

# Weighted uniform consistency of kernel density estimators with general bandwidth sequences



Department of Mathematics, Free University of Brussels (VUB)  
 Pleinlaan 2, B-1050 Brussels, Belgium  
 jdonoy@vub.ac.be  
 MSC: 60B12, 60F, 62G07



## 1. Introduction

Let  $* X_1, X_2, \dots$  be  $d$ -dimensional and i.i.d random vectors with common density  $f$ ,

\*  $K : \mathbb{R}^d \rightarrow \mathbb{R}$  be a measurable and bounded *kernel*, i.e  $\int K(s)ds = 1$  and  $\|K\|_\infty =: \kappa < \infty$ ,

\*  $f_{n,h}(t)$  be the *kernel density estimator* for  $f(t)$  based on the sample  $X_1, \dots, X_n$  and bandwidth  $0 < h < 1$ , i.e

$$f_{n,h}(t) := \frac{1}{nh} \sum_{i=1}^n K\left(\frac{t - X_i}{h^{1/d}}\right), \quad t \in \mathbb{R}^d.$$

**!!! The choice of  $h$  is crucial for the consistency, and it should be chosen dependent on the data !!!**

$\hookrightarrow$  if  $h_n \searrow 0$  and  $\frac{nh_n}{\log \log n} \rightarrow \infty$ , we get a *strongly consistent* estimator of  $f$ , i.e  $f_{n,h_n}(t) \xrightarrow{a.s.} f(t)$ .

*When do we have  $\|f_{n,h} - f\|_\infty \rightarrow 0$  and at which rate?*

## 2. Known consistency results

### Uniform consistency

*E. Giné and A. Guillou (2002)*

For a bandwidth  $h_n \searrow 0$  satisfying  $\frac{nh_n}{\log n} \rightarrow \infty$  and  $\frac{|\log h_n|}{\log \log n} \rightarrow \infty$ , we have

$$\|f_{n,h_n} - \mathbb{E}f_{n,h_n}\|_\infty = O\left(\sqrt{\frac{|\log h_n|}{nh_n}}\right).$$

### Uniform in bandwidth consistency

*U. Einmahl and D. M. Mason (2005)*

If  $f$  is bounded, we have with probability one,

$$\limsup_{n \rightarrow \infty} \sup_{\frac{c \log n}{n} \leq h \leq 1} \frac{\sqrt{nh} \|f_{n,h} - \mathbb{E}f_{n,h}\|_\infty}{\sqrt{|\log h|} \sqrt{\log \log n}} < \infty.$$

$\hookrightarrow$  the consistency rate of a kernel density estimator based on either a statistical or deterministic bandwidth remains the same!

### Weighted uniform consistency

*E. Giné, V. Koltchinskii and J. Zinn (2004)*

Under some extra assumptions on  $f$ , we can infer that

$$\limsup_{n \rightarrow \infty} \sqrt{\frac{nh_n}{|\log h_n|}} \|\psi(f_{n,h_n} - \mathbb{E}f_{n,h_n})\|_\infty < \infty,$$

where  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  is an unbounded weight function satisfying among other properties  $\|\psi f^\beta\|_\infty < \infty$  for some  $0 < \beta < 1/2$ , and where  $h_n$  is chosen to be  $h_n = n^{-\alpha} L(n)$  for some  $0 < \alpha < 1$  and slowly varying function  $L$ .

## 3. New consistency results

*What about weighted uniform in bandwidth consistency?*

Assume  $f$  to be continuous and bounded, and for *regularly varying* functions  $a_n < b_n \rightarrow 0$ , consider

$$\Delta_n := \sup_{a_n \leq h \leq b_n} \sqrt{\frac{nh}{|\log h|}} \|\psi(f_{n,h} - \mathbb{E}f_{n,h})\|_\infty, \quad n \geq 1.$$

Then under assumptions similar to those in the weighted uniform consistency case (GKZ, 2004), we can infer:

### Theorem 1 (Stochastic boundedness)

$\Delta_n$  is stochastically bounded



$$\limsup_{t \rightarrow \infty} t \mathbb{P} \left\{ \psi(X) > \sqrt{ta_t |\log a_t|} \right\} < \infty.$$

### Theorem 2 (Almost sure convergence)

$$\limsup_{n \rightarrow \infty} \Delta_n \leq C < \infty, \quad a.s.$$



$$\int_1^\infty \mathbb{P} \left\{ \psi(X) > \sqrt{ta_t |\log a_t|} \right\} dt < \infty$$

#### References

- [1] Dony J. and Einmahl U. (2006). *Weighted uniform consistency of kernel density estimators with general bandwidth sequences*. Submitted.
- [2] Giné E. and Guillou A. (2002). *Rates of strong uniform consistency for multivariate kernel density estimators*. Ann. Inst. H. Poincaré Probab. Statist. **38**(6), 907-921.
- [3] Giné E., Koltchinskii V. and Zinn J. (2004). *Weighted uniform consistency of kernel density estimators*. Ann. Probab. **32**, 2570-2605.
- [4] Einmahl U. and Mason D.M. (2005). *Uniform in bandwidth consistency of kernel function-type estimators*. Ann. Stat. **33**, 1601-1624.