When uniform weak convergence fails: Empirical processes for dependence functions and residuals via epi- and hypographs

Axel Bücher, Johan Segers and Stanislav Volgushev

Université catholique de Louvain and Ruhr-Universität Bochum

Van Dantzig Seminar, Mathematical Institute, Leiden University, 11 Apr 2014

InterUniversity Attraction Poles, Phase VII, 2012-2017

P7/06 StUDyS

Developing crucial Statistical methods for Understanding major complex Dynamic Systems in natural, biomedical and social sciences

SFB 823

Motivation

Uniform convergence of bounded functions

Strong implications vs. **Restricted applicability**

Motivation

Uniform convergence of bounded functions

Strong implications vs. **Restricted applicability**

- Implies pointwise, continuous, L^p-convergence . . .

Motivation

Uniform convergence of bounded functions

Strong implications vs. **Restricted applicability**

- Implies pointwise, continuous, L^p-convergence ...
- Well-developed weak convergence theory

Great success story in mathematical statistics

[Van der Vaart and Wellner (1996): Weak convergence and empirical processes]

Motivation

Uniform convergence of bounded functions

Strong implications vs. **Restricted applicability**

- Implies pointwise, continuous, L^p-convergence ...
- Well-developed weak convergence theory

  Great success story in mathematical statistics

  [Van der Vaart and Wellner (1996): Weak convergence and empirical processes]
- Many applications through the continuous mapping theorem and the functional delta method

Motivation

Uniform convergence of bounded functions

Strong implications vs. **Restricted applicability**

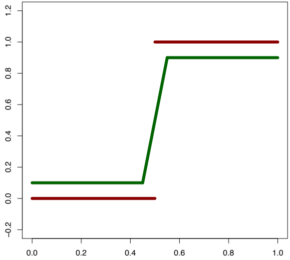
- Implies pointwise, continuous, L^p-convergence ...
- Well-developed weak convergence theory

Great success story in mathematical statistics

[Van der Vaart and Wellner (1996): Weak convergence and empirical processes]

- Many applications through the continuous mapping theorem and the functional delta method

- Continuous functions cannot converge to jump functions



Motivation

Uniform convergence of bounded functions

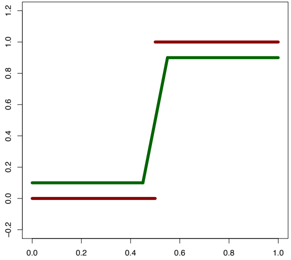
Strong implications vs. **Restricted applicability**

- Implies pointwise, continuous, L^p-convergence ...
- Well-developed weak convergence theory

  Great success story in mathematical statistics

  [Van der Vaart and Wellner (1996): Weak convergence and empirical processes]
- Many applications through the continuous mapping theorem and the functional delta method

- Continuous functions cannot converge to jump functions



- **Questions:** Weaker metric? Weak convergence theory? Applications?

Empirical processes via epi- and hypographs

The empirical copula process

Weak convergence with respect to the uniform metric

Non-smooth copulas: when weak convergence fails

The hypi-semimetric and weak convergence

Applications

Empirical processes via epi- and hypographs

Copulas

- A d-variate copula C is a d-variate distribution function with uniform $(0, 1)$ margins.

Copulas

- A d-variate copula C is a d-variate distribution function with uniform $(0, 1)$ margins.
- Sklar's (1959) theorem: If F is a d-variate distribution function with margins $F_1, \ldots, F_d$, then there exists a copula C such that

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big)$$

Copulas

- A d-variate copula C is a d-variate distribution function with uniform $(0,1)$ margins.
- Sklar's (1959) theorem: If F is a d-variate distribution function with margins $F_1, \ldots, F_d$, then there exists a copula C such that

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big)$$

- Moreover, if the margins are continuous, then C is unique and is given by the distribution function of $(F_1(X_1), \ldots, F_d(X_d))$, with $(X_1, \ldots, X_d) \sim F$:

$$\begin{aligned} C(u_1, \ldots, u_d) &= P[F_1(X_1) \le u_1, \ldots, F_d(X_d) \le u_d] \\ &= P[X_1 \le F_1^{-}(u_1), \ldots, X_d \le F_d^{-}(u_d)] \\ &= F\big(F_1^{-}(u_1), \ldots, F_d^{-}(u_d)\big) \end{aligned}$$

with $F_j^{-}(u) = \inf\{x : F_j(x) \ge u\}$ the generalized inverse (quantile function)

Copulas

- A d-variate copula C is a d-variate distribution function with uniform $(0,1)$ margins.
- Sklar's (1959) theorem: If F is a d-variate distribution function with margins $F_1, \ldots, F_d$, then there exists a copula C such that

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big)$$

- Moreover, if the margins are continuous, then C is unique and is given by the distribution function of $(F_1(X_1), \ldots, F_d(X_d))$, with $(X_1, \ldots, X_d) \sim F$:

$$\begin{aligned} C(u_1, \ldots, u_d) &= P[F_1(X_1) \le u_1, \ldots, F_d(X_d) \le u_d] \\ &= P[X_1 \le F_1^-(u_1), \ldots, X_d \le F_d^-(u_d)] \\ &= F\big(F_1^-(u_1), \ldots, F_d^-(u_d)\big) \end{aligned}$$

  with $F_j^-(u) = \inf\{x : F_j(x) \ge u\}$ the generalized inverse (quantile function)

- Usage: Modelling dependence between components $X_1, \ldots, X_d$, irrespective of their marginal distributions

The empirical copula

- Situation: $(\mathbf{X}_i)_{i=1,\ldots,n}$ i.i.d. rvs, $\mathbf{X}_i \sim F = C(F_1, \ldots, F_d)$, continuous marginals F_j.

[hence $C(\mathbf{u}) = F\{F_1^-(u_1), \ldots, F_d^-(u_d)\}$ with the generalized inverse $F_j^-(u) = \inf\{x : F_j(x) \geq u\}$]

The empirical copula

- Situation: $(\mathbf{X}_i)_{i=1,\dots,n}$ i.i.d. rvs, $\mathbf{X}_i \sim F = C(F_1, \dots, F_d)$, continuous marginals F_j.

  [hence $C(\mathbf{u}) = F\{F_1^-(u_1), \dots, F_d^-(u_d)\}$ with the generalized inverse $F_j^-(u) = \inf\{x : F_j(x) \geq u\}$]

- Goal: Estimate C nonparametrically.

The empirical copula

- Situation: $(\mathbf{X}_i)_{i=1,\dots,n}$ i.i.d. rvs, $\mathbf{X}_i \sim F = C(F_1, \dots, F_d)$, continuous marginals F_j.

  [hence $C(\mathbf{u}) = F\{F_1^-(u_1), \dots, F_d^-(u_d)\}$ with the generalized inverse $F_j^-(u) = \inf\{x : F_j(x) \geq u\}$]

- Goal: Estimate C nonparametrically.

- Simple plug-in estimation: empirical cdfs

$$F_n(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(X_{i1} \leq x_1, \dots, X_{id} \leq x_d), \quad F_{nj}(x_j) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(X_{ij} \leq x_j).$$

  yield the empirical copula

$$C_n(\mathbf{u}) = F_n\{F_{n1}^-(u_1), \dots, F_{nd}^-(u_d)\} = n^{-1}\sum_{i=1}^{n} \mathbb{I}\{X_{i1} \leq F_{n1}^-(u_1), \dots, X_{id} \leq F_{nd}^-(u_d)\}$$

The empirical copula

- Situation: $(\mathbf{X}_i)_{i=1,\ldots,n}$ i.i.d. rvs, $\mathbf{X}_i \sim F = C(F_1, \ldots, F_d)$, continuous marginals F_j.

  [hence $C(\mathbf{u}) = F\{F_1^-(u_1), \ldots, F_d^-(u_d)\}$ with the generalized inverse $F_j^-(u) = \inf\{x : F_j(x) \geq u\}$]

- Goal: Estimate C nonparametrically.

- Simple plug-in estimation: empirical cdfs

$$F_n(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(X_{i1} \leq x_1, \ldots, X_{id} \leq x_d), \quad F_{nj}(x_j) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(X_{ij} \leq x_j).$$

  yield the empirical copula

$$C_n(\mathbf{u}) = F_n\{F_{n1}^-(u_1), \ldots, F_{nd}^-(u_d)\} = n^{-1}\sum_{i=1}^{n} \mathbb{I}\{X_{i1} \leq F_{n1}^-(u_1), \ldots, X_{id} \leq F_{nd}^-(u_d)\}$$

$$= n^{-1}\sum_{i=1}^{n} \mathbb{I}\big(\hat{U}_{i1} \leq u_1, \ldots, \hat{U}_{id} \leq u_d\big) + O(n^{-1})$$

  [where $\hat{U}_{ij} = \text{rank}(X_{ij})/n$ are 'pseudo-observations' of C (rescaled ranks)]

The empirical copula process

$\mathbf{u} \mapsto \mathbb{C}_n(\mathbf{u}) = \sqrt{n}\{C_n(\mathbf{u}) - C(\mathbf{u})\} \in \ell^\infty([0,1]^d)$ is called empirical copula process.

[$\ell^\infty([0,1]^d)$ the space of bounded functions on $[0,1]^d$.]

The empirical copula process

$$\mathbf{u} \mapsto \mathbb{C}_n(\mathbf{u}) = \sqrt{n}\{C_n(\mathbf{u}) - C(\mathbf{u})\} \in \ell^\infty([0,1]^d) \quad \text{is called empirical copula process.}$$

[$\ell^\infty([0,1]^d)$ the space of bounded functions on $[0,1]^d$.]

Many applications.

- Testing for structural assumptions. Example: symmetry [Genest, Nešlehová, Quessy (2012)]. Null hypothesis: $C(u,v) = C(v,u)$ for all u, v.

$$T_n = n \int \{C_n(u,v) - C_n(v,u)\}^2 \, du \, dv \stackrel{H_0}{=} \int \{\mathbb{C}_n(u,v) - \mathbb{C}_n(v,u)\}^2 \, du \, dv$$

The empirical copula process

$$\mathbf{u} \mapsto \mathbb{C}_n(\mathbf{u}) = \sqrt{n}\{C_n(\mathbf{u}) - C(\mathbf{u})\} \in \ell^\infty([0,1]^d) \quad \text{is called empirical copula process.}$$

[$\ell^\infty([0,1]^d)$ the space of bounded functions on $[0,1]^d$.]

Many applications.

- Testing for structural assumptions. Example: symmetry [Genest, Nešlehová, Quessy (2012)]. Null hypothesis: $C(u,v) = C(v,u)$ for all u, v.

$$T_n = n \int \{C_n(u,v) - C_n(v,u)\}^2 \, du \, dv \overset{H_0}{=} \int \{\mathbb{C}_n(u,v) - \mathbb{C}_n(v,u)\}^2 \, du \, dv$$

- Minimum-distance estimators of parametric copulas [Tsukahara (2005)]. $\{C_\theta \mid \theta \in \Theta\}$ class of parametric candidate models. Estimator:

$$\hat{\theta} := \operatorname{argmin}_\theta \int \{C_\theta(u,v) - C_n(u,v)\}^2 \, du \, dv.$$

The empirical copula process

$$\mathbf{u} \mapsto \mathbb{C}_n(\mathbf{u}) = \sqrt{n}\{C_n(\mathbf{u}) - C(\mathbf{u})\} \in \ell^\infty([0,1]^d) \quad \text{is called empirical copula process.}$$

[$\ell^\infty([0,1]^d)$ the space of bounded functions on $[0,1]^d$.]

Many applications.

- Testing for structural assumptions. Example: symmetry [Genest, Nešlehová, Quessy (2012)]. Null hypothesis: $C(u,v) = C(v,u)$ for all u, v.

$$T_n = n \int \{C_n(u,v) - C_n(v,u)\}^2 \, du \, dv \overset{H_0}{=} \int \{\mathbb{C}_n(u,v) - \mathbb{C}_n(v,u)\}^2 \, du \, dv$$

- Minimum-distance estimators of parametric copulas [Tsukahara (2005)]. $\{C_\theta \mid \theta \in \Theta\}$ class of parametric candidate models. Estimator:

$$\hat{\theta} := \operatorname{argmin}_\theta \int \{C_\theta(u,v) - C_n(u,v)\}^2 \, du \, dv.$$

- Goodness-of fit tests, Asymptotics of estimators for Pickands dep. fct. ...

The empirical copula process

$$\mathbf{u} \mapsto \mathbb{C}_n(\mathbf{u}) = \sqrt{n}\{C_n(\mathbf{u}) - C(\mathbf{u})\} \in \ell^\infty([0,1]^d) \quad \text{is called empirical copula process.}$$

[$\ell^\infty([0,1]^d)$ the space of bounded functions on $[0,1]^d$.]

Many applications.

- Testing for structural assumptions. Example: symmetry [Genest, Nešlehová, Quessy (2012)]. Null hypothesis: $C(u,v) = C(v,u)$ for all u, v.

$$T_n = n \int \{C_n(u,v) - C_n(v,u)\}^2 \, du \, dv \overset{H_0}{=} \int \{\mathbb{C}_n(u,v) - \mathbb{C}_n(v,u)\}^2 \, du \, dv$$

- Minimum-distance estimators of parametric copulas [Tsukahara (2005)]. $\{C_\theta \mid \theta \in \Theta\}$ class of parametric candidate models. Estimator:

$$\hat{\theta} := \operatorname{argmin}_\theta \int \{C_\theta(u,v) - C_n(u,v)\}^2 \, du \, dv.$$

- Goodness-of fit tests, Asymptotics of estimators for Pickands dep. fct. ...

Derivation of asymptotic distributions: Process convergence of $\mathbb{C}_n$

The empirical copula process

Weak convergence with respect to the uniform metric

Non-smooth copulas: when weak convergence fails

The hypi-semimetric and weak convergence

Applications

Key quantities

Vector of quantile functions:

$$\mathbf{F}^{-}(\mathbf{u}) = \big(F_1^{-}(u_1), \dots, F_d^{-}(u_d)\big)$$
$$\mathbf{F}_n^{-}(\mathbf{u}) = \big(F_{n,1}^{-}(u_1), \dots, F_{n,d}^{-}(u_d)\big)$$

Copula and empirical copula:

$$C(\mathbf{u}) = F\big(\mathbf{F}^{-}(\mathbf{u})\big)$$
$$C_n(\mathbf{u}) = F_n\big(\mathbf{F}_n^{-}(\mathbf{u})\big)$$

Empirical process:

$$\alpha_n(\mathbf{x}) = \sqrt{n}\{F_n(\mathbf{x}) - F(\mathbf{x})\}$$

Standard empirical process theory

- Since the empirical copula is rank-based, we can without loss of generality assume that margins are uniform, hence $F = C$.
- Classical empirical process theory yields

$$\begin{aligned}\alpha_n(\mathbf{u}) &= \sqrt{n}\{F_n(\mathbf{u}) - C(\mathbf{u})\} \\ &\rightsquigarrow \mathbb{B}_C(\mathbf{u}) \qquad \text{in } \left(\ell^\infty([0,1]^d), \|\cdot\|_\infty\right)\end{aligned}$$

  a C-Brownian bridge.
- The Bahadur–Kiefer theorem links the empirical quantile and distribution functions:

$$\begin{aligned}\sqrt{n}\{F_{n,j}^{-}(u_j) - u_j\} &= -\sqrt{n}\{F_{n,j}(u_j) - u_j\} + o_{\mathbb{P}}(1) \\ &\rightsquigarrow -\mathbb{B}_{C,j}(u_j) = \mathbb{B}_C(1,\ldots,1,u_j,1,\ldots,1)\end{aligned}$$

Decomposition of the empirical copula process

Fundamental decomposition:

$$\begin{aligned}\mathbb{C}_n(\mathbf{u}) &= \sqrt{n}\big(C_n(\mathbf{u}) - C(\mathbf{u})\big) \\ &= \sqrt{n}\big\{F_n\big(\mathbf{F}_n^{-}(\mathbf{u})\big) - F\big(\mathbf{F}^{-}(\mathbf{u})\big)\big\} \\ &= \sqrt{n}\big\{F_n\big(\mathbf{F}_n^{-}(\mathbf{u})\big) - F\big(\mathbf{F}_n^{-}(\mathbf{u})\big)\big\} + \sqrt{n}\big\{F\big(\mathbf{F}_n^{-}(\mathbf{u})\big) - F\big(\mathbf{F}^{-}(\mathbf{u})\big)\big\}\end{aligned}$$

Recall $F = C$ (uniform margins). We find

$$\mathbb{C}_n(\mathbf{u}) = \alpha_n\big(\mathbf{F}_n^{-}(\mathbf{u})\big) + \sqrt{n}\big\{C\big(\mathbf{F}_n^{-}(\mathbf{u})\big) - C(\mathbf{u})\big\}$$

Treat each of the two terms separately:

$$\alpha_n\big(\mathbf{F}_n^{-}(\mathbf{u})\big) = \alpha_n(\mathbf{u}) + o_{\mathbb{P}}(1)$$

$$\sqrt{n}\big\{C\big(\mathbf{F}_n^{-}(\mathbf{u})\big) - C(\mathbf{u})\big\} = \sum_{j=1}^{d} \dot{C}_j(\mathbf{u})\,\sqrt{n}\{F_{n,j}^{-}(u_j) - u_j\} + o_{\mathbb{P}}(1)$$

Weak convergence of the empirical copula process in the topology of uniform convergence

Theorem [Weak uniform convergence of $\mathbb{C}_n$]
Suppose that

(S_1) $\dot{C}_j = \frac{\partial}{\partial u_j} C$ exists and is continuous for $\mathbf{u} \in [0,1]^d$ with $u_j \in (0,1)$.

Then, in $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$,

$$\sqrt{n}(C_n - C)(\mathbf{u}) \rightsquigarrow \mathbb{C}_C(\mathbf{u}) := \mathbb{B}_C(\mathbf{u}) - \sum_{j=1}^{d} \dot{C}_j(\mathbf{u}) \mathbb{B}_{C,j}(u_j)$$

where $\mathbb{B}_C$ is a C-brownian bridge and $\mathbb{B}_{C,j}(u_j) = \mathbb{B}_C(1, ..., 1, u_j, 1..., 1)$.

Weak convergence of the empirical copula process in the topology of uniform convergence

Theorem [Weak uniform convergence of $\mathbb{C}_n$]
Suppose that

(S_1) $\dot{C}_j = \frac{\partial}{\partial u_j} C$ exists and is continuous for $\mathbf{u} \in [0,1]^d$ with $u_j \in (0,1)$.

Then, in $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$,

$$\sqrt{n}(C_n - C)(\mathbf{u}) \rightsquigarrow \mathbb{C}_C(\mathbf{u}) := \mathbb{B}_C(\mathbf{u}) - \sum_{j=1}^{d} \dot{C}_j(\mathbf{u})\mathbb{B}_{C,j}(u_j)$$

where $\mathbb{B}_C$ is a C-brownian bridge and $\mathbb{B}_{C,j}(u_j) = \mathbb{B}_C(1, ..., 1, u_j, 1..., 1)$.

Discussion

- Dating back to Rüschendorf (1976), Gaenssler and Stute (1987)
- Assumption (S_1) due to S. (2012)
- Possible relaxation: stationary and short range dependent instead of i.i.d.

'Non-smooth' copulas: examples

- (Un)fortunately: The assumption

$$(\mathrm{S}_1)\ \dot{C}_j \text{ exists and is continuous for } \mathbf{u} \in [0,1]^d \text{ with } u_j \in (0,1)$$

is satisfied by many, but not by all interesting copulas.

'Non-smooth' copulas: examples

- (Un)fortunately: The assumption

$$(\mathrm{S}_1)\ \dot{C}_j \text{ exists and is continuous for } \mathbf{u} \in [0,1]^d \text{ with } u_j \in (0,1)$$

is satisfied by many, but not by all interesting copulas.

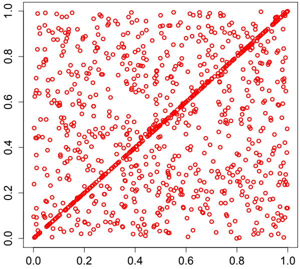
- Example:

$$C(\mathbf{u}) := \lambda u_1 u_2 + (1-\lambda) min(u_1, u_2)$$

Here

$$\dot{C}_1(\mathbf{u}) = \lambda u_2 + (1-\lambda) 1_{\{u_1 < u_2\}},$$
$$\dot{C}_2(\mathbf{u}) = \lambda u_1 + (1-\lambda) 1_{\{u_1 > u_2\}},$$

for $u_1 \neq u_2$ and the partial derivatives do not exist for $u_1 = u_2$.



'Non-smooth' copulas: examples

- (Un)fortunately: The assumption

$$(\mathrm{S}_1)\ \dot{C}_j \text{ exists and is continuous for } \mathbf{u} \in [0,1]^d \text{ with } u_j \in (0,1)$$

  is satisfied by many, but not by all interesting copulas.

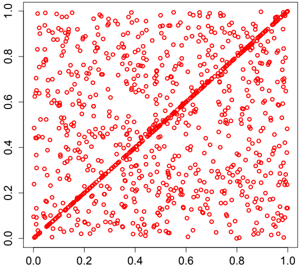
- Example:

$$C(\mathbf{u}) := \lambda u_1 u_2 + (1-\lambda) min(u_1, u_2)$$

  Here

$$\dot{C}_1(\mathbf{u}) = \lambda u_2 + (1-\lambda) 1_{\{u_1 < u_2\}},$$
$$\dot{C}_2(\mathbf{u}) = \lambda u_1 + (1-\lambda) 1_{\{u_1 > u_2\}},$$

  for $u_1 \neq u_2$ and the partial derivatives do not exist for $u_1 = u_2$.



- Other examples
  - Extreme-value copulas with non-differentiable Pickands dependence function
  - Marshall-Olkin copulas
  - Archimedean copulas with non-smooth generators
  - ...

Non-smooth copulas: pointwise vs. functional weak convergence

- Pointwise limit for the previous example:

$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}_C^*(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \dot{C}_1(\mathbf{u})\,\mathbb{B}_C(u_1, 1) - \dot{C}_2(\mathbf{u})\,\mathbb{B}_C(1, u_2),$$

apart from the diagonal and

$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}_C^*(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \lambda u\{\mathbb{B}_C(u,1) + \mathbb{B}_C(1,u)\} - (1-\lambda)\max\{\mathbb{B}_C(u,1), \mathbb{B}_C(1,u)\}$$

on the diagonal.

Non-smooth copulas: pointwise vs. functional weak convergence

- Pointwise limit for the previous example:

$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}_C^*(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \dot{C}_1(\mathbf{u})\,\mathbb{B}_C(u_1, 1) - \dot{C}_2(\mathbf{u})\,\mathbb{B}_C(1, u_2),$$

  apart from the diagonal and

$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}_C^*(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \lambda u\{\mathbb{B}_C(u, 1) + \mathbb{B}_C(1, u)\} - (1 - \lambda)\max\{\mathbb{B}_C(u, 1), \mathbb{B}_C(1, u)\}$$

  on the diagonal.

- Question: Can we have: $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C^*$ in $(\ell^\infty([0, 1]^2), \|\cdot\|_\infty)$?

Non-smooth copulas: pointwise vs. functional weak convergence

- Pointwise limit for the previous example:

$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}_C^*(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \dot{C}_1(\mathbf{u})\,\mathbb{B}_C(u_1,1) - \dot{C}_2(\mathbf{u})\,\mathbb{B}_C(1,u_2),$$

  apart from the diagonal and

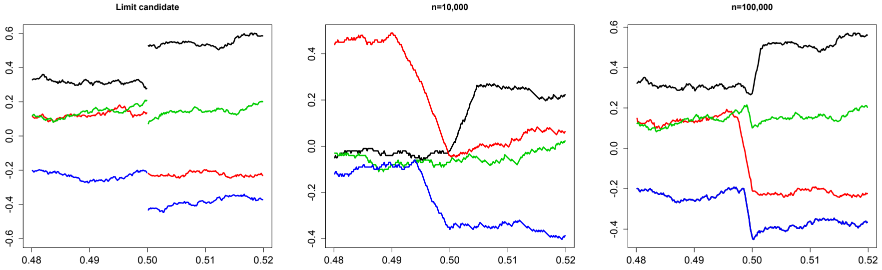
$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}_C^*(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \lambda u\{\mathbb{B}_C(u,1) + \mathbb{B}_C(1,u)\} \\ - (1-\lambda)\max\{\mathbb{B}_C(u,1), \mathbb{B}_C(1,u)\}$$

  on the diagonal.

- Question: Can we have: $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C^*$ in $(\ell^\infty([0,1]^2), \|\cdot\|_\infty)$?
- Answer: Lemma [Bücher, Segers, Volgushev, 2013]:
  If $\mathbb{C}_n$ converges weakly with respect to $\|\cdot\|_\infty$, then the limit must have continuous trajectories, a.s.

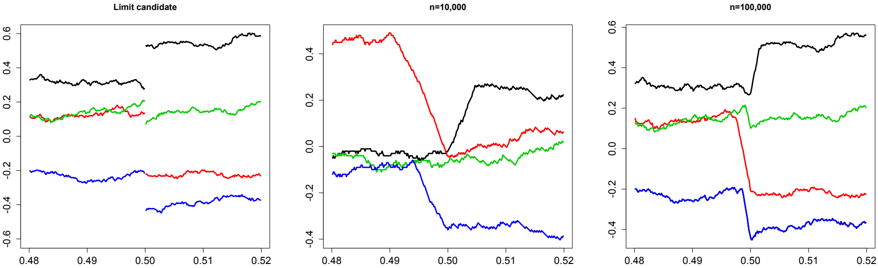
  This is not the case here!

Lack of uniform convergence.



- Left: sample paths of candidate limit process (based on $n = 100,000$) on $[-0.48, 0.52] \times \{0.5\}$.
- Middle and right: 'typical realizations' of the empirical copula process, $n = 10,000$ and $n = 100,000$.

Lack of uniform convergence.



- Left: sample paths of candidate limit process (based on $n = 100,000$) on $[-0.48, 0.52] \times \{0.5\}$.
- Middle and right: 'typical realizations' of the empirical copula process, $n = 10,000$ and $n = 100,000$.

Suggestion: Weak convergence may hold with respect to a metric, for which jump functions can be 'close' to continuous functions. $\rightsquigarrow$ Generalize Skorohod's M_2 metric.

Empirical processes via epi- and hypographs

Painlevé–Kuratowski convergence

Sequence of sets A_n in a metric space $(\mathbb{T}, d)$.

$$\liminf_{n\to\infty} A_n = \{x \in \mathbb{T} \mid \exists x_n \in A_n : x_n \to x\}$$
$$\limsup_{n\to\infty} A_n = \{x \in \mathbb{T} \mid \exists x_{n_k} \in A_{n_k} : x_{n_k} \to x\}$$

Painlevé–Kuratowski convergence

Sequence of sets A_n in a metric space $(\mathbb{T}, d)$.

$$\liminf_{n\to\infty} A_n = \{x \in \mathbb{T} \mid \exists x_n \in A_n : x_n \to x\}$$
$$\limsup_{n\to\infty} A_n = \{x \in \mathbb{T} \mid \exists x_{n_k} \in A_{n_k} : x_{n_k} \to x\}$$

Painlevé–Kuratowski convergence: $A_n \to A$ if

$$A = \liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n$$

Painlevé–Kuratowski convergence

Sequence of sets A_n in a metric space $(\mathbb{T}, d)$.

$$\liminf_{n\to\infty} A_n = \{x \in \mathbb{T} \mid \exists x_n \in A_n : x_n \to x\}$$
$$\limsup_{n\to\infty} A_n = \{x \in \mathbb{T} \mid \exists x_{n_k} \in A_{n_k} : x_{n_k} \to x\}$$

Painlevé–Kuratowski convergence: $A_n \to A$ if

$$A = \liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n$$

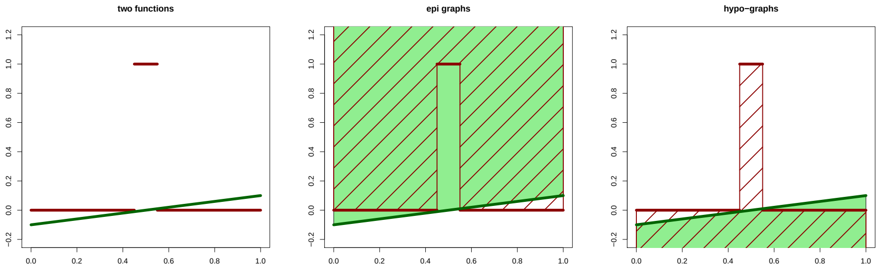
Properties:

- Necessarily, A is closed.
- $A_n \to A$ iff $\mathrm{cl}(A_n) \to A$.
- Metrizable if $(\mathbb{T}, d)$ is locally compact and separable: Fell topology
- If $(\mathbb{T}, d)$ is compact, then PK convergence is convergence in the Hausdorff metric.

Introducing hypi-convergence

- Epi- and hypograph of a function $f \in \ell^\infty([0,1]^d)$:

$$\text{epi}\, f := \{(\mathbf{u}, t) \in [0,1]^d \times \mathbb{R} \mid f(\mathbf{u}) \leq t\}$$
$$\text{hypo}\, f := \{(\mathbf{u}, t) \in [0,1]^d \times \mathbb{R} \mid f(\mathbf{u}) \geq t\}$$



Introducing hypi-convergence

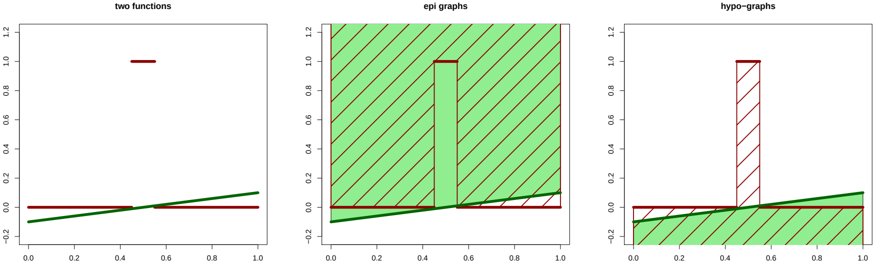
- Epi- and hypograph of a function $f \in \ell^\infty([0,1]^d)$:

$$\operatorname{epi} f := \{(\mathbf{u}, t) \in [0,1]^d \times \mathbb{R} \mid f(\mathbf{u}) \le t\}$$
$$\operatorname{hypo} f := \{(\mathbf{u}, t) \in [0,1]^d \times \mathbb{R} \mid f(\mathbf{u}) \ge t\}$$

- The hypi-semimetric is defined as

$$d_{\text{hypi}}(f, g) = \max\{d_{\mathcal{F}}(\operatorname{cl}(\operatorname{epi} f), \operatorname{cl}(\operatorname{epi} g)), d_{\mathcal{F}}(\operatorname{cl}(\operatorname{hypo} f), \operatorname{cl}(\operatorname{hypo} g))\}.$$

where $d_{\mathcal{F}}$ is a metric on closed sets inducing Painlevé–Kuratowski convergence.



Introducing hypi-convergence

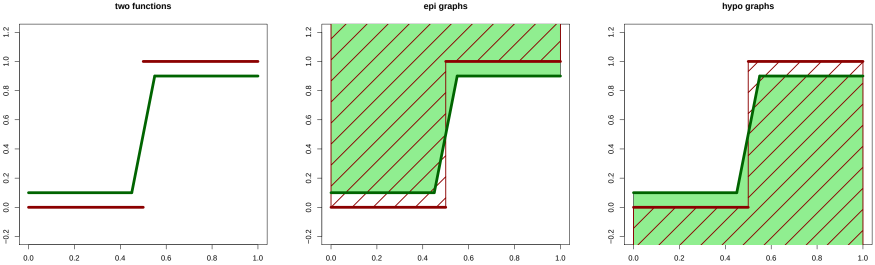
- Epi- and hypograph of a function $f \in \ell^\infty([0,1]^d)$:

$$\operatorname{epi} f := \{(\mathbf{u}, t) \in [0,1]^d \times \mathbb{R} \mid f(\mathbf{u}) \leq t\}$$
$$\operatorname{hypo} f := \{(\mathbf{u}, t) \in [0,1]^d \times \mathbb{R} \mid f(\mathbf{u}) \geq t\}$$

- The hypi-semimetric is defined as

$$d_{\text{hypi}}(f, g) = \max\{d_{\mathcal{F}}(\operatorname{cl}(\operatorname{epi} f), \operatorname{cl}(\operatorname{epi} g)), d_{\mathcal{F}}(\operatorname{cl}(\operatorname{hypo} f), \operatorname{cl}(\operatorname{hypo} g))\}.$$

where $d_{\mathcal{F}}$ is a metric on closed sets inducing Painlevé–Kuratowski convergence.



Point-wise criteria for hypi-convergence

Define lower and upper semicontinuous hulls of f:

$$f_\wedge(x) = \sup_{\varepsilon>0} \inf\{f(x') : \|x' - x\| < \varepsilon\}$$

$$f_\vee(x) = \inf_{\varepsilon>0} \sup\{f(x') : \|x' - x\| < \varepsilon\}$$

Point-wise criteria for hypi-convergence

Define lower and upper semicontinuous hulls of f:

$$f_\wedge(x) = \sup_{\varepsilon>0} \inf\{f(x') : \|x' - x\| < \varepsilon\}$$
$$f_\vee(x) = \inf_{\varepsilon>0} \sup\{f(x') : \|x' - x\| < \varepsilon\}$$

Then $d_{\text{hypi}}(f_n, f) \to 0$ iff the following two conditions hold:

1. $f_\wedge$ and $f_\vee$ provide asymptotic bounds for f_n:

$$\forall x \in [0,1]^d : \forall x_n \to x : f_\wedge(x) \leq \liminf_{n\to\infty} f_n(x_n) \leq \limsup_{n\to\infty} f_n(x_n) \leq f_\vee(x)$$

Point-wise criteria for hypi-convergence

Define lower and upper semicontinuous hulls of f:

$$f_\wedge(x) = \sup_{\varepsilon>0} \inf\{f(x') : \|x' - x\| < \varepsilon\}$$
$$f_\vee(x) = \inf_{\varepsilon>0} \sup\{f(x') : \|x' - x\| < \varepsilon\}$$

Then $d_{\text{hypi}}(f_n, f) \to 0$ iff the following two conditions hold:

1. $f_\wedge$ and $f_\vee$ provide asymptotic bounds for f_n:

$$\forall x \in [0,1]^d : \forall x_n \to x : f_\wedge(x) \leq \liminf_{n\to\infty} f_n(x_n) \leq \limsup_{n\to\infty} f_n(x_n) \leq f_\vee(x)$$

2. $f_\wedge$ and $f_\vee$ are asymptotically attainable by f_n:

$$\forall x \in [0,1]^d : \begin{cases} \exists x_n \to x : \liminf_{n\to\infty} f_n(x_n) = f_\wedge(x), \\ \exists x_n \to x : \limsup_{n\to\infty} f_n(x_n) = f_\vee(x) \end{cases}$$

Hypi-convergence: Useful at all?

Theorem [Handy implications of hypi-convergence]

Let $f_n, f \in \ell^\infty([0,1]^d)$ and $d_{\text{hypi}}(f_n, f) \to 0$.

- Let μ be a finite measure on $[0,1]^d$. If $\mu(\text{discontinuity points of } f) = 0$, then $\|f_n - f\|_{L^p(\mu)} \to 0$ for any $p \in [1, \infty)$.

Hypi-convergence: Useful at all?

Theorem [Handy implications of hypi-convergence]

Let $f_n, f \in \ell^\infty([0,1]^d)$ and $d_{\text{hypi}}(f_n, f) \to 0$.

- Let μ be a finite measure on $[0,1]^d$. If $\mu(\text{discontinuity points of } f) = 0$, then $\|f_n - f\|_{L^p(\mu)} \to 0$ for any $p \in [1, \infty)$.
- $\sup f_n \to \sup f$ and $\inf f_n \to \inf f$

Hypi-convergence: Useful at all?

Theorem [Handy implications of hypi-convergence]

Let $f_n, f \in \ell^\infty([0,1]^d)$ and $d_{\text{hypi}}(f_n, f) \to 0$.

- Let μ be a finite measure on $[0,1]^d$. If $\mu(\text{discontinuity points of } f) = 0$, then $\|f_n - f\|_{L^p(\mu)} \to 0$ for any $p \in [1, \infty)$.
- $\sup f_n \to \sup f$ and $\inf f_n \to \inf f$
- If f is continuous in x, then $f_n(x_n) \to f(x)$ whenever $x_n \to x$.
  Also uniformly over compact sets.

Hypi-convergence: Useful at all?

Theorem [Handy implications of hypi-convergence]

Let $f_n, f \in \ell^\infty([0,1]^d)$ and $d_{\text{hypi}}(f_n, f) \to 0$.

- Let μ be a finite measure on $[0,1]^d$. If $\mu(\text{discontinuity points of } f) = 0$, then $\|f_n - f\|_{L^p(\mu)} \to 0$ for any $p \in [1, \infty)$.
- $\sup f_n \to \sup f$ and $\inf f_n \to \inf f$
- If f is continuous in x, then $f_n(x_n) \to f(x)$ whenever $x_n \to x$. Also uniformly over compact sets.

Interpretation:
d_{hypi} is 'between' $\|\cdot\|_\infty$ and $\|\cdot\|_p$ with $p < \infty$.
It adapts to regularity of the limit function.

Comments on hypi-convergence

- hypi = epi + hypo:

$$d_{\text{hypi}}(f_n, f) \iff \begin{cases} f_n \text{ epi-converges to } f_\wedge, & \text{i.e., epi } f_n \to \text{epi } f_\wedge \\ f_n \text{ hypo-converges to } f_\vee, & \text{i.e., hypo } f_n \to \text{hypo } f_\vee \end{cases}$$

  Epi- and hypoconvergence have a long history in the analysis of minimizers and maximizers of functions (Rockafeller & Wets 1998, Molchanov 2005)

Comments on hypi-convergence

- hypi = epi + hypo:

$$d_{\text{hypi}}(f_n, f) \iff \begin{cases} f_n \text{ epi-converges to } f_\wedge, & \text{i.e., epi } f_n \to \text{epi } f_\wedge \\ f_n \text{ hypo-converges to } f_\vee, & \text{i.e., hypo } f_n \to \text{hypo } f_\vee \end{cases}$$

  Epi- and hypoconvergence have a long history in the analysis of minimizers and maximizers of functions (Rockafeller & Wets 1998, Molchanov 2005)

- Only defines a semi-metric:

$$d_{\text{hypi}}(f, g) = 0 \iff \begin{cases} f_\wedge = g_\wedge \\ f_\vee = g_\vee \end{cases}$$

  Care must be taken when considering weak convergence.

Comments on hypi-convergence

- hypi = epi + hypo:

$$d_{\text{hypi}}(f_n, f) \iff \begin{cases} f_n \text{ epi-converges to } f_\wedge, & \text{i.e., epi } f_n \to \text{epi } f_\wedge \\ f_n \text{ hypo-converges to } f_\vee, & \text{i.e., hypo } f_n \to \text{hypo } f_\vee \end{cases}$$

  Epi- and hypoconvergence have a long history in the analysis of minimizers and maximizers of functions (Rockafeller & Wets 1998, Molchanov 2005)

- Only defines a semi-metric:

$$d_{\text{hypi}}(f, g) = 0 \iff \begin{cases} f_\wedge = g_\wedge \\ f_\vee = g_\vee \end{cases}$$

  Care must be taken when considering weak convergence.

- Addition is not continuous! Extra work needed to deal with convergence of sums.

Comments on hypi-convergence

- hypi = epi + hypo:

$$d_{\text{hypi}}(f_n, f) \iff \begin{cases} f_n \text{ epi-converges to } f_\wedge, & \text{i.e., epi } f_n \to \text{epi } f_\wedge \\ f_n \text{ hypo-converges to } f_\vee, & \text{i.e., hypo } f_n \to \text{hypo } f_\vee \end{cases}$$

  Epi- and hypoconvergence have a long history in the analysis of minimizers and maximizers of functions (Rockafeller & Wets 1998, Molchanov 2005)

- Only defines a semi-metric:

$$d_{\text{hypi}}(f, g) = 0 \iff \begin{cases} f_\wedge = g_\wedge \\ f_\vee = g_\vee \end{cases}$$

  Care must be taken when considering weak convergence.

- Addition is not continuous! Extra work needed to deal with convergence of sums.

- Can be generalized to functions on locally compact, separable metric spaces.

Weak hypi-convergence of the empirical copula process

Theorem [Bücher, Segers, Volgushev, 2013] Let

$\mathcal{D}(C) := \{\mathbf{u} \in [0,1]^d \mid \dot{C}_j(\mathbf{u})$ does not exist or is not continuous for some $1 \le j \le d\}$

and suppose that (S_2) $\mathcal{D}(C)$ is a Lebesgue-null set. Then,

$$[\mathbb{C}_n]_{d_{\text{hypi}}} = [\sqrt{n}(C_n - C)]_{d_{\text{hypi}}} \rightsquigarrow [\mathbb{C}_C]_{d_{\text{hypi}}}$$

in $(L^\infty([0,1]^d), d_{\text{hypi}})$, where

$$\mathbb{C}_C(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) + dC_{(-\mathbb{B}_{C,1},\ldots,-\mathbb{B}_{C,d})}(\mathbf{u})$$

and where, for $a = (a_1, \ldots, a_d)$ with $a_j : [0,1] \to \mathbb{R}$ continuous,

$$dC_a(\mathbf{u}) = \lim_{\varepsilon \to 0} \inf \left\{ \sum_{j=1}^{d} \dot{C}_j(\mathbf{v})\, a_j(v_j) : \mathbf{v} \in [0,1]^d \setminus \mathcal{D}(C),\ |\mathbf{v} - \mathbf{u}| < \varepsilon \right\}.$$

- Recall (S_1): $\dot{C}_j(\mathbf{u})$ exists and is continuous for $\mathbf{u}$ with $u_j \in (0,1)$.
- Recall $\mathbb{C}_C(\mathbf{u}) := \mathbb{B}_C(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u})\mathbb{B}_{C,j}(u_j)$.

Consequences of hypi-convergence of the empirical copula process

Consequences for $\mathbb{C}_n$ through the continuous mapping theorem:

- Hypi-convergence implies uniform convergence if the limit is continuous
  $\Rightarrow$ Retrieve usual weak convergence result under (S_1)

Consequences of hypi-convergence of the empirical copula process

Consequences for $\mathbb{C}_n$ through the continuous mapping theorem:

- Hypi-convergence implies uniform convergence if the limit is continuous
  $\Rightarrow$ Retrieve usual weak convergence result under (S_1)
- Hypi-convergence implies L^p convergence for $p < \infty$
  $\Rightarrow$ Weak convergence with respect to L^p
  $\Rightarrow$ Cramér–von Mises type statistics

Consequences of hypi-convergence of the empirical copula process

Consequences for $\mathbb{C}_n$ through the continuous mapping theorem:

- Hypi-convergence implies uniform convergence if the limit is continuous
  $\Rightarrow$ Retrieve usual weak convergence result under (S_1)

- Hypi-convergence implies L^p convergence for $p < \infty$
  $\Rightarrow$ Weak convergence with respect to L^p
  $\Rightarrow$ Cramér–von Mises type statistics

- Hypi-convergence implies convergence of infima and suprema
  $\Rightarrow$ Weak convergence of and Kolmogorov–Smirnov statistics.

Empirical processes via epi- and hypographs

The empirical copula process

Weak convergence with respect to the uniform metric

Non-smooth copulas: when weak convergence fails

The hypi-semimetric and weak convergence

Applications

Comparing test statistics via local power curves

- Test for
$$H_0 : C = C_0, \quad \text{where } C_0 \text{ is a given copula (e.g, } C_0 = \Pi).$$

Comparing test statistics via local power curves

- Test for

$$H_0 : C = C_0, \quad \text{where } C_0 \text{ is a given copula (e.g, } C_0 = \Pi).$$

- Two competing test statistics

$$S_n = n \int \{C_n - C_0\}^2 \, d\Pi \qquad \text{Cramér–von Mises}$$

$$T_n = \sqrt{n} \| C_n - C_0 \|_\infty \qquad \text{Kolmogorov–Smirnov}$$

Comparing test statistics via local power curves

- Test for

$$H_0 : C = C_0, \quad \text{where } C_0 \text{ is a given copula (e.g, } C_0 = \Pi\text{).}$$

- Two competing test statistics

$$S_n = n \int \{C_n - C_0\}^2 \, d\Pi \qquad \text{Cramér–von Mises}$$

$$T_n = \sqrt{n} \| C_n - C_0 \|_\infty \qquad \text{Kolmogorov–Smirnov}$$

- Comparing the quality of tests: Local power curves
  How well does a test detect alternatives that converge to the null hypothesis?

Local power curves of simple goodness-of-fit tests

- Local alternatives in direction Λ:
  Let $(\mathbf{X}_i^{(n)})_{i=1,\ldots,n}$ be row-wise i.i.d. with copula $C^{(n)}$. Assume

$$\Delta_n = \sqrt{n}(C^{(n)} - C_0) \to \delta\Lambda$$

  uniformly, $\delta > 0, \Lambda \not\equiv 0$.

Local power curves of simple goodness-of-fit tests

- Local alternatives in direction Λ:
  Let $(\mathbf{X}_i^{(n)})_{i=1,\ldots,n}$ be row-wise i.i.d. with copula $C^{(n)}$. Assume

$$\Delta_n = \sqrt{n}(C^{(n)} - C_0) \to \delta\Lambda$$

  uniformly, $\delta > 0, \Lambda \not\equiv 0$.

- **Proposition.** If C_0 satisfies (S_2), then

$$\sqrt{n}(C_n - C_0) \rightsquigarrow \mathbb{C}_{C_0} + \delta\Lambda \qquad \text{in } (L^\infty([0,1]^d), d_{\text{hypi}}).$$

  Consequence: limit distribution of the test statistics under the local alternatives

$$S_n \to \mathcal{S}_\delta = \int \{\mathbb{C}_{C_0} + \delta\Lambda\}^2 \, d\Pi$$

$$T_n \to \mathcal{T}_\delta = \|\mathbb{C}_{C_0} + \delta\Lambda\|_\infty.$$

Local power curves of simple goodness-of-fit tests

- Local alternatives in direction Λ:
  Let $(\mathbf{X}_i^{(n)})_{i=1,\ldots,n}$ be row-wise i.i.d. with copula $C^{(n)}$. Assume

$$\Delta_n = \sqrt{n}(C^{(n)} - C_0) \to \delta\Lambda$$

  uniformly, $\delta > 0, \Lambda \not\equiv 0$.

- **Proposition.** If C_0 satisfies (S_2), then

$$\sqrt{n}(C_n - C_0) \rightsquigarrow \mathbb{C}_{C_0} + \delta\Lambda \qquad \text{in } (L^\infty([0,1]^d), d_{\text{hypi}}).$$

  Consequence: limit distribution of the test statistics under the local alternatives

$$S_n \to \mathcal{S}_\delta = \int \{\mathbb{C}_{C_0} + \delta\Lambda\}^2 \, d\Pi$$

$$T_n \to \mathcal{T}_\delta = \|\mathbb{C}_{C_0} + \delta\Lambda\|_\infty.$$

- Local power curves in direction Λ:
  '$\delta \mapsto$ asymptotic power(δ)' at significance level α

$$\delta \mapsto \Pr\{\mathcal{S}_\delta > q_{\mathcal{S}_0}(1-\alpha)\},$$

$$\delta \mapsto \Pr\{\mathcal{T}_\delta > q_{\mathcal{T}_0}(1-\alpha)\}$$

Minimum L^2-distance estimators à la Tsukahara

- Let $\{C_\theta \mid \theta \in \Theta \subset \mathbb{R}^p\}$ be a class of parametric candidate models. Estimator:

$$\hat{\theta} := \operatorname{argmin}_\theta \int (C_\theta - C_n)^2 \, d\mu$$

Minimum L^2-distance estimators à la Tsukahara

- Let $\{C_\theta \mid \theta \in \Theta \subset \mathbb{R}^p\}$ be a class of parametric candidate models. Estimator:

$$\hat{\theta} := \text{argmin}_\theta \int (C_\theta - C_n)^2 \, d\mu$$

Proposition (Asymptotic normality of $\hat{\theta}$):
Suppose that (S_2) holds and that $\mu(\mathcal{D}(C)) = 0$. Under usual regularity conditions on the model:

(i) for correctly specified models (θ_0 is the 'true' parameter):

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightsquigarrow \left\{ \int \nabla C_{\theta_0} \nabla C_{\theta_0}^T \, d\mu \right\}^{-1} \int \nabla C_{\theta_0} \mathbb{C}_C \, d\mu,$$

Minimum L^2-distance estimators à la Tsukahara

- Let $\{C_\theta \mid \theta \in \Theta \subset \mathbb{R}^p\}$ be a class of parametric candidate models. Estimator:

$$\hat{\theta} := \operatorname{argmin}_\theta \int (C_\theta - C_n)^2 \, d\mu$$

Proposition (Asymptotic normality of $\hat{\theta}$):
Suppose that (S_2) holds and that $\mu(\mathcal{D}(C)) = 0$. Under usual regularity conditions on the model:

(i) for correctly specified models (θ_0 is the 'true' parameter):

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightsquigarrow \left\{ \int \nabla C_{\theta_0} \nabla C_{\theta_0}^T \, d\mu \right\}^{-1} \int \nabla C_{\theta_0} \mathbb{C}_C \, d\mu,$$

(ii) for incorrectly specified models:

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightsquigarrow \left\{ \int \nabla C_{\theta_0} \nabla C_{\theta_0}^T + (C_{\theta_0} - C) J_{\theta_0} \, d\mu \right\}^{-1} \int \nabla C_{\theta_0} \mathbb{C}_C \, d\mu,$$

where $\theta_0 = \arg\min \int (C_\theta - C)^2 \, d\mu$.

Beyond copulas. . .

Helpful for different problems?

- The hypi-semimetric can be defined for real-valued, locally bounded functions on a compact, separable, metrizable domain

Beyond copulas. . .

Helpful for different problems?

- The hypi-semimetric can be defined for real-valued, locally bounded functions on a compact, separable, metrizable domain
- Might help whenever a (pointwise) candidate limit has discontinuities that are not exactly matched for finite n

Empirical processes of residuals (measurement error in the ordinates)

Conclusion

- Weak convergence w.r.t. topology of uniform convergence:
  great success story in mathematical statistics
- Occasionally, it fails: continuous functions cannot converge to functions with jumps
- Alternative: weak convergence with respect to a new topology:

$$\text{hypi} = \text{epi} \cap \text{hypo}$$

  - implies uniform convergence for continuous limits
  - implies convergence of infima and suprema
  - adapts to limit functions with jumps
  - stronger than L^p convergence
- Potentially useful for empirical processes based on estimated data
  Examples: empirical copula processes, empirical processes of regression residuals

Thank you!

A. Bücher, J. Segers & S. Volgushev (2013)

When uniform weak convergence fails:
Empirical processes for dependence functions via epi- and hypographs

Submitted for publication, arXiv:1305.6408