is symmetric about $1/2$ and decreasing on $(0,1/2]$.
- For $\beta \in (0,1)$ and a u-shaped function r, define

$$r_\beta(t) = \begin{cases} r(\beta u) & \text{if} \quad 0 < u \leq 1/2; \\ r(1-\beta(1-u)) & \text{if} \quad 1/2 < u \leq 1. \end{cases}$$

  If, for every $\beta > 0$ in a neighborhood of 0, there exists a constant M_β, such that $r_\beta < M_\beta . r$ on $(0,1)$, then r is called a reproducing u-shaped function.
- We denote by $\mathcal{R}$ the set of univariate reproducing u-shaped functions. The set $\mathcal{R}_d$ is the set of functions $r : (0,1)^d \to \mathbb{R}^+$, $r(\boldsymbol{u}) = \prod_{k=1}^d r_k(u_k)$, and $r_k \in \mathcal{R}$ for every k.

Typically, $r(u) = C_r u^{-a}(1-u)^{-a}$, for some positive constants a and C_r (Tsukahara 2005).

Assumption 4

There exist some functions $r, \tilde{r}_1, \ldots, \tilde{r}_d$ *in* $\mathcal{R}_d$ *s.t., for every* $\boldsymbol{u} \in (0,1)^d$,

$$\sup_{\theta \in \Theta} |\nabla_\theta \ln c_\theta(\boldsymbol{u})| \leq r(\boldsymbol{u}),\ E\left[r(\boldsymbol{U}_{\boldsymbol{Z}})\mathbf{1}(\boldsymbol{Z} \in \mathcal{Z})\right] < \infty,$$

$$\sup_{\theta \in \Theta} |\partial_{u_k} \ln c_\theta(\boldsymbol{u})| \leq \tilde{r}_k(\boldsymbol{u}),\ \textit{for every } k = 1, \ldots, d,\ \textit{with}$$

$$E\left[U_k(1 - U_k)\tilde{r}_k(\boldsymbol{U}_{\boldsymbol{Z}})\mathbf{1}(\boldsymbol{Z} \in \mathcal{Z})\right] < \infty.$$

Consistency

Theorem 3

Under the assumptions 1-4, the estimator $\hat{\beta}$ given by (7) tends to β_0 in probability, when n tends to the infinity.

Example : the Gaussian copula model

$$C_{\beta_0}(\boldsymbol{u}|\boldsymbol{Z}=\boldsymbol{z}) = C^G_{\Sigma(\boldsymbol{z})}(\boldsymbol{u}) = \Phi_{\Sigma(\boldsymbol{z})}\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_d)\right),$$

where the correlation matrix depends on the index $\beta_0'\boldsymbol{z}$:

$$\Sigma(\boldsymbol{z}) = \psi(\beta_0, \beta_0'\boldsymbol{z}) = [\theta_{k,l}(\boldsymbol{z})]_{1\leq k,l\leq d},$$

Example : the Gaussian copula model

$$C_{\beta_0}(\boldsymbol{u}|\boldsymbol{Z}=\boldsymbol{z}) = C^G_{\Sigma(\boldsymbol{z})}(\boldsymbol{u}) = \Phi_{\Sigma(\boldsymbol{z})}\left(\Phi^{-1}(u_1),\dots,\Phi^{-1}(u_d)\right),$$

where the correlation matrix depends on the index $\beta_0'\boldsymbol{z}$:

$$\Sigma(\boldsymbol{z}) = \psi(\beta_0, \beta_0'\boldsymbol{z}) = [\theta_{k,l}(\boldsymbol{z})]_{1\le k,l\le d},$$

$$\theta_{k,l}(\boldsymbol{z}) = \sin(\frac{\pi}{2}\tau_{k,l}(\beta_0'\boldsymbol{z}))$$

$\tau_{k,l}(y)$: the conditional Kendall's tau that is associated to (X_k, X_l), knowing $\beta_0'\boldsymbol{Z} = y$, that can be estimated easily by standard nonparametric techniques, as in Gijbels et al. (2011).

Example : the Gaussian copula model

$$C_{\beta_0}(\boldsymbol{u}|\boldsymbol{Z}=\boldsymbol{z}) = C^G_{\Sigma(\boldsymbol{z})}(\boldsymbol{u}) = \Phi_{\Sigma(\boldsymbol{z})}\left(\Phi^{-1}(u_1),\dots,\Phi^{-1}(u_d)\right),$$

where the correlation matrix depends on the index $\beta_0'\boldsymbol{z}$:

$$\Sigma(\boldsymbol{z}) = \psi(\beta_0, \beta_0'\boldsymbol{z}) = [\theta_{k,l}(\boldsymbol{z})]_{1\leq k,l\leq d},$$

$$\theta_{k,l}(\boldsymbol{z}) = \sin(\frac{\pi}{2}\tau_{k,l}(\beta_0'\boldsymbol{z}))$$

$\tau_{k,l}(y)$: the conditional Kendall's tau that is associated to (X_k, X_l), knowing $\beta_0'\boldsymbol{Z} = y$, that can be estimated easily by standard nonparametric techniques, as in Gijbels et al. (2011).

$$\hat{\psi}(\beta, \beta'\boldsymbol{z}) = [\sin(\frac{\pi}{2}\hat{\tau}_{k,l}(\beta'\boldsymbol{z}))]_{1\leq k,l\leq d},$$

$$\hat{\tau}_{k,l}(t) := 4\int \hat{\hat{C}}_{k,l}(u,v|\beta'\boldsymbol{Z}=t)\,\hat{\hat{C}}_{k,l}(du,dv|\beta'\boldsymbol{Z}=t) - 1,$$

for some estimator $\hat{\hat{C}}_{k,l}(\cdot|\beta'\boldsymbol{z})$ of the copula of (X_k, X_l) given $\beta'\boldsymbol{Z}$.

Example : the Gaussian copula cont'd

The marginal cdfs' $\hat{U}_k$, $k = 1, \ldots, d$: standard univariate kernel-based conditional distributions $\hat{U}_{i,k} := \hat{F}_k(X_{i,k}|\boldsymbol{Z}_i)$.

Example : the Gaussian copula cont'd

The marginal cdfs' $\hat{U}_k$, $k = 1, \ldots, d$: standard univariate kernel-based conditional distributions $\hat{U}_{i,k} := \hat{F}_k(X_{i,k}|\boldsymbol{Z}_i)$.

For a large choice of bandwiths, the distance between $\hat{\boldsymbol{U}}_i$ and $\boldsymbol{U}_i$ is of order $\sqrt{\ln(n)}/\sqrt{nh}$ uniformly (Einmahl and Mason 2005):

$$\sup_i |\hat{\boldsymbol{U}}_i - \boldsymbol{U}_i|.\mathbf{1}(\boldsymbol{Z}_i \in \mathcal{Z}) = O_P(\sqrt{\ln(n)}/\sqrt{nh^p} + h^{p\pi}).$$

Example : the Gaussian copula cont'd

The marginal cdfs' $\hat{U}_k$, $k = 1, \ldots, d$: standard univariate kernel-based conditional distributions $\hat{U}_{i,k} := \hat{F}_k(X_{i,k}|\boldsymbol{Z}_i)$.

For a large choice of bandwiths, the distance between $\hat{\boldsymbol{U}}_i$ and $\boldsymbol{U}_i$ is of order $\sqrt{\ln(n)}/\sqrt{nh}$ uniformly (Einmahl and Mason 2005):

$$\sup_i |\hat{\boldsymbol{U}}_i - \boldsymbol{U}_i| . \mathbf{1}(\boldsymbol{Z}_i \in \mathcal{Z}) = O_P(\sqrt{\ln(n)}/\sqrt{nh^p} + h^{p\pi}).$$

$\Rightarrow$ Assumption 3 is easily satisfied.

Example : the Gaussian copula cont'd

Assumption 2: to check (10), we can require

$$\inf_{\boldsymbol{z}\in\mathcal{Z}} \inf_{\beta\in\mathcal{B}} \lambda_{\min}(\psi(\beta,\beta'\boldsymbol{z})) \geq \underline{\lambda} > 0, \tag{13}$$

In this case, it is easy to bound the log-density of $\boldsymbol{U}$ (conditionally to $\boldsymbol{Z}$) from above, and to satisfy (10).

Example : the Gaussian copula cont'd

Assumption 2: to check (10), we can require

$$\inf_{\boldsymbol{Z}\in\mathcal{Z}} \inf_{\beta\in\mathcal{B}} \lambda_{\min}(\psi(\beta,\beta'\boldsymbol{z})) \geq \underline{\lambda} > 0, \tag{13}$$

In this case, it is easy to bound the log-density of $\boldsymbol{U}$ (conditionally to $\boldsymbol{Z}$) from above, and to satisfy (10).

Assumption 4 is satisfied, as in most usual copula families: choose $r(\boldsymbol{u}) \propto \prod_{k=1}^d u_k^{-a}(1-u_k)^{-a}$ for some $a>0$, and

$$\tilde{r}_k(\boldsymbol{u}) \propto u_k^{-a-1}(1-u_k)^{-a-1} \prod_{l=1,l\neq k}^{d} u_l^{-a}(1-u_l)^{-a}.$$

$\Rightarrow$ $\hat{\beta}$ is consistent under a Gaussian copula framework.

Asymptotic normality

Notation: $\psi_i = \psi(\beta_0, \beta_0' \boldsymbol{Z}_i)$ and $\hat{\psi}_i = \hat{\psi}(\beta_0, \beta_0' \boldsymbol{Z}_i)$.

Assumption 5

For every $\boldsymbol{z} \in \mathcal{Z}$, assume that $\psi_{\boldsymbol{z}} : \mathcal{B} \to \Theta, \beta \mapsto \psi(\beta, \beta' \boldsymbol{z})$ is two times continuously differentiable. Moreover, for every $\theta \in \Theta$, assume that $\ln c_\theta : (0,1)^d \to \mathbb{R}, \boldsymbol{u} \mapsto \ln c_\theta(\boldsymbol{u})$ is two times continuously differentiable.

Assumption 6

Let the functions on $(0,1)^d \times \mathcal{Z}$ defined by

$$f(\boldsymbol{u}, \boldsymbol{z}) = \frac{\nabla_\theta c_\theta}{c_\theta}_{|\theta = \psi(\beta_0, \beta_0' \boldsymbol{z})}(\boldsymbol{u}), \text{ and } \hat{f}(\boldsymbol{u}, \boldsymbol{z}) = \frac{\nabla_\theta c_\theta}{c_\theta}_{|\theta = \hat{\psi}(\beta_0, \beta_0' \boldsymbol{z})}(\boldsymbol{u}).$$

For almost every realization, the functions f and $\hat{f}$ belong to a Donsker class for the underlying law of $(\boldsymbol{X}, \boldsymbol{Z})$.

Assumption 7

Let the functions on $\mathcal{Z}$ defined by

$$p : \boldsymbol{z} \to p(\boldsymbol{z}) = \nabla_\beta \psi(\beta, \beta' \boldsymbol{z})_{|\beta=\beta_0}, \quad \textit{and}$$

$$\hat{p} : \boldsymbol{z} \to \hat{p}(\boldsymbol{z}) = \nabla_\beta \hat{\psi}(\beta, \beta' \boldsymbol{z})_{\beta=\beta_0}.$$

For almost every realization, the functions p and $\hat{p}$ belong to a Donsker class for the underlying law of $(\boldsymbol{X}, \boldsymbol{Z})$.

Assumption 8

Assume that, for every $(\boldsymbol{u}, \boldsymbol{u}') \in (0,1)^{2d}$*, we have*

$$\left|\nabla_\theta \ln c_\theta(\boldsymbol{u}) - \nabla_\theta \ln c_{\theta'}(\boldsymbol{u})\right| \leq \Phi(\boldsymbol{u}).|\theta - \theta'|, \quad (14)$$

$$\left|\nabla^2_\theta \ln c_\theta(\boldsymbol{u}) - \nabla^2_\theta \ln c_{\theta'}(\boldsymbol{u})\right| \leq \Phi(\boldsymbol{u}).|\theta - \theta'|, \quad (15)$$

for some function Φ *s.t.* $E[\Phi(\boldsymbol{U})] < \infty$.

Assumption 9

Assume that, for every $(\beta_1, \beta_2) \in \mathcal{B}^2$ *and* $j = 1, 2$,

$$\sup_{\boldsymbol{z} \in \mathcal{Z}} |\nabla^j_\beta \psi(\beta_1, \beta_1' \boldsymbol{z}) - \nabla^j_\beta \psi(\beta_2, \beta_2' \boldsymbol{z})| \leq C.|\beta_1 - \beta_2|,$$

where C *is a finite constant.*

Asymptotic normality

Assumption 10

Assume that

$$\sup_{\beta\in\mathcal{B},\boldsymbol{z}\in\mathcal{Z}} \left|\psi(\beta,\beta'\boldsymbol{z}) - \hat{\psi}(\beta,\beta'\boldsymbol{z})\right| = o_P(1), \quad (16)$$

$$\sup_{\beta\in\mathcal{B},\boldsymbol{z}\in\mathcal{Z}} \left|\nabla_\beta\psi(\beta,\beta'\boldsymbol{z}) - \nabla_\beta\hat{\psi}(\beta,\beta'\boldsymbol{z})\right| = o_P(1), \quad (17)$$

$$\sup_{\beta\in\mathcal{B},\boldsymbol{z}\in\mathcal{Z}} \left|\nabla^2_\beta\psi(\beta,\beta'\boldsymbol{z}) - \nabla^2_\beta\hat{\psi}(\beta,\beta'\boldsymbol{z})\right| = o_P(1). \quad (18)$$

Asymptotic normality

Assumption 11

$$\sup_{\boldsymbol{z}\in\mathcal{Z}} \sup_{k} \|\hat{F}_k(\cdot|\boldsymbol{z}) - F_k(\cdot|\boldsymbol{z})\|_\infty = O_P(\varepsilon_n),$$

with $\varepsilon_n = o(n^{-1/4})$.

Assumption 12

Let Assume that

$$\begin{aligned} \sup_{\boldsymbol{z}\in\mathcal{Z}} |\hat{\psi}(\beta_0, \beta_0'\boldsymbol{z}) - \psi(\beta_0, \beta_0'\boldsymbol{z})| &= O_P(\eta_{1n}), \\ \sup_{\boldsymbol{z}\in\mathcal{Z}} |\nabla_\beta\hat{\psi}(\beta_0, \beta_0'\boldsymbol{z}) - \nabla_\beta\psi(\beta_0, \beta_0'\boldsymbol{z})| &= O_P(\eta_{2n}), \end{aligned}$$

with $\varepsilon_n\eta_{jn} = o(n^{-1/2})$, *for* $j = 1, 2$, *and* $\eta_{1n}\eta_{2n} = o(n^{-1/2})$.

Assumption 13

Assume that

$$\nabla_\theta \ln c_\theta(\boldsymbol{u}) - \nabla_\theta \ln c_\theta(\boldsymbol{u}') = \Lambda_\theta(\boldsymbol{u}).(\boldsymbol{u}-\boldsymbol{u}') + \rho_\theta(\boldsymbol{u}^*).(\boldsymbol{u}-\boldsymbol{u}')^{(2)},$$

for some $\boldsymbol{u}^*$ *s.t.* $|\boldsymbol{u}-\boldsymbol{u}^*| < |\boldsymbol{u}-\boldsymbol{u}'|$, *and, for every* $k = 1, \ldots, d$, *there exists a constance* $\alpha \in (0,1)$ *s.t.*

$$\sup_\theta |\nabla_\theta(\Lambda_\theta(\boldsymbol{u}))_k| \leq \Gamma_k(\boldsymbol{u}), \; E\left[U_k^\alpha(1-U_k)^\alpha \Gamma_k\left(\boldsymbol{U}_{\boldsymbol{Z}}\right)\right] < \infty.$$

Moreover, for every $k, l = 1, \ldots, d$, *there exists a function* $\bar{r}_{k,l}$ *in* $\mathcal{R}_d$ *s.t., for every* $\boldsymbol{u} \in (0,1)^d$,

$$\sup_{\theta\in\Theta} |(\rho_\theta(\boldsymbol{u}))_{k,l}| \leq \bar{r}_{k,l}(\boldsymbol{u}), \; \textit{and}$$

$$E\left[U_k^\gamma(1-U_k)^\gamma U_l^\gamma(1-U_l)^\gamma \bar{r}_{k,l}(\boldsymbol{U}_{\boldsymbol{Z}})\right] < \infty, \; \textit{for some } \gamma \in (0,1)$$

Assumption 14

Assume that $\beta \mapsto M(\beta)$ *is twice continuously differentiable. Its Hessian matrix at point* β_0 *is denoted by* $\Sigma = \nabla^2_\beta M(\beta_0)$, *and is invertible.*

Asymptotic normality

Assumption 15

For any $\boldsymbol{u} \in \mathbb{R}^d$, *set*

$$g(\boldsymbol{u}, \boldsymbol{z}) := \sup_{\theta \in B(\theta_0(\boldsymbol{Z}), \eta_{1,n})} \sup_{\boldsymbol{v} \in B(\boldsymbol{u}, \delta_n)} |\nabla_\theta \ln c_\theta(\boldsymbol{v})|,$$

where $B(\boldsymbol{u}, \delta)$ *(resp.* $B(\theta, \eta)$*) denotes the closed ball of center* $\boldsymbol{u}$ *(resp.* θ*) and radius* δ *(resp.* η*). Assume*

$$\sup_{k=1,\ldots,d} E[g(\boldsymbol{U}_i, \boldsymbol{Z}_i) \cdot \mathbf{1}(\boldsymbol{Z}_i \in \mathcal{Z}, |U_{i,k} - \nu_n| < \delta_n)] = o(n^{-1/2}), \quad (19)$$

and similarly after having replaced ν_n *by* $1 - \nu_n$.

Broadly speaking, it means that

$$\delta_n \int \nabla_\theta c_\theta(\boldsymbol{u}_{-k}, \nu_n | \boldsymbol{z})_{|\theta = \theta_0(\boldsymbol{z})} \cdot \mathbf{1}(\boldsymbol{z} \in \mathcal{Z}) \, d\boldsymbol{u}_{-k} \, d\mathbb{P}_{\boldsymbol{Z}}(\boldsymbol{z}) = o(n^{-1/2}),$$

and the same replacing ν_n by $1 - \nu_n$.

Asymptotic normality

Theorem 4

Under Assumptions 1 to 15,

$$(\hat{\beta} - \beta_0) = -\Sigma^{-1} \cdot \frac{1}{n} \sum_{i=1}^{n} \omega_{i,n} \frac{\nabla_\theta c_\theta}{c_\theta}_{|\theta=\psi_i} (\hat{\boldsymbol{U}}_i) \nabla_\beta \psi(\beta, \beta' \boldsymbol{Z}_i)_{|\beta=\beta_0}$$

$$+ \quad o_P(n^{-1/2}).$$

Asymptotic normality

Assumption 16

For every $k = 1, \dots, d$, $x \in \mathbb{R}$ *and* $\boldsymbol{z} \in \mathcal{Z}$, *we can write*

$$\hat{F}_k(x|\boldsymbol{z}) - F_k(x|\boldsymbol{z}) = \frac{1}{n}\sum_{j=1}^{n} a_{k,n}(\boldsymbol{X}_j, \boldsymbol{Z}_j, x, \boldsymbol{z}) + r_n(x|\boldsymbol{z}), \tag{20}$$

for some particular functions $a_{k,n}$ *and for some sequence* (r_n) *s.t.*

$$\sup_{x\in\mathbb{R}} \sup_{\boldsymbol{z}\in\mathcal{Z}} |r_n(x,\boldsymbol{z})| = o_P(n^{-1/2}).$$

$$\hat{U}_{i,k} - U_{i,k} = \frac{1}{n}\sum_{j=1}^{n} a_{k,n}(\boldsymbol{X}_j, \boldsymbol{Z}_j, X_{i,k}, \boldsymbol{Z}_i) + r_{n,i}, \;\; n^{1/2}\sup_i |r_{n,i}| = o_P(1).$$

Denote $\boldsymbol{a}_n(\boldsymbol{X}_j, \boldsymbol{Z}_j, \boldsymbol{X}_i, \boldsymbol{Z}_i)$ (or even $\boldsymbol{a}_{i,j}$) the d-vector whose components are $a_{k,n}(\boldsymbol{X}_j, \boldsymbol{Z}_j, X_{i,k}, \boldsymbol{Z}_i)$, $k = 1, \dots, d$.

Asymptotic normality

Assumption 17

Assume that there exists a function W such that

$$\sup_{x\in\mathbb{R},\boldsymbol{Z}\in\mathcal{Z}} \left|E\left[a_n(\boldsymbol{X}_j,\boldsymbol{Z}_j,\boldsymbol{x},\boldsymbol{z})\right] - W(\boldsymbol{z},\boldsymbol{x})\right| = o(n^{-1/2}),$$

and such that

$$E\left[\left\{\Lambda_{\psi(\beta_0,\beta_0'\boldsymbol{Z}_i)}(\boldsymbol{U}_i).W(\boldsymbol{Z},\boldsymbol{X})\nabla_\beta\psi(\beta,\beta'\boldsymbol{Z}_i)_{|\beta=\beta_0}\right\}^2\right] < \infty.$$

Hopefully, W is often the null function...

Corollary 5

Under the Assumptions of Theorem 4 and Assumptions 16 to 17, we have

$$n^{1/2}\left\{\Sigma.(\hat{\beta}-\beta_0)+b_n\right\} \Longrightarrow \mathcal{N}(0,S),$$

where $S = E[\omega_1\mathcal{M}_1\mathcal{M}_1']$, *where*

$$\mathcal{M}_1 = \frac{\nabla_\theta c_\theta}{c_\theta}_{|\theta=\psi_1}(\boldsymbol{U}_1)\nabla_\beta\psi(\beta,\beta'\boldsymbol{Z}_1)_{|\beta=\beta_0} + \Lambda_{\psi(\beta_0,\beta_0'\boldsymbol{Z}_1)}(\boldsymbol{U}_1).W(\boldsymbol{Z}_1,\boldsymbol{X}_1)\nabla_\beta\psi(\beta,\beta'\boldsymbol{Z}_1)_{|\beta=\beta_0},$$

$$b_n = E[\omega_{1,n}\mathcal{M}_1] = E[\mathbf{1}(\boldsymbol{U}_1\in\mathcal{E}_n,\boldsymbol{Z}_1\in\mathcal{Z})\mathcal{M}_1].$$

Asymptotic normality

In general, the bias b_n cannot be removed, even if $E[\boldsymbol{a}_{i,j}] = 0$: the trimming part $E[\omega_{i,n}\mathcal{M}_i] \sim \delta_n$, that is not $o(n^{-1/2})$ in general.

Asymptotic normality

In general, the bias b_n cannot be removed, even if $E[\boldsymbol{a}_{i,j}] = 0$: the trimming part $E[\omega_{i,n}\mathcal{M}_i] \sim \delta_n$, that is not $o(n^{-1/2})$ in general.

Nonetheless, if

$$E\Big[\Lambda_{\psi(\beta_0,\beta_0'\boldsymbol{Z}_1)}(\boldsymbol{U}_1).W(\boldsymbol{Z}_1,\boldsymbol{X}_1)\nabla_\beta\psi(\beta,\beta'\boldsymbol{Z}_1)_{|\beta=\beta_0} \cdot \quad \{\mathbf{1}(|U_{k,1}-\nu_n|<\delta_n)+\mathbf{1}(|1-U_{k,1}-\nu_n|<\delta_n)\}\Big] = o(n^{-1/2}),$$

for every $k = 1, \ldots, d$, then $n^{1/2}b_n = o(1)$ and

$$n^{1/2}(\hat{\beta}-\beta_0) \Longrightarrow \mathcal{N}(0, \Sigma^{-1}S\Sigma^{-1}).$$

- Einmahl, U. and Mason, D. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.* **33**, 1380-1403.
- J-D. Fermanian and M. Wegkamp (2012). Time-dependent copulae, *Journal of Multivariate Analysis* **110**, 19-29.
- Gijbels, I., Veraverbeke, N. and Omelka, M. (2011). Conditional copulae, association measures and their applications. *Computational Stat. Data Anal.* **55**, 1919-1932.
- Schmid, F. and R. Schmidt. (2007). Multivariate extensions of Spearman rho and related statistics, *Statistics & Probability Letters*, vol. 77 (4), 407-416.
- Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canadian Journal of Statistics* **33**, 357-375.