Central and non-central limit theorems for statistical functionals based on weakly and strongly dependent data

Eric Beutner, Maastricht University

Tokyo, September 3

Overview

Motivation

Quasi-Hadamard differentiability

Applications

Continuous mapping approach to U- and V-statistics

Motivation

Statistical functionals

- Let $\theta = T(F)$ be some characteristic of the distribution function F (df).

- Let $\theta = T(F)$ be some characteristic of the distribution function F (df).
- Typical examples are
  - $T(F) = -\int_{-\infty}^{0} K(F(x))\,dx + \int_{0}^{\infty}(1 - K(F(x))\,dx$ so called L-statistics. Distortion risk measures which are quite popular are also of this form.

- Let $\theta = T(F)$ be some characteristic of the distribution function F (df).
- Typical examples are
  - $T(F) = -\int_{-\infty}^{0} K(F(x))\,dx + \int_0^{\infty}(1 - K(F(x))\,dx$ so called L-statistics. Distortion risk measures which are quite popular are also of this form.
  - $T(F) = \int\int g(x_1, x_2)\,dF(x_1)dF(x_2)$ so called U- or V-statistic (of degree 2).

Statistical functionals

- Let $\theta = T(F)$ be some characteristic of the distribution function F (df).
- Typical examples are
  - $T(F) = -\int_{-\infty}^{0} K(F(x))\,dx + \int_0^\infty (1 - K(F(x))\,dx$ so called L-statistics. Distortion risk measures which are quite popular are also of this form.
  - $T(F) = \int\int g(x_1, x_2)\,dF(x_1)dF(x_2)$ so called U- or V-statistic (of degree 2).
  - Z-estimators.

- Let $\theta = T(F)$ be some characteristic of the distribution function F (df).

- Typical examples are
  - $T(F) = -\int_{-\infty}^{0} K(F(x))\,dx + \int_0^\infty (1 - K(F(x))\,dx$ so called L-statistics. Distortion risk measures which are quite popular are also of this form.
  - $T(F) = \int\int g(x_1, x_2)\,dF(x_1)dF(x_2)$ so called U- or V-statistic (of degree 2).
  - Z-estimators.

- Given n observations $X_1, \dots, X_n$ with df F a natural estimator is then $T(F_n)$ with F_n the empirical distribution function.

- Well-known: If T is Hadamard differentiable at F, then the asymptotic distribution of $T(F_n)$ follows immediately by the (functional) delta method from the asymptotic distribution of $F_n - F$.

Functional delta method

- Well-known: If T is Hadamard differentiable at F, then the asymptotic distribution of $T(F_n)$ follows immediately by the (functional) delta method from the asymptotic distribution of $F_n - F$.

- Thus, delta method leads asymptotic distribution of $T(F_n) - T(F)$ whenever we have weak convergence of the empirical process.

- Well-known: If T is Hadamard differentiable at F, then the asymptotic distribution of $T(F_n)$ follows immediately by the (functional) delta method from the asymptotic distribution of $F_n - F$.
- Thus, delta method leads asymptotic distribution of $T(F_n) - T(F)$ whenever we have weak convergence of the empirical process.
- Many results on weak convergence of the empirical process (iid, short-memory like α-mixing or β-mixing, long memory, etc.).

Quasi-Hadamard differentiability

Illustrative example: sample mean

- If K corresponds to Lebesgue measure on $[0,1]$, then

$$-\int_{-\infty}^{0} K(F_n(x))\,dx + \int_0^{\infty} (1 - K(F_n(x))\,dx,$$

corresponds to the sample mean.

Illustrative example: sample mean

- If K corresponds to Lebesgue measure on $[0,1]$, then

$$-\int_{-\infty}^{0} K(F_n(x))\,dx + \int_{0}^{\infty} (1 - K(F_n(x))\,dx,$$

corresponds to the sample mean.

- Proving Hadamard differentiability of this L-statistic (at F) in the direction of V boils down to prove that

$$\left| \int [V_n(x) - V(x)]\,dx \right| \to 0,$$

whenever $||V_n - V||_\infty \to 0$. $||\cdot||_\infty$ denotes sup-norm.

Illustrative example: sample mean

- If K corresponds to Lebesgue measure on $[0,1]$, then

$$-\int_{-\infty}^{0} K(F_n(x))\,dx + \int_{0}^{\infty} (1 - K(F_n(x))\,dx,$$

  corresponds to the sample mean.
- Proving Hadamard differentiability of this L-statistic (at F) in the direction of V boils down to prove that

$$\left| \int [V_n(x) - V(x)]\,dx \right| \to 0,$$

  whenever $||V_n - V||_\infty \to 0$. $||\cdot||_\infty$ denotes sup-norm.
- $\Rightarrow$ The simplest L-statistic the sample mean is not Hadamard differentiable w.r.t. the sup-norm.

Illustrative example: V-statistic

- Proving Hadamard differentiability of a V-statistic (at F) in the direction of V, involves among other things showing that $||V_n - V||_\infty \to 0$ implies

$$\int V_n(x_2)\,|dg_F|(x_2) \to \int V(x_2)\,|dg_F|(x_2).$$

where $|dg_F|$ is the absolute measure generated by $g_F(x_2) = \int g(x_1, x_2) dF(x_1)$.

- Proving Hadamard differentiability of a V-statistic (at F) in the direction of V, involves among other things showing that $||V_n - V||_\infty \to 0$ implies

$$\int V_n(x_2)\,|dg_F|(x_2) \to \int V(x_2)\,|dg_F|(x_2).$$

  where $|dg_F|$ is the absolute measure generated by $g_F(x_2) = \int g(x_1, x_2) dF(x_1)$.

- If g_F generates a finite (signed) measure, then $||V_n - V||_\infty$ this implication indeed holds.

- Proving Hadamard differentiability of a V-statistic (at F) in the direction of V, involves among other things showing that $||V_n - V||_\infty \to 0$ implies

$$\int V_n(x_2)\,|dg_F|(x_2) \to \int V(x_2)\,|dg_F|(x_2).$$

where $|dg_F|$ is the absolute measure generated by $g_F(x_2) = \int g(x_1, x_2) dF(x_1)$.

- If g_F generates a finite (signed) measure, then $||V_n - V||_\infty$ this implication indeed holds.

- For $g(x_1, x_2) = (1/2)(x_1 - x_2)^2$ (the variance kernel) the measure dg_F has density $(x_2 - c)\,dx_2$.

Illustrative example: V-statistic

- Proving Hadamard differentiability of a V-statistic (at F) in the direction of V, involves among other things showing that $||V_n - V||_\infty \to 0$ implies

$$\int V_n(x_2)\,|dg_F|(x_2) \to \int V(x_2)\,|dg_F|(x_2).$$

where $|dg_F|$ is the absolute measure generated by $g_F(x_2) = \int g(x_1,x_2)dF(x_1)$.

- If g_F generates a finite (signed) measure, then $||V_n - V||_\infty$ this implication indeed holds.

- For $g(x_1,x_2) = (1/2)(x_1 - x_2)^2$ (the variance kernel) the measure dg_F has density $(x_2 - c)\,dx_2$.
$\Rightarrow$ The simplest V-statistic the variance is not Hadamard differentiable w.r.t. the sup-norm.

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1 + |x|)^\lambda$, $\lambda > 0$, converges uniformly to zero,

Way out & Problems

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1 + |x|)^\lambda$, $\lambda > 0$, converges uniformly to zero, then we only need $\int (1 + |x|)^{-\lambda} \, |dg_F|(x) < \infty$.

Way out & Problems

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1+|x|)^\lambda$, $\lambda > 0$, converges uniformly to zero, then we only need $\int (1+|x|)^{-\lambda} \, |dg_F|(x) < \infty$.
- Similar, for the sample mean.

Way out & Problems

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1 + |x|)^\lambda$, $\lambda > 0$, converges uniformly to zero, then we only need $\int (1 + |x|)^{-\lambda} \, |dg_F|(x) < \infty$.

- Similar, for the sample mean.

- However, Hadamard differentiability is defined as

  "Let B_1 and B_2 be normed spaces. Then $\phi : B_1 \to B_2$ is Hadamard differentiable at $b_1 \in B_1$ if ..."

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1 + |x|)^\lambda$, $\lambda > 0$, converges uniformly to zero, then we only need $\int (1 + |x|)^{-\lambda} \, |dg_F|(x) < \infty$.
- Similar, for the sample mean.
- However, Hadamard differentiability is defined as

  "Let B_1 and B_2 be normed spaces. Then $\phi : B_1 \to B_2$ is Hadamard differentiable at $b_1 \in B_1$ if ..."
- But for an arbitrary df F we have $\|F\|_\lambda := \|F(x)(1 + |x|)^\lambda\| = \infty$.

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1+|x|)^\lambda$, $\lambda > 0$, converges uniformly to zero, then we only need $\int (1+|x|)^{-\lambda} \, |dg_F|(x) < \infty$.
- Similar, for the sample mean.
- However, Hadamard differentiability is defined as

  "Let B_1 and B_2 be normed spaces. Then $\phi : B_1 \to B_2$ is Hadamard differentiable at $b_1 \in B_1$ if ..."
- But for an arbitrary df F we have $\|F\|_\lambda := \|F(x)(1+|x|)^\lambda\| = \infty$.
- Hence, with a weighted sup-norm $\|\cdot\|_\lambda$ Hadamard differentiability at F cannot be shown.

- If we require not only that $||V_n - V||_\infty \to 0$ but that $(V_n(x) - V(x))(1 + |x|)^\lambda$, $\lambda > 0$, converges uniformly to zero, then we only need $\int (1 + |x|)^{-\lambda} |dg_F|(x) < \infty$.
- Similar, for the sample mean.
- However, Hadamard differentiability is defined as

  "Let B_1 and B_2 be normed spaces. Then $\phi : B_1 \to B_2$ is Hadamard differentiable at $b_1 \in B_1$ if ..."
- But for an arbitrary df F we have $\|F\|_\lambda := \|F(x)(1 + |x|)^\lambda\| = \infty$.
- Hence, with a weighted sup-norm $\| \cdot \|_\lambda$ Hadamard differentiability at F cannot be shown.
- $\Rightarrow$: (Functional) Delta method cannot be applied.

- **Definition:** *(Quasi-Hadamard differentiability)*
  Let $\mathbf{V}$ be a vector space, and $\mathbf{V}_0 \subset \mathbf{V}$ be equipped with a norm $\|\cdot\|_{\mathbf{V}_0}$. Let $(\mathbf{V}', \|\cdot\|_{\mathbf{V}'})$ be a normed vector space, and $T : \mathbf{V}_T \to \mathbf{V}', \mathbf{V}_T \subset \mathbf{V}$.

Quasi-Hadamard differentiability

- **Definition:** *(Quasi-Hadamard differentiability)*
  Let $\mathbf{V}$ be a vector space, and $\mathbf{V}_0 \subset \mathbf{V}$ be equipped with a norm $\|\cdot\|_{\mathbf{V}_0}$. Let $(\mathbf{V}', \|\cdot\|_{\mathbf{V}'})$ be a normed vector space, and $T : \mathbf{V}_T \to \mathbf{V}'$, $\mathbf{V}_T \subset \mathbf{V}$.

  Then T is said to be quasi-Hadamard differentiable at $\theta \in \mathbf{V}_T$ tangentially to $\mathbb{C}_0$, $\mathbb{C}_0 \subset \mathbf{V}_0$, if for a continuous map $D^{\mathrm{Had}}_{\theta;T} : \mathbb{C}_0 \to \mathbf{V}'$

$$\lim_{n\to\infty} \left\| D^{\mathrm{Had}}_{\theta;T}(v) - \frac{T(\theta + h_n v_n) - T(\theta)}{h_n} \right\|_{\mathbf{V}'} = 0$$

  holds for each triplet $(v, (v_n), (h_n))$ with $h_n \to 0$, and

  $v \in \mathbb{C}_0$, $(v_n) \subset \mathbf{V}_0$ satisfying $\|v_n - v\|_{\mathbf{V}_0} \to 0$.

- **Definition:** *(Quasi-Hadamard differentiability)*
  Let $\mathbf{V}$ be a vector space, and $\mathbf{V}_0 \subset \mathbf{V}$ be equipped with a norm $\|\cdot\|_{\mathbf{V}_0}$. Let $(\mathbf{V}', \|\cdot\|_{\mathbf{V}'})$ be a normed vector space, and $T : \mathbf{V}_T \to \mathbf{V}', \mathbf{V}_T \subset \mathbf{V}$.

  Then T is said to be quasi-Hadamard differentiable at $\theta \in \mathbf{V}_T$ tangentially to $\mathbb{C}_0$, $\mathbb{C}_0 \subset \mathbf{V}_0$, if for a continuous map $D^{\mathrm{Had}}_{\theta;T} : \mathbb{C}_0 \to \mathbf{V}'$

$$\lim_{n\to\infty} \left\| D^{\mathrm{Had}}_{\theta;T}(v) - \frac{T(\theta + h_n v_n) - T(\theta)}{h_n} \right\|_{\mathbf{V}'} = 0$$

  holds for each triplet $(v, (v_n), (h_n))$ with $h_n \to 0$, and

$$v \in \mathbb{C}_0, \ (v_n) \subset \mathbf{V}_0 \text{ satisfying } \|v_n - v\|_{\mathbf{V}_0} \to 0.$$

- With this definition we find

Quasi-Hadamard differentiability (cont.)

- **Theorem:** The distortion risk measure (L-statistic) $\int x\,dK(F(x))$ is quasi-Hadamard differentiable if *K is continuous and piecewise differentiable, and K' is bounded above by some constant $M > 0$.*

Quasi-Hadamard differentiability (cont.)

- **Theorem:** The distortion risk measure (L-statistic) $\int x\,dK(F(x))$ is quasi-Hadamard differentiable if *K is continuous and piecewise differentiable, and K' is bounded above by some constant $M > 0$.*

- **Theorem:** The V-statistic $\int\int g(x_1, x_2)\,dF(x_1)dF(x_2)$ is quasi-Hadamard differentiable if for some $\lambda > \lambda' \geq 0$

  (a) For every $x_2 \in \mathbb{R}$ fixed, the function $g_{x_2}(\cdot) := g(\,\cdot\,, x_2)$ lies in $\mathbb{BV}_{\text{loc,rc}}$ and $g_{x_2}(x_1)(1 + |x_1|)^{-\lambda'}$ is uniformly bounded.

  (b) The function $g_F(\cdot) := \int g(x_1,\,\cdot\,)dF(x_1)$ lies in $\mathbb{BV}_{\text{loc,rc}}$, and $\int \phi_{-\lambda}(x)\,|dg_F|(x) < \infty$.

Quasi-Hadamard differentiability (cont.)

- **Theorem:** The distortion risk measure (L-statistic) $\int x\,dK(F(x))$ is quasi-Hadamard differentiable if *K is continuous and piecewise differentiable, and K' is bounded above by some constant $M > 0$.*

- **Theorem:** The V-statistic $\int\int g(x_1, x_2)\,dF(x_1)dF(x_2)$ is quasi-Hadamard differentiable if for some $\lambda > \lambda' \geq 0$

  (a) For every $x_2 \in \mathbb{R}$ fixed, the function $g_{x_2}(\cdot) := g(\,\cdot\,, x_2)$ lies in $\mathbb{BV}_{\text{loc,rc}}$ and $g_{x_2}(x_1)(1 + |x_1|)^{-\lambda'}$ is uniformly bounded.

  (b) The function $g_F(\cdot) := \int g(x_1,\,\cdot\,)dF(x_1)$ lies in $\mathbb{BV}_{\text{loc,rc}}$, and $\int \phi_{-\lambda}(x)\,|dg_F|(x) < \infty$.

- Notice sample mean and variance (with $\lambda' = 2$) are quasi-Hadamard differentiable. However, results might be completely useless.

- Fortunately, they are not. Because

- Fortunately, they are not. Because

- **Theorem:** *(Modified functional delta method)*
  Let $T, \theta, \mathbf{V}, \mathbf{V}_f, \mathbf{V}_0, \mathbb{C}_0$ be as above. If:

  (i) T is quasi-Hadamard differentiable at θ tangentially to $\mathbb{C}_0$ with quasi-Hadamard derivative $D^{\mathrm{Had}}_{\theta;T}$,

  (ii) $X_n - \theta$ takes values only in $\mathbf{V}_0$ and satisfies

$$a_n(X_n - \theta) \stackrel{d}{\to} V \qquad \text{(in } (\mathbf{V}_0, \mathcal{V}_0, \|\cdot\|_{\mathbf{V}_0})\text{)},$$

  V a random element of $(\mathbf{V}_0, \mathcal{V}_0)$ taking values only in $\mathbb{C}_0$.

  Then

$$a_n(T(X_n) - T(\theta)) \stackrel{d}{\to} D^{\mathrm{Had}}_{\theta;T}(V) \qquad \text{(in } (\mathbf{V}', \mathcal{V}', \|\cdot\|_{\mathbf{V}'})\text{)}.$$

Applications

Weakly dependent data

- The quasi-Hadamard derivative of a U-statistic is given by

$$\dot{U}_F(B^\circ) := -2 \int B^\circ(x) dg_F(x).$$

Weakly dependent data

- The quasi-Hadamard derivative of a U-statistic is given by

$$\dot{U}_F(B^\circ) := -2\int B^\circ(x)\,dg_F(x).$$

- Let (X_i) be α-mixing with $\alpha(n) = \mathcal{O}(n^{-\theta})$ for some $\theta > 1+\sqrt{2}$. If F has finite γ-moment for some $\gamma > \frac{2\theta\lambda}{\theta-1}$, then (Shao and Yu (1996)) with $\mathbb{D}_\lambda$ càdlàg functions with finite weighted sup-norm

$$\sqrt{n}(F_n - F) \xrightarrow{d} \widetilde{B}_F^\circ \qquad \text{(in } (\mathbb{D}_\lambda, \mathcal{D}_\lambda, \|\cdot\|_\lambda)).$$

- The quasi-Hadamard derivative of a U-statistic is given by

$$\dot{U}_F(B^\circ) := -2 \int B^\circ(x) dg_F(x).$$

- Let (X_i) be α-mixing with $\alpha(n) = \mathcal{O}(n^{-\theta})$ for some $\theta > 1 + \sqrt{2}$. If F has finite γ-moment for some $\gamma > \frac{2\theta\lambda}{\theta-1}$, then (Shao and Yu (1996)) with $\mathbb{D}_\lambda$ càdlàg functions with finite weighted sup-norm

$$\sqrt{n}(F_n - F) \xrightarrow{d} \widetilde{B}_F^\circ \qquad \text{(in } (\mathbb{D}_\lambda, \mathcal{D}_\lambda, \|\cdot\|_\lambda)).$$

- Asymptotic distribution of $\sqrt{n}(U(F_n) - U(F))$ follows then for every df with finite γ-moment for some $\gamma > \frac{2\theta\lambda}{\theta-1}$.

Weakly dependent data

- The quasi-Hadamard derivative of a U-statistic is given by

$$\dot{U}_F(B^\circ) := -2\int B^\circ(x)dg_F(x).$$

- Let (X_i) be α-mixing with $\alpha(n) = \mathcal{O}(n^{-\theta})$ for some $\theta > 1+\sqrt{2}$. If F has finite γ-moment for some $\gamma > \frac{2\theta\lambda}{\theta-1}$, then (Shao and Yu (1996)) with $\mathbb{D}_\lambda$ càdlàg functions with finite weighted sup-norm

$$\sqrt{n}(F_n - F) \stackrel{d}{\to} \widetilde{B}_F^\circ \qquad (\text{in } (\mathbb{D}_\lambda, \mathcal{D}_\lambda, \|\cdot\|_\lambda)).$$

- Asymptotic distribution of $\sqrt{n}(U(F_n) - U(F))$ follows then for every df with finite γ-moment for some $\gamma > \frac{2\theta\lambda}{\theta-1}$.

- For the variance the assumptions are weaker than in Dehling and Wendler (2010) whenever $\gamma < \frac{7+8\sqrt{2}}{2\sqrt{2}-1}$.

Strongly dependent data

- Consider the process $X_t := \sum_{s=0}^{\infty} a_s \, \varepsilon_{t-s}, t \in \mathbb{N}_0,$, where $(\varepsilon_i)_{i\in\mathbb{Z}}$ iid random variables with zero mean and finite variance, and $\sum_{s=0}^{\infty} a_s^2 < \infty$.

Strongly dependent data

- Consider the process $X_t := \sum_{s=0}^{\infty} a_s \, \varepsilon_{t-s}, t \in \mathbb{N}_0$,, where $(\varepsilon_i)_{i \in \mathbb{Z}}$ iid random variables with zero mean and finite variance, and $\sum_{s=0}^{\infty} a_s^2 < \infty$.
- If $\mathbb{C}\mathrm{ov}(X_0, X_m) = m^{1-2\beta}$, $\beta \in (0.5, 1)$, then $\sum_{m=1}^{\infty} \mathbb{C}\mathrm{ov}(X_0, X_m)$ is not absolute summable, and the process (X_t) is called a long-memory process.

- Consider the process $X_t := \sum_{s=0}^{\infty} a_s \, \varepsilon_{t-s}, t \in \mathbb{N}_0$,, where $(\varepsilon_i)_{i \in \mathbb{Z}}$ iid random variables with zero mean and finite variance, and $\sum_{s=0}^{\infty} a_s^2 < \infty$.
- If $\mathbb{C}\mathrm{ov}(X_0, X_m) = m^{1-2\beta}$, $\beta \in (0.5, 1)$, then $\sum_{m=1}^{\infty} \mathbb{C}\mathrm{ov}(X_0, X_m)$ is not absolute summable, and the process (X_t) is called a long-memory process.
- No general result for L- and V-statistics of such processes.

Strongly dependent data (cont.)

- Because quasi-Hadamard differentiability already established, to apply the Modified Functional Delta Method, we only have

Strongly dependent data (cont.)

- Because quasi-Hadamard differentiability already established, to apply the Modified Functional Delta Method, we only have to prove weak convergence of weighted empirical processes based on long-memory sequences.

Strongly dependent data (cont.)

- Because quasi-Hadamard differentiability already established, to apply the Modified Functional Delta Method, we only have to prove weak convergence of weighted empirical processes based on long-memory sequences.

- **Theorem** Let $\lambda \geq 0$, $\beta \in (0.5, 1)$, and assume that

  $\mathbb{E}[|\varepsilon_0|^{2+2\lambda}] < \infty$, *the df* G *of* ε_0 *is twice differentiable, and* $\sum_{j=1}^{2} \int |G^{(j)}(x)|^2 (1 + |x|^{2\lambda})\, dx < \infty$.

  Then

  $$n^{\beta-1/2}\big(F_n(\cdot) - F(\cdot)\big) \stackrel{\mathrm{d}}{\longrightarrow} -c_{1,\beta}\, f(\cdot) Z \qquad (\text{in } \mathbb{D}_\lambda),$$

  where f is the density of X_0 and Z is normally distributed with mean 0 and variance 1.

Continuous mapping approach to U- and V-statistics

- For the variance & long memory the above approach leads that the asymptotic distribution of the sample variance multiplied by the rate of the empirical process equals

$$-2\int B^\circ(x-)\,dg_F(x) = 2Z_{1,\beta}\int f(x-)\,(x-\mathbb{E}[X_1])\,dx = 0$$

- For the variance & long memory the above approach leads that the asymptotic distribution of the sample variance multiplied by the rate of the empirical process equals

$$-2\int B^\circ(x-)\,dg_F(x) = 2Z_{1,\beta}\int f(x-)\,(x-\mathbb{E}[X_1])\,dx = 0$$

- In particular for long memory the follwong representation for U-statistics turns out to be useful

$$\begin{aligned} a_n\big(V_g(F_n) - V_g(F)\big) &= 2\Phi_{1,g}\big(a_n(F_n - F)\big) \\ &\quad + \Phi_{2,g}\big(\sqrt{a_n}(F_n - F)\big), \end{aligned}$$

where $\Phi_{1,g}(f) := -\int f(x-)\,dg_F(x)$ and $\Phi_{2,g}(f) := \iint f(x_1-)f(x_2-)\,dg(x_1,x_2)$ are continuous mappings for appropriate weigthed sup-norms.

Expansion

- For the following expansion (with $p \geq 1$)

$$F_n(\cdot) - F(\cdot) - \sum_{j=1}^{p} (-1)^j F^{(j)}(\cdot) \left(\frac{1}{n} \sum_{i=1}^{n} A_{j;F}(X_i) \right),$$

where $A_{j;F}$ denotes the jth order Appell polynomial associated with F and $F^{(j)}$ is the jth derivative of F, weak convergence at the rate $n^{p(\beta-1/2)}$ to $(-1)^p F^{(p)}(\cdot) Z_{p,\beta}$ in a weighted sup-norm can be shown.

- For the following expansion (with $p \geq 1$)

$$F_n(\cdot) - F(\cdot) - \sum_{j=1}^{p} (-1)^j F^{(j)}(\cdot) \left(\frac{1}{n} \sum_{i=1}^{n} A_{j;F}(X_i) \right),$$

where $A_{j;F}$ denotes the jth order Appell polynomial associated with F and $F^{(j)}$ is the jth derivative of F, weak convergence at the rate $n^{p(\beta-1/2)}$ to $(-1)^p F^{(p)}(\cdot) Z_{p,\beta}$ in a weighted sup-norm can be shown.

- Then we can introduce the following statistic

$$
\begin{aligned}
\mathcal{V}_{n,g;p,q,r}(F_n) &:= V_g(F_n) - V_g(F) \\
&+ \sum_{\ell=1}^{2}\sum_{j=1}^{p-1}(-1)^j\Big(\frac{1}{n}\sum_{i=1}^{n}A_{j;F}(X_i)\Big)\int F^{(j)}(x-)\,dg_{\ell,F}(x) \\
&- \sum_{j=1}^{q-1}(-1)^j\Big(\frac{1}{n}\sum_{i=1}^{n}A_{j;F}(X_i)\Big) \\
&\qquad\times\iint F^{(j)}(x_1-)\,(F_n(x_2-)-F(x_2-))\,dg(x_1,x_2) \\
&- \sum_{k=1}^{r-1}(-1)^k\Big(\frac{1}{n}\sum_{i=1}^{n}A_{k;F}(X_i)\Big) \\
&\qquad\times\iint (F_n(x_1-)-F(x_1-))\,F^{(k)}(x_2-)\,dg(x_1,x_2) \\
&+ \sum_{j=1}^{q-1}\sum_{k=1}^{r-1}(-1)^{j+k}\Big(\frac{1}{n}\sum_{i=1}^{n}A_{j;F}(X_i)\Big)\Big(\frac{1}{n}\sum_{i=1}^{n}A_{k;F}(X_i)\Big) \\
&\qquad\times\iint F^{(j)}(x_1-)\,F^{(k)}(x_2-)\,dg(x_1,x_2).
\end{aligned}
$$

Using the continuous mapping approach and the above result:

- (i) Assume $q + r > p$, then $n^{p(\beta-1/2)}\,\mathcal{V}_{n,g;p,q,r}(F_n)$ converges in distribution to

$$(-1)^p\, Z_{p,\beta} \sum_{\ell=1}^{2} \int F^{(p)}(x-)\, dg_{\ell,F}(x).$$

Using the continuous mapping approach and the above result:

- (i) Assume $q + r > p$, then $n^{p(\beta-1/2)}\,\mathcal{V}_{n,g;p,q,r}(F_n)$ converges in distribution to

$$(-1)^p\, Z_{p,\beta} \sum_{\ell=1}^{2} \int F^{(p)}(x-)\, dg_{\ell,F}(x).$$

- (ii) Assume $q + r = p$, then $n^{p(\beta-1/2)}\,\mathcal{V}_{n,g;p,q,r}(F_n)$ converges in distribution to

$$(-1)^p\, Z_{p,\beta} \sum_{\ell=1}^{2} \int F^{(p)}(x-)\, dg_{\ell,F}(x) +$$
$$(-1)^p\, Z_{q,\beta} Z_{r,\beta} \iint F^{(q)}(x_1-) F^{(r)}(x_2-)\, dg(x_1, x_2).$$

- Consider kernel $g(x_1, x_2) = x_1(|x_2| - 1)$, and suppose that $F^{(1)}$ is symmetric about zero and that $\mathbb{E}[|X_1|] = 1$.

- Consider kernel $g(x_1, x_2) = x_1(|x_2| - 1)$, and suppose that $F^{(1)}$ is symmetric about zero and that $\mathbb{E}[|X_1|] = 1$.
- Taking $n^{2(\beta-(1/2))}$ leads to:

$$n^{2\beta-1}\,\mathcal{V}_{n,g;2,1,1}(F_n) = n^{2\beta-1}\left(V_g(F_n) - V_g(F)\right)$$
$$\stackrel{\mathsf{d}}{\longrightarrow} Z_{1,\beta}^2 \iint F^{(1)}(x_1-)F^{(1)}(x_2-)\,dg(x_1,x_2) = 0.$$

- Consider kernel $g(x_1, x_2) = x_1(|x_2| - 1)$, and suppose that $F^{(1)}$ is symmetric about zero and that $\mathbb{E}[|X_1|] = 1$.
- Taking $n^{2(\beta-(1/2))}$ leads to:

$$n^{2\beta-1}\,\mathcal{V}_{n,g;2,1,1}(F_n) = n^{2\beta-1}\left(V_g(F_n) - V_g(F)\right)$$
$$\xrightarrow{\mathsf{d}} Z_{1,\beta}^2 \iint F^{(1)}(x_1-)F^{(1)}(x_2-)\,dg(x_1,x_2) = 0.$$

- However, with $n^{3(\beta-(1/2))}$ we have

$$n^{3(\beta-1/2)}\,\mathcal{V}_{n,g;3,1,2}(F_n) = n^{3(\beta-1/2)}\left(V_g(F_n) - V_g(F)\right)$$
$$\xrightarrow{\mathsf{d}} -Z_{1,\beta}Z_{2,\beta} \iint F^{(1)}(x_1-)F^{(2)}(x_2-)\,dg(x_1,x_2)$$
$$= -2\,Z_{1,\beta}Z_{2,\beta}\int_0^\infty F^{(2)}(x_2)\,dx_2.$$

- Consider the test statistic

$$T_n := \int_0^\infty \Big(\hat{F}_n(-t) - \big[1 - \hat{F}_n(t-)\big]\Big)^2 \, dt.$$

- Consider the test statistic

$$T_n := \int_0^\infty \Big(\hat{F}_n(-t) - \big[1 - \hat{F}_n(t-)\big]\Big)^2 \, dt.$$

- Taking $n^{3(\beta-1/2)}$ leads to:

$$n^{3(\beta-1/2)}\, \mathcal{V}_{n,g;3,1,2}(F_n)$$
$$\xrightarrow{\mathsf{d}} Z_{1,\beta} Z_{2,\beta} \left(\int F^{(1)}(x) F^{(2)}(x) - F^{(1)}(x) F^{(2)}(-x)\, dx \right)$$
$$= 0.$$

- Consider the test statistic
$$T_n := \int_0^\infty \Big(\hat{F}_n(-t) - \big[1 - \hat{F}_n(t-)\big]\Big)^2 \, dt.$$
- Taking $n^{3(\beta-1/2)}$ leads to:

$$n^{3(\beta-1/2)}\,\mathcal{V}_{n,g;3,1,2}(F_n)$$
$$\xrightarrow{\mathsf{d}} Z_{1,\beta} Z_{2,\beta} \left(\int F^{(1)}(x) F^{(2)}(x) - F^{(1)}(x) F^{(2)}(-x) \, dx \right)$$
$$= 0.$$

- However, with $n^{4(\beta-1/2)}$ we find

$$n^{4(\beta-1/2)}\,\mathcal{V}_{n,g;4,2,2}(F_n)$$
$$\xrightarrow{\mathsf{d}} Z_{2,\beta} Z_{2,\beta} \Big(\int F^{(2)}(x) F^{(2)}(x) - F^{(2)}(x) F^{(2)}(-x) \, dx \Big)$$

- Beutner, E. and Zähle, H. (2010). A modified functional delta method and its application to the estimation of risk functionals. *J. Multivariate Anal.* 101, 2452–2463.
- Beutner, E. and Zähle, H. (2012). Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes, *Bernoulli* 18, 803–822.
- Beutner, E., Wu, W.B. and Zähle, H. (2012). Asymptotics for statistical functionals of long-memory sequences. *Stochastic Process. Appl.* 122, 910–929.
- Beutner, E. and Zähle, H. Continuous mapping approach to the asymptotics of U- and V-statistics, *Bernoulli*, to appear.

- Dehling, H. and Wendler, M. (2010). Central limit theorem and the bootstrap for U-statistics of strongly mixing data. *Journal of Multivariate Analysis*, 101(1), 126–137.
- Sen, P.K. (1996). Statistical functionals, Hadamard differentiability and martingales. In *A Festschrift for J. Medhi* (Eds. Borthakur, A.C, and Chaudhury, H.), New Age Press, Delhi, 29–47.
- Shao, Q.-M. and Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* 24, 2098–2127.
- Wu, W.B. (2003). Empirical processes of long-memory sequences. *Bernoulli*, 9, 809–831.