Donsker results for the smoothed empirical process of dependent observations

Eric Beutner, Maastricht University

Joint work with Henryk Zähle

Besancon, May 13

Overview

Introduction

Extending the benchmark

Free lunch? Yes free lunch!

One more free lunch? No (too much asked).

Introduction

Introduction

- Let $Z_1, \ldots, Z_n$ be identically distributed with cdf F_0 and density f_0 and consider the kernel based estimator for f_0

$$\hat{f}_n(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_n} K\left(\frac{z - Z_i}{h_n}\right)$$

Introduction

- Let $Z_1, \ldots, Z_n$ be identically distributed with cdf F_0 and density f_0 and consider the kernel based estimator for f_0

$$\hat{f}_n(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_n} K\left(\frac{z - Z_i}{h_n}\right) = \int \frac{1}{h_n} K\left(\frac{z - y}{h_n}\right) \, d\widehat{F}_n(y),$$

  where $\widehat{F}_n$ is the empirical distribution function based on $Z_1, \ldots, Z_n$ and h_n is the bandwidth.

Introduction

- Let $Z_1, \ldots, Z_n$ be identically distributed with cdf F_0 and density f_0 and consider the kernel based estimator for f_0

$$\hat{f}_n(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_n} K\left(\frac{z - Z_i}{h_n}\right) = \int \frac{1}{h_n} K\left(\frac{z-y}{h_n}\right)\, d\widehat{F}_n(y),$$

  where $\widehat{F}_n$ is the empirical distribution function based on $Z_1, \ldots, Z_n$ and h_n is the bandwidth.
- Let the kernel K be symmetric and for some $C > 0$

$$\int K(y)\, dy = 1, \int yK(y)\, dy = 0, \text{ and } \int y^2 K(y)\, dy = C.$$

Introduction

- Let $Z_1, \dots, Z_n$ be identically distributed with cdf F_0 and density f_0 and consider the kernel based estimator for f_0

$$\hat{f}_n(z) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_n} K\left(\frac{z - Z_i}{h_n}\right) = \int \frac{1}{h_n} K\left(\frac{z - y}{h_n}\right)\, d\widehat{F}_n(y),$$

  where $\widehat{F}_n$ is the empirical distribution function based on $Z_1, \dots, Z_n$ and h_n is the bandwidth.

- Let the kernel K be symmetric and for some $C > 0$

$$\int K(y)\, dy = 1, \int yK(y)\, dy = 0, \text{ and } \int y^2 K(y)\, dy = C.$$

- For f_0 twice continuously differentiable, $h_n = n^{-1/5}$ is optimal for

$$\text{MISE}(\hat{f}_n) = \int \mathbb{E}[(\hat{f}_n(z) - f_0(z))^2]\, dz$$

  i.e. the mean integrated squared error.

Introduction (cont'd)

- Ideally, we could use this estimate for f_0 to estimate, for instance, moments of F_0. The second moment plug-in estimate would be

$$\int z^2 \hat{f}_n(z)\, dz = \sum_{i=1}^{n} \int z^2 \frac{1}{nh_n} K\left(\frac{z - Z_i}{h_n}\right) dz.$$

Introduction (cont'd)

- Ideally, we could use this estimate for f_0 to estimate, for instance, moments of F_0. The second moment plug-in estimate would be

$$\int z^2 \hat{f}_n(z)\,dz = \sum_{i=1}^{n} \int z^2 \frac{1}{nh_n} K\left(\frac{z - Z_i}{h_n}\right)\,dz.$$

- Substituting z by $zh_n + Z_i$ this equals

$$\frac{1}{n}\sum_{i=1}^{n} \int z^2 h_n^2 K(z)\,dz + \frac{2}{n}\sum_{i=1}^{n} Z_i h_n \int zK(z)\,dz + \frac{1}{n}\sum_{i=1}^{n} Z_i^2.$$

Introduction (cont'd)

- Ideally, we could use this estimate for f_0 to estimate, for instance, moments of F_0. The second moment plug-in estimate would be

$$\int z^2 \hat{f}_n(z)\, dz = \sum_{i=1}^{n} \int z^2 \frac{1}{nh_n} K\left(\frac{z - Z_i}{h_n}\right) dz.$$

- Substituting z by $zh_n + Z_i$ this equals

$$\frac{1}{n}\sum_{i=1}^{n} \int z^2 h_n^2 K(z)\, dz + \frac{2}{n}\sum_{i=1}^{n} Z_i h_n \int zK(z)\, dz + \frac{1}{n}\sum_{i=1}^{n} Z_i^2.$$

- The difference between this estimate and $\int z^2 f_0(z)\, dz$ equals

$$\frac{1}{n}\sum_{i=1}^{n} Z_i^2 - \int z^2 f_0(z)\, dz + Ch_n^2.$$

Introduction (cont'd)

- Ideally, we could use this estimate for f_0 to estimate, for instance, moments of F_0. The second moment plug-in estimate would be

$$\int z^2 \hat{f}_n(z)\, dz = \sum_{i=1}^{n} \int z^2 \frac{1}{nh_n} K\left(\frac{z - Z_i}{h_n}\right) dz.$$

- Substituting z by $zh_n + Z_i$ this equals

$$\frac{1}{n}\sum_{i=1}^{n} \int z^2 h_n^2 K(z)\, dz + \frac{2}{n}\sum_{i=1}^{n} Z_i h_n \int zK(z)\, dz + \frac{1}{n}\sum_{i=1}^{n} Z_i^2.$$

- The difference between this estimate and $\int z^2 f_0(z)\, dz$ equals

$$\frac{1}{n}\sum_{i=1}^{n} Z_i^2 - \int z^2 f_0(z)\, dz + Ch_n^2.$$

- $\Rightarrow \int z^2 \hat{f}_n(z)\, dz$ not $\sqrt{n}$ consistent for $\mathbb{E}[Z_1^2]$ as $\sqrt{n}h_n = n^{1/10} \to \infty$.

Introduction (cont'd)

- Can we find K such that we have $\sqrt{n}$ consistency for the plug-in estimator

$$\hat{f}_n \to \int z^2 \hat{f}_n(z)\,dz,$$

i.e.

$$\sqrt{n}\left(\int z^2 \hat{f}_n(z)\,dz - \int z^2 f_0(z)\,dz\right) = O_P(1),$$

and MISE-optimality at the same time?

Introduction (cont'd)

- Can we find K such that we have $\sqrt{n}$ consistency for the plug-in estimator

$$\hat{f}_n \to \int z^2 \hat{f}_n(z)\,dz,$$

i.e.

$$\sqrt{n}\left(\int z^2 \hat{f}_n(z)\,dz - \int z^2 f_0(z)\,dz\right) = O_P(1),$$

and MISE-optimality at the same time?

- More generally, for which functions g and kernels K can we have MISE-optimality and

$$\sqrt{n}\left(\int g(z)\hat{f}_K(z)\,dz - \int g(z) f_0(z)\,dz\right) = O_P(1)$$

at the same time?

Introduction (cont'd)

- Comparing this last question to what is known for the plug-in estimator based on $\widehat{F}_n$ where even

$$\sup_{\tilde{g}\in\tilde{\mathcal{G}}} \sqrt{n}\left(\int \tilde{g}(z)\,d\widehat{F}_n(z) - \int \tilde{g}(z)\,dF_0(z)\right) = O_P(1)$$

for several classes of functions $\tilde{\mathcal{G}}$ (even converges weakly)

Introduction (cont'd)

- Comparing this last question to what is known for the plug-in estimator based on $\widehat{F}_n$ where even

$$\sup_{\tilde{g}\in\tilde{\mathcal{G}}} \sqrt{n}\left(\int \tilde{g}(z)\, d\widehat{F}_n(z) - \int \tilde{g}(z)\, dF_0(z)\right) = O_P(1)$$

for several classes of functions $\tilde{\mathcal{G}}$ (even converges weakly)

- we may ask if we can have for some kernels MISE-optimality and

$$\sup_{\tilde{g}\in\tilde{\mathcal{G}}} \sqrt{n}\left(\int \tilde{g}(z) \hat{f}_K(z) dz - \int \tilde{g}(z) f_0(z) dz\right) = O_P(1)$$

for the same classes of functions (or even weak convergence).

Introduction (cont'd)

- Already known: If $Z_1, \dots, Z_n$ are iid it holds, for instance, that for $\tilde{\mathcal{G}} = \mathbb{BV} = \{\tilde{g} : \mathbb{R} \to \mathbb{R} \mid \text{ variation of } \tilde{g} \leq c\}$ we have

$$\sup_{\tilde{g} \in \mathbb{BV}} \sqrt{n} \left(\int \tilde{g}(z)\, d\widehat{F}_n(z) - \int \tilde{g}(z)\, dF_0(z) \right) = O_P(1).$$

- Giné and Nickl (2008) proved that for this class of functions and f_0 bounded and m-times continuously differentiable one has with $h_n = n^{-\frac{1}{2m+1}}$ (i.e.MISE optimality)

$$\sup_{\tilde{g} \in \mathbb{BV}} \sqrt{n} \left(\int \tilde{g}(z) \hat{f}_K(z) dz - \int \tilde{g}(z) f_0(z) dz \right) = O_P(1),$$

Introduction (cont'd)

- Already known: If $Z_1, \ldots, Z_n$ are iid it holds, for instance, that for $\tilde{\mathcal{G}} = \mathbb{BV} = \{\tilde{g} : \mathbb{R} \to \mathbb{R} \mid \text{ variation of } \tilde{g} \leq c\}$ we have

$$\sup_{\tilde{g} \in \mathbb{BV}} \sqrt{n} \left(\int \tilde{g}(z)\, d\widehat{F}_n(z) - \int \tilde{g}(z)\, dF_0(z) \right) = O_P(1).$$

- Giné and Nickl (2008) proved that for this class of functions and f_0 bounded and m-times continuously differentiable one has with $h_n = n^{-\frac{1}{2m+1}}$ (i.e.MISE optimality)

$$\sup_{\tilde{g} \in \mathbb{BV}} \sqrt{n} \left(\int \tilde{g}(z) \hat{f}_K(z) dz - \int \tilde{g}(z) f_0(z) dz \right) = O_P(1),$$

- if $\sqrt{n} h_n^{m+k} \to 0$ and if with $r = 1, \ldots, m$, and $k > 1/2$

$$\int K(z)\, dz = 1, \int z^r K(z)\, dz = 0, \int |z|^{m+k} |K(z)|\, dz < \infty.$$

Introduction (cont'd)

- Giné and Nickl (2008) proved two more results in the same spirit with $\tilde{G} \subset \mathbf{C}(\mathbb{R})$ where

$$\mathbf{C}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \,|\, |f(x)| \leq M, x \in \mathbb{R},\ f \text{ continuous}\}.$$

- Clearly, our example from the beginning $g(z) = z^2$ does not belong to $\mathbf{C}(\mathbb{R})$ nor is it of bounded variation on $\mathbb{R}$.

Introduction (cont'd)

- Giné and Nickl (2008) proved two more results in the same spirit with $\tilde{G} \subset \mathbf{C}(\mathbb{R})$ where

$$\mathbf{C}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \,|\, |f(x)| \leq M, x \in \mathbb{R},\ f \text{ continuous}\}.$$

- Clearly, our example from the beginning $g(z) = z^2$ does not belong to $\mathbf{C}(\mathbb{R})$ nor is it of bounded variation on $\mathbb{R}$.
- First, one might ask whether the above results can be extended, for instance, to the set

$$\mathbb{BV}_{loc} = \{\tilde{g} : \mathbb{R} \to \mathbb{R} \,|\, \tilde{g} \text{ is locally of bounded variation}\}.$$

Introduction (cont'd)

- Giné and Nickl (2008) proved two more results in the same spirit with $\tilde{G} \subset \mathbf{C}(\mathbb{R})$ where

$$\mathbf{C}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \,|\, |f(x)| \leq M, x \in \mathbb{R}, \ f \text{ continuous}\}.$$

- Clearly, our example from the beginning $g(z) = z^2$ does not belong to $\mathbf{C}(\mathbb{R})$ nor is it of bounded variation on $\mathbb{R}$.
- First, one might ask whether the above results can be extended, for instance, to the set

$$\mathbb{BV}_{loc} = \{\tilde{g} : \mathbb{R} \to \mathbb{R} \,|\, \tilde{g} \text{ is locally of bounded variation}\}.$$

- Second, one might ask whether we can extend the results to time series settings, where

$$Z_i \text{ is ARMA}(p, q) \text{ or } Z_i \text{ is GARCH}(p, q).$$

Extending the benchmark

Locally bounded variation: iid

- First recall the benchmark for the kernel based density estimator

$$\sup_{\tilde{g} \in \mathbb{BV}} \sqrt{n} \left(\int \tilde{g}(z)\, d\widehat{F}_n(z) - \int \tilde{g}(z)\, dF(z) \right) = O_P(1).$$

- Hence, we first need to extend this result into two directions:

1. Replace $\sup$ w.r.t. $\tilde{g} \in \mathbb{BV}$ by $\sup$ w.r.t. $\tilde{g} \in \mathbb{BV}_{\text{loc}}$;
2. Replace $\widehat{F}_n$ based on iid $Z_1, \ldots, Z_n$ by Z_i is ARMA(p, q) or Z_i is GARCH(p, q) (or more generally, by some weak dependence concept).

Locally bounded variation: iid

- Let $\phi : \mathbb{R} \to [1, \infty)$ be a weight function and put

$$\mathbb{BV}_{(1/\phi),\leq c} = \left\{ \tilde{g} \in \mathbb{BV}_{\text{loc}} : \int \frac{1}{\phi(z)} \, |d\tilde{g}|(z) \leq c \right\}.$$

Locally bounded variation: iid

- Let $\phi : \mathbb{R} \to [1, \infty)$ be a weight function and put

$$\mathbb{BV}_{(1/\phi),\leq c} = \left\{ \tilde{g} \in \mathbb{BV}_{\text{loc}} : \int \frac{1}{\phi(z)} |d\tilde{g}|(z) \leq c \right\}.$$

- Note that for $\phi \equiv 1$ we get the functions of bounded variation.

Locally bounded variation: iid

- Let $\phi : \mathbb{R} \to [1, \infty)$ be a weight function and put

$$\mathbb{BV}_{(1/\phi),\leq c} = \left\{ \tilde{g} \in \mathbb{BV}_{\text{loc}} : \int \frac{1}{\phi(z)} \, |d\tilde{g}|(z) \leq c \right\}.$$

- Note that for $\phi \equiv 1$ we get the functions of bounded variation.
- If we take $\phi(z) = (1 + |z|)^{2+\epsilon}$, $\epsilon > 0$, then our example from the beginning $g(z) = z^2$ ($dg(z) = 2z$) is included in

$$\mathbb{BV}_{(1/(1+|z|)^{2+\epsilon}),\leq c} = \left\{ \tilde{g} \in \mathbb{BV}_{\text{loc}} : \int \frac{1}{(1 + |z|)^{2+\epsilon}} \, |d\tilde{g}|(z) \leq c \right\}.$$

Locally bounded variation: iid

- Let $\phi : \mathbb{R} \to [1, \infty)$ be a weight function and put

$$\mathbb{BV}_{(1/\phi),\leq c} = \left\{ \tilde{g} \in \mathbb{BV}_{\text{loc}} : \int \frac{1}{\phi(z)} \, |d\tilde{g}|(z) \leq c \right\}.$$

- Note that for $\phi \equiv 1$ we get the functions of bounded variation.
- If we take $\phi(z) = (1 + |z|)^{2+\epsilon}$, $\epsilon > 0$, then our example from the beginning $g(z) = z^2$ ($dg(z) = 2z$) is included in

$$\mathbb{BV}_{(1/(1+|z|)^{2+\epsilon}),\leq c} = \left\{ \tilde{g} \in \mathbb{BV}_{\text{loc}} : \int \frac{1}{(1 + |z|)^{2+\epsilon}} \, |d\tilde{g}|(z) \leq c \right\}.$$

- **Theorem:** Let $Z_1, \ldots, Z_n$ be iid and ϕ be a weight function. Then for F_0 with $\int \phi^2 dF_0 < \infty$, we have

$$\sup_{\tilde{g} \in \mathbb{BV}_{(1/\phi),\leq c}} \sqrt{n} \left(\int \tilde{g}(z) \, d\widehat{F}_n(z) - \int \tilde{g}(z) \, dF_0(z) \right) = O_P(1).$$

Dependent data

- **Theorem:** $Z_1, \ldots, Z_n$ be strictly stationary and α-mixing with $\alpha(n) = \mathcal{O}(n^{-\theta})$ for some $\theta > 1 + \sqrt{2}$. Let $\phi_\lambda(x) := (1 + |x|)^\lambda$, $\lambda \geq 0$ and assume that $\int_{\mathbb{R}} |x|^\gamma \, dF_0(x) < \infty$ where $\gamma > \frac{2\theta\lambda}{\theta - 1}$. Then we have

$$\sup_{\tilde{g} \in \mathbb{BV}_{(1/\phi), \leq c}} \sqrt{n} \left(\int \tilde{g}(z) \, d\widehat{F}_n(z) - \int \tilde{g}(z) \, dF_0(z) \right) = O_P(1).$$

Dependent data

- **Theorem:** $Z_1, \dots, Z_n$ be strictly stationary and α-mixing with $\alpha(n) = \mathcal{O}(n^{-\theta})$ for some $\theta > 1 + \sqrt{2}$. Let $\phi_\lambda(x) := (1 + |x|)^\lambda$, $\lambda \geq 0$ and assume that $\int_{\mathbb{R}} |x|^\gamma \, dF_0(x) < \infty$ where $\gamma > \frac{2\theta\lambda}{\theta - 1}$. Then we have

$$\sup_{\tilde{g} \in \mathbb{BV}_{(1/\phi), \leq c}} \sqrt{n} \left(\int \tilde{g}(z) \, d\widehat{F}_n(z) - \int \tilde{g}(z) \, dF_0(z) \right) = O_P(1).$$

- **Theorem:** Let $Z_t := \sum_{s=0}^{\infty} a_s \, \varepsilon_{t-s}$, $t \in \mathbb{N}$, with $(\varepsilon_i)_{i \in \mathbb{Z}}$ i.i.d. Assume $a_s = s^{-\beta} \, \ell(s)$, $\beta \in (\frac{1}{2}, 1)$, $s \in \mathbb{N}$. Then $\mathbb{C}\mathrm{ov}(X_0, X_k) \sim k^{1-2\beta}$, hence non-summable, thus long-memory.
With ϕ_λ as above and $\mathbb{E}[|\varepsilon_0|^{2+2\lambda}] < \infty$, we have

$$\sup_{\tilde{g} \in \mathbb{BV}_{(1/\phi), \leq c}} r_n \left(\int \tilde{g}(z) \, d\widehat{F}_n(z) - \int \tilde{g}(z) \, dF_0(z) \right) = O_P(1),$$

where $r_n = n^{\beta - 1/2}$.

Free lunch? Yes free lunch!

Intro

- First note that we have

$$\int \tilde{g}(z)\hat{f}_K(z)dz - \int \tilde{g}(z)f_0(z)dz$$

$$= \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)d\widehat{F}_n(y)dz - \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dF_0(y)dz$$

$$+ \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{x-y}{h_n}\right)dF_0(y)\,dz - \int \tilde{g}(z)dF_0(z)$$

- Rewriting the second line as

$$\int \underbrace{\int \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dz}_{\bar{g}_n(y)}\,d\widehat{F}_n(y) - \int \underbrace{\int \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dz}_{\bar{g}_n(y)}\,dF_0(y)$$

we see that the above benchmark result applies if all the $\bar{g}_n$ belong to $\mathbb{BV}_{(1/\phi),\leq c}$

Intro

- First note that we have

$$\int \tilde{g}(z)\hat{f}_K(z)dz - \int \tilde{g}(z)f_0(z)dz$$
$$= \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)d\widehat{F}_n(y)dz - \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dF_0(y)dz$$
$$+ \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{x-y}{h_n}\right)dF_0(y)\,dz - \int \tilde{g}(z)dF_0(z)$$

- Rewriting the second line as

$$\int \underbrace{\int \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dz}_{\bar{g}_n(y)}\,d\widehat{F}_n(y) - \int \underbrace{\int \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dz}_{\bar{g}_n(y)}\,dF_0(y)$$

Intro

- First note that we have

$$
\begin{aligned}
&\int \tilde{g}(z)\hat{f}_K(z)dz - \int \tilde{g}(z)f_0(z)dz \\
&= \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)d\widehat{F}_n(y)dz - \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dF_0(y)dz \\
&\quad + \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{x-y}{h_n}\right)dF_0(y)\,dz - \int \tilde{g}(z)dF_0(z)
\end{aligned}
$$

- Rewriting the second line as

$$
\int \underbrace{\int \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dz}_{\bar{g}_n(y)}\,d\widehat{F}_n(y) - \int \underbrace{\int \tilde{g}(z)\frac{1}{h_n}K\left(\frac{z-y}{h_n}\right)dz}_{\bar{g}_n(y)}\,dF_0(y)
$$

we see that the above benchmark result applies if all the $\bar{g}_n$ belong to $\mathbb{BV}_{(1/\phi),\leq c}$

Intro

- Hence, it only remains to consider

$$\sup_{\tilde{g}\in\tilde{\mathcal{G}}} \iint \tilde{g}(z)\frac{1}{h_n}K\left(\frac{x-y}{h_n}\right)\, dF_0(y)\, dz - \int \tilde{g}(z)dF_0(z).$$

- Consider this for

$$\tilde{g} \in \tilde{\mathcal{G}} := \{g_x : \mathbb{R} \to \mathbb{R} | g_x(z) = \phi(x)\mathbb{1}_{(-\infty,x)}(z), x \leq 0, \\ \text{and } g_x(z) = -\phi(x)\mathbb{1}_{[x,\infty)}(z), x > 0\}.$$

- Then the above becomes

$$\sup_x \phi(x) \int K(z)\left(F_0(x+zh_n) - F_0(x)\right)\, dz.$$

Free lunch

- If the benchmark result holds for F_0 we have

$$\phi(x)F_0(x) \to 0 \text{ for } x \to -\infty \text{ and } \phi(x)(1 - F_0(x)) \text{ for } x \to \infty.$$

Free lunch

- If the benchmark result holds for F_0 we have

$$\phi(x)F_0(x) \to 0 \text{ for } x \to -\infty \text{ and } \phi(x)(1 - F_0(x)) \text{ for } x \to \infty.$$

- Hence, for K compact and x small (or large)

$$F_0(x + zh_n) - F_0(x)$$

will also be small for all z even when compared to $\phi(x)$.

Free lunch

- If the benchmark result holds for F_0 we have

$$\phi(x)F_0(x) \to 0 \text{ for } x \to -\infty \text{ and } \phi(x)(1 - F_0(x)) \text{ for } x \to \infty.$$

- Hence, for K compact and x small (or large)

$$F_0(x + zh_n) - F_0(x)$$

  will also be small for all z even when compared to $\phi(x)$.
- **Theorem:** We have the following extension of the above result

$$\sup_{\tilde{g} \in \mathbb{BV}_{(1/\phi), \leq c}} \sqrt{n} \left(\int \tilde{g}(z) \hat{f}_K(z) dz - \int \tilde{g}(z) f_0(z) dz \right) = O_P(1),$$

Free lunch

- If the benchmark result holds for F_0 we have

$$\phi(x)F_0(x) \to 0 \text{ for } x \to -\infty \text{ and } \phi(x)(1 - F_0(x)) \text{ for } x \to \infty.$$

- Hence, for K compact and x small (or large)

$$F_0(x + zh_n) - F_0(x)$$

will also be small for all z even when compared to $\phi(x)$.

- **Theorem:** We have the following extension of the above result

$$\sup_{\tilde{g} \in \mathbb{BV}_{(1/\phi), \leq c}} \sqrt{n} \left(\int \tilde{g}(z) \hat{f}_K(z) dz - \int \tilde{g}(z) f_0(z) dz \right) = O_P(1),$$

if $\sqrt{n} h_n^{m+k} \to 0$, if $\sup_z \phi(z) f_0^{(m)}(z) \leq C$ and if with $r = 1, \ldots, m$, and $k > 1/2$

$$\int K(z)\, dz = 1, \int z^r K(z)\, dz = 0, \int |z|^{m+k} |K(z)|\, dz < \infty.$$

One more free lunch? No (too much asked).

Intro

- Now consider K non-compact. Then the above reasoning that

$$F_0(x + zh_n) - F_0(x)$$

will be small if $F_0(x)$ is small does not apply anymore (we can make the first (almost) equal to 1 by making z large).

Intro

- Now consider K non-compact. Then the above reasoning that

$$F_0(x + zh_n) - F_0(x)$$

will be small if $F_0(x)$ is small does not apply anymore (we can make the first (almost) equal to 1 by making z large).

- Yet, intuitively, if K puts not too much weight on these z then

$$\int K(z)\left(F_0(x + zh_n) - F_0(x)\right) dz$$

should still be small if x is small.

Intro

- Now consider K non-compact. Then the above reasoning that

$$F_0(x + zh_n) - F_0(x)$$

will be small if $F_0(x)$ is small does not apply anymore (we can make the first (almost) equal to 1 by making z large).

- Yet, intuitively, if K puts not too much weight on these z then

$$\int K(z)\,(F_0(x + zh_n) - F_0(x))\; dz$$

should still be small if x is small.

- **Impose the following**: f_0 is m-times continuously differentiable and for all $t \in [0, 1]$ and all $x, y \in \mathbb{R}$ we have

$$\sup_x |\phi(x) f^{(m)}(x + ty)| \leq L_2 |y|^{p(y)},$$

where p is a bounded function.

Too much asked

- **Theorem:** Let f_0 be as above and K be non-compact. Then

$$\sup_{\tilde{g}\in\mathbb{BV}_{(1/\phi),\leq c}} \sqrt{n}\left(\int \tilde{g}(z)\hat{f}_K(z)dz - \int \tilde{g}(z)f_0(z)dz\right) = O_P(1),$$

Too much asked

- **Theorem:** Let f_0 be as above and K be non-compact. Then

$$\sup_{\tilde{g}\in\mathbb{BV}_{(1/\phi),\leq c}} \sqrt{n}\left(\int \tilde{g}(z)\hat{f}_K(z)dz - \int \tilde{g}(z)f_0(z)dz\right) = O_P(1),$$

if $\sqrt{n}h_n^{m+s} \to 0$, and if with $r = 1, \ldots, \lfloor m+s \rfloor$

$$\int K(z)\,dz = 1, \int z^r K(z)\,dz = 0, \int |z|^{m+s}|K(z)|\,dz < \infty,$$

where $s = \sup_y p(y)$.

Example

- K compact: Then for f_0 density of the double exponential we clearly have that

$$\sup_z \phi(z) f_0(z)$$

is finite for any polynomial weight function.

Example

- K compact: Then for f_0 density of the double exponential we clearly have that

$$\sup_z \phi(z) f_0(z)$$

is finite for any polynomial weight function.

- K non-compact and same f_0: Then for a polynomial weight of the form $(1+|z|)^\lambda$ the above condition holds with $p > \lambda$.

- Thus, we have to increase the order of the kernel beyond what is needed by the smoothness and that increase is a function of the weight.

That’s all.