Journal of the Korean Statistical Society (2002), **31**: *2, pp 261–276*

Nonparametric Discontinuity Point Estimation in Density or Density Derivatives

Jib Huh[1]

Abstract

Probability density or its derivatives may have a discontinuity/change point at an unknown location. We propose a method of estimating the location and the jump size of the discontinuity point based on kernel type density or density derivatives estimators with one-sided equivalent kernels. The rates of convergence of the proposed estimators are derived, and the finite-sample performances of the methods are illustrated by simulated examples.

Keywords. Kernel density estimation, one-sided kernel, rate of convergence, two-sided Brownian motion, weak convergence.
AMS 2000 subject classifications. Primary 62G05, 62G07; Secondary 62E20.

1. Introduction

We observe independent and identically distributed X_i, for $1 \leq i \leq n$, and want to estimate their common probability density function f. Nonparametric density estimators are employed in order to estimate a smooth density with no parametric assumptions. To do so, one usually assumes that the density has at least two continuous derivatives. In practice, however, we are often interested in estimating a density function which has some discontinuity points in itself or in its derivatives. The usual nonparametric approaches suffer from poor practical and theoretical performance in such situations. In kernel-based approach, Cline and Hart (1991) gave expressions for the asymptotic mean integrated squared error in case the density or its derivatives has simple discontinuity points. When a density of survival times or its derivatives has a discontinuity point, Müller and Wang (1990) proposed a kernel type method to estimate the discontinuity point.

Received January 2002; accepted May 2002.
[1]Department of Mathematics, Kyung Hee University, Yongin, Kyungki-do 449-701, Korea

Figure 1.1 displays a truncated double exponential density represented by the solid curve. The density is discontinuous at the point 0.5. To show the effect of the discontinuity point in kernel density estimation, we generate a pseudo sample of size 500 from the density. The ordinary kernel density estimate, denoted by dotted curve in Figure 1.1, has large bias near the discontinuity point. Schuster (1985) used the data symmetrized about a boundary point to improve the density estimate. When the density has a known location of the discontinuity point, Cline and Hart (1991) gave a slightly generalized version of Schuster's estimator to reduce the bias near the discontinuity point. The dashed curve in Figure 1.1 is the estimated density by their method. Note that this method needs the knowledge of the location of the discontinuity point to reduce the bias. In this paper, our interest is to estimate both the location and the jump size of the discontinuity point in the νth derivative of f.

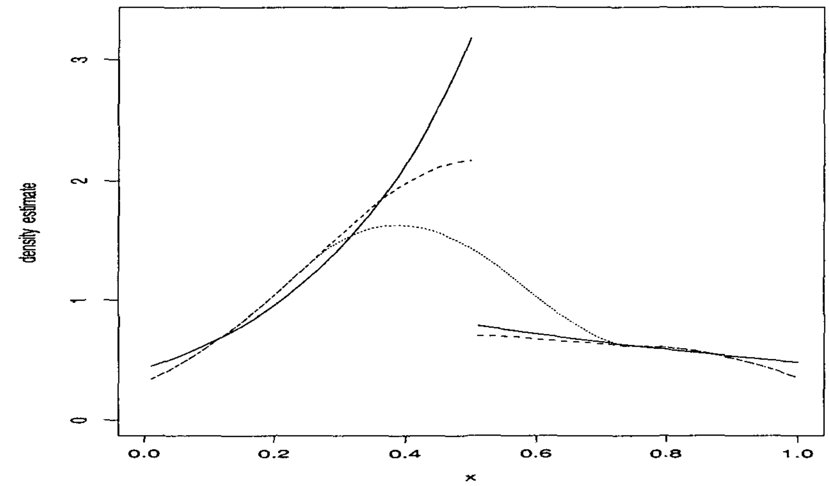

FIGURE 1.1 *Ordinary kernel density estimate (dotted curve) and Cline and Hart's estimate (dashed curve) for a pseudo sample of size 500 from a truncated double exponential density defined in Section 3 (bandwidth=0.25, Epanechnikov kernel)*

We first introduce the density derivative function having a discontinuity point. For the sake of definiteness, suppose that the density f is supported on $[0, 1]$. We assume that a discontinuity point exists for the νth derivative of f, denoted by $f^{(\nu)}$, at some point τ in the interior of the support of f, as given in the following assumption:

(A.1) There exists a constant L such that

$$|f^{(\nu)}(x) - f^{(\nu)}(y)| \leq L|x - y| \text{ whenever } (x - \tau)(y - \tau) > 0. \tag{1.1}$$

The jump size at the discontinuity point τ in the νth derivative of f is given by $\Delta_\nu = f_+^{(\nu)}(\tau) - f_-^{(\nu)}(\tau)$ where $f_+^{(\nu)}(\tau) = \lim_{x\to\tau+} f^{(\nu)}(x)$ and $f_-^{(\nu)}(\tau) = \lim_{x\to\tau-} f^{(\nu)}(x)$. We assume $0 < \Delta_\nu < \infty$. The case of $-\infty < \Delta_\nu < 0$ can be treated analogously.

The motivation of detecting the location of the discontinuity point is to calculate the difference between right and left side estimators of $f^{(\nu)}(x)$ for all x. We can define the point, which maximizes the differences, as the estimator of the location of the discontinuity point. Müller (1992), Loader (1996), and Huh and Park (2001) used this idea with some one-sided kernels in nonparametric estimation of discontinuity point for a regression function itself or its derivatives. In Section 2, the estimators for the location of the discontinuity point and for the corresponding jump size are proposed, and their asymptotic properties are given. The finite sample performances of the methods are investigated in Section 3 through two simulated examples. The proofs of the asymptotic results are given in Section 4.

2. Discontinuity Point Estimations

Here, we take a one-sided kernel K described in Loader (1996), and Huh and Park (2001) with support $[0, 1]$ as follows:

(A.2) The function K satisfies $\int_0^1 K(u)du = 1$, and $K(0) > 0, K(u) \geq 0$ for $0 < u \leq 1$, and its first derivative K' satisfies the Lipschitz condition of order 1.

Huh and Park (2001) used the local polynomial fits related to equivalent kernels to estimate the location of the discontinuity point in the νth derivative regression function. Equivalent kernels discussed in Müller (1987) and in Fan and Gijbels (1996, p.103) satisfy a moment condition which insures the consistency and improves the order of bias as well. Note that these equivalent kernels can be viewed as boundary kernels. See, for examples, Fan and Gijbels (1996, Section 3.2.5), and Wand and Jones (1995, p.47). We can build the equivalent kernels to estimate right and left side estimators of $f^{(\nu)}$ using the kernel function K satisfying the assumption (A.2).

Let $\boldsymbol{S}$ and $\boldsymbol{T}$ be $(\nu+1)\times(\nu+1)$ matrices having their (i,j)th entries equal to $\int_0^1 K(u)u^{i+j}du$ and $\int_0^1 K(u)(-u)^{i+j}du$, respectively. Here and below, we use the convention that indices for the entries of a matrix or a vector count from 0. Define $\boldsymbol{e}_\nu = (0,\cdots,0,1)^T$ with 1 appearing at the νth position. Now we define

$$K_\nu^+(u) = \boldsymbol{e}_\nu^T \boldsymbol{S}^{-1}(1,u,\ldots,u^\nu)^T K(u), \quad K_\nu^-(u) = \boldsymbol{e}_\nu^T \boldsymbol{T}^{-1}(1,-u,\ldots,(-u)^\nu)^T K(u)$$

as the one-sided equivalent kernels. They satisfy the following moment conditions (see, for example, Fan and Gijbels, 1996, p. 103):

$$\int u^l K_\nu^+(u) = \delta_{\nu,l} \text{ and } \int (-u)^l K_\nu^-(u) = \delta_{\nu,l}, \quad 0 \le l \le \nu, \tag{2.1}$$

where δ is the Kronecker delta function.

We define $\widehat{f}_+^{(\nu)}(x)$ and $\widehat{f}_-^{(\nu)}(x)$ as the right and the left side estimators of $f^{(\nu)}(x)$ respectively as follows:

$$\widehat{f}_+^{(\nu)}(x) = \frac{\nu!}{nh^{\nu+1}}\sum_{i=1}^n K_\nu^+\left(\frac{X_i - x}{h}\right), \quad \widehat{f}_-^{(\nu)}(x) = \frac{\nu!}{nh^{\nu+1}}\sum_{i=1}^n K_\nu^-\left(\frac{x - X_i}{h}\right) \tag{2.2}$$

where $h = h_n$ is a sequence of bandwidths, which satisfies the following assumption:

(A.3) $h \to 0$, $nh^{2\nu+1} \to \infty$, and $nh^{2\nu+3} \to 0$, as $n \to \infty$.

The estimators $\widehat{f}_+^{(\nu)}(x)$ and $\widehat{f}_-^{(\nu)}(x)$ are based on the one-sided data at the right and the left of x, respectively. We estimate the jump size at a point x by taking the differences of these two estimators: $\widehat{\Delta}_\nu(x) = \widehat{f}_+^{(\nu)}(x) - \widehat{f}_-^{(\nu)}(x)$. A reasonable estimator $\widehat{\tau}$ of τ is the value of x that maximizes $\widehat{\Delta}_\nu(x)$. Let $Q \subset (0,1)$ be a closed interval such that $\tau \in Q$. Define

$$\widehat{\tau} = \inf\left\{z \in Q : \widehat{\Delta}_\nu(z) = \sup_{x\in Q}\widehat{\Delta}_\nu(x)\right\}$$

for the location of the discontinuity point τ. An estimator of the jump size Δ_ν may be obtained by

$$\widehat{\Delta}_\nu(\widehat{\tau}) = \widehat{f}_+^{(\nu)}(\widehat{\tau}) - \widehat{f}_-^{(\nu)}(\widehat{\tau}). \tag{2.3}$$

It is difficult to show the asymptotic properties of $\widehat{\tau}$ and $\widehat{\Delta}_\nu(\widehat{\tau})$ directly since we cannot have the explicit formulas of the estimators. First, we describe in the following theorem weak convergence of the sequence of the process $\{\varphi_{n,\nu}(z) : -M \le z \le M\}$ where

$$\varphi_{n,\nu}(z)=\begin{cases} (nh^{2\nu+1})^{\frac{\nu+1}{2\nu+1}} \left\{\widehat{\Delta}_\nu\left(\tau+\frac{h}{(nh^{2\nu+1})^{1/(2\nu+1)}}z\right) - \widehat{\Delta}_\nu(\tau)\right\}, & \nu : \text{even}, \\ (nh^{2\nu+1})^{\frac{\nu+1}{2\nu}} \left\{\widehat{\Delta}_\nu\left(\tau+\frac{h}{(nh^{2\nu+1})^{1/2\nu}}z\right) - \widehat{\Delta}_\nu(\tau)\right\}, & \nu : \text{odd}, \end{cases} \tag{2.4}$$

and $M < \infty$. Existence of the unique maximizer of the limit of the process $\varphi_{n,\nu}$ will be discussed later on. The process $\varphi_{n,\nu}$ lies in the space, denoted by $\mathcal{C}[-M, M]$, of continuous functions defined on $[-M, M]$. To obtain the theorem, consider the following additional assumption:

(A.4) $K_\nu^-(0) > 0$.

Let $\xrightarrow{\mathcal{W}}$ denote weak convergence in the space $\mathcal{C}([-M, M])$, and denote the derivatives of $K_\nu^\pm$ by $K_\nu^{'\pm}$.

Theorem 2.1. *Suppose that the assumptions (A.1)-(A.4) are satisfied.*

(i) *If ν is even, then*

$$\varphi_{n,\nu}(z) \xrightarrow{\mathcal{W}} \varphi_\nu(z) = -\frac{\Delta_\nu}{\nu+1}K_\nu^-(0)|z|^{\nu+1} + \sigma_1 W(z) \tag{2.5}$$

where $W(z)$ is a two-sided Brownian motion defined in Bhattacharya and Brockwell (1976), and

$$\sigma_1 = \begin{cases} 2\nu!\sqrt{f_+(\tau)}K_\nu^-(0), & z \ge 0, \\ 2\nu!\sqrt{f_-(\tau)}K_\nu^-(0), & z < 0. \end{cases} \tag{2.6}$$

(ii) *If ν is odd, then*

$$\varphi_{n,\nu}(z) \xrightarrow{\mathcal{W}} \varphi_\nu(z) = -\frac{\Delta_\nu}{\nu+1}K_\nu^-(0)|z|^{\nu+1} + zU \tag{2.7}$$

where U is a normal random variable with mean 0 and variance

$$\sigma_2^2 = (\nu!)^2\{f_+(\tau) + f_-(\tau)\} \int_0^1 \{K_\nu^{'-}(u)\}^2 du. \tag{2.8}$$

Note that Huh and Park (2001) showed the following relation

$$K_\nu^+(u) = (-1)^\nu K_\nu^-(u) \tag{2.9}$$

for $0 \le u \le 1$. This relation makes it possible to state the theorems in terms of K_ν^- and $K_\nu^{'-}$ only.

Now we describe the asymptotic distribution of $\hat{\tau}$. Let Z_ν be the maximizer of the process φ_ν. By (A.4) the limit process φ_ν in (2.5) has a unique maximizer with probability one when ν is even (see Remark 5.3 in Bhattacharya and Brockwell, 1976). For odd ν, the continuous Gaussian process φ_ν in (2.7) has a unique maximum at

$$Z_\nu = \left[\frac{U}{\Delta_\nu K_\nu^-(0)}\right]^{1/\nu}.$$

Let $Z_{n,\nu}$ be the maximizer of $\varphi_{n,\nu}$. By construction,

$$\hat{\tau} = \begin{cases} \tau + Z_{n,\nu}\dfrac{h}{(nh^{2\nu+1})^{1/(2\nu+1)}}, & \nu : \text{even}, \\ \tau + Z_{n,\nu}\dfrac{h}{(nh^{2\nu+1})^{1/2\nu}}, & \nu : \text{odd}. \end{cases}$$

By Theorem 5 in Whitt (1970), the weak convergence in Theorem 2.1 can be extended to the space $\mathcal{C}(-\infty,\infty)$. Theorem 3 in Bhattacharya and Brockwell (1976) then gives $Z_{n,\nu} \xrightarrow{\mathcal{D}} Z_\nu$, where $Z_{n,\nu}$ is the global maximizer of $\varphi_{n,\nu}$ on $(-\infty,\infty)$. Therefore, we have the following corollary.

Corollary 2.1. *Suppose that the assumptions in Theorem 2.1 are satisfied.*

(i) *If ν is even, then*

$$n^{1/(2\nu+1)}(\hat{\tau}-\tau) \xrightarrow{\mathcal{D}} \underset{z\in(-\infty,\infty)}{\operatorname{argmax}} \left\{-\frac{\Delta_\nu}{\nu+1}K_\nu^-(0)|z|^{\nu+1} + \sigma_1 W(z)\right\}.$$

(ii) *If ν is odd, then*

$$\sqrt{nh}\,(\hat{\tau}-\tau)^\nu \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\sigma_2^2}{\Delta_\nu^2(K_\nu^-(0))^2}\right).$$

According to Corollary 2.1, the rate of convergence of $\hat{\tau}$ decays rapidly as ν increases. The proposed estimator achieves the rate $n^{-1/(2\nu+1)}$ when ν is even. On the other hand, from (2.9), we have

$$K_\nu^+(0) = (-1)^\nu K_\nu^-(0), \tag{2.10}$$

which makes that the rate of $\hat{\tau}$ is faster than the rate $n^{-1/(2\nu+1)}$ when ν is odd. The result (ii) in Corollary 2.1 shows that the rate is $(1/nh)^{1/2\nu}$. If we choose the bandwidth $h \sim n^{-1/(2\nu+2)}$, which satisfies (A.3), $\widehat{\tau}$ achieves the rate $n^{-\alpha}$ where $\alpha = (2\nu+1)/\{2\nu(2\nu+2)\}$. For instance, a discontinuity point of the density function itself can be estimated at the rate $O_p(n^{-1})$, whereas a discontinuity point of the first derivative is detectable at the rate of $O_p(n^{-3/8})$. Note that Huh and Park (2001) showed their estimator of the location of the discontinuity point in a regression function or its derivatives has the same rate of convergence.

The following theorem describes the asymptotic distribution of the estimator $\widehat{\Delta}_\nu(\widehat{\tau})$ for the jump size defined in (2.3).

Theorem 2.2. *Under the assumptions of Theorem 2.1,*

$$\sqrt{nh^{2\nu+1}}(\widehat{\Delta}_\nu(\widehat{\tau}) - \Delta_\nu) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, (\nu!)^2\{f_+(\tau) + f_-(\tau)\} \int_0^1 \{K_\nu^-(u)\}^2 du\right).$$

If we are interested in a test whether a discontinuity point exists or not, the following test statistic can be used for the hypothesis $H_0 : \Delta_\nu = 0$ versus $H_1 : \Delta_\nu > 0$. Under H_0, the test statistic is

$$\frac{\sqrt{nh^{2\nu+1}}\widehat{\Delta}_\nu(\widehat{\tau})}{\nu!\sqrt{2\widehat{f}(\widehat{\tau}) \int_0^1 \{K_\nu^-(u)\}^2 du}}$$

where $\widehat{f}(\widehat{\tau})$ is the ordinary kernel density estimate at the estimated location $\widehat{\tau}$. By Theorem 2.2, the test statistic tends in distribution to the standard normal distribution $\mathcal{N}(0,1)$.

3. Numerical Experiments

To investigate the performance of the proposed estimators, simulation studies are carried out. We consider two probability density functions. The first density displayed in Figure 1.1 is

$$f_1(x) = p_1\left\{\frac{\lambda_1}{2}\exp(-\lambda_1|x-0.5|)1_{[0\le x<0.5]} + \frac{\lambda_2}{2}\exp(-\lambda_2|x-0.5|)1_{[0.5\le x\le 1]}\right\}$$

where $\lambda_1 = 4$, $\lambda_2 = 1$ and p_1 is the normalizing constant to make f_1 a proper density. Then we have the discontinuity point $\tau = 0.5$ and the jump size $\Delta_0 = p_1(\lambda_2 - \lambda_1)/2$. The second example concerns with a density which has a discontinuity point in the first derivative. The density f_2 is given by

$$f_2(x) = p_2\frac{\lambda}{2}\exp(-\lambda|x-0.5|)1_{[0\le x\le 1]}$$

where $\lambda = 7$ and p_2 is the normalizing constant to make f_2 a proper density. The first derivative of the density f_2 has a discontinuity point at $\tau = 0.5$ and its jump size $\Delta_1 = -p_2\lambda^2$.

We consider the one-sided kernel function $K(x) = \frac{3}{2}(1 - x^2)1_{[0 \le x \le 1]}$ which makes the one-sided equivalent kernels $K_1^{\pm}$ for estimations of τ and Δ_1 in the density f_2. Table 3.1 contains the results of the simulation based on 1000 pseudo samples of size 500. To estimate the location of the discontinuity point, we first compute the jump sizes at $x_k = k/100$, $k = 1, \cdots, 100$, and then choose a point which maximizes the absolute value of the calculated jump sizes over the interval Q. As suggested in Müller (1992), we take $Q = [h, 1 - h]$ for our simulation settings. We compute the Monte Carlo estimates of the mean squared errors (MSE) of the proposed estimators for various values of bandwidth h. But, only the minimum MSEs of $\widehat{\Delta}$ are reported here with their corresponding bandwidths, averages of $\widehat{\Delta}$ and $\widehat{\tau}$, and MSEs of $\widehat{\tau}$.

TABLE 3.1 *The Monte Carlo estimates of the MSEs and the averages with standard errors in parentheses for the discontinuity point estimators*

Density	*h*	*Average of* $\widehat{\Delta}$	*MSE of* $\widehat{\Delta}$	*Average of* $\widehat{\tau}$	*MSE of* $\widehat{\tau}$
f_1	0.08	−2.055786	0.231873	0.499630	5.1700×10^{-5}
		(0.011128)	(0.021141)	(0.000227)	(2.9847×10^{-5})
f_2	0.11	−40.884433	222.547819	0.499610	0.000501
		(0.359990)	(16.911937)	(0.000708)	(0.000068)

Table 3.2 and 3.3 show the frequencies with which discontinuities identified by the 1000 replications using the bandwidths in Table 3.1. Here the integer k denoted by the index for the point x_k which maximizes the absolute value of the estimated jump sizes.

TABLE 3.2 *Discontinuity point identification frequency in the case of* f_1

k (location)	35	36	47	48	49	50	51	52
frequency	1	1	2	3	28	936	26	3

According to Table 3.1, and Table 3.2 and 3.3, the performances of the estimators of the density f_2 seem to be inferior to those of the density f_1 in these simulation settings since the convergent rates of the proposed estimators are rapidly decreases as ν increases, which is described in Corollary 2.1.

TABLE 3.3 *Discontinuity point identification frequency in the case of f_2*

k (location)	36	42	43	44	45	46	47	48	49	50
frequency	1	1	1	2	10	9	43	105	216	290
k (location)	51	52	53	54	55	56	57	58	59	
frequency	168	89	38	14	8	2	1	1	1	

4. Proofs

We need the following two lemmas to prove Theorem 2.1.

Lemma 4.1. *Suppose that the assumptions in Theorem 2.1 are satisfied. Then,*

$$E(\varphi_{n,\nu}(z)) = -\frac{\Delta_\nu}{\nu+1}K_\nu^-(0)|z|^{\nu+1} + o(1)$$

uniformly in $z \in [-M, M]$.

Proof. For simplicity, define

$$a_{n,\nu} = \begin{cases} (nh^{2\nu+1})^{1/(2\nu+1)}, & \text{when } \nu \text{ is even}, \\ (nh^{2\nu+1})^{1/2\nu}, & \text{when } \nu \text{ is odd}, \end{cases}$$

and $z_{n,\nu} = (h/a_{n,\nu})z$. Let

$$C_{n,\nu}^+(u,z) = K_\nu^+\left(\frac{u-\tau-z_{n,\nu}}{h}\right) - K_\nu^+\left(\frac{u-\tau}{h}\right)$$
$$C_{n,\nu}^-(u,z) = K_\nu^-\left(\frac{\tau+z_{n,\nu}-u}{h}\right) - K_\nu^-\left(\frac{\tau-u}{h}\right).$$

Then, the process $\varphi_{n,\nu}$ can be written as

$$\varphi_{n,\nu}(z) = a_{n,\nu}^{\nu+1}\frac{\nu!}{nh^{\nu+1}}\sum_{i=1}^{n}\{C_{n,\nu}^+(X_i,z) - C_{n,\nu}^-(X_i,z)\}$$

for all $z \in [-M, M]$.

We prove the lemma for $z \geq 0$ only. The other case is analogous. Note that

$$E[C_{n,\nu}^\pm(X_1,z)] = h\int K_\nu^\pm(u)\{f(\tau+z_{n,\nu}\pm hu) - f(\tau\pm hu)\}du. \tag{4.1}$$

For all $0 \le u \le 1$,

$$f(\tau + z_{n,\nu} + hu) = \sum_{l=0}^{\nu} \frac{f_+^{(l)}(\tau)}{l!}(z_{n,\nu} + hu)^l + \frac{f_+^{(\nu)}(\tau_+') - f_+^{(\nu)}(\tau)}{\nu!}(z_{n,\nu} + hu)^\nu, \quad (4.2)$$

$$f(\tau \pm hu) = \sum_{l=0}^{\nu} \frac{f_\pm^{(l)}(\tau)}{l!}(\pm hu)^l + \frac{f_\pm^{(\nu)}(\tau_\pm) - f_\pm^{(\nu)}(\tau)}{\nu!}(\pm hu)^\nu \quad (4.3)$$

where τ_+' lies between τ and $\tau + z_{n,\nu} + hu$, and $\tau_\pm$ lie between τ and $\tau \pm hu$, respectively. By the difference (4.2) and (4.3), and the moment condition of K_ν^+, $E\big[C_{n,\nu}^+(X_1, z)\big] = O(h^{\nu+3}/a_{n,\nu})$. However, the case $E\big[C_{n,\nu}^-(X_1, z)\big]$ is slightly different since the discontinuity point τ lies between $\tau + z_{n\nu} - hu$ and $\tau + z_{n,\nu}$. We divide the interval of integration into two parts. Note that, for $0 \le u \le z/a_{n,\nu}$,

$$f(\tau + z_{n,\nu} - hu) = \sum_{l=0}^{\nu} \frac{f_+^{(l)}(\tau)}{l!}(z_{n,\nu} - hu)^l + \frac{f_+^{(\nu)}(\tau_+'') - f_+^{(\nu)}(\tau)}{\nu!}(z_{n,\nu} - hu)^\nu \quad (4.4)$$

where τ_+'' lies between τ and $\tau + z_{n,\nu} - hu$. However, for $z/a_{n,\nu} < u \le 1$,

$$f(\tau + z_{n,\nu} - hu) = \sum_{l=0}^{\nu} \frac{f_-^{(l)}(\tau)}{l!}(z_{n,\nu} - hu)^l + \frac{f_-^{(\nu)}(\tau_-') - f_-^{(\nu)}(\tau)}{\nu!}(z_{n,\nu} - hu)^\nu \quad (4.5)$$

where τ_-' lies between τ and $\tau + z_{n,\nu} - hu$. The first term of the right-hand side in (4.4) can be written as

$$\sum_{l=0}^{\nu} \frac{f_-^{(l)}(\tau)}{l!}(z_{n,\nu} - hu)^l + \frac{\Delta_\nu}{\nu!}(z_{n,\nu} - hu)^\nu. \quad (4.6)$$

By (4.3), (4.4), (4.5) and (4.6), $E\big[C_{n,\nu}^-(X_1, z)\big]$ equals

$$\begin{aligned}
&h\int_0^{z/a_{n,\nu}} K_\nu^-(u)\frac{\Delta_\nu}{\nu!}(z_{n,\nu} - hu)^\nu du \\
&+ h\int_0^1 K_\nu^-(u)\sum_{l=0}^{\nu}\frac{f_-^{(l)}(\tau)}{l!}\left\{(z_{n,\nu} - hu)^l - (-hu)^l\right\} du \\
&+ h\int_0^{z/a_{n,\nu}} K_\nu^-(u)\frac{f_+^{(\nu)}(\tau_+'') - f_+^{(\nu)}(\tau)}{\nu!}(z_{n,\nu} - hu)^\nu du \\
&+ h\int_{z/a_{n,\nu}}^1 K_\nu^-(u)\frac{f_-^{(\nu)}(\tau_-') - f_-^{(\nu)}(\tau)}{\nu!}(z_{n,\nu} - hu)^\nu du \\
&- h\int K_\nu^-(u)\frac{f_-^{(\nu)}(\tau_-) - f_-^{(\nu)}(\tau)}{\nu!}(-hu)^\nu du
\end{aligned}$$

uniformly in z. The second integral is zero by the moment condition of K_ν^-. Then,

$$E\big[C_{n,\nu}^-(X_1,z)\big] = h\int_0^{z/a_{n,\nu}} K_\nu^-(u)\frac{\Delta_\nu}{\nu!}\,(z_{n,\nu}-hu)^\nu\,du \;+O\left(\frac{h^{\nu+2}}{a_{n,\nu}}\right) \qquad (4.7)$$

uniformly in z. Therefore, we have that

$$E(\varphi_{n,\nu}(z)) = -a_{n,\nu}^{\nu+1}\Delta_\nu\int_0^{z/a_{n,\nu}} K_\nu^-(u)\Big(\frac{z}{a_{n,\nu}}-u\Big)^\nu du + O\left(a_{n,\nu}^\nu h\right)$$

uniformly in $z\in[0,M]$. Since $K_\nu^-(u) = K_\nu^-(0)(1+o(1))$ uniformly for $u\in[0,M/a_{n,\nu}]$, the result follows. □

Lemma 4.2. *Suppose that the assumptions in Theorem 2.1 are satisfied.*

(i) *If ν is even, then*

$$Cov(\varphi_{n,\nu}(z_1),\varphi_{n,\nu}(z_2)) = \begin{cases} 4(\nu!)^2 f_+(\tau)\min(|z_1|,|z_2|)\{K_\nu^-(0)\}^2+o(1), & z_1,z_2\ge 0,\\ 4(\nu!)^2 f_-(\tau)\min(|z_1|,|z_2|)\{K_\nu^-(0)\}^2+o(1), & z_1,z_2<0,\\ o(1), & elsewhere\end{cases}$$

uniformly in $z_1,z_2\in[-M,M]$.

(ii) *If ν is odd, then*

$$Cov(\varphi_{n,\nu}(z_1),\varphi_{n,\nu}(z_2)) = (\nu!)^2\big\{f_+(\tau)+f_-(\tau)\big\}z_1z_2\int_0^1\{K_\nu^{'-}(u)\}^2du+o(1)$$

uniformly in $z_1,z_2\in[-M,M]$.

Proof. We prove the lemma for $z_1,z_2>0$ first. By Lemma 1,

$$\begin{aligned} &Cov(\varphi_{n,\nu}(z_1),\varphi_{n,\nu}(z_2))\\ &= n(\nu!)^2\frac{a_{n,\nu}^{2(\nu+1)}}{(nh^{\nu+1})^2}E\Big[C_{n,\nu}^+(X_1,z_1)C_{n,\nu}^+(X_1,z_2) - C_{n,\nu}^+(X_1,z_1)C_{n,\nu}^-(X_1,z_2)\\ &\quad - C_{n,\nu}^-(X_1,z_1)C_{n,\nu}^+(X_1,z_2) + C_{n,\nu}^-(X_1,z_1)C_{n,\nu}^-(X_1,z_2)\Big] + O\Big(\frac{1}{n}\Big). \end{aligned} \qquad (4.8)$$

Define $z_{\min} = \min(z_1,z_2)$, $z_{\max}=\max(z_1,z_2)$, $\tau_{n,\nu}^{\min} = \tau+(h/a_{n,\nu})z_{\min}$ and $\tau_{n,\nu}^{\max} = \tau+(h/a_{n,\nu})z_{\max}$. Since $C_{n,\nu}^{\pm}(u,z) = O(a_{n,\nu}^{-1})$ uniformly in u and z by (A.2), we have that

$$E\Big[C^{+}_{n,\nu}(X_1,z_1)C^{+}_{n,\nu}(X_1,z_2)\Big] \tag{4.9}$$

$$= \Bigg[\int_{\tau}^{\tau^{\min}_{n,\nu}} \left\{K^{+}_{\nu}\left(\frac{u-\tau}{h}\right)\right\}^2 + \int_{\tau^{\min}_{n,\nu}}^{\tau^{\max}_{n,\nu}} C^{+}_{n,\nu}(u,z_{\min})\left\{-K^{+}_{\nu}\left(\frac{u-\tau}{h}\right)\right\}$$

$$+ \int_{\tau^{\max}_{n,\nu}}^{\tau^{\max}_{n,\nu}+h} C^{+}_{n,\nu}(u,z_{\min})C^{+}_{n,\nu}(u,z_{\max})\Bigg] f(u)du$$

$$= hf_{+}(\tau)\left[\{K^{+}_{\nu}(0)\}^2\frac{z_{\min}}{a_{n,\nu}} + K^{+}_{\nu}(0)K'^{+}_{\nu}(0)\frac{z_{\min}z_{\max}}{a^2_{n,\nu}}(1+o(1))\right](1+O(h))$$

$$+ \int_{\tau^{\max}_{n,\nu}}^{\tau^{\max}_{n,\nu}+h} C^{+}_{n,\nu}(u,z_{\min})C^{+}_{n,\nu}(u,z_{\max})f(u)du \tag{4.10}$$

uniformly in z_1 and z_2. Next, consider the second term in the square bracket at (4.8) for the case $z_{\min} = z_1$. The other cases can be dealt in a similar way. Since $C_{n,\nu}(u,z) = 0$ for $u < \tau$,

$$E\Big[C^{+}_{n,\nu}(X_1,z_1)C^{-}_{n,\nu}(X_1,z_2)\Big]$$

$$= \Bigg[\int_{\tau}^{\tau^{\min}_{n,\nu}} \left\{-K^{+}_{\nu}\left(\frac{u-\tau}{h}\right)\right\} K^{-}_{\nu}\left(\frac{\tau^{\max}_{n,\nu}-u}{h}\right)$$

$$+ \int_{\tau^{\min}_{n,\nu}}^{\tau^{\max}_{n,\nu}} C^{+}_{n,\nu}(u,z_{\min})K^{-}_{\nu}\left(\frac{\tau^{\max}_{n,\nu}-u}{h}\right)\Bigg] f(u)du$$

$$= -hf_{+}(\tau)\Bigg[K^{+}_{\nu}(0)K^{-}_{\nu}(0)\frac{z_{\min}}{a_{n,\nu}} + \frac{1}{2}\left\{K^{+}_{\nu}(0)K'^{-}_{\nu}(0) + K'^{+}_{\nu}(0)K^{-}_{\nu}(0)\right\}$$

$$\times \left\{2\frac{z_{\min}z_{\max}}{a^2_{n,\nu}} - \left(\frac{z_{\min}}{a_{n,\nu}}\right)^2\right\}(1+o(1))\Bigg](1+O(h)) \tag{4.11}$$

uniformly in z_1 and z_2. Analogously,

$$E\Big[C^{-}_{n,\nu}(X_1,z_1)C^{+}_{n,\nu}(X_1,z_2)\Big]$$

$$= -hf_{+}(\tau)\Bigg[K^{-}_{\nu}(0)K^{+}_{\nu}(0)\frac{z_{\min}}{a_{n,\nu}} + \frac{1}{2}\left\{K^{-}_{\nu}(0)K'^{+}_{\nu}(0) + K'^{-}_{\nu}(0)K^{+}_{\nu}(0)\right\}$$

$$\times \left(\frac{z_{\min}}{a_{n,\nu}}\right)^2(1+o(1))\Bigg](1+O(h)) \tag{4.12}$$

uniformly in z_1 and z_2. And,

$$E\Big[C^{-}_{n,\nu}(X_1,z_1)C^{-}_{n,\nu}(X_1,z_2)\Big]$$

$$= \Bigg[\int_{\tau}^{\tau^{\min}_{n,\nu}} K^{-}_{\nu}\left(\frac{\tau^{\min}_{n,\nu}-u}{h}\right)K^{-}_{\nu}\left(\frac{\tau^{\max}_{n,\nu}-u}{h}\right)$$

$$+ \int_{\tau-h}^{\tau} C_{n,\nu}^{-}(u, z_{\min}) C_{n,\nu}^{-}(u, z_{\max}) \Big] f(u) du$$
$$= h f_{+}(\tau) \Big[\{K_{\nu}^{-}(0)\}^2 \frac{z_{\min}}{a_{n,\nu}} + K_{\nu}^{-}(0) K_{\nu}^{\prime -}(0) \frac{z_{\min} z_{\max}}{a_{n,\nu}^2} (1 + o(1)) \Big] (1 + O(h))$$
$$+ \int_{\tau-h}^{\tau} C_{n,\nu}^{-}(u, z_{\min}) C_{n,\nu}^{-}(u, z_{\max}) f(u) du \tag{4.13}$$

uniformly in z_1 and z_2. These (4.9), (4.11), (4.12) and (4.13) with the relation (2.10) imply Lemma 4.2–(i) for the case $z_1, z_2 > 0$.

Next, we prove the second part of the lemma for the case $z_1, z_2 > 0$. If ν is odd, all terms in the square brackets at (4.9), (4.11), (4.12) and (4.13) are canceled due to the relation (2.10). Thus, it is enough to consider the integral terms in (4.9) and (4.13). By (A.2) and (A.4), the integral term in (4.9) can be written as

$$f_{+}(\tau) \frac{z_1 z_2}{a_{n,\nu}^2} \int_{\tau_{n,\nu}^{\max}}^{\tau_{n,\nu}^{\max}+h} \left\{ K_{\nu}^{\prime +} \left(\frac{u-\tau}{h} \right) \right\}^2 du (1 + o(1))$$
$$= h f_{+}(\tau) \frac{z_1 z_2}{a_{n,\nu}^2} \int_0^1 \{K_{\nu}^{\prime +}(u)\}^2 du (1 + o(1)) \tag{4.14}$$

uniformly in z_1 and z_2. Similarly, we can easily show that the integral term in (4.13) equals

$$f_{-}(\tau) \frac{z_1 z_2}{a_{n,\nu}^2} \int_{\tau-h}^{\tau} \left\{ K_{\nu}^{\prime -} \left(\frac{\tau-u}{h} \right) \right\}^2 du (1 + o(1))$$
$$= h f_{-}(\tau) \frac{z_1 z_2}{a_{n,\nu}^2} \int_0^1 \{K_{\nu}^{\prime -}(u)\}^2 du (1 + o(1)) \tag{4.15}$$

uniformly in z_1 and z_2. From (2.9), (4.14) and (4.15), Lemma 4.2-(ii) follows for the case $z_1, z_2 > 0$. It can be shown in a similar way that the lemma follows for the case $z_1, z_2 < 0$ too.

Now, consider the case of $z_1 > 0$, $z_2 < 0$. Following the lines in the proof for the case $z_1, z_2 > 0$, we obtain

$$\begin{aligned} E\big[C_{n,\nu}^{+}(X_1, z_1) C_{n,\nu}^{+}(X_1, z_2)\big] &= h f_{+}(\tau) \frac{z_1 z_2}{a_{n,\nu}^2} \Big[K_{\nu}^{+}(0) K_{\nu}^{\prime +}(0) + \int_0^1 \{K_{\nu}^{\prime +}(u)\}^2 du \Big] \\ &\quad \times (1 + o(1)), \\ E\big[C_{n,\nu}^{+}(X_1, z_1) C_{n,\nu}^{-}(X_1, z_2)\big] &= 0, \\ E\big[C_{n,\nu}^{-}(X_1, z_1) C_{n,\nu}^{+}(X_1, z_2)\big] &= -h \frac{z_1 z_2}{a_{n,\nu}^2} \Big[f_{-}(\tau) K_{\nu}^{\prime -}(0) K_{\nu}^{+}(0) \\ &\quad + f_{+}(\tau) K_{\nu}^{-}(0) K_{\nu}^{\prime +}(0) \Big] (1 + o(1)), \end{aligned}$$

$$E[C^-_{n,\nu}(X_1,z_1)C^-_{n,\nu}(X_1,z_2)] = hf_-(\tau)\frac{z_1z_2}{a^2_{n,\nu}}\left[K^-_\nu(0)K^{'-}_\nu(0)+\int_0^1\{K^{'-}_\nu(u)\}^2du\right] \times(1+o(1)) \tag{4.16}$$

uniformly in z_1 and z_2. Here, the second identity follows from the fact that

$$C^+_{n,\nu}(u,z_1)C^-_{n,\nu}(u,z_2)=0$$

for all w. The leading terms in (4.9) are $O(ha^{-2}_{n,\nu})$ when ν is even. For odd ν, the facts (2.9), (2.10) and (4.9) imply the result (ii) immediately. □

Proof of Theorem 2.1. To prove the theorem, we need to check the Lyapounov's condition and the tightness of $\psi_{n,\nu}(\cdot) = \varphi_{n,\nu}(\cdot) - E(\varphi_{n,\nu}(\cdot))$. The proofs follow lines similar to those in Huh and Park (2001). These can imply that $\psi_{n,\nu}(z)$, for fixed $z \in [-M,M]$, converges weakly to a normal distribution. Furthermore, by the Cramer-Wold device we may show that for fixed $z_1,\ldots,z_l,z_i \in [-M,M]$,

$$(\psi_{n,\nu}(z_1),\ldots,\psi_{n,\nu}(z_l)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\Sigma)$$

where Σ is the asymptotic covariance described in Lemma 4.2. This concludes the proof. See Theorem 8.1 and 12.3 of Billingsley (1968). □

Proof of Theorem 2.2. Theorem 2.1 shows that

$$\sqrt{nh^{2\nu+1}}(\widehat{\Delta}_\nu(\tau+(h/a_{n,\nu})Z_{n,\nu})-\widehat{\Delta}_\nu(\tau)) \xrightarrow{p} 0.$$

This implies that $\sqrt{nh^{2\nu+1}}(\widehat{\Delta}_\nu(\widehat{\tau})-\widehat{\Delta}_\nu(\tau)) \xrightarrow{p} 0$. Now,

$$\begin{aligned}&\sqrt{nh^{2\nu+1}}(\widehat{\Delta}_\nu(\widehat{\tau})-\Delta_\nu)\\ &=\sqrt{nh^{2\nu+1}}(\widehat{\Delta}_\nu(\widehat{\tau})-\widehat{\Delta}_\nu(\tau))+\sqrt{nh^{2\nu+1}}(\widehat{\Delta}_\nu(\tau)-\Delta_\nu).\end{aligned} \tag{4.17}$$

Then, we will show the asymptotic normality of the second term of the right-hand side in (4.17). One can easily show that

$$\begin{aligned}\sqrt{nh^{2\nu+1}}E[\widehat{\Delta}_\nu(\tau)] &= \nu!\frac{\sqrt{nh^{2\nu+1}}}{h}\int\{K^+_\nu(u)f(\tau+hu)-K^-_\nu(u)f(\tau-hu)\}du\\ &=\Delta_\nu+O(\sqrt{nh^{2\nu+1}}h).\end{aligned}$$

By the assumption $nh^{2\nu+3}\to 0$ in (A.5), the last term is $o(1)$. Now, since the support of K is $[0,1]$ and $E[K^\pm_\nu(\pm(X_1-\tau)/h)] = O(h^{\nu+1})$,

$$nh^{2\nu+1}\quad Var[\widehat{\Delta}_\nu(\tau)] = \frac{(\nu!)^2}{h}\left[Var\left\{K^+_\nu\left(\frac{X_1-\tau}{h}\right)\right\}+Var\left\{K^-_\nu\left(\frac{\tau-X_1}{h}\right)\right\}\right]$$

$$= (\nu!)^2 \left\{ f_+(\tau) \int \{K_\nu^+(u)\}^2 du + f_-(\tau) \int \{K_\nu^-(u)\}^2 du \right\} (1 + O(h)) + O(h^{2\nu+1}).$$

The Lyapounov's condition for $\sqrt{nh^{2\nu+1}}\widehat{\Delta}_\nu(\tau)$ can be easily verified. These together with (4.17) imply Theorem 2.2. □

REFERENCES

Bhattacharya, P. K. and Brockwell, P. J. (1976). "The minimum of an additive process with applications to signal estimation and storage theory", *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **37**, 51–75.

Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

Cline, D. B. H. and Hart, J. D. (1991). "Kernel estimation of densities with discontinuities or discontinuous derivatives", *Statistics*, **22**, 69–84.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Application*, Chapman and Hall, London.

Huh, J. and Park, B. U. (2001). "Detection of change point with local polynomial fits for random design case", *manuscript.*

Loader, C. R. (1996). "Change point estimation using nonparametric regression", *The Annals of Statistics*, **24**, 1667–1678.

Müller, H. G. (1987). "Weighted local regression and kernel methods for nonparametric curve fitting", *Journal of the American Statistical Association*, **82**, 231–238.

Müller, H. G. (1992). "Change-points in nonparametric regression analysis", *The Annals of Statistics*, **20**, 737–761.

Müller, H. G. and Wang, J. L. (1990). "Nonparametric analysis of changes in hazard rates for censored survival data: An alternative to change-point models", *Biometrika*, **77**, 305–314.

Schuster, E. F. (1985). "Incorporating support constraints into nonparametric estimators of densities", *Communications in Statistics-Theory and Methods*, **14**, 1123–1136.

Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing*, Chapman and Hall, London.

Whitt, W. (1970). "Weak convergence of probability measures on the function space $C([0,\infty])$", *Annals of Mathematical Statistics*, **41**, 939–944.