Convergence of empirical Kantorovitch contrasts and surfaces, for symmetric and asymmetric convex costs

Philippe Berthet
Institut Mathématique de Toulouse

Besançon - Workshop On Empirical Processes
13 mai 2019

Introduction

Plan

1. Motivations
2. Real Quantiles
3. q_α-surfaces (and depth fields)
4. Kantorovitch contrasts/distances K_c
5. A CLT for $K_c(F, G)$ if $F \neq G$
6. K_c-surfaces
7. Weak limits for $K_c(\mathbb{F}_n, \mathbb{G}_n)$ if $F = G$ and non symmetric cost

Motivations

Motivation 1 - Sensitivity analysis

Meta-model :

- $X^1, \ldots, X^d$ input r.v.'s (no restriction)
- $Y = \varphi(X^1, \ldots, X^d)$ output r.v. (real valued)

Motivation 1 - Sensitivity analysis

Meta-model :

- $X^1, \ldots, X^d$ input r.v.'s (no restriction)
- $Y = \varphi(X^1, \ldots, X^d)$ output r.v. (real valued)

Mean influence of X^j :

- indexes based on $\mathbb{E}(Y|X^j)$
- typically $\mathbb{V}(\mathbb{E}(Y|X^j))$ or α-quantiles $q_\alpha(\mathbb{E}(Y|X^j))$

Motivation 1 - Sensitivity analysis

Meta-model :

- $X^1, \ldots, X^d$ input r.v.'s (no restriction)
- $Y = \varphi(X^1, \ldots, X^d)$ output r.v. (real valued)

Mean influence of X^j :

- indexes based on $\mathbb{E}(Y|X^j)$
- typically $\mathbb{V}(\mathbb{E}(Y|X^j))$ or α-quantiles $q_\alpha(\mathbb{E}(Y|X^j))$

Influence of X^j on the distribution $\mathbb{P}^Y$:

- use instead : $\cdot \longrightarrow K_c(\mathbb{P}^{Y|X^j=\cdot}, \mathbb{P}^Y) = \int_0^1 c(q_\alpha(\mathbb{P}^{Y|X^j=\cdot}), q_\alpha(\mathbb{P}^Y))d\alpha$
- K_c contrast sensible to quantiles and tails, cost c
- deduce K_c-indexes $\mathbb{E}_{\mathrm{X^j}}(K_c(\mathbb{P}^{Y|X^j=\mathrm{X^j}}, \mathbb{P}^Y))$ or $\mathbb{V}_{\mathrm{X^j}}$ or quantiles

Meta-model :

- $X^1, \ldots, X^d$ input r.v.'s (no restriction)
- $Y = \varphi(X^1, \ldots, X^d)$ output r.v. (real valued)

Mean influence of X^j :

- indexes based on $\mathbb{E}(Y|X^j)$
- typically $\mathbb{V}(\mathbb{E}(Y|X^j))$ or α-quantiles $q_\alpha(\mathbb{E}(Y|X^j))$

Influence of X^j on the distribution $\mathbb{P}^Y$:

- use instead : $\cdot \longrightarrow K_c(\mathbb{P}^{Y|X^j=\cdot}, \mathbb{P}^Y) = \int_0^1 c(q_\alpha(\mathbb{P}^{Y|X^j=\cdot}), q_\alpha(\mathbb{P}^Y))d\alpha$
- K_c contrast sensible to quantiles and tails, cost c
- deduce K_c-indexes $\mathbb{E}_{\mathrm{X^j}}(K_c(\mathbb{P}^{Y|X^j=\mathrm{X^j}}, \mathbb{P}^Y))$ or $\mathbb{V}_{\mathrm{X^j}}$ or quantiles

jointe estimation of K_c-indices :

- by using : $\cdot \longrightarrow Y_i^j(\cdot) = \varphi(X_i^1, ..., X_i^{j-1}, \cdot, X_i^{j+1}, ..., X_i^d), i \leqslant n$
- then : $\cdot = X_k^j$ whence a correlation between $(Y_i^j(X_k^j), Y_i)$!

Motivation 2 - Comparison of processes

Functional data :

- X_i a random path, projection of X_i on a base
- ex: Karhunen-Loëve-Hotelling-Mercer-Haar-Fejer ...
- typical dimension reduction : d first coordinates
- i.i.d. cloud $\mathbb{P}_n$ of n points in $\mathbb{R}^d$

Motivation 2 - Comparison of processes

Functional data :

- X_i a random path, projection of X_i on a base
- ex: Karhunen-Loëve-Hotelling-Mercer-Haar-Fejer ...
- typical dimension reduction : d first coordinates
- i.i.d. cloud $\mathbb{P}_n$ of n points in $\mathbb{R}^d$

Comparison :

- to other sampled coordinates $\mathbb{Q}_n$ from a reference process
- to some exact distribution P on $\mathbb{R}^d$ (goodness-of-fit)

Motivation 2 - Comparison of processes

Functional data :

- X_i a random path, projection of X_i on a base
- ex: Karhunen-Loëve-Hotelling-Mercer-Haar-Fejer ...
- typical dimension reduction : d first coordinates
- i.i.d. cloud $\mathbb{P}_n$ of n points in $\mathbb{R}^d$

Comparison :

- to other sampled coordinates $\mathbb{Q}_n$ from a reference process
- to some exact distribution P on $\mathbb{R}^d$ (goodness-of-fit)

Directional quantiles and integrated contrasts :

- projection p_u direction $u \in S_{d-1}$ (uniform or random)
- q_α-surface : $u \longrightarrow q_{\alpha,u}(\mathbb{P}_n) = q_\alpha(\mathbb{P}_n \circ p_u^{-1}) \in \mathbb{R}$
- K_c-surface : $u \longrightarrow K_{c,u}(\mathbb{P}_n, \mathbb{Q}_n) = K_c(\mathbb{P}_n \circ p_u^{-1}, \mathbb{Q}_n \circ p_u^{-1}) \leqslant 0$

Motivation 2 - Comparison of processes

Functional data :

- X_i a random path, projection of X_i on a base
- ex: Karhunen-Loëve-Hotelling-Mercer-Haar-Fejer ...
- typical dimension reduction : d first coordinates
- i.i.d. cloud $\mathbb{P}_n$ of n points in $\mathbb{R}^d$

Comparison :

- to other sampled coordinates $\mathbb{Q}_n$ from a reference process
- to some exact distribution P on $\mathbb{R}^d$ (goodness-of-fit)

Directional quantiles and integrated contrasts :

- projection p_u direction $u \in S_{d-1}$ (uniform or random)
- q_α-surface : $u \longrightarrow q_{\alpha,u}(\mathbb{P}_n) = q_\alpha(\mathbb{P}_n \circ p_u^{-1}) \in \mathbb{R}$
- K_c-surface : $u \longrightarrow K_{c,u}(\mathbb{P}_n, \mathbb{Q}_n) = K_c(\mathbb{P}_n \circ p_u^{-1}, \mathbb{Q}_n \circ p_u^{-1}) \leqslant 0$

$$\int_0^1 \int_{S_{d-1}} c(q_{\alpha,u}(\mathbb{P}_n), q_{\alpha,u}(\mathbb{Q}_n))\,du\,d\alpha = \int_{S_{d-1}} K_{c,u}(\mathbb{P}_n, \mathbb{Q}_n)\,du$$

Motivation 3 - Multiple tests, multicriteria analysis

Goodness-of-fit, direct :

- $H_0 : P = P_0$
- multiple test based on $K_{c_k}(\mathbb{P}_n, P_0)$, $k \leqslant \kappa$
- cost c_k charges quantiles of P_0 in various ways
- joint distributions required : only one sample $X_1, ..., X_n$!

Motivation 3 - Multiple tests, multicriteria analysis

Goodness-of-fit, direct :

- $H_0 : P = P_0$
- multiple test based on $K_{c_k}(\mathbb{P}_n, P_0)$, $k \leqslant \kappa$
- cost c_k charges quantiles of P_0 in various ways
- joint distributions required : only one sample $X_1, ..., X_n$!

Goodness-of-fit, indirect :

- $H_0 : P = P_0$ with P_0 known relatively to Q_k, $k \leqslant \kappa$
- test statistics $K_{c_k}(P_0, Q_k)$, $P_0 \neq Q_k$
- cost c_k well suited to the tail of Q_k
- joint distributions required : $K_{c_k}(\mathbb{P}_n, Q_k)$ correlated
- centering : compute $K_{c_k}(P_0, Q_k)$ or simulate $K_{c_k}(P_0, \mathbb{Q}_{k,m})$
- robustness : average over (d_{c_k}, Q_k), $k \leqslant \kappa$

Motivation 4 - Depth indexes, mass localization

Vector field of α-depth :

- $1/2 \leqslant \alpha < 1$, any observation point O
- vector $D_{\alpha,n}(O)$ of α-depth at O generated by the q_α-surface at O
- directed toward mass seen from O, turns as α changes
- long/stable if O outsider, short/oscillating if O insider
- each O contains the whole field information : $\alpha \to D_{\alpha,n}(O)$
- more informative than Tukey depth at points

Motivation 4 - Depth indexes, mass localization

Vector field of α-depth :

- $1/2 \leqslant \alpha < 1$, any observation point O
- vector $D_{\alpha,n}(O)$ of α-depth at O generated by the q_α-surface at O
- directed toward mass seen from O, turns as α changes
- long/stable if O outsider, short/oscillating if O insider
- each O contains the whole field information : $\alpha \to D_{\alpha,n}(O)$
- more informative than Tukey depth at points

Use in multivariate statistics :

- descriptive tool : draw contours as a gradient field
- always exists (not Tukey contours)
- mode detection as $\alpha \to 1/2$
- symmetry test : spherical, pointwise, axial, orthotropy
- non-connected component detection
- α-depth fields and contours to compare distributions on $\mathbb{R}^d$

Motivation 5 - Quantile process of a process

Process Z indexed by $t \in T$:

- n realizations $Z_1(t), \ldots, Z_n(t)$
- empirical process indexed by $T \times \mathbb{R}$:

$$\begin{aligned}\alpha_n(t,x) &= \sqrt{n}(F_{Z(t),n}(x) - F_{Z(t)}(x)) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(1_{\{Z_i(t) \leqslant x\}} - \mathbb{P}(Z(t) \leqslant x)\right)\end{aligned}$$

- Ex : fractional Brownian motion, $0 < H < 1, t_n < t < T_0$ (Kevei-Mason, coupling using Berthet-Mason, bracketing entropy using Landau-Shepp)
- empirical quantile process indexed by $T \times (0,1)$:

$$\begin{aligned}\beta_n(\alpha, t) &= \sqrt{n}\left(F^{-1}_{Z(t),n}(\alpha) - F^{-1}_{Z(t)}(\alpha)\right) \\ &= \sqrt{n}\left(q_\alpha(\mathbb{P}_n^{Z(t)}) - q_\alpha(\mathbb{P}^{Z(t)})\right)\end{aligned}$$

- strong correlation between s, t for any fixed α

Motivation 6 - General setting

A class $\mathcal{F}$ of transport maps :

- consider simultaneously $\phi(X_1), ..., \phi(X_n)$ for $\phi \in \mathcal{F}$
- find the joint distribution (in $\alpha \in (0,1)$ and $\phi \in \mathcal{F}$) of empirical quantiles or contrasts $q_{\alpha,\phi}(\mathbb{P}_n) = q_\alpha(\mathbb{P}_n \circ \phi^{-1})$
- includes : q_α-surfaces (directions $\phi = p_u$), marginal distributions of $X \in \mathbb{R}^d$ ($\phi = p_{u_k}, k \leqslant d$), quantiles of processes, $K_c(P_n, P)$

Motivation 7 - Data "distribution"

Examples of data "distribution" :

- an algorithm outputs a distribution on $\mathbb{R}$
- groups of pixels in image analysis, when comparing to merge or separate
- clustering sample points themselves samples (different sizes) seen as empirical distributions. To classify use a contrast $K(\mathbb{P}_n, \mathbb{Q}_m)$ between distributions.

Motivation 8 - Level sets M-estimators

Level set :

- density f on $(\mathbb{R}^d, \mu)$ at level $\lambda > 0$ such that

$$L = \left\{x \in \mathbb{R}^d : f(x) \geqslant \lambda\right\}, \quad \alpha = \mathbb{P}(L), \quad v = \mu(L),$$

  is convex, compact, non empty interior, topologic boundary $\left\{x \in \mathbb{R}^d : f(x) = \lambda\right\}$.
- then

$$\begin{aligned} L = &\ \underset{A \in \mathcal{A}}{\arg\max} \left\{P(A) - \lambda\mu(A)\right\} \\ = &\ \underset{A \in \mathcal{A}}{\arg\min} \left\{\mu(A) : P(A) \geqslant \alpha\right\} \\ = &\ \underset{A \in \mathcal{A}}{\arg\max} \left\{P(A) : \mu(A) \leqslant v\right\}, \end{aligned}$$

  with unique maximiser/minimiser.
- let $\mathcal{A} \subset \mathcal{B}(\mathbb{R}^d)$ be a Donsker class of convex compact sets, $L \in \mathcal{A}$

Motivation 8 - Level sets M-estimators

Level set M-estimators :

- $X_1, ..., X_n$ i.i.d. density $f \in \mathbb{R}^d$
- if λ known, use the excess mass set

$$L_{1,n} \in \arg\max_{A \in \mathcal{A}} \{P_n(A) - \lambda\mu(A)\}$$

- if α known, use the minimum volume set

$$L_{2,n} \in \arg\min_{A \in \mathcal{A}} \{\mu(A) : P_n(A) \geqslant \alpha\}$$

- if v known use the maximum probability set

$$L_{3,n} \in \arg\max_{A \in \mathcal{A}} \{P_n(A) : \mu(A) \leqslant v\}$$

("modal set" for Polonik and Tsybakov).

Motivation 8 - Level sets M-estimators

Level set M-estimators :

- $X_1, \ldots, X_n$ i.i.d. density $f \in \mathbb{R}^d$
- if λ known, use the excess mass set

$$L_{1,n} \in \underset{A \in \mathcal{A}}{\arg\max} \{P_n(A) - \lambda\mu(A)\}$$

- if α known, use the minimum volume set

$$L_{2,n} \in \underset{A \in \mathcal{A}}{\arg\min} \{\mu(A) : P_n(A) \geqslant \alpha\}$$

- if v known use the maximum probability set

$$L_{3,n} \in \underset{A \in \mathcal{A}}{\arg\max} \{P_n(A) : \mu(A) \leqslant v\}$$

  (”modal set” for Polonik and Tsybakov).
- if not unique, all are equivalent asymptotically
- jointly : $L_{2,n}, L_{3,n}$ have same limit distribution $L_{1,n}$ differs

Motivation 8 - Level sets M-estimators

- shapes of neighbours on the cylinder space of ∂L
- local empirical process (Einmahl-Khmaladze), geometry, dimension free concentration, Wiener process $-$ Drift

Motivation 8 - Level sets M-estimators

- shapes of neighbours on the cylinder space of ∂L
- local empirical process (Einmahl-Khmaladze), geometry, dimension free concentration, Wiener process − Drift

Theorem (P. Berthet - J.H.J. Einmahl)

- Under some assumptions, on a common probability space there exists a triangular array $X_{n,1}, \ldots, X_{n,n}$, $n \in \mathbb{N}$, of vectors with density f on $\mathbb{R}^d$ rowwise independent, and a version of

$$Z(\mathcal{B}') = \operatorname*{arg\,max}_{B \in \mathcal{B}'} \left\{ \sqrt{\lambda} W(B) - D(B) \right\}$$

such that, as $n \to \infty$,

$$n^{1/3} \mu \left((L_{1,n} \Delta L_\lambda) \Delta \tau^{-1}_{n^{-1/3}}(Z(\mathcal{B})) \right) \xrightarrow{\mathbb{P}} 0$$

$$n^{1/3} \mu \left((L_{2,n} \Delta L_\lambda) \Delta \tau^{-1}_{n^{-1/3}}(Z(\mathcal{B}^*)) \right) \xrightarrow{\mathbb{P}} 0$$

$$n^{1/3} \mu \left((L_{3,n} \Delta L_\lambda) \Delta \tau^{-1}_{n^{-1/3}}(Z(\mathcal{B}^*)) \right) \xrightarrow{\mathbb{P}} 0$$

Real quantiles

Quantile transform

Density quantile function $h_X = f \circ F^{-1}$:

- $X_1, \ldots, X_n$ i.i.d. density $f \in \mathcal{C}_1$ on $\mathbb{R}$
- $0 < \inf_\Delta h_X < \sup_\Delta h_X < +\infty$, $\Delta \subset (0,1)$

Quantile transform

Density quantile function $h_X = f \circ F^{-1}$:

- $X_1, \ldots, X_n$ i.i.d. density $f \in \mathcal{C}_1$ on $\mathbb{R}$
- $0 < \inf_\Delta h_X < \sup_\Delta h_X < +\infty$, $\Delta \subset (0, 1)$

Back to uniform quantiles :

- if $X_i = F^{-1}(U_i)$ then $F_n^{-1}(\alpha) = F^{-1}(F_{U,n}^{-1}(\alpha))$ thus

$$\sup_{\alpha \in \Delta} \left| \left(F_n^{-1}(\alpha) - F^{-1}(\alpha)\right) h_X(\alpha) - \left(F_{U,n}^{-1}(\alpha) - \alpha\right) \right|$$

$$= \sup_{\alpha \in \Delta} \left| \left(F^{-1}(\alpha + F_{U,n}^{-1}(\alpha) - \alpha) - F^{-1}(\alpha)\right) h_X(\alpha) - \left(F_{U,n}^{-1}(\alpha) - \alpha\right) \right|$$

$$\sim \sup_{\alpha \in \Delta} \left| \alpha (1 - \alpha) \frac{h_X'(\alpha)}{h_X(\alpha)} \right| \left(\frac{F_{U,n}^{-1}(\alpha) - \alpha}{\sqrt{2\alpha (1 - \alpha)}} \right)^2 = O_{p.s.} \left(\frac{\log \log n}{n} \right)$$

Quantile transform

Density quantile function $h_X = f \circ F^{-1}$:

- $X_1, ..., X_n$ i.i.d. density $f \in \mathcal{C}_1$ on $\mathbb{R}$
- $0 < \inf_\Delta h_X < \sup_\Delta h_X < +\infty$, $\Delta \subset (0,1)$

Back to uniform quantiles :

- if $X_i = F^{-1}(U_i)$ then $F_n^{-1}(\alpha) = F^{-1}(F_{U,n}^{-1}(\alpha))$ thus

$$\sup_{\alpha\in\Delta} \left| \left(F_n^{-1}(\alpha) - F^{-1}(\alpha)\right) h_X(\alpha) - \left(F_{U,n}^{-1}(\alpha) - \alpha\right) \right|$$

$$= \sup_{\alpha\in\Delta} \left| \left(F^{-1}(\alpha + F_{U,n}^{-1}(\alpha) - \alpha) - F^{-1}(\alpha)\right) h_X(\alpha) - \left(F_{U,n}^{-1}(\alpha) - \alpha\right) \right|$$

$$\sim \sup_{\alpha\in\Delta} \left| \alpha\,(1-\alpha) \frac{h_X'(\alpha)}{h_X(\alpha)} \right| \left(\frac{F_{U,n}^{-1}(\alpha) - \alpha}{\sqrt{2\alpha\,(1-\alpha)}} \right)^2 = O_{p.s.}\left(\frac{\log\log n}{n}\right)$$

- $\mathbb{B}$ standard Brownian bridge

$$\left\{\sqrt{n}\left(F_{U,n}^{-1}(\alpha) - \alpha\right) : \alpha \in (0,1)\right\} \to_{weak} \{\mathbb{B}(\alpha) : \alpha \in (0,1)\}$$

$$\left\{\sqrt{n}\left(F_n^{-1}(\alpha) - F^{-1}(\alpha)\right) h_X(\alpha) : \alpha \in \Delta\right\} \to_{weak} \{\mathbb{B}(\alpha) : \alpha \in \Delta\}$$

Approximation

- if $f > 0$ at $\alpha_1, ..., \alpha_k$

$$\sqrt{n}\begin{pmatrix} F_n^{-1}(\alpha_1) - F^{-1}(\alpha_1) \\ ... \\ F_n^{-1}(\alpha_k) - F^{-1}(\alpha_k) \end{pmatrix} \rightarrow_{weak} \mathcal{N}(0_k, \Sigma), \quad \Sigma_{i,j} = \frac{\alpha_i \wedge \alpha_j - \alpha_i \alpha_j}{h_X(\alpha_i) h_X(\alpha_j)}.$$

Approximation

- if $f > 0$ at $\alpha_1, ..., \alpha_k$

$$\sqrt{n}\begin{pmatrix} F_n^{-1}(\alpha_1) - F^{-1}(\alpha_1) \\ \dots \\ F_n^{-1}(\alpha_k) - F^{-1}(\alpha_k) \end{pmatrix} \to_{weak} \mathcal{N}(0_k, \Sigma), \quad \Sigma_{i,j} = \frac{\alpha_i \wedge \alpha_j - \alpha_i\alpha_j}{h_X(\alpha_i)h_X(\alpha_j)}.$$

- uniform quantiles approximated by a KMT Brownian sequence

$$\mathbb{P}\left(\sup_{\alpha\in(0,1)}\left|\sqrt{n}\left(F_{U,n}^{-1}(\alpha) - \alpha\right) - \mathbb{B}_n(\alpha)\right| \geqslant c_1(t + \log n)\right) \leqslant c_2 e^{-c_3 t}$$

Approximation

- if $f > 0$ at $\alpha_1, \ldots, \alpha_k$

$$\sqrt{n}\begin{pmatrix} F_n^{-1}(\alpha_1) - F^{-1}(\alpha_1) \\ \ldots \\ F_n^{-1}(\alpha_k) - F^{-1}(\alpha_k) \end{pmatrix} \to_{weak} \mathcal{N}(0_k, \Sigma), \quad \Sigma_{i,j} = \frac{\alpha_i \wedge \alpha_j - \alpha_i\alpha_j}{h_X(\alpha_i)h_X(\alpha_j)}.$$

- uniform quantiles approximated by a KMT Brownian sequence

$$\mathbb{P}\left(\sup_{\alpha\in(0,1)} \left|\sqrt{n}\left(F_{U,n}^{-1}(\alpha) - \alpha\right) - \mathbb{B}_n(\alpha)\right| \geqslant c_1(t + \log n)\right) \leqslant c_2 e^{-c_3 t}$$

- approximation of $F_n^{-1}(\alpha)$ over $\Delta_n = [\varepsilon_n, 1-\varepsilon_n]$, $\varepsilon_n = (\log\log n)/n$ via $F_{U,n}^{-1}(\alpha)$ hence $\mathbb{B}_n(\alpha)/h_X(\alpha)$ but tedious satements

Approximation

- if $f > 0$ at $\alpha_1, \dots, \alpha_k$

$$\sqrt{n}\begin{pmatrix} F_n^{-1}(\alpha_1) - F^{-1}(\alpha_1) \\ \dots \\ F_n^{-1}(\alpha_k) - F^{-1}(\alpha_k) \end{pmatrix} \to_{weak} \mathcal{N}(0_k, \Sigma), \quad \Sigma_{i,j} = \frac{\alpha_i \wedge \alpha_j - \alpha_i\alpha_j}{h_X(\alpha_i)h_X(\alpha_j)}.$$

- uniform quantiles approximated by a KMT Brownian sequence

$$\mathbb{P}\left(\sup_{\alpha\in(0,1)} \left|\sqrt{n}\left(F_{U,n}^{-1}(\alpha) - \alpha\right) - \mathbb{B}_n(\alpha)\right| \geqslant c_1(t + \log n)\right) \leqslant c_2 e^{-c_3 t}$$

- approximation of $F_n^{-1}(\alpha)$ over $\Delta_n = [\varepsilon_n, 1-\varepsilon_n]$, $\varepsilon_n = (\log\log n)/n$ via $F_{U,n}^{-1}(\alpha)$ hence $\mathbb{B}_n(\alpha)/h_X(\alpha)$ but tedious satements
- if no density at $x_\alpha = F^{-1}(\alpha)$ and regular variation of $F(x) - \alpha$, CLT rate n^θ, approximation rate changes

Approximation

- if $f > 0$ at $\alpha_1, ..., \alpha_k$

$$\sqrt{n}\begin{pmatrix} F_n^{-1}(\alpha_1) - F^{-1}(\alpha_1) \\ \cdots \\ F_n^{-1}(\alpha_k) - F^{-1}(\alpha_k) \end{pmatrix} \to_{weak} \mathcal{N}(0_k, \Sigma), \quad \Sigma_{i,j} = \frac{\alpha_i \wedge \alpha_j - \alpha_i\alpha_j}{h_X(\alpha_i)h_X(\alpha_j)}.$$

- uniform quantiles approximated by a KMT Brownian sequence

$$\mathbb{P}\left(\sup_{\alpha\in(0,1)}\left|\sqrt{n}\left(F_{U,n}^{-1}(\alpha) - \alpha\right) - \mathbb{B}_n(\alpha)\right| \geqslant c_1(t + \log n)\right) \leqslant c_2 e^{-c_3 t}$$

- approximation of $F_n^{-1}(\alpha)$ over $\Delta_n = [\varepsilon_n, 1-\varepsilon_n]$, $\varepsilon_n = (\log\log n)/n$ via $F_{U,n}^{-1}(\alpha)$ hence $\mathbb{B}_n(\alpha)/h_X(\alpha)$ but tedious satements
- if no density at $x_\alpha = F^{-1}(\alpha)$ and regular variation of $F(x) - \alpha$, CLT rate n^θ, approximation rate changes
- approximation $O_{p.s.}(n^{-1/2+\varepsilon})$ by iterated Kiefer processes

Approximation par le processus empirique

Using the empirical process itself, by KMT,

$$\sqrt{n}(F_n(x) - F(x)) = \mathbb{B}_n^X(x) + \xi_n(x), \quad \sup_{x\in\mathbb{R}} |\xi_n(x)| \leqslant \frac{12 \log n}{\sqrt{n}}$$

Approximation par le processus empirique

Using the empirical process itself, by KMT,

$$\sqrt{n}(F_n(x) - F(x)) = \mathbb{B}_n^X(x) + \xi_n(x), \quad \sup_{x\in\mathbb{R}} |\xi_n(x)| \leqslant \frac{12 \log n}{\sqrt{n}}$$

we have, for $x_\alpha = F^{-1}(\alpha)$ and $\mathbb{B}_n^X(x) = \mathbb{B}_n(F(x))$

$$\begin{aligned}
F_n^{-1}(\alpha) &= \inf \{x : F_n(x) \geqslant \alpha\} \\
&= \inf \left\{x : F(x) + \frac{\mathbb{B}_n^X(x) + \xi_n(x)}{\sqrt{n}} \geqslant F(x_\alpha)\right\} \\
&= \inf \left\{x : \mathbb{B}_n^X(x) \geqslant \sqrt{n}\,(F(x_\alpha) - F(x)) - \xi_n(x)\right\} \\
&= \inf \left\{x : \mathbb{B}_n^X(x) \geqslant -\sqrt{n}(f(x_\alpha) + o(1))(x - x_\alpha) - \xi_n(x)\right\} \\
&= \inf \left\{x : \mathbb{B}_n^X(x) \text{ franchit une droite de pente } -\sqrt{n}h_X(\alpha)\right\}
\end{aligned}$$

Approximation par le processus empirique

Using the empirical process itself, by KMT,

$$\sqrt{n}(F_n(x) - F(x)) = \mathbb{B}_n^X(x) + \xi_n(x), \quad \sup_{x\in\mathbb{R}} |\xi_n(x)| \leqslant \frac{12 \log n}{\sqrt{n}}$$

we have, for $x_\alpha = F^{-1}(\alpha)$ and $\mathbb{B}_n^X(x) = \mathbb{B}_n(F(x))$

$$\begin{aligned}
F_n^{-1}(\alpha) &= \inf \{x : F_n(x) \geqslant \alpha\} \\
&= \inf \left\{x : F(x) + \frac{\mathbb{B}_n^X(x) + \xi_n(x)}{\sqrt{n}} \geqslant F(x_\alpha)\right\} \\
&= \inf \left\{x : \mathbb{B}_n^X(x) \geqslant \sqrt{n}\,(F(x_\alpha) - F(x)) - \xi_n(x)\right\} \\
&= \inf \left\{x : \mathbb{B}_n^X(x) \geqslant -\sqrt{n}(f(x_\alpha) + o(1))(x - x_\alpha) - \xi_n(x)\right\} \\
&= \inf \left\{x : \mathbb{B}_n^X(x) \text{ franchit une droite de pente } -\sqrt{n}h_X(\alpha)\right\}
\end{aligned}$$

then $\sqrt{n}(F_n^{-1}(\alpha) - x_\alpha)h_X(\alpha) = \mathbb{B}_n^X(x) + \xi_n'(\alpha)$ with error

$$\sup_{\alpha\in\boldsymbol{\Delta}_n} \left|\xi_n'(\alpha)\right| = O_{p.s.}\left(\frac{(\log n)^{1/2}(\log\log n)^{1/4}}{n^{1/4}}\right)$$

since $\mathbb{B}_n$ oscillates on $\left[F_n^{-1}(\alpha), x_\alpha\right]$ within range $\sqrt{(\log\log n)/n}$.

Bahadur-Kiefer

- $\mathbb{U}_n(\alpha) = \sqrt{n}(F_{U,n}(\alpha) - \alpha)$, $\mathbb{Q}_n(\alpha) = \sqrt{n}(F_{U,n}^{-1}(\alpha) - \alpha)$
- Compensation

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{(\log\log n)^{3/4}} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = \left(\frac{32}{27}\alpha(1-\alpha)\right)^{1/4} \quad p.s.$$

Bahadur-Kiefer

- $\mathbb{U}_n(\alpha) = \sqrt{n}(F_{U,n}(\alpha) - \alpha)$, $\mathbb{Q}_n(\alpha) = \sqrt{n}(F_{U,n}^{-1}(\alpha) - \alpha)$
- Compensation

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{(\log\log n)^{3/4}} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = \left(\frac{32}{27}\alpha(1-\alpha)\right)^{1/4} \quad p.s.$$

- Uniform version

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{(\sup_{\alpha\in(0,1)} |\mathbb{U}_n(\alpha)| \log n)^{1/2}} \sup_{\alpha\in(0,1)} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = 1 \quad p.s.$$

Bahadur-Kiefer

- $\mathbb{U}_n(\alpha) = \sqrt{n}(F_{U,n}(\alpha) - \alpha)$, $\mathbb{Q}_n(\alpha) = \sqrt{n}(F_{U,n}^{-1}(\alpha) - \alpha)$
- Compensation

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{(\log\log n)^{3/4}} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = \left(\frac{32}{27}\alpha(1-\alpha)\right)^{1/4} \quad p.s.$$

- Uniform version

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{(\sup_{\alpha\in(0,1)} |\mathbb{U}_n(\alpha)| \log n)^{1/2}} \sup_{\alpha\in(0,1)} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = 1 \quad p.s.$$

qui implique

$$\limsup_{n\to+\infty} \frac{n^{1/4}}{(\log n)^{1/2}(\log\log n)^{1/4}} \sup_{\alpha\in(0,1)} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = \left(\frac{1}{2}\right)^{1/4} \quad p.s.$$

$$\liminf_{n\to+\infty} \frac{n^{1/4}(\log\log n)^{1/4}}{(\log n)^{1/2}} \sup_{\alpha\in(0,1)} \left|\mathbb{U}_n(\alpha) + \mathbb{Q}_n(\alpha)\right| = \left(\frac{\pi}{\sqrt{8}}\right)^{1/2} \quad p.s.$$

q_α-surfaces and depth fields : blackboard

Non asymptotic empirical process coupling

Empirical process $\alpha_n(f) = \sqrt{n}\,(P_n f - Pf)$ where

$$P_n f = \frac{1}{n}\sum_{i=1}^{n} f(X_i), \quad Pf = \mathbb{E}f(X), \quad f \in \mathcal{F}$$

and the $(P, \mathcal{F})$-Brownian bridge

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = P(fg) - (Pf)(Pg)$$

Non asymptotic empirical process coupling

Empirical process $\alpha_n(f) = \sqrt{n}\,(P_n f - Pf)$ where

$$P_n f = \frac{1}{n}\sum_{i=1}^{n} f(X_i), \quad Pf = \mathbb{E}f(X), \quad f \in \mathcal{F}$$

and the $(P, \mathcal{F})$-Brownian bridge

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = P(fg) - (Pf)(Pg)$$

Entropy numbers $N(\varepsilon, \mathcal{F}, L_2(P)) \leqslant \phi(\varepsilon)$ and associated functions

$$J_\phi(\varepsilon) = \int_0^\varepsilon \sqrt{\log \phi(\cdot)}\, d\cdot \qquad \Psi_\phi(\varepsilon) = \frac{\varepsilon}{\phi(\varepsilon)^{5/2}}$$

satisfy $\varepsilon \ll J_\phi \circ \Psi_\phi^{-1}(\varepsilon) \ll 1$, as $\varepsilon \to 0$.

Non asymptotic empirical process coupling

Empirical process $\alpha_n(f) = \sqrt{n}\,(P_n f - Pf)$ where

$$P_n f = \frac{1}{n}\sum_{i=1}^{n} f(X_i), \quad Pf = \mathbb{E}f(X), \quad f \in \mathcal{F}$$

and the $(P, \mathcal{F})$-Brownian bridge

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = P(fg) - (Pf)(Pg)$$

Entropy numbers $N(\varepsilon, \mathcal{F}, L_2(P)) \leqslant \phi(\varepsilon)$ and associated functions

$$J_\phi(\varepsilon) = \int_0^\varepsilon \sqrt{\log \phi(\cdot)}d\,\cdot \quad \Psi_\phi(\varepsilon) = \frac{\varepsilon}{\phi(\varepsilon)^{5/2}}$$

satisfy $\varepsilon \ll J_\phi \circ \Psi_\phi^{-1}(\varepsilon) \ll 1$, as $\varepsilon \to 0$.

Theorem (P. Berthet) Given $\theta > 0$ there exists (c_θ, n_θ) and a probability space with versions of $(\alpha_n, \mathbb{G}_n)$ such that

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n\| \geqslant c_\theta\, J_\phi \circ \Psi_\phi^{-1}\left(\sqrt{\frac{\log n}{n}}\right)\right) \leqslant \frac{1}{n^\theta}, \quad \text{all } n > n_\theta.$$

Kantorovitch contrasts K_c

Wasserstein distance

The simplest Kantorovich contrast is a distance for a symmetric cost.

Definition

Let F, G be dist. funct. on $\mathbb{R}$. The p-Wasserstein distance W_p is

$$W_p^p(F, G) = \min_{X \sim F, Y \sim G} \mathbb{E}|X - Y|^p,$$

where $X \sim F, Y \sim G$ means that (X, Y) has marginals F, G.

Wasserstein distance

The simplest Kantorovitch contrast is a distance for a symmetric cost.

Definition

Let F, G be dist. funct. on $\mathbb{R}$. The p-Wasserstein distance W_p is

$$W_p^p(F, G) = \min_{X \sim F, Y \sim G} \mathbb{E}|X - Y|^p,$$

where $X \sim F, Y \sim G$ means that (X, Y) has marginals F, G.

Property

If F,G continuous, the optimal joint law is unique : the deterministic optimal transportation $(X, Y) = (X, T(X))$ of quantiles of F to quantiles of G, $T(x) = G^{-1}(F(x))$.

Wasserstein distance

The simplest Kantorovitch contrast is a distance for a symmetric cost.

Definition

Let F, G be dist. funct. on $\mathbb{R}$. The p-Wasserstein distance W_p is

$$W_p^p(F,G) = \min_{X \sim F, Y \sim G} \mathbb{E}|X - Y|^p,$$

where $X \sim F, Y \sim G$ means that (X, Y) has marginals F, G.

Property

If F,G continuous, the optimal joint law is unique : the deterministic optimal transportation $(X, Y) = (X, T(X))$ of quantiles of F to quantiles of G, $T(x) = G^{-1}(F(x))$.

Corollary

From $(X, Y) = (F^{-1}(U), G^{-1}(U))$ with U uniform on $(0, 1)$ we have

$$W_p^p(F,G) = \int_0^1 |F^{-1}(u) - G^{-1}(u)|^p du = \mathbb{E}|F^{-1}(U) - G^{-1}(U)|^p$$

Kantorovitch type distances

Theorem (Cambanis, Simon, Stout zwvg76)

$$c(x', y') - c(x', y) - c(x, y') + c(x, y) \leqslant 0, \quad x \leqslant x', \quad y \leqslant y'$$

implies

$$K_c(F, G) = \min_{X \sim F, Y \sim G} \mathbb{E} c(X, Y) = \int_0^1 c(F^{-1}(u), G^{-1}(u)) du.$$

Theorem (Cambanis, Simon, Stout zwvg76)

$$c(x', y') - c(x', y) - c(x, y') + c(x, y) \leqslant 0, \quad x \leqslant x', \quad y \leqslant y'$$

implies

$$K_c(F, G) = \min_{X \sim F, Y \sim G} \mathbb{E}c(X, Y) = \int_0^1 c(F^{-1}(u), G^{-1}(u))du.$$

Conditions (C) (Berthet, Fort, Klein ihp19)

Notations for a convex symmetric cost c.

$$\begin{array}{ll} (C1) & c(x,x) = 0,\ c(x,y) \geqslant 0, \quad c \in \mathcal{C}_1([-m,m] \times \mathbb{R} \cup \mathbb{R} \times [-m,m]) \\ (C2) & c(x,y) = \exp(l(|x-y|)), \quad x, y \in (m, +\infty), \quad l \in RV_2^+(\gamma) \\ (C3) & \left|c\left(x', y'\right) - c\left(x, y\right)\right| \leqslant d(\tau)\left(\left|x' - x\right| + \left|y' - y\right|\right) \quad x, y, x', y' \in D(\tau) \end{array}$$

where $d(\tau) \to 0$ if $\tau \to 0$ and
$D(\tau) = \{(x, y) : \max(|x|, |y|) \leqslant m, |x - y| \leqslant \tau\}$.

Regular variation

Regularity (C2)

In $c(x,y) = \exp(l(|x-y|))$:

- either $l(x) = L(x)$ is slowly varying

$$L'(x) = \frac{\varepsilon_1(x)L(x)}{x}, \quad \varepsilon_1(x) \to 0 \text{ as } x \to +\infty. \tag{1}$$

  and

$$L'(x) \geqslant \frac{l_1}{x}, \quad l_1 \geqslant 1.$$

- or $l(x) = x^\gamma L(x)$ is regularly varying $\gamma > 0$, and L' obeys (1).

Regular variation

Regularity (C2)

In $c(x,y) = \exp(l(|x-y|))$:

- either $l(x) = L(x)$ is slowly varying

$$L'(x) = \frac{\varepsilon_1(x)L(x)}{x}, \quad \varepsilon_1(x) \to 0 \text{ as } x \to +\infty. \tag{1}$$

  and

$$L'(x) \geqslant \frac{l_1}{x}, \quad l_1 \geqslant 1.$$

- or $l(x) = x^\gamma L(x)$ is regularly varying $\gamma > 0$, and L' obeys (1).

Examples

Conditions (C) are satisfied, for $\alpha > 1$, $\beta > 1$,

$$c(x,y) = |x-y|^\alpha, \gamma = 0$$

$$c(x,y) = \exp((\log(1+|x-y|))^\beta) - 1, \ \gamma = 0$$

$$c(x,y) = \exp(|x-y|^\beta) - 1, \ \gamma = \beta$$

Conditions FG

Tails on $(m, +\infty)$

(X, F, f) has the heavier right hand tail compared to (Y, G, g),

$$\psi_X(x) = -\log \mathbb{P}(X > x) \quad > \quad \psi_Y(x) = -\log \mathbb{P}(Y > x), \quad x \in (m, +\infty).$$

To their "density quantile" $h_X = f \circ F^{-1}$, $h_Y = g \circ G^{-1}$ associate

$$H_X(u) = \frac{1-u}{F^{-1}(u) h_X(u)}, \quad H_Y(u) = \frac{1-u}{G^{-1}(u) h_Y(u)}.$$

Conditions FG

Tails on $(m, +\infty)$

(X, F, f) has the heavier right hand tail compared to (Y, G, g),

$$\psi_X(x) = -\log \mathbb{P}(X > x) \quad > \quad \psi_Y(x) = -\log \mathbb{P}(Y > x), \quad x \in (m, +\infty).$$

To their "density quantile" $h_X = f \circ F^{-1}$, $h_Y = g \circ G^{-1}$ associate

$$H_X(u) = \frac{1-u}{F^{-1}(u) h_X(u)}, \quad H_Y(u) = \frac{1-u}{G^{-1}(u) h_Y(u)}.$$

Conditions FG

For $\overline{u} = \max(F(m), G(m)) > 1/2$,

$$(FG1) \quad F, G \in \mathcal{C}_2(\mathbb{R}_+), \quad f, g > 0 \text{ on } \mathbb{R}_+.$$

$$(FG2) \quad (1-u)\left|(\log h(u))'\right| \text{ bounded on } (\bar{u}, 1), \quad h = h_X, h_Y.$$

$$(FG3) \quad H_X, H_Y \text{ bounded on } (\bar{u}, 1).$$

$$(FG4) \quad \tau(u) = F^{-1}(u) - G^{-1}(u) \geqslant \tau_0 > 0, \quad u \geqslant \overline{u}.$$

Hypotheses FG

$$(FG2+3) \quad \sup_{x>m} \frac{1-F(x)}{f(x)}\left(\frac{1}{x}+\frac{|f'(x)|}{f(x)}\right) < +\infty, \text{ idem } g, G$$

Hypotheses FG

$$(FG2+3) \quad \sup_{x>m} \frac{1-F(x)}{f(x)} \left(\frac{1}{x} + \frac{|f'(x)|}{f(x)} \right) < +\infty, \text{ idem } g, G$$

Proposition

If $\psi_X \in RV_2^+(0,m)$ then F satisfies (FG).
If $\psi_X \in RV_2(\gamma_1, m)$ for $\gamma_1 > \gamma_0 > 0$ and, whenever $\gamma_1 = 1$ assume $\psi_X(x) = xL(x)$ with $L' \in RV_1(-1,m)$ and (1), then F satisfies (FG) and $(FG3)$ is strenghtened into

$$H_X(u) \leqslant \frac{1}{\gamma_0 \log(1/(1-u))}, \quad u > \overline{u}.$$

All usual distributions satisfy (FG).

Conditions FG

Examples

- Heavy tail, Pareto with parameter $p > 0$,

$$\psi_X(x) = p \log x,\ F^{-1}(u) = \frac{1}{(1-u)^{1/p}},\ H_X(u) = \frac{1}{p},$$
$$h_X(u) = p(1-u)^{1+1/p},\ (1-u)\left|(\log h_X(u))'\right| = \frac{1}{p}.$$

Conditions FG

Examples

- Heavy tail, Pareto with parameter $p > 0$,

$$\psi_X(x) = p \log x,\ F^{-1}(u) = \frac{1}{(1-u)^{1/p}},\ H_X(u) = \frac{1}{p},$$
$$h_X(u) = p(1-u)^{1+1/p},\ (1-u)\left|(\log h_X(u))'\right| = \frac{1}{p}.$$

- Light tail, Weibull with parameter $q > 0$,

$$\psi_X(x) = x^q,\ F^{-1}(u) = (\log(1/(1-u)))^{1/q},\ H_X(u) = \frac{1}{q\log(1/(1-u))}$$
$$h_X(u) = q(1-u)(\log(1/(1-u)))^{1-1/q},\ (1-u)\left|(\log h_X(u))'\right| \sim \frac{1}{q}\ (u \to 1).$$

that is log-convex if $q < 1$, log-concave if $q > 1$.

Conditions CFG

Let

$$\limsup_{x \to +\infty} \frac{\log(x l'(x))}{\log l(x)} = \theta_1 \in [0, 1] \, .$$

Thus $\theta_1 = 0$ for small costs as Wasserstein $l(x) = \alpha \log x$.

Condition cost-tail

For $\theta > 1 + \theta_1$,

$$(CFG) \left(\psi_X \circ l^{-1}\right)' (x) \geqslant 2 + \frac{2\theta}{x}, \quad x > m.$$

Conditions CFG

Let

$$\limsup_{x\to+\infty} \frac{\log(xl'(x))}{\log l(x)} = \theta_1 \in [0,1]\,.$$

Thus $\theta_1 = 0$ for small costs as Wasserstein $l(x) = \alpha \log x$.

Condition cost-tail

For $\theta > 1 + \theta_1$,

$$(CFG)\ \left(\psi_X \circ l^{-1}\right)'(x) \geqslant 2 + \frac{2\theta}{x}, \quad x > m.$$

Sufficient condition

If, for $\zeta > 2$,

$$\mathbb{P}\left(X > x\right) \leqslant \frac{1}{\exp(\zeta l(x))}, \quad x > m$$

then $\psi_X(x) \geqslant \zeta l(x)$ and (CFG) is true for all θ.

Consequences of conditions CFG

- F has heavier tail, hence G also satisfies (CFG).

Consequences of conditions CFG

- F has heavier tail, hence G also satisfies (CFG).
- by integrating (*CFG*),

$$\psi_X \circ l^{-1}(x) \geqslant 2x + 2\theta \log x + K, \quad x > m,$$

Consequences of conditions CFG

- F has heavier tail, hence G also satisfies (CFG).
- by integrating (*CFG*),

$$\psi_X \circ l^{-1}(x) \geqslant 2x + 2\theta \log x + K, \quad x > m,$$

- this implies

$$\mathbb{P}\left(\exp l(X) > x\right) = \exp\left(-\psi_X \circ l^{-1}(\log x)\right) \leqslant \frac{K}{x^2 (\log x)^{2\theta}}$$

Consequences of conditions CFG

- F has heavier tail, hence G also satisfies (CFG).
- by integrating (*CFG*),

$$\psi_X \circ l^{-1}(x) \geqslant 2x + 2\theta \log x + K, \quad x > m,$$

- this implies

$$\mathbb{P}\left(\exp l(X) > x\right) = \exp\left(-\psi_X \circ l^{-1}(\log x)\right) \leqslant \frac{K}{x^2 (\log x)^{2\theta}}$$

- thus, since $\theta > 1$,

$$\int_m^{+\infty} \sqrt{\mathbb{P}\left(\exp l(X) > x\right)} dx < +\infty$$

that is strictly minimal for the existence of the limiting variance below, and for a CLT (see Bobkov, Ledoux ams19 for W_1).

Conditions CFG

Examples

- Wasserstein cost $l(x) = \alpha \log x$, $\alpha > 1$, and Pareto, $\psi_X(x) = p \log x$. Then (CFG) reads

  $$\alpha x/p < x/2 - \theta \log x$$

  is satisfied if $p > 2\alpha$.
- over-exponential cost $l(x) = x^\beta$, $\beta > 1$, (CFG) true if Weibul

  $$\mathbb{P}(X > x) \leqslant \exp(-2x^\beta - \delta \log x), \quad \delta > 4(1-\beta).$$

- Normal distributions satisfy (CFG) for costs

  $$exp(l(x)) = \exp(ax^\gamma), \quad \gamma < 2, \quad a > 0,$$

  and if $\gamma = 2$ the variance of X should be less than $a/4$. Moreover G can also be normal with smaller variance or the same but smaller expectation.

Estimation of $K_c(F, G)$ for $F \neq G$

A simple nonparametric estimator

Estimator : quick to compute !

We observe (X_i, Y_i) with marginals F, G and joint distribution H.

- let $\mathbb{F}_n, \mathbb{G}_n$ be empirical distr. funct. (random, non indep.) and

$$K_c(\mathbb{F}_n, \mathbb{G}_n) = \frac{1}{n}\sum_{i=1}^{n} c(X_{(i)}, Y_{(i)}).$$

- this estimates $K_c(F, G)$ with an explicit CLT even if not observing data $(X, G^{-1}(F(X))$ driven by the optimal transportation map !
- studied by Munk and co-autors but by truncating the integral $K_c(F, G)$, indep. case and assuming that remaining terms vanish in probability (false in general) to apply delta-method.
- very difficult problem without truncating: aeras distribution dependant, integration of diverging random Taylor remainder processes, diagonal explosion on $c(x, x)$ everywhere (0 and ∞), dependent samples, bias, work below convergence in probability

Convergences

Consistency

If $0 \leqslant c(x,y) \leqslant D(x)+D(y)$ with D continuous positive such that $\mathbb{E}D(X) < +\infty$ and $\mathbb{E}D(Y) < +\infty$ then

$$\lim_{n\to+\infty} K_c(\mathbb{F}_n, \mathbb{G}_n) = K_c(F,G) < +\infty \quad a.s.$$

Notation

For $u, v \in (0,1)$,

$$\begin{aligned}
\Pi(u,v) &= \mathbb{P}\left(X \leqslant F^{-1}(u), Y \leqslant G^{-1}(v)\right) \\
\Sigma(u,v) &= \begin{pmatrix} \frac{\min(u,v)-uv}{h_X(u)h_X(v)} & \frac{\Pi(u,v)-uv}{h_X(u)h_Y(v)} \\ \frac{\Pi(v,u)-uv}{h_X(v)h_Y(u)} & \frac{\min(u,v)-uv}{h_Y(v)h_Y(u)} \end{pmatrix} \\
\nabla(u) &= \left(\frac{\partial}{\partial x}c(F^{-1}(u), G^{-1}(u)), \frac{\partial}{\partial y}c(F^{-1}(u), G^{-1}(u))\right) \\
&= (\nabla_x(u), \nabla_y(u))
\end{aligned}$$

Theorem (Berthet Fort Klein aihp19)

Under (C), (FG), (CFG) we have

$$\sqrt{n}\left(K_c(\mathbb{F}_n, \mathbb{G}_n) - K_c(F,G)\right) \to_{weak} \mathcal{N}(0, \sigma^2(H,c))$$

with

$$\sigma^2(H,c) = \int_0^1 \int_0^1 \nabla(u)\Sigma(u,v)\nabla(v)dudv < +\infty.$$

Convergences

Theorem (Berthet Fort Klein aihp19)

Under (C), (FG), (CFG) we have

$$\sqrt{n}\left(K_c(\mathbb{F}_n, \mathbb{G}_n) - K_c(F, G)\right) \rightarrow_{weak} \mathcal{N}(0, \sigma^2(H, c))$$

with

$$\sigma^2(H, c) = \int_0^1 \int_0^1 \nabla(u)\Sigma(u, v)\nabla(v)dudv < +\infty.$$

Order of the critical extreme quantile

$$K_n \to +\infty, \quad \frac{K_n}{\log\log n} \to 0, \quad k_n = \frac{\sqrt{n}}{K_n \exp\left(l \circ \psi_X^{-1}(\log n + K_n)\right)}.$$

Corollary (truncated version asympt. unbiased)

Under (C), (FG), (CFG) for any $\varepsilon_n \leqslant k_n/n$ and $\varepsilon_n^- \leqslant k_n^-/n$ (left and right tails),

$$K_{c,n}(\mathbb{F}_n, \mathbb{G}_n) = \int_{\varepsilon_n^-}^{1-\varepsilon_n} c(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(u))du$$

also satisfies

$$\sqrt{n}\,(K_{c,n}(\mathbb{F}_n, \mathbb{G}_n) - K_c(F, G)) \rightarrow_{weak} \mathcal{N}(0, \sigma^2(H, c)).$$

For all $\varepsilon_n, \varepsilon_n^- \rightarrow 0$,

$$\sqrt{n}\,(K_{c,n}(\mathbb{F}_n, \mathbb{G}_n) - K_{c,n}(F, G)) \rightarrow_{weak} \mathcal{N}(0, \sigma^2(H, c)).$$

Convergences

We can also control the distance to a given distribution G : by sampling indepenently, or by the following

Corollary (one known marginal)

Under (C), (FG), (CFG) with G fixed,

$$\sqrt{n}\,(W_c(\mathbb{F}_n, G) - W_c(F,G)) \to_{weak} \mathcal{N}(0, \sigma_x^2(F,c)),$$
$$\sqrt{n}\,(W_c(F, \mathbb{G}_n) - W_c(F,G)) \to_{weak} \mathcal{N}(0, \sigma_y^2(G,c)),$$

where

$$\sigma_x^2(F,c) = \int_0^1\int_0^1 \nabla_x(u)\nabla_x(v)\frac{\min(u,v)-uv}{h_X(u)h_X(v)}dudv \leqslant \sigma^2(H,c) < +\infty,$$
$$\sigma_y^2(G,c) = \int_0^1\int_0^1 \nabla_y(u)\nabla_y(v)\frac{\min(u,v)-uv}{h_Y(u)h_Y(v)}dudv \leqslant \sigma^2(H,c) < +\infty.$$

(These convergences are joint toward a Gaussian vector with explicit variance).

Convergences

Corollary (Wasserstein W_2)

If the samples are independent then under (FG) and

$$\mathbb{P}(X > x) \leqslant 1/x^{4+\varepsilon}, \ \varepsilon > 0$$

we get

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_{(i)} - Y_{(i)})^2 - W_2^2(F,G)\right) \rightarrow_{weak} \mathcal{N}(0, \sigma^2(H, c_2))$$

with

$$\sigma^2(H, c_2) = 4\int_0^1\int_0^1 \left(\frac{\min(u,v) - uv}{h_X(u)h_X(v)} + \frac{\min(u,v) - uv}{h_Y(u)h_Y(v)}\right)\dots$$
$$\dots(F^{-1}(u) - G^{-1}(u))(F^{-1}(v) - G^{-1}(v))dudv.$$

Corollary (semi-parametric familly)

- assume F symmetric, $var(Z) = 1$, $V_4 = var(Z^2) < +\infty$ for Z distr. F.
- for $a > 0$, $b \in \mathbb{R}$ define $F_{a,b}(x) = F((x-b)/a)$, $x \in \mathbb{R}$.
- if $(a,b) \neq (a',b')$ and $H(x,y) = F_{a,b}(x)F_{a',b'}(y)$ then

$$\sigma^2(H, c_2) = 4(a^2 + a'^2)\left((b-b')^2 + \frac{V_4}{4}(a-a')^2\right).$$

Gaussian familly

For two indép. Gaussian samples with distribution $\mathcal{N}(\nu, \zeta^2)$ and $\mathcal{N}(\mu, \xi^2)$ we recover the limiting variance

$$2(\zeta^2 + \xi^2)(2(\nu-\mu)^2 + (\zeta-\xi)^2)$$

of Theorem 2.2 in Rippl, Munk, Sturm jma16 which proves that our nonparametric estimator performs as efficiently as their parametric MLE and plug-in estimator!

Brownian approximation of joint quantiles

Quantile processes

$$\beta_n^X = \sqrt{n}(\mathbb{F}_n^{-1} - F^{-1}), \quad \beta_n^Y = \sqrt{n}(\mathbb{G}_n^{-1} - G^{-1})$$

Let P_H with bivariate D.F. H. We construct a sequence $\mathbb{B}_n$ of P_H-Brownian bridges indexed by marginal half-plans of X and Y, $\mathcal{H}_{x_0} = \{(x,y) : x \leqslant x_0\}$ et $\mathcal{H}^{y_0} = \{(x,y) : y \leqslant y_0\}$,

$$cov(\mathbb{B}_n(A), \mathbb{B}_n(B)) = P_H(A \cap B) - P_H(A)P_H(B),$$

Write

$$B_n^X(u) = \mathbb{B}_n\left(\mathcal{H}_{F^{-1}(u)}\right) \quad \text{et} \quad B_n^Y(u) = \mathbb{B}_n\left(\mathcal{H}^{G^{-1}(u)}\right), \quad u \in [0,1]$$

then, for $u, v \in [0,1]$,

$$cov(B_n^X(u), B_n^X(v)) = P_H(\mathcal{H}_{F^{-1}(u)} \cap \mathcal{H}_{F^{-1}(v)}) - uv = \min(u,v) - uv,$$

$$cov(B_n^Y(u), B_n^Y(v)) = P_H(\mathcal{H}^{G^{-1}(u)} \cap \mathcal{H}^{G^{-1}(v)}) - uv = \min(u,v) - uv,$$

$$cov(B_n^X(u), B_n^Y(v)) = P_H(\mathcal{H}_{F^{-1}(u)} \cap \mathcal{H}^{G^{-1}(v)}) - uv = L(u,v) - uv,$$

copula $L(u,v) = H(F^{-1}(u), G^{-1}(v)) = \mathbb{P}(F(X) \leqslant u, G(Y) \leqslant v)$.

Brownian approximation of joint quantiles

Brownian coupling

- combine strong approximation on $\mathcal{D}$ of
$$\Delta_n = \sqrt{n}(P_{H_n}(A) - P_H(A)), \quad A \in \mathcal{D}, \quad P_{H_n} = \frac{1}{n}\sum_{i \leqslant n} \delta_{(X_i, Y_i)},$$
with quantile transform and KMT.
- marginals have to satisfy $(FG1)$ and $(FG2)$.
- remains true for d marginal quantile processes of a distribution in $\mathbb{R}^d$.
- with Berthet, Mason (2006-) since $\mathcal{D}$ is γ-VC, we can build a probability space supporting P_H-Brownian bridges $\mathbb{B}_n$,
$$\mathbb{P}\left(\sup_{A \in \mathcal{D}} |\Delta_n(A) - \mathbb{B}_n(A)| > c_\theta v_n\right) < \frac{1}{n^\theta}, \quad v_n = \frac{(\log n)^{\cdots}}{n^{1/(5\gamma+2)}}.$$

Brownian approximation of joint quantiles

Theorem

- Si F, G vérifient $(FG1)$, $(FG2)$, on the same probability space we build $\{(X_n, Y_n)\}$ and versions of $\left\{(B_n^X(u), B_n^Y(u)) : u \in \mathcal{I}_n\right\}$ st

$$\beta_n^X(u) = \frac{B_n^X(u) + Z_n^X(u)}{h_X(u)}, \quad \beta_n^Y(u) = \frac{B_n^Y(u) + Z_n^Y(u)}{h_Y(u)},$$

satisfy, for a $\xi > 0$,

$$\lim_{n \to +\infty} n^\xi \sup_{u \in \mathcal{I}_n} \left|Z_n^X(u)\right| = \lim_{n \to +\infty} n^\xi \sup_{u \in \mathcal{I}_n} \left|Z_n^Y(u)\right| = 0 \quad a.s.$$

Brownian approximation of joint quantiles

Theorem

- Si F, G vérifient $(FG1)$, $(FG2)$, on the same probability space we build $\{(X_n, Y_n)\}$ and versions of $\left\{(B_n^X(u), B_n^Y(u)) : u \in \mathcal{I}_n\right\}$ st

$$\beta_n^X(u) = \frac{B_n^X(u) + Z_n^X(u)}{h_X(u)}, \quad \beta_n^Y(u) = \frac{B_n^Y(u) + Z_n^Y(u)}{h_Y(u)},$$

  satisfy, for a $\xi > 0$,

$$\lim_{n \to +\infty} n^{\xi} \sup_{u \in \mathcal{I}_n} \left| Z_n^X(u) \right| = \lim_{n \to +\infty} n^{\xi} \sup_{u \in \mathcal{I}_n} \left| Z_n^Y(u) \right| = 0 \quad a.s.$$

- We can take $(B_n^X(u), B_n^Y(u)) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (G_k^X(u), G_k^Y(u))$ with $\left\{(G_k^X(u), G_k^Y(u)) : u \in (0,1)\right\}$ i.i.d. Brownian bridges (G^X, G^Y) such that $cov(G^X(u), G^Y(v)) = L(u,v) - uv$.

K_c-surfaces

Multivariate setting : surfaces

Two samples

- $\{X_i\}$ and $\{Y_i\}$ on $\mathbb{R}^d$
- with marginal distributions P and $Q \neq P$
- (X_i, Y_i) joint distribution Π on $\mathbb{R}^d \times \mathbb{R}^d$
- $u \in S_{d-1}$ unit sphere, directional projections $u \to (\langle X_i, u \rangle, \langle Y_i, u \rangle)$
- correlations in directions

Multivariate setting : surfaces

Two samples

- $\{X_i\}$ and $\{Y_i\}$ on $\mathbb{R}^d$
- with marginal distributions P and $Q \neq P$
- (X_i, Y_i) joint distribution Π on $\mathbb{R}^d \times \mathbb{R}^d$
- $u \in S_{d-1}$ unit sphere, directional projections $u \to (\langle X_i, u\rangle, \langle Y_i, u\rangle)$
- correlations in directions

K_c-surfaces

- true K_c-surface :
  $u \to K_c(u) = K_c(F_{\langle X,u\rangle}, G_{\langle Y,u\rangle}) = K_{c,u}(P, Q) = K_c(P \circ p_u^{-1}, Q \circ p_u^{-1}) \geqslant 0$
- empirical $d_{c,n}$-surface :

$$u \to K_{c,n}(u) = K_c(F_{n,\langle X,u\rangle}, G_{n,\langle Y,u\rangle}) = K_{c,u}(\mathbb{P}_n, \mathbb{Q}_n) = K_c(\mathbb{P}_n \circ p_u^{-1}, \mathbb{Q}_n \circ p_u^{-1})$$

Multivariate setting : surfaces

Goal

- estimate directional transport, simultaneously
- compare two or more (correlated/independent) samples
- goodness of fit $H_0 : P = P_0$

Multivariate setting : surfaces

Goal

- estimate directional transport, simultaneously
- compare two or more (correlated/independent) samples
- goodness of fit $H_0 : P = P_0$

Theorem

- under directional assumptions on $F_{\langle X,u\rangle} \neq G_{\langle Y,u\rangle}$
- under directional hypotheses on tails versus c
- we have the weak convergence to a Gaussian process

$$\left\{\sqrt{n}(K_{c,n}(u) - K_c(u)) : u \in S_{d-1}\right\} \rightarrow_{weak} \{B(u) : u \in S_{d-1}\}$$

Multivariate setting : surfaces

Goal

- estimate directional transport, simultaneously
- compare two or more (correlated/independent) samples
- goodness of fit $H_0 : P = P_0$

Theorem

- under directional assumptions on $F_{\langle X,u\rangle} \neq G_{\langle Y,u\rangle}$
- under directional hypotheses on tails versus c
- we have the weak convergence to a Gaussian process

$$\left\{\sqrt{n}(K_{c,n}(u) - K_c(u)) : u \in S_{d-1}\right\} \to_{weak} \{B(u) : u \in S_{d-1}\}$$

Limiting G

G determined by the underlying Π-Brownian bridge B_Π indexed by the VC class of functions $(\cdot, y)$ or $(x, \cdot) \to 1_{\{\langle \cdot, u\rangle < a\}}$ on $\mathbb{R}^d$ and not $\mathbb{R}^{2d}$.

Underlying correlations

$$cov(1_{\{\langle X,u_1\rangle<x\}},1_{\{\langle Y,u_2\rangle<y\}}) = \Pi(1_{\{\langle X,u_1\rangle<x\}}1_{\{\langle Y,u_2\rangle<y\}}) - F_{\langle X,u_1\rangle}(x)G_{\langle Y,u_2\rangle}(y).$$

Limiting covariance function

Underlying correlations

$$cov(1_{\{\langle X,u_1\rangle<x\}}, 1_{\{\langle Y,u_2\rangle<y\}}) = \Pi(1_{\{\langle X,u_1\rangle<x\}}1_{\{\langle Y,u_2\rangle<y\}}) - F_{\langle X,u_1\rangle}(x)G_{\langle Y,u_2\rangle}(y).$$

Directional notation

- density quantile $h_{\langle X,u\rangle} = f_{\langle X,u\rangle} \circ F^{-1}_{\langle X,u\rangle}$ and $h_{\langle Y,u\rangle} = f_{\langle Y,u\rangle} \circ G^{-1}_{\langle Y,u\rangle}$
- gradient $\nabla_u(\alpha) = (\nabla_{x,u}(\alpha), \nabla_{y,u}(\alpha))$

$$\begin{aligned}\nabla_{x,u}(\alpha) &= \frac{\partial}{\partial x} c(F^{-1}_{\langle X,u\rangle}(\alpha), G^{-1}_{\langle Y,u\rangle}(\alpha))\\ \nabla_{y,u}(\alpha) &= \frac{\partial}{\partial y} c(F^{-1}_{\langle X,u\rangle}(\alpha), G^{-1}_{\langle Y,u\rangle}(\alpha))\end{aligned}$$

Limiting covariance function

Underlying correlations

$$
\begin{aligned}
P(\alpha_1,\alpha_2,u_1,u_2) &= \mathbb{P}\left(\langle X,u_1\rangle < F^{-1}_{\langle X,u_1\rangle}(\alpha_1) \cap \langle X,u_2\rangle < F^{-1}_{\langle X,u_2\rangle}(\alpha_2)\right)\\
Q(\alpha_1,\alpha_2,u_1,u_2) &= \mathbb{P}\left(\langle Y,u_1\rangle < G^{-1}_{\langle Y,u_1\rangle}(\alpha_1) \cap \langle Y,u_2\rangle < G^{-1}_{\langle Y,u_2\rangle}(\alpha_2)\right)\\
\Pi(\alpha_1,\alpha_2,u_1,u_2) &= \mathbb{P}\left(\langle X,u_1\rangle < F^{-1}_{\langle Y,u_1\rangle}(\alpha_1) \cap \langle Y,u_2\rangle < G^{-1}_{\langle Y,u_2\rangle}(\alpha_2)\right)
\end{aligned}
$$

Limiting covariance function

Underlying correlations

$$
\begin{aligned}
P(\alpha_1, \alpha_2, u_1, u_2) &= \mathbb{P}\left(\langle X, u_1\rangle < F^{-1}_{\langle X, u_1\rangle}(\alpha_1) \cap \langle X, u_2\rangle < F^{-1}_{\langle X, u_2\rangle}(\alpha_2)\right) \\
Q(\alpha_1, \alpha_2, u_1, u_2) &= \mathbb{P}\left(\langle Y, u_1\rangle < G^{-1}_{\langle Y, u_1\rangle}(\alpha_1) \cap \langle Y, u_2\rangle < G^{-1}_{\langle Y, u_2\rangle}(\alpha_2)\right) \\
\Pi(\alpha_1, \alpha_2, u_1, u_2) &= \mathbb{P}\left(\langle X, u_1\rangle < F^{-1}_{\langle Y, u_1\rangle}(\alpha_1) \cap \langle Y, u_2\rangle < G^{-1}_{\langle Y, u_2\rangle}(\alpha_2)\right)
\end{aligned}
$$

Covariance matrix

$$
\Sigma(\alpha_1, \alpha_2, u_1, u_2) = \begin{pmatrix} \frac{P(\alpha_1,\alpha_2,u_1,u_2)-\alpha_1\alpha_2}{h_{\langle X,u_1\rangle}(\alpha_1)h_{\langle X,u_2\rangle}(\alpha_2)} & \frac{\Pi(\alpha_1,\alpha_2,u_1,u_2)-\alpha_1\alpha_2}{h_{\langle X,u_1\rangle}(\alpha_1)h_{\langle Y,u_2\rangle}(\alpha_2)} \\ \frac{\Pi(\alpha_2,\alpha_1,u_2,u_1)-\alpha_1\alpha_2}{h_{\langle Y,u_1\rangle}(\alpha_1)h_{\langle X,u_2\rangle}(\alpha_2)} & \frac{Q(\alpha_1,\alpha_2,u_1,u_2)-\alpha_1\alpha_2}{h_{\langle Y,u_1\rangle}(\alpha_1)h_{\langle Y,u_2\rangle}(\alpha_2)} \end{pmatrix}
$$

$$
cov(B(u_1), B(u_2)) = \int_0^1 \int_0^1 \nabla_{u_1}(\alpha_1)\Sigma(\alpha_1, \alpha_2, u_1, u_2)\nabla_{u_2}'(\alpha_2) d\alpha_1 \alpha_2
$$

Weak limits for $K_c(\mathbb{F}_n, \mathbb{G}_n)$ if $F = G$

Non symmetric cost

Different behavior at 0 and ∞ (Berthet, Fort spa19)

For $0 < x_0 < y_0 < +\infty$,

$$
\begin{aligned}
&(C0) \quad c(z,z') = \rho_c(z-z') \geqslant 0, \quad z,z' \in \mathbb{R}, \quad c(0,0)=0, \quad \rho_c \text{ is convex.}\\
&(C1) \quad \rho_c(z) = \rho_-(-z)1_{z\leqslant 0} + \rho_+(z)1_{z\geqslant 0}, \quad z \in \mathbb{R}, \quad \rho_\pm \in \mathcal{C}_2((0,+\infty)).\\
&(C2) \quad \rho_+(x) = x^{b_+}L_+(x) > 0, \quad 0 < x \leqslant x_0, \quad \rho_+ \in RV_2(0,b_+), \quad b_+ \geqslant 1,\\
&\qquad\quad \rho_-(x) = x^{b_-}L_-(x) > 0, \quad 0 < x \leqslant x_0, \quad \rho_- \in RV_2(0,b_-), \quad b_- \geqslant 1.\\
&(C3) \quad \rho_+(y) = \exp(l_+(y)), \quad y \geqslant y_0, \quad l_+ \in RV_2(+\infty,\gamma_+), \quad \gamma_+ \geqslant 0,\\
&\qquad\quad \rho_-(y) = \exp(l_-(y)), \quad y \geqslant y_0, \quad l_- \in RV_2(+\infty,\gamma_-), \quad \gamma_- \geqslant 0.
\end{aligned}
$$

Thus $\rho_\pm(0) = 0$ and $\rho_\pm$ are positive, continuous, convex increasing on $\mathbb{R}_+$. Let $\rho(x) = \max(\rho_+(x), \rho_-(x))$ and $b = \min(b_+, b_-)$. For $0 \leqslant x \leqslant x_0$,

$$
\rho(x) = x^b L(x), \quad L(x) = \begin{cases} L_+(x) & \text{if } b_+ < b_-,\\ L_-(x) & \text{if } b_- < b_+,\\ \max(L_+(x), L_-(x)) & \text{if } b_+ = b_-. \end{cases}
$$

Non symmetric cost

Stability at 0 :

$$(C4) \qquad \lim_{x \to 0} \frac{\rho_+(x)}{\rho(x)} = \pi_+, \qquad \lim_{x \to 0} \frac{\rho_-(x)}{\rho(x)} \to \pi_-, \qquad \pi_+, \pi_- \in [0,1] \, .$$

Example

- For $a = (a_-, a_+)$ st $a_\pm > 0$ and $b = (b_-, b_+)$ st $b_\pm \geqslant 1$,

  $$c_{a,b}(z,z') = a_- \left(z' - z\right)^{b_-} 1_{z<z'} + a_+ \left(z - z'\right)^{b_+} 1_{z'<z}$$

  satisfies (C).
- this includes Wasserstein W_p^p, $1 \leqslant p < 2$ for $a = (1,1)$ and $b = (p,p)$.
- allowed to mix Wasserstein costs W_p^p, $1 \leqslant p < 2$ at 0 and W_q^q, $p \leqslant q < +\infty$ beyond 0.

To work at 0, we consider

Conditions FG

Keep $(FG1,2,3)$, replace $(FG4)$ with cases :

Diagonal support

- On partitionne $(0,1)$ en

$$E = \left\{u : F^{-1}(u) = G^{-1}(u)\right\}, \quad D = \left\{u : F^{-1}(u) \neq G^{-1}(u)\right\}.$$

- if u moves i.o. from E to D the stochastic integral $K_c(\mathbb{F}_n, \mathbb{G}_n)$ is hard to control. Finitely many crossings :
- For $\kappa \geqslant 2$, $0 = u_0 < u_1 < ... < u_\kappa = 1$, $A_k = (u_{k-1}, u_k)$ we have

$$(FG0) \quad F^{-1}(u_k) = G^{-1}(u_k) \text{ and } A_k \subset E \text{ or } A_k \subset D,\ k = 1, ..., \kappa.$$

- Three generic cases :
  - $E = (0,1)$, fast rates
  - $D = (0,1)$, CLT
  - at least one interval is included in D with $E \neq \varnothing$.

Coupling Gaussian

Two joint Gaussian processes

$$\mathbb{B}^X(u) = \frac{B^X(u)}{h_X(u)}, \quad \mathbb{B}^Y(u) = \frac{B^Y(u)}{h_Y(u)}, \quad u \in (0,1)$$

where (B^X, B^Y) are standard bridges

$$cov(B^X(u), B^X(v)) = cov(B^Y(u), B^Y(v)) = \min(u,v) - uv, \quad u, v \in (0,1)\,,$$

with cross-covariance

$$cov(B^X(u), B^Y(v)) = H(F^{-1}(u), G^{-1}(v)) - uv, \quad u, v \in (0,1)\,.$$

The limiting behaviour is ruled by

$$\mathbb{B}(u) = \mathbb{B}^X(u) - \mathbb{B}^Y(u), \quad u \in (0,1).$$

Notice that $\mathbb{B}^Y, \mathbb{B}^Y$ indep. iif X, Y indep. (and $\mathbb{B} = 0$ iif $X = Y$).

Case $F = G$, $E = (0,1)$

Condition CFG_E case $1 < b < 2$ or $b = 1$

$$l \circ \psi^{-1}(y) \leqslant \left(1 - \frac{b}{2}\right) y + \log L\left(\exp(-y/2)\right) - 2\log\psi^{-1}(y) - \theta_2 \log y,$$

$$l \circ \psi^{-1}(y) \leqslant \frac{y}{2} - 2\log\psi^{-1}(y) - \theta_2 \log y.$$

Consequences

- (CFG_E) and $(FG3)$ imply, for $1 \leqslant b < b' < 2$

$$\int_0^1 \left(\frac{\sqrt{u(1-u)}}{h_X(u)}\right)^{b'} du \leqslant \int_0^1 \left(\frac{\left|F^{-1}(u)\right|}{\sqrt{u(1-u)}}\right)^{b'} du < +\infty$$

- if $L(x) = 1$ then (FG) and (CFG_E) imply

$$\mathbb{P}\left(X > y\right) \leqslant \left(\frac{1}{y^2 \rho(y)}\right)^{2/(2-b)}, \quad y > y_0.$$

Case $F = G$, $E = (0,1)$

Rate

$$K\sqrt{n} \leqslant v_n = \frac{1}{\rho\left(1/\sqrt{n}\right)} = \frac{n^{b/2}}{L\left(1/\sqrt{n}\right)} = o(n).$$

Theorem (case $F = G$)

Under (FG), (C), $E = (0,1)$ and (CFG_E) we have

$$v_n K_c(\mathbb{F}_n, \mathbb{G}_n) \to_{weak} \pi_- \int_0^1 1_{\{\mathbb{B}(u)<0\}} \left|\mathbb{B}(u)\right|^{b-} du + \pi_+ \int_0^1 1_{\{\mathbb{B}(u)>0\}} \left|\mathbb{B}(u)\right|^{b+} du$$

Limiting r.v. positive, finite

$$\mathbb{P}\left(\int_0^1 \left|\mathbb{B}^X(u)\right|^b du < +\infty\right) = 1 \text{ iif } \int_0^1 \left(\frac{\sqrt{u(1-u)}}{h_X(u)}\right)^b du < +\infty.$$

Case $F \neq G$, $D = (0,1)$

Condition CFG_D

(CFG) strengthened if $b_- < 2$ and $b_+ < 2$ to avoid $(FG4)$. If

$$\liminf_{u \to 1} \left| F^{-1}(u) - G^{-1}(u) \right| = 0 \quad \text{or} \quad \liminf_{u \to 0} \left| F^{-1}(u) - G^{-1}(u) \right| = 0$$

then for (l, ψ)=$(l_+, \psi_X^+), (l_-, \psi_Y^+)$ or $(l, \psi) = (l_-, \psi_X^-), (l_+, \psi_Y^-)$,

$$l \circ \psi^{-1}(y) \leqslant \frac{y}{2} - 2 \log \psi^{-1}(y) - \theta_2 \log y, \quad y > y_0.$$

Theorem (case $F \neq G$)

Under (FG), (C), $D = (0,1)$ and (CFG_D) we have

$$\sqrt{n} \left(K_c(\mathbb{F}_n, \mathbb{G}_n) - K_c(F, G) \right) \to_{weak} \mathcal{N}\left(0, \sigma^2(c, H)\right)$$

with

$$\sigma^2(c, H) = \mathbb{E}\left(\left(\int_0^1 \rho_c'(F^{-1}(u) - G^{-1}(u)) \mathbb{B}(u) du \right)^2 \right) < +\infty.$$

Mixted case

If there is a point where $F \neq G$ the rate is $\sqrt{n}$.

Condition(CFG_{ED})

- assume (CFG_D).
- if $A_1 \subset E$ then (CFG_E) for $(l, \psi) = (l_-, \psi_X^-), (l_+, \psi_X^-)$.
- if $A_\kappa \subset E$ then (CFG_E) for $(l, \psi) = (l_-, \psi_X^+), (l_+, \psi_X^+)$.

Theorem (case $1 < b < 2$)

Under (FG), (C), $D \neq \emptyset$ and (CFG_{ED}) we have

$$\sqrt{n}\,(K_c(\mathbb{F}_n, \mathbb{G}_n) - K_c(F, G)) \to_{weak} \mathcal{N}\left(0, \sigma_D^2\right)$$

with

$$\sigma_D^2 = \mathbb{E}\left(\left(\int_D \rho_c'(F^{-1}(u) - G^{-1}(u))\mathbb{B}(u)du\right)^2\right) < +\infty.$$

Mixted case, semi-W_1

We assume for

$$\rho(x) = x^b L(x), \quad L(x) = \begin{cases} L_+(x) & \text{if } b_+ < b_-, \\ L_-(x) & \text{if } b_- < b_+, \\ \max(L_+(x), L_-(x)) & \text{if } b_+ = b_-. \end{cases}$$

the existence of

$$\lim_{x \searrow 0} L(x) = L(0) \in \mathbb{R}_+.$$

Theorem (case $b = 1$)

Under (FG), (C), $D \neq \emptyset$ and (CFG_{ED}) we have

$$\begin{aligned} \sqrt{n}\,(K_c(\mathbb{F}_n, \mathbb{G}_n) - K_c(F, G)) \rightarrow_{weak} & \int_D \rho_c'(F^{-1}(u) - G^{-1}(u))\mathbb{B}(u)du \\ & + 1_{\{b_- = 1\}} L_-(0) \int_E 1_{\{\mathbb{B}(u) < 0\}} \left|\mathbb{B}(u)\right| du \\ & + 1_{\{b_+ = 1\}} L_+(0) \int_E 1_{\{\mathbb{B}(u) > 0\}} \left|\mathbb{B}(u)\right| du. \end{aligned}$$

Case W_1 asymmetric

Let $c_{a,1}(z,z') = a_- \,(z'-z)\, 1_{z<z'} + a_+ \,(z-z')\, 1_{z'<z}$.

Theorem (case $b = 1$, includes W_1)

Under (FG), (C) and (CFG_{ED}) we have

$$\sqrt{n}\left(d_{c_{a,1}}(\mathbb{F}_n,\mathbb{G}_n) - d_{c_{a,1}}(F,G)\right) \to_{weak} \int_D \left(a_- 1_{\{F(u)<G(u)\}} + a_+ 1_{\{F(u)>G(u)\}}\right)\mathbb{B}(u)d$$

$$+ \int_E \left(a_- 1_{\{\mathbb{B}(u)<0\}} + a_+ 1_{\{\mathbb{B}(u)>0\}}\right)|\mathbb{B}(u)|\,du$$

and, in particular for $a_- = a_+ = 1$,

$$\sqrt{n}\,(W_1(\mathbb{F}_n,\mathbb{G}_n) - W_1(F,G)) \to_{weak} \int_D \mathbb{B}(u)du + \int_E |\mathbb{B}(u)|\,du.$$

$$\sqrt{n}\left(\int_{-\infty}^{+\infty} |\mathbb{F}_n(t) - \mathbb{G}_n(t)|\,dt - \int_{F^{-1}(D)} |F(t)-G(t)|\,dt\right)$$

$$\to_{weak} \int_D \mathbb{B}(u)du + \int_E |\mathbb{B}(u)|\,du$$

Theorem (breaking case $b = 2$, includes W_2)

Assume $E = (0,1)$, $(FG1)$, $(FG2)$ and

$$\lim_{u\to 0} \frac{u}{h(u)} = \lim_{u\to 1} \frac{1-u}{h(u)} = 0, \quad \int_0^1 \frac{u(1-u)}{h^2(u)} du < +\infty.$$

If moreover $(C0)$ with $\rho_c(x) = x^2$ for $|x| \leqslant x_0$. Then

$$nK_c(\mathbb{F}_n, \mathbb{G}_n) \to_{weak} \int_0^1 \mathbb{B}(u)^2 du.$$

Remarque

- Gaussians excluded, from a few.
- Weibull $w > 0$, $h(u) = w(1-u)\left(\log(1/(1-u))\right)^{1-1/w}$ and

$$(1-u)/h^2(u) = 1/w\left((1-u)\log(1/(1-u))\right)^{2(1-1/w)}$$

thus conditions require $w > 1$ and $w > 2$ resp.

Case with one sample

Corollaries

Adapting the proof of these theorems provides similar results for $\sqrt{n}\,(K_c(\mathbb{F}_n, G) - K_c(F, G))$ and $v_n K_c(\mathbb{F}_n, F)$ by replacing $\mathbb{B}(u) = \mathbb{B}^X(u) - \mathbb{B}^Y(u)$ with $\mathbb{B}^X(u) = B^X(u)/h_X(u)$.

Case W_p for $1 \leqslant p < 2$

If F satisfies (FG) with tails lighter than Pareto of index $2(p+2)/(2-p)$ we have

$$n^{p/2} W_p^p(\mathbb{F}_n, F) \to_{weak} \int_0^1 \left|\mathbb{B}^X(u)\right|^p du$$

and the limiting r.v. is positive and finite.

Thanks for your attention!

Any questions ?