Root-n consistent density estimators of convolutions in weighted L_1-norms

Anton Schick and Wolfgang Wefelmeyer

Abstract. It is known that the convolution of a smooth density with itself can be estimated at the root-n rate by a convolution of an appropriate density estimator with itself. We show that this remains true even for discontinuous densities as long as they are of bounded variation. The assumption of bounded variation can be relaxed. We consider convergence in weighted L_1-norms and show that the standardized convolution estimator converges in distribution to a centered Gaussian process.

1. Introduction

Let $X_1, \dots, X_n$ be independent and identically distributed random variables with distribution function F and density f. Let g denote the convolution $f * f$ of f with itself,

$$g(x) = f * f(x) = \int f(x-y) f(y)\, dy, \quad x \in \mathbb{R}.$$

We are interested in obtaining root-n consistent estimators of g under weak assumptions. Estimators of $g = f * f$ can in particular be used to test whether f belongs to a given family of densities that is closed under convolutions, like the normal densities.

There are two estimators for g in the literature. Frees (1994) introduced the U-statistic kernel-type estimator

$$\hat{g}(x) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} k_b(x - X_i - X_j), \quad x \in \mathbb{R},$$

where $k_b(x) = k(x/b)/b$ for some density k and some bandwidth $b = b_n$. Saavedra and Cao (2000) studied the convolution-type plug-in estimator $\hat{g}_* = \hat{f} * \hat{f}$ with $\hat{f}$ a kernel estimator of f,

$$\hat{f}(x) = \frac{1}{n} \sum_{j=1}^{n} k_b(x - X_j), \quad x \in \mathbb{R}.$$

Let $\tilde{k} = k * k$ and $\tilde{k}_b(x) = \tilde{k}(x/b)/b$. Then $\tilde{k}_b = k_b * k_b$, and it is easy to see that $\hat{g}_*(x)$ can be expressed as the von Mises statistic

$$\hat{g}_*(x) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{k}_b(x - X_i - X_j), \quad x \in \mathbb{R}.$$

Anton Schick was supported by NSF Grant DMS 0405791.

This estimator is closely related to the Frees estimator with the special kernel $k = \tilde{k}$,

$$\tilde{g}(x) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \tilde{k}_b(x - X_i - X_j), \quad x \in \mathbb{R}.$$

Frees (1994) and Saavedra and Cao (2000) showed that their estimators converge pointwise at the root-n rate. We are interested in minimal assumptions under which $\hat{g}$ and $\tilde{g}$ have root-n rates in L_1-spaces. It follows from Schick and Wefelmeyer (2004) that under appropriate conditions on f and for proper choices of kernel and bandwidth, the estimator $\hat{g}_*$ is root-n consistent in the L_1-norm, and that the scaled process $\sqrt{n}(\hat{g}_* - g)$ converges in distribution in the space L_1 to a centered Gaussian processes. More precisely, the following result can be derived from their Theorem 2.

THEOREM 1. *Suppose that $\psi = F^{1/2}(1-F)^{1/2}$ is integrable and that there is an integrable function f' such that $f(x) = \int_{-\infty}^{x} f'(u)\,du$, $x \in \mathbb{R}$. Let $nb^2 \to \infty$ and $nb^4 \to 0$. Let k be uniformly continuous with mean zero and finite variance. Then $\sqrt{n}(\hat{g}_* - g)$ converges in distribution (in the space L_1) to $2W * f'$ with $W = B_0 \circ F$ and B_0 a standard Brownian bridge.*

The goal of this paper is to study what results are possible under weaker smoothness assumptions on the density f. The above theorem does not apply to densities with jumps such as uniform densities or exponential densities. These latter densities are of bounded variation. For such densities, root-n consistency and weak convergence still hold.

THEOREM 2. *Suppose f is of bounded variation and $\int |x|^\alpha f(x)\,dx$ is finite for some $\alpha > 1$. Let $nb^2 \to \infty$ and $nb^4 \to 0$. Let k have mean zero and finite variance and be bounded. Then $\sqrt{n}(\hat{g} - g)$ and $\sqrt{n}(\hat{g}_* - g)$ converge in distribution (in the space L_1) to the same centered Gaussian process whose covariance structure matches that of $2f(\cdot - X_1)$.*

The moment assumption in Theorem 2 is weaker than the integrability of $\psi = F^{1/2}(1-F)^{1/2}$ required in Theorem 1. Indeed, Schick and Wefelmeyer (2004) have shown in the proof of their Theorem 2 that the latter condition implies that f has a finite moment of order $3/2$. If one repeats their argument, one can derive that f has finite moments of all orders less than 2 if ψ is integrable.

We refer to Giné and Mason (2005) for general functional central limit theorems for the variance term $\sqrt{n}(\hat{g} - E[\hat{g}])$ in L_p, $1 \le p \le \infty$, in the general setting of Frees (1994), and uniformly in the bandwidth.

The conclusions of Theorem 2 hold even for some functions of unbounded variation such as Gamma densities with shape parameter between $1/2$ and 1. More precisely, we shall prove the following result which is formulated in terms of the smoothness of g. For this we need the following terminology. We say a measurable function h is *L_1-Hölder with exponent* γ (with $0 < \gamma \le 1$) if there is a finite constant C_h such that $\int |h(x-t) - h(x)|\,dx \le C_h |t|^\gamma$ for $t \in \mathbb{R}$.

THEOREM 3. *Suppose $\int (1+|x|)^\alpha (f(x) + f^2(x))\,dx$ is finite for some $\alpha > 1$. Let k have mean zero and a finite variance and be bounded. Let $nb \to \infty$. Then $\sqrt{n}(\hat{g} - g)$ and $\sqrt{n}(\hat{g}_* - g)$ converge in distribution (in the space L_1) to the same centered Gaussian process if one of the following three conditions hold.*

(1) *The function g is L_1-Hölder with exponent $\gamma > 1/2$, and $nb^{2\gamma} \to 0$.*
(2) *The function g is absolutely continuous with integrable a.e.-derivative, and $nb^2 = O(1)$.*
(3) *The function g is absolutely continuous with an integrable a.e.-derivative that is also L_1-Hölder with exponent γ, and $nb^{2+2\gamma} \to 0$.*

The covariance structure of the limiting process matches that of $2f(\cdot - X_1)$.

If g is of bounded variation, then g is L_1-Hölder with exponent $\gamma = 1$; see Corollary 2 in Section 4. If f is of bounded variation, then g satisfies (3) of Theorem 3 with $\gamma = 1$ as shown in Corollary 2 below, and since f is bounded the integrability required of f in Theorem 3 follows from that required in Theorem 2. Thus Theorem 2 follows from Theorem 3.

In Section 2 we prove an extension of Theorem 3 to function spaces with stronger norms, namely V-norms $\|h\|_V = \int V(x)|h(x)|\,dx$ with V such that

$$\|h_1 * h_2\|_V \le \|h_1\|_V \|h_2\|_V \tag{1.1}$$

and show that the above theorems are special cases of this result. Functional convergence in the stronger L_V distance rather than in L_1 is needed if we want to estimate expectations $E[a(X_1 + X_2)] = \int a(x)g(x)\,dx$ by $\int a(x)\hat{g}_*(x)\,dx$, where a is a function bounded by V but not by a constant, for example moments of $X_1 + X_2$. Especially when g is smooth, the estimator $\int a(x)\hat{g}_*(x)\,dx$ can be better for small samples than the von Mises statistic

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} a(X_i + X_j).$$

In Section 3 we characterize compact subsets of the space L_V of functions with finite V-norm and prove equi-continuity of $\int h(\cdot - bt)k(t)\,dt$ at $b = 0$ in L_V with h ranging over compact subsets of L_V. Section 4 gives expansions of shifts and related transformations in L_V.

2. The main result

Let V be a positive measurable function. For a measurable function h, the V-*norm* is defined by

$$\|h\|_V = \int V(x)|h(x)|\,dx.$$

If we take $V = 1$, this is the usual L_1-norm. Let L_V denote the (separable) Banach space of all (equivalence classes of) measurable functions with finite V-norm. We impose the following assumptions on V.

Assumption 1. The function V is continuous at 0 with $V(0) = 1$ and satisfies

$$V(x + y) \le V(x)V(y), \quad x, y \in \mathbb{R}; \tag{2.1}$$

$$V(sx) \le V(x), \quad |s| \le 1,\ x \in \mathbb{R}. \tag{2.2}$$

Possible choices for V are $V(x) = (1 + \log(1 + |x|))^r$, $V(x) = (1 + |x|)^r$ and $V(x) = \exp(r|x|)$ with $r \ge 0$. In each case, the choice $r = 0$ gives again the L_1-norm. The condition (2.1) meshes well with shifts and convolutions. Indeed, (2.1) yields

$$\|h(\cdot - s)\|_V = \int V(x + s)|h(x)|\,dx \le V(s)\|h\|_V, \quad s \in \mathbb{R}, \tag{2.3}$$

and this and (2.2) imply

$$\sup_{|w| \le 1} \|h(\cdot - wt)\|_V \le V(t)\|h\|_V, \quad t \in \mathbb{R}. \tag{2.4}$$

In particular, we obtain (1.1).

REMARK 1. It follows from (2.2) that $V(x) \geq V(0) = 1$ for all x in $\mathbb{R}$, that V is symmetric in the sense that $V(x) = V(-x)$ for all x in $\mathbb{R}$, and that $V(x) \geq V(y)$ if $|x| \geq |y|$. These properties and (2.1) yield

$$|V(x+s) - V(x)| \leq V(x)(V(s)-1), \quad x, s \in \mathbb{R}. \tag{2.5}$$

The following lemma shows that $\tilde{g}$ and $\hat{g}_*$ are asymptotically equivalent. We can therefore concentrate on the estimator $\hat{g}$.

LEMMA 1. *Suppose Assumption 1 holds. Let* $\int V(x)k(x)\,dx$ *and* $\int V^2(x)f(x)\,dx$ *be finite. Then*

$$\|\hat{g}_* - \tilde{g}\|_V = O_p(n^{-1}).$$

PROOF. It is easy to check that $n(\hat{g}_* - \tilde{g}) = \hat{f}_2 - \tilde{g}$ with $\hat{f}_2(x) = \frac{1}{n}\sum_{j=1}^n \tilde{k}_b(x - 2X_j)$, $x \in \mathbb{R}$, a kernel estimator of the density of $2X_1$. Thus we only need to show that $\|\hat{f}_2\|_V = O_p(1)$ and $\|\tilde{g}\|_V = O_p(1)$. For $b < 1$, we have

$$\|\hat{f}_2\|_V \leq \frac{1}{n}\sum_{j=1}^n \int V(bx + 2X_j)\tilde{k}(x)\,dx \leq \frac{1}{n}\sum_{j=1}^n V^2(X_j)\|k * k\|_V$$

and

$$\|\tilde{g}\|_V \leq \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \int V(bx + X_i + X_j)\tilde{k}(x)\,dx \leq \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} V(X_i)V(X_j)\|k * k\|_V.$$

By the assumptions on f and k we see that $\|\hat{f}_2\|_V = O_p(1)$ and $\|\tilde{g}\|_V = O_p(1)$. □

The Hoeffding decomposition yields

$$\hat{g}(x) = k_b * g(x) + 2\mathbb{K}_n(x) + \mathbb{U}_n(x), \quad x \in \mathbb{R},$$

where $\mathbb{K}_n(x)$ is the average

$$\mathbb{K}_n(x) = \frac{1}{n}\sum_{j=1}^n \big(k_b * f(x - X_j) - k_b * g(x)\big)$$

and $\mathbb{U}_n(x)$ is the degenerate U-statistic

$$\mathbb{U}_n(x) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \big(k_b(x - X_i - X_j) - k_b * f(x - X_i) - k_b * f(x - X_j) + k_b * g(x)\big).$$

One can express $\mathbb{K}_n$ as $\mathbb{H}_n * k_b$, where

$$\mathbb{H}_n(x) = \frac{1}{n}\sum_{j=1}^n \big(f(x - X_j) - E[f(x - X_j)]\big) = \frac{1}{n}\sum_{j=1}^n \big(f(x - X_j) - g(x)\big), \quad x \in \mathbb{R}.$$

Thus we arrive at the representation

$$\hat{g} - g = 2\mathbb{H}_n + g * k_b - g + 2(\mathbb{H}_n * k_b - \mathbb{H}_n) + \mathbb{U}_n. \tag{2.6}$$

Our approach will be to show that norms of the bias $g * k_b - g$ and of the terms $\mathbb{H}_n * k_b - \mathbb{H}_n$ and $\mathbb{U}_n$ are of order $o_p(n^{-1/2})$, and that $\sqrt{n}\mathbb{H}_n$ converges in distribution in L_V.

We first consider the remainder term $\|\mathbb{U}_n\|_V$. For this and for later use, we introduce, for $\alpha \geq 0$, the function W_α by

$$W_\alpha(x) = \int (1 + |x|)^\alpha V^2(x), \quad x \in \mathbb{R}.$$

This function has the same properties as V. In particular we have

$$\|h_1 * h_2\|_{W_\alpha} \le \|h_1\|_{W_\alpha}\|h_2\|_{W_\alpha}. \tag{2.7}$$

An application of the Cauchy–Schwarz inequality yields that for measurable h and $\alpha > 1$,

$$\|h\|_V^2 \le K_\alpha \|h^2\|_{W_\alpha} \tag{2.8}$$

with $K_\alpha = \int (1+|x|)^{-\alpha}\,dx$. We also need to impose the following condition on the kernel k.

(K) The density k is bounded and has mean zero, and $\|k\|_{W_2}$ is finite.

LEMMA 2. *Suppose Assumption 1 and (K) hold and $\|f\|_{W_\alpha}$ is finite for some $\alpha > 1$. Then*

$$\|\mathbb{U}_n\|_V = O_p(n^{-1}b^{-1/2}).$$

PROOF. We may assume that $\alpha < 2$. It follows from (K) that $\|k^2\|_{W_\alpha}$ is finite. We calculate

$$E[\mathbb{U}_n^2(x)] \le \frac{2}{n(n-1)} E[k_b^2(x - X_1 - X_2)] \le \frac{2}{n(n-1)} k_b^2 * g(x).$$

From this we obtain that

$$E[\|\mathbb{U}_n^2\|_{W_\alpha}] \le \|E[\mathbb{U}_n^2]\|_{W_\alpha} \le \frac{2}{n(n-1)} \|k_b^2 * g\|_{W_\alpha}.$$

Since W_α inherits the properties of V, it follows that

$$\|k_b^2 * g\|_{W_\alpha} = b^{-1} \iiint W_\alpha(x+y+bz) f(x) f(y) k^2(z)\,dx\,dy\,dz \le b^{-1}\|f\|_{W_\alpha}^2 \|k^2\|_{W_\alpha}$$

for $b < 1$. This shows that $E[\|\mathbb{U}_n^2\|_{W_\alpha}] = O(n^{-2}b^{-1})$. □

To deal with $\mathbb{H}_n$, we rely on the following central limit theorem in L_1-spaces; see Ledoux and Talagrand (1991, Theorem 10.10) or van der Vaart and Wellner (1996, page 92).

THEOREM 4. *Let μ be a σ-finite measure on the Borel-σ-field on $\mathbb{R}$. Let $Z_1, Z_2, \dots$ be independent and identically distributed zero-mean random elements in $L_1(\mu)$. Then the sequence $n^{-1/2}\sum_{i=1}^n Z_i$ converges in distribution (in $L_1(\mu)$) to a centered Gaussian process if and only if*

$$\lim_{t\to\infty} t^2 P\Big(\int |Z_1(x)|\,\mu(dx) > t\Big) = 0 \quad \textit{and} \quad \int E[Z_1^2(x)]^{1/2}\,\mu(dx) < \infty.$$

LEMMA 3. *Suppose Assumption 1 holds and $\|f\|_{W_\alpha}$ and $\|f^2\|_{W_\alpha}$ are finite for some $\alpha > 1$. Then $\|\mathbb{H}_n * k_b - \mathbb{H}_n\|_V = o_p(n^{-1/2})$ and $\sqrt{n}\mathbb{H}_n$ converges in distribution in the space L_V to a centered Gaussian process whose covariance structure matches that of $f(\cdot - X_1)$.*

PROOF. The second conclusion implies tightness of $\sqrt{n}\mathbb{H}_n$ and hence the first conclusion in view of Remark 5 below. To prove the second conclusion, we apply the previous theorem with $\mu(dx) = V(x)\,dx$ and $Z_i(x) = f(x - X_i) - g(x)$. Using (2.8) we find that

$$E\Big[\Big(\int |Z_1(x)| V(x)\,dx\Big)^2\Big] = E[\|Z_1\|_V^2] \le K_\alpha \int W_\alpha(x) E[Z_1^2(x)]\,dx$$

and

$$\Big(\int E[Z_1^2(x)]^{1/2} V(x)\,dx\Big)^2 \le K_\alpha \int W_\alpha(x) E[Z_1^2(x)]\,dx.$$

Thus we only need to show that $\int W_\alpha(x)E[Z_1^2(x)]\,dx$ is finite. Since $E[Z_1^2(x)] \le E[f^2(x-X_1)] = f^2*f(x)$, this follows from $\|f^2 * f\|_{W_\alpha} < \infty$. But $\|f^2 * f\|_{W_\alpha} \le \|f^2\|_{W_\alpha}\|f\|_{W_\alpha}$ is finite by the assumptions on f. □

To deal with the bias term, we introduce the following terminology.

DEFINITION 1. We say that h is *V-Hölder with exponent* γ for some $0 < \gamma \le 1$ if there is a finite constant C such that

$$\int V(x)|h(x-t) - h(x)|\,dx \le C|t|^\gamma V(t), \quad t \in \mathbb{R}.$$

We say that a function h is *V-smooth* if h is absolutely continuous (on compacts) and if its almost everywhere derivative h' has finite V-norm. We say h is *V-smooth of order* $1+\gamma$ if h is V-smooth and h' is V-Hölder with exponent γ.

Lemma 7 below now gives the desired extension of Theorem 3 to the space L_V. The choice $V = 1$ yields Theorem 3.

THEOREM 5. *Let Assumption 1 and (K) hold and let $nb \to \infty$. Let $\|f\|_{W_\alpha}$ and $\|f^2\|_{W_\alpha}$ be finite for some $\alpha > 1$. Then $\sqrt{n}(\hat{g} - g)$ converges in distribution (in the space L_V) to some centered Gaussian process if one of the following three conditions holds.*

(1) *The function g is V-Hölder with exponent $\gamma > 1/2$, and $nb^{2\gamma} \to 0$.*
(2) *The function g is V-smooth, and $nb^2 = O(1)$.*
(3) *The function g is V-smooth of order $1+\gamma$, and $nb^{2+2\gamma} \to 0$.*

The covariance structure of the limiting process matches that of $2f(\cdot - X_1)$.

REMARK 2. The requirement $\gamma > 1/2$ in part (1) of Theorem 5 is needed to guarantee the existence of a bandwidth that satisfies $nb \to \infty$ and $nb^{2\gamma} \to 0$. The choice $b = (\log n)/n$ works for all three cases, while the choice $b = n^{-1/2}$ works in cases (2) and (3).

REMARK 3. If k satisfies condition (K), so does $k * k$. Thus Theorem 5 holds with $\hat{g}$ replaced by $\hat{g}_*$ and thus, by Lemma 1, also with $\hat{g}$ replaced by $\tilde{g}$.

REMARK 4. For the following discussion let us take $V(x) = (1+|x|)^r$ for some non-negative r. Let $f = f_a$, with f_a the Gamma density with positive shape parameter a:

$$f_a(x) = \mathbf{1}(x > 0)x^{a-1}e^{-x}/\Gamma(a), \quad x \in \mathbb{R}.$$

If $a \ge 1$, then f_a is of bounded variation and Theorem 3 applies. If $a < 1$, then f_a is no longer of bounded variation. Let us now show that the assumptions of Theorem 5 can be met if $a > 1/2$. Actually, $\|f\|_{W_\alpha}$ is finite for all $a > 0$, but $\|f^2\|_{W_\alpha}$ is finite for $a > 1/2$ only. We have $f_a * f_a = f_{2a}$, and f_{2a} is absolutely continuous with almost everywhere derivative $f'_{2a} = (2a-1)(\Gamma(2a-1)/\Gamma(2a))f_{2a-1} - f_{2a}$. Since f_{2a-1} and f_{2a} have finite moments of all order, we see that part (2) of Theorem 5 can be met with $b = n^{-1/2}$. A more careful analysis shows that f_{2a} is V-smooth of order $1+\gamma$ for any $0 < \gamma < 2a-1$.

Let h be an integrable function of bounded variation. Then we can write h as the difference of two bounded non-decreasing functions which vanish at $-\infty$. Without loss of generality we may assume that these functions are right-continuous as this changes h only on a countable set. This shows that

$$h(x) = \mu_1((-\infty, x]) - \mu_2((-\infty, x]), \quad x \in \mathbb{R}, \tag{2.9}$$

for two finite measures μ_1 and μ_2. We call $\nu_h = \mu_1 + \mu_2$ the *measure of variation* of h.

If f is of bounded variation, then f is bounded and $\|f\|_{W_\alpha} < \infty$ implies $\|f^2\|_{W_\alpha} < \infty$. If also $\int V\, d\nu_f$ is finite, then g is V-smooth of order 2 by Lemma 8 below. Thus we have the following corollary.

COROLLARY 1. *Suppose Assumption 1 and (K) hold and $\|f\|_{W_\alpha}$ is finite for some $\alpha > 1$. Let the density f be of bounded variation with $\int V\, d\nu_f$ finite. Let $nb \to \infty$ and $nb^4 \to 0$. Then $\sqrt{n}(\hat{g} - g)$ converges in distribution (in the space L_V) to some centered Gaussian process.*

Since the assumptions of this corollary become those of Theorem 2 if $V = 1$, Theorem 2 is a special case of Corollary 1.

3. Compact subsets of the space L_V.

In this section we study compact subsets of the (Banach) space of all (equivalence classes of) measurable functions h with finite V-norm $\|h\|_V = \int V(x)|h(x)|\, dx$. We shall do this under the following condition, which by Remark 1 is a consequence of Assumption 1.

ASSUMPTION 2. The function V satisfies $V \geq 1$ and

$$\sup_{x\in\mathbb{R}} \frac{|V(x+t) - V(x)|}{V(x)} \to 0 \quad \text{as } t \to 0. \tag{3.1}$$

It is easy to see that a subset H of L_V is compact if and only if $VH = \{Vh : h \in H\}$ is compact in L_1. A characterization of compact subsets of L_1 is given by the Fréchet–Kolmogorov theorem; see Yosida (1980, p. 275). From this theorem and the properties of V we obtain the following characterization of compact subsets of L_V.

LEMMA 4. *Let Assumption 2 hold. Then a closed subset H of L_V is compact if and only if*

$$\sup_{h\in H} \|h\|_V < \infty, \tag{3.2}$$

$$\lim_{t\to 0} \sup_{h\in H} \int V(x)|h(x-t) - h(x)|\, dx = 0, \tag{3.3}$$

$$\lim_{K\uparrow\infty} \sup_{h\in H} \int_{|x|>K} V(x)|h(x)|\, dx = 0. \tag{3.4}$$

PROOF. By the Fréchet–Kolmogorov theorem, the set VH is compact if and only if

$$\sup_{h\in H} \int |(Vh)(x)|\, dx < \infty, \tag{3.5}$$

$$\lim_{t\to 0} \sup_{h\in H} \int \big|(Vh)(x-t) - (Vh)(x)\big|\, dx = 0, \tag{3.6}$$

$$\lim_{K\uparrow\infty} \sup_{h\in H} \int_{|x|>K} |(Vh)(x)|\, dx = 0. \tag{3.7}$$

Clearly, (3.2) and (3.5) are equivalent, and so are (3.4) and (3.7). It remains to show that (3.3) is equivalent to (3.6) under (3.2) and (3.1). For a measurable function h we have

$$\Big|V(x)|h(x-t) - h(x)| - |(Vh)(x-t) - (Vh)(x)|\Big| \leq |V(x) - V(x-t)||h(x-t)|,$$

so that the substitution $x = y + t$ and (2.5) yield

$$\int \Big| V(x)|h(x-t) - h(x)| - |(Vh)(x-t) - (Vh)(x)| \Big| \, dx \le D_t \|h\|_V,$$

with D_t the left-hand side of (3.1). This shows that (3.6) and (3.3) are equivalent under (3.2) and (3.1). □

LEMMA 5. *Let Assumption 1 hold. Let H be a compact subset of L_V and $k \in L_V$. Then*

$$\sup_{h \in H} \int V(y) \Big| \int (h(y-bt) - h(y))k(t)\, dt \Big| \, dy \to 0 \quad \text{as } b \to 0. \tag{3.8}$$

PROOF. Without loss of generality we may assume that $|b| \le 1$. For such a b let ψ_b denote the function defined by

$$\psi_b(t) = \sup_{h \in H} \int V(y)|h(y-bt) - h(y)|\, dy, \quad t \in \mathbb{R}.$$

Then the left-hand side of (3.8) is bounded by $\int \psi_b(t)|k(t)|\, dt$. Since H is compact, we obtain from (3.3) that $\psi_b(t) \to 0$ as $b \to 0$, for each $t \in \mathbb{R}$. We also have $\psi_b(t) \le \sup_{h \in H} \|h\|_V (1+V(t))$ for all $t \in \mathbb{R}$. Thus the desired result follows from an application of the Lebesgue dominated convergence theorem. □

REMARK 5. Let $k_b(x) = k(x/b)/b$, $x \in \mathbb{R}$, $b > 0$, for some density k with finite V-norm. Then the lemma implies that $\|h * k_b - h\|_V \to 0$ as $b \to 0$ uniformly in h over compact subsets of L_V. In particular, if $\mathbb{G}_n$ is a tight sequence in L_V, then $\|\mathbb{G}_n * k_b - \mathbb{G}_n\|_V = o_p(1)$ if $b = b_n \to 0$.

4. Expansions in L_V

In this section we study continuity and Taylor expansions of shifts $h(\cdot - t)$ in L_V.

LEMMA 6. *Let Assumption 1 hold. Let h be V-smooth and $t \in \mathbb{R}$. Set $h_t(x) = h(x-t)$ and $h'_t(x) = h'(x-t)$. Then*

$$\|h_t - h\|_V \le \|h'\|_V |t| V(t), \tag{4.1}$$

$$\|h_t - h + th'\|_V \le |t| \sup_{|s| \le |t|} \|h'_s - h'\|_V. \tag{4.2}$$

If h is V-smooth of order $1+\gamma$, we have

$$\|h_t - h + th'\|_V \le C|t|^{1+\gamma} V(t). \tag{4.3}$$

PROOF. Since h is absolutely continuous, we can write

$$h(x-t) - h(x) = -\int_0^1 th'(x-ut)\, du, \quad x \in \mathbb{R}. \tag{4.4}$$

From this and (2.4) we obtain that the left-hand side of (4.1) is bounded by

$$|t| \int_0^1 \int V(x)|h'(x-ut)|\, dx\, du \le |t| V(t) \|h'\|_V.$$

The left-hand side of (4.2) is bounded by

$$|t| \int_0^1 \|h'_{ut} - h'\|_V \, du \le |t| \sup_{|s| \le |t|} \|h'_s - h'\|_V.$$

If h' is V-Hölder with exponent γ, then we have $\|h'_s - h'\|_V \le CV(s)|s|^\gamma \le CV(t)|t|^\gamma$ for $|s| \le |t|$, and (4.3) follows from this and (4.2). $\square$

LEMMA 7. *Let Assumption 1 hold. Let $k_b(x) = k(x/b)/b$ for $b > 0$ and some density k with zero mean. Then the following are true.*

(1) *If h is V-Hölder with exponent γ and $\int |t|^\gamma V(t)\,dt$ is finite, we have $\|h * k_b - h\|_V = O(b^\gamma)$.*
(2) *If h is V-smooth and $\int |t|V(t)k(t)\,dt$ is finite, we have $\|h * k_b - h\|_V = o(b)$.*
(3) *If h is V-smooth of order $1+\gamma$ and $\int |t|^{1+\gamma}V(t)k(t)\,dt$ is finite, we have $\|h*k_b - h\|_V = O(b^{1+\gamma})$.*

PROOF. Part (1) is immediate. Since $\int tk(t)\,dt = 0$, we can write

$$h * k_b(x) - h(x) = \int \big(h(x-bt) - h(x) + bth'(x)\big)k(t)\,dt.$$

Part (2) is now an easy consequence of (4.2) and the Lebesgue dominated convergence theorem; note that the supremum in (4.2) is bounded by $(1 + V(t))\|h'\|_V$. Part (3) follows from (4.3). $\square$

LEMMA 8. *Let Assumption 1 hold. Let h in L_V be of bounded variation and let $\int V\,d\nu_h$ be finite with ν_h the measure of variation of h. Then h is V-Hölder with exponent 1 and $h * h$ is Lipschitz and V-smooth of order 2.*

PROOF. Let us first show that

$$\int V(x)|h(x-t) - h(x)|\,dx \le |t|V(t)\int V(y)\,d\nu_h(y), \quad t \in \mathbb{R}. \tag{4.5}$$

This bound shows that h is V-Hölder with exponent 1. The bound is clear for $t = 0$. For $t > 0$, we can bound the left-hand side of (4.5) by

$$\int V(x)\nu_h((x-t,x])\,dx = \int \Big(\int_{[y,y+t)} V(x)dx\Big)\,d\nu_h(y) \le tV(t)\int V(y)\,d\nu_h(y).$$

Here we used that $V(x) \le V(y)V(t)$ for $y \le x \le x+t$. The same arguments can be used to obtain the bound with $tV(t)$ replaced by $-tV(t)$ for negative t.

Now set $\chi = \chi_1 - \chi_2$, where

$$\chi_i(x) = \int h(x-y)\,d\mu_i(y), \quad x \in \mathbb{R},\ i = 1,2.$$

Then straightforward calculations show that

$$\|\chi\|_V \le \int V(x+y)|h(x)|\,dx\,d\nu_h(y) \le \|h\|_V \int V\,d\nu_h.$$

Next, using the substitution $u = x - y$ we calculate for $i = 1,2$ that

$$\begin{aligned}\int_{-\infty}^{z} \chi_i(x)\,dx &= \iint_{x<z} h(x-y)\,d\mu_i(y)\,dx = \iint_{u<z-y} h(u)\,d\mu_i(y)\,du \\ &= \int \mu_i((-\infty, z-u])h(u)\,du, \quad z \in \mathbb{R}.\end{aligned}$$

This shows that $h * h$ is absolutely continuous with almost everywhere derivative χ. As χ has finite V-norm, $h * h$ is V-smooth. Since h is bounded, so is χ. More precisely, $\|\chi\|_\infty \le \|h\|_\infty \nu_h(\mathbb{R})$. This shows that $h * h$ is Lipschitz. Straightforward calculations show that

$$\chi(x) = \mu_1 * \mu_1((-\infty,x]) + \mu_2 * \mu_2((-\infty,x]) - 2\mu_1 * \mu_2((-\infty,x]), \quad x \in \mathbb{R}.$$

This shows that χ is of bounded variation with measure of variation $\nu_\chi = \mu_1 * \mu_1 + \mu_2 * \mu_2 + 2\mu_1 * \mu_2 = \nu_h * \nu_h$. We have $\int V\,d\nu_h * \nu_h \le \int V\,d\nu_h \int V\,d\nu_h$, so that $\int V\,d\nu_\chi$ is finite. Thus, by what we have already shown, χ is V-Hölder with exponent 1. This shows that $h * h$ is V-smooth of order 2. $\square$

For $V = 1$, we have the following result.

Corollary 2. *Let h be an integrable function of bounded variation. Then h is L_1-Hölder with exponent one, and $h*h$ is Lipschitz with an a.e.-derivative that is integrable and L_1-Hölder with exponent one.*

References

Frees, E. W. (1994). Estimating densities of functions of observations. *J. Amer. Statist. Assoc.* **89**, 517–525.

Giné, E. and and Mason, D. M. (2005). On local U-statistic processes and the estimation of densities of functions of several variables. Technical Report, Department of Mathematics, University of Connecticut.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces. Isoperimetry and Processes.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3) 23, Springer, Berlin.

Saavedra, A. and Cao, R. (2000). On the estimation of the marginal density of a moving average process. *Canad. J. Statist.* **28**, 799–815.

Schick, A. and Wefelmeyer, W. (2004). Root n consistent density estimators for sums of independent random variables. *J. Nonparametr. Statist.* **16**, 925–935.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics.* Springer Series in Statistics. Springer, New York.

Yosida, K. (1980). *Functional Analysis*, 6th ed. Grundlehren der mathematischen Wissenschaften 123, Springer, Berlin.

Anton Schick
Binghamton University
Department of Mathematical Sciences
Binghamton, NY 13902-6000, USA
E-mail address: anton@math.binghamton.edu
URL: http://math.binghamton.edu/anton/index.html

Wolfgang Wefelmeyer
Universität zu Köln
Mathematisches Institut
Weyertal 86-90
50931 Köln, Germany
E-mail address: wefelm@math.uni-koeln.de
URL: http://www.mi.uni-koeln.de/~wefelm/