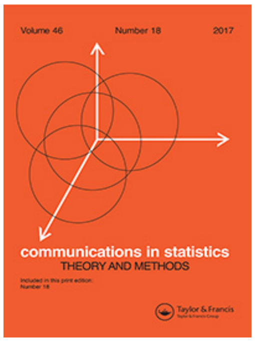

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

L_p distance for kernel density estimator in length-biased data

Vahid Fakoor & Raheleh Zamini

To cite this article: Vahid Fakoor & Raheleh Zamini (2017) L_p distance for kernel density estimator in length-biased data, Communications in Statistics - Theory and Methods, 46:18, 9247-9264, DOI: 10.1080/03610926.2016.1205621

To link to this article: https://doi.org/10.1080/03610926.2016.1205621

Accepted author version posted online: 03 Sep 2016.
Published online: 02 Jun 2017.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2017, VOL. 46, NO. 18, 9247–9264
https://doi.org/10.1080/03610926.2016.1205621

L_p distance for kernel density estimator in length-biased data

Vahid Fakoor[a] and Raheleh Zamini[b]

[a]Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran;
[b]Faculty of Mathematical Sciences and Computer, Department of Mathematics, Kharazmi University, Tehran, Iran

ABSTRACT

In this article we prove a central limit theorem for the L_p distance $I_n(p) = \int_{\mathbb{R}} |f_n(x) - f(x)|^p d\mu(x), 1 \leq p < \infty$, where μ is a weight function and f_n is the kernel density estimator proposed by Jones (1991) for length-biased data. The approach is based on the invariance principle for the empirical processes proved by Horváth (1985). We study the difference $I_n(p)$ with its approximation in terms of its rates of convergence to zero. We subsequently present a central limit theorem for approximation of $I_n(p)$.

ARTICLE HISTORY
Received 18 February 2016
Accepted 20 June 2016

KEYWORDS
Central limit theorem; kernel density estimator; L_p distance; length-biased data.

MATHEMATICS SUBJECT CLASSIFICATION
62G07; 62G20

1. Introduction

Consider the kernel estimate f_n of a real univariate density f introduced by Rosenblatt (1956):

$$f_n(t) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h_n}\right)$$

where $X_1, \ldots, X_n$ are independent observations from f, K is a kernel function, and h_n a bandwidth. A common stochastic measure of the global performance of f_n is defined by the L_p distance

$$I_n(p) = \int_{\mathbb{R}} |f_n(x) - f(x)|^p d\mu(x), \quad 1 \leq p < \infty$$

Wegman (1972) used $I_n(2)$ to compare the performance of estimators in Monte Carlo trials. Steele (1978) identified the need to determine the relationship between various measures of accuracy in density estimation. One such measure, the order of $I_n(2) - EI_n(2)$, is particularly important in statistics. Hall (1982) first began addressing the issues raised in Steele (1978) by computing the exact order of convergence of $I_n(2) - EI_n(2)$ to zero using the strong approximation technique developed by Komlós et al. (1975) for the standard empirical process. Central limit theorems for the $I_n(2)$, based on the Karhunen-Loève expansion, are proved in Bickel and Rosenblatt (1973). By using martingale techniques, central limit theorems for $I_n(2)$ have been obtained by Hall (1984).

A remarkable central limit theorem for the L_1 distance of Grenander's maximum likelihood estimate for monotone densities concentrated on a bounded interval is due to Groeneboom

CONTACT Vahid Fakoor fakoor@math.um.ac.ir Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, P. O. Box 1159-91775, Iran.

(1985), which has been discussed in Devroye and Györfi (1985). Their book contains necessary and sufficient conditions for $I_n(1) \to 0$ in probability or almost surely converges when $\mu(t) = t$. The choice of the L_1 distance is motivated by invariance under monotone transformations of the coordinate axes and the fact that it is well defined for $\mu(t) = t$. A central limit theorem for the L_p distance of Grenander-type estimators for monotone functions was proved by Durot (2007).

Central limit theorems for $I_n(p)$, for any $p \geq 1$, have been established by Csörgő and Horváth (1988), by which a test of hypothesis can be carried out for the density function of f. The results in Csörgő and Horváth (1988) are not restricted to the case of $\mu(t) = t$, and do not assume the finiteness of the support of μ. Horváth (1991) has extended the work of Csörgő and Horváth (1988) to the multivariate case. In the random censorship model, Csörgő et al. (1991) obtained central limit theorems for L_p distances ($1 \leq p < \infty$) of kernel estimators. Mojirsheibani (2009) presented two approximations for L_p distances ($1 \leq p < \infty$) of kernel estimators with central limit theorems for them on complete samples.

The aim of this article is to study a central limit theorem for $I_n(p)$ in the length-biased setting. The practical applications of biased sampling range from social sciences, economics, and quality control to biological and epidemiological studies. A case of particular interest is the so-called length-biased sampling, also known as stock sampling in labor force studies. In length-biased sampling subjects are recruited with a probability proportional to their "length". The resulting distribution is called the length-biased distribution.

There are studies in the literature on length-biased data at least as old as Wicksel (1925). The phenomenon of length-bias was systematically studied by McFadden (1962), Blumenthal (1967), and later by Cox (1969) in the context of estimating the distribution of fiber lengths in a fabric. Vardi (1982, 1985), Gill et al. (1988), and Vardi (1989) laid down the theoretical foundation of biased sampling. Furthermore, an invariance principle for the empirical processes was proven by Horváth (1985).

An interesting overview of non parametric contributions to the literature on estimating problems when the observations are taken from weighted distributions can be found in Cristóbal and Alcalá (2001).

Kernel density estimation for length-biased data has been investigated by Bhattacharyya et al. (1988) and Jones (1991). Jones' estimator proved to possess various advantages over the former. It is a probability density function, which is particularly better behaved near zero, that has better asymptotic mean integrated squared error properties and is more readily extendable to related problems such as density derivative estimation. From another perspective, the asymptotic results on sharp minimax density estimation for length-biased data were derived by Efromovich (2004).

More recently, based on invariance principles for empirical processes, the strong uniform consistency and asymptotic normality of the kernel density estimator proposed by Jones (1991) has been proved by Ajami et al. (2013).

In this article, based on the invariance principle for the empirical processes proven by Horváth (1985), we first establish a central limit theorem for $I_n(p)$ in the length-biased setting. Much like the approximation of L_p distance in Mojirsheibani (2009), we obtain an approximation for $I_n(p)$ for which we prove a central limit theorem.

The layout of this article is as follows: In Sec. 2, after a review on the length-biased distribution we introduce our notation and present some preliminaries. In Sec. 3, we present the main results. In order to prove the main theorems, some auxiliary results are needed, which are included in the Appendix.

2. Preliminaries

The random variable Y has a length-biased distribution of G, if for a given distribution function F, the d.f. of Y is defined by

$$G(t) = \frac{1}{\mu}\int_0^t x dF(x), \quad t \geq 0 \tag{1}$$

where $\mu = \int_0^\infty x dF(x)$, and is assumed to be finite. Throughout this article we assume that G is continuous on $\mathbb{R}^+ = [0, \infty)$. From this it can be concluded that F is also continuous. Let F and G have density functions of f and g, respectively. Using Equation (1) the density of Y is given by

$$g(t) = \frac{tf(t)}{\mu}, \quad t \geq 0$$

An elementary calculation shows that F is determined uniquely by G, as follows:

$$F(t) = \mu \int_0^t y^{-1} dG(y), \quad t \geq 0$$

Let $Y_1, \ldots, Y_n$ be a sample of the independently and identically distributed from G. The empirical estimator of F can thus be written in the form of

$$F_n(t) = \mu_n \int_0^t y^{-1} dG_n(y) \tag{2}$$

where

$$\mu_n^{-1} = \int_0^\infty y^{-1} dG_n(y) \tag{3}$$

G_n is an empirical estimator of G given by

$$G_n(t) = \frac{1}{n}\sum_{i=1}^n I(Y_i < t)$$

where $I(A)$ denotes the indicator of the event A.

Based on a random sample $Y_1, \ldots, Y_n$, Jones (1991) proposed the following estimator for the density function of f:

$$f_n(t) = \frac{1}{h_n}\int_{\mathbb{R}} K\left(\frac{t-u}{h_n}\right) dF_n(u) \tag{4}$$

where K is a kernel function and h_n is a sequence of (positive) bandwidths tending to zero as $n \to \infty$.

In this article based on (4), we study the L_p distance

$$I_n(p) = I_n(T, p) = \int_0^T \left|f_n(x) - f(x)\right|^p d\mu(x) \quad 1 \leq p < \infty \tag{5}$$

where $0 < T < \infty$ and μ is a measure on the Borel sets on $\mathbb{R}$.

Before stating our results, we introduce further notations and then list all the assumptions used in this article.

Assumptions

C(1). $d\mu(t) = w(t)dt$, where $w(t) \geq 0$ and continuous on $[0, \tau]$, where $T < \tau < \infty$ and $\tau = \sup\{x, G(x) < 1\}$.

K. Assumptions on the kernel K:

K(1). There is a finite interval such that K is continuous and bounded on it and vanishes outside of this interval.

K(2). $\int_{\mathbb{R}} K^2(t)dt > 0$.

K(3). K is of bounded variation.

K(4). $K^{'}$ exists and is bounded.

K(5). $\int_{\mathbb{R}} K(t)dt = 1$.

K(6). $\int_{\mathbb{R}} tK(t)dt = 0, \quad \text{and} \int_{\mathbb{R}} |t|^3 K(t)dt < \infty$.

Assumptions on the density f:

F(1). f is uniformly bounded (a.s.) on the $[0, \tau]$.

F(2). $\left|\frac{f^{'}(x)}{x^{\frac{1}{2}} f^{\frac{1}{2}}(x)}\right|$ and $\left|\frac{f^{\frac{1}{2}}(x)}{x^{\frac{3}{2}}}\right|$ are uniformly bounded (a.s.) on the $[0, \tau]$.

F(3). $f^{'''}$ exists and is uniformly bounded (a.s.) on the $[0, \tau]$.

Assumptions on the distributions function G:

G(1). $(G(x))^{\frac{1}{r}} x^{-2}$ is uniformly bounded (a.s.) on the $(0, \tau)$ for some $r > 2$.

Throughout this article $N = N(0, 1)$ stands for a standard normal random variable. Let

$$\sigma(t) = \mu^2 \int_0^t y^{-2} dG(y)$$

$$\sigma = \lim_{t\to\infty} \sigma(t) = \mu^2 \int_0^\infty y^{-2} dG(y)$$

$$P(x) = (\sigma(x)^{'})^{\frac{1}{2}}$$

$$m(p) = m(T, p) = E|N|^p \left(\int_{\mathbb{R}} K^2(t)dt\right)^{\frac{p}{2}} \int_0^T (\sigma^{'}(t))^{\frac{p}{2}} d\mu(t)$$

$$r(t) = \frac{\int_{\mathbb{R}} K(u)K(t+u)du}{\int_{\mathbb{R}} K^2(u)du}$$

$$\sigma_1^2 = (2\pi)^{-1} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |xy|^p (1 - r^2(u))^{-1/2} \times \exp\left(-\frac{1}{2(1-r^2(u))}(x^2 - 2xyr(u) + y^2)\right) dxdy - (E|N|^p)^2 \right) du$$

$$\sigma^2(p) = \sigma^2(T, p) = \sigma_1^2 \int_o^T (\sigma^{'}(t))^p w^2(t)dt \left(\int_{\mathbb{R}} K^2(t)dt\right)^p$$

A natural choice for $w(t)dt = d\mu(t)$ is

$$w(t)dt = d\mu(t) = (\sigma^{'}(x))^{-\frac{p}{2}} f^{\frac{p}{2}} dF = \left(\frac{x}{\mu}\right)^{\frac{p}{2}} dF \tag{6}$$

Namely by (6) and the above definitions of $m(p)$ and $\sigma^2(p)$ and $I_n(p)$ in (5), we have

$$m(p) = E|N|^p \left(\int_{\mathbb{R}} K^2(t)dt\right)^{\frac{p}{2}} \int_0^T f^{\frac{p+2}{2}}(t)d(t)$$

$$\sigma^2 = \sigma_1^2 \int_0^T f^{p+2}(t)dt \left(\int_{\mathbb{R}} K^2(t)dt\right)^p$$

and

$$I_n(p) = \int_0^T |f_n(x) - f(x)|^p \left(\frac{x}{\mu}\right)^{\frac{p}{2}} dF(x) \tag{7}$$

We propose an approximation to (7) by

$$J_n(p) = J_n(T, p) = \int_{\mathbb{R}} \left|f_n(x) - f(x)\right|^p \left(\frac{x}{\mu_n}\right)^{\frac{p}{2}} dF_n(x) \tag{8}$$

Observe that $J_n(p)$ corresponds to the approximation of $I_n(p)$ by replacing F and μ with F_n and μ_n given by (2) and (3), respectively. We study the difference between $J_n(p)$ and $I_n(p)$ as $n \to \infty$. As a by-product of our findings, we will also state a central limit theorem for the properly standardized version of $J_n(p)$.

3. Main results

In this section we present our main results. The following inequality will be used in the proof of the main theorems. Let $1 \le p < \infty$, then for functions q and u in L_p, we have

$$\begin{aligned}
&\int_0^\infty \left||q(t)|^p - |u(t)|^p\right| d\mu(t) \\
&\quad \le p2^{p-1} \int_0^\infty \left|q(t) - u(t)\right|^p d\mu(t) \\
&\qquad + p2^{p-1} \left(\int_0^\infty |u(t)|^p d\mu(t)\right)^{1-\frac{1}{p}} \left(\int_0^\infty |q(t) - u(t)|^p d\mu(t)\right)^{\frac{1}{p}}
\end{aligned} \tag{9}$$

In the following theorem, we study the asymptotic normality of $I_n(p)$.

Theorem 1. *Let* $1 < p < \infty$. *Assume* $\mathbf{K(1)}$–$\mathbf{K(3)}$, $\mathbf{K(5)} - \mathbf{K(6)}$, $\mathbf{F(2)} - \mathbf{F(3)}$, *and* $\mathbf{C(1)}$ *hold. If as* $n \to \infty$,

$$h_n \to 0, \quad h_n^{-1} n^{-B} \to 0, \quad h_n^4 n \to 0$$

(for any $0 < B < \frac{1}{2} - \frac{1}{r}$, $(r > 4)$*), then*

$$(h_n \sigma^2(p))^{-\frac{1}{2}} \{(nh_n)^{\frac{p}{2}} I_n(p) - m(p)\} \xrightarrow{\mathcal{D}} N(0, 1)$$

Proof. $\mathbf{K(5)}-\mathbf{K(6)}$ and $\mathbf{F(3)}$ with the two-term Taylor expansion

$$f(t - uh_n) - f(t) = -uh_n f'(t) + \frac{1}{2}u^2 h_n^2 f''(t) - \frac{1}{6}u^3 h_n^3 f'''(t^*),$$

where $t^* \in \big(t \wedge (t - uh_n), t \vee (t - uh_n)\big)$, imply that

$$\begin{aligned}
f_{(n)}(t) - f(t) &= \int_{\mathbb{R}} \Big(f(t - uh_n) - f(t)\Big) K(u) du \\
&= \frac{1}{2} h_n^2 f''(t) \int_{\mathbb{R}} u^2 K(u) du - \frac{1}{6} h_n^3 f'''(t^*) \int_{\mathbb{R}} u^3 K(u) du \\
&= O_p(h_n^2)
\end{aligned} \tag{10}$$

where

$$f_{(n)}(t) = (h_n)^{-1} \int_{\mathbb{R}} K\left(\frac{t - x}{h_n}\right) dF(x) \tag{11}$$

Hence by Lemma 4 and (10), we obtain

$$\begin{aligned}
&\int_0^T \Big| |f_n(t) - f_{(n)}(t)|^p - |f_n(t) - f(t)|^p \Big| d\mu(t) \\
&\le p2^{p-1} \int_0^T |f_{(n)}(t) - f(t)|^p d\mu(t) \\
&\quad + p2^{p-1} \left(\int_0^T |f_n(t) - f_{(n)}(t)|^p d\mu(t) \right)^{1-\frac{1}{p}} \\
&\quad \times \left(\int_0^T |f_{(n)}(t) - f(t)|^p d\mu(t) \right)^{\frac{1}{p}} \\
&= O_p(h_n^{2p}) + O_p\Big((nh_n)^{-\frac{p}{2}(1-\frac{1}{p})} \Big) O_p(h_n^2)
\end{aligned}$$

Therefore,

$$\begin{aligned}
(\sigma^2(p)h_n)^{-\frac{1}{2}}\{(nh_n)^{\frac{p}{2}} I_n(p) - m(p)\} &= n^{\frac{p}{2}} O_p\left(h_n^{\frac{5p-1}{2}} \right) + O_p\Big(h_n^2 n^{\frac{1}{2}} \Big) \\
&\quad + (\sigma^2(p)h_n)^{-\frac{1}{2}}\{(nh_n)^{\frac{p}{2}} \hat{I}_n(p) - m(p)\}
\end{aligned}$$

The condition $\lim_{n\to\infty} nh_n^4 = 0$, Slutsky theorem, and Lemma 4 complete the proof. □

Since F is unknown, Theorem 1 is not practically useful. In Lemma 6, we approximate $I_n(p)$ with $J_n(p)$ in (8). Next in Theorem 2, we state a central limit theorem for $J_n(p)$.

Theorem 2. *Let $p > 1$ be an even integer and assume that* $\mathbf{K(1)} - \mathbf{K(3)}$, $\mathbf{K(5)} - \mathbf{K(6)}$, $\mathbf{F(2)} - \mathbf{F(3)}$, $\mathbf{C(1)}$, *and* $\mathbf{G(1)}$ *hold. If, as* $n \to \infty$,

$$h_n \to 0, \quad nh_n^4 \to 0,$$

$$n^{-B} h_n^{-1} \to 0, \quad \frac{nh_n^3}{\log\log n} \to \infty$$

(for any $0 < B < \frac{1}{2} - \frac{1}{r}$, $(r > 4)$), then

$$(h_n \sigma^2(p))^{-\frac{1}{2}} \Big[(nh_n)^{\frac{p}{2}} J_n(p) - m(p) \Big] \xrightarrow{\mathcal{D}} N(0, 1)$$

Proof. The proof follows from Lemma 6, Theorem 1, and the fact that, under the conditions of the theorem (on n and h_n),

$$(h_n \sigma^2(p))^{-\frac{1}{2}} n^{\frac{p}{2}} h_n^{\frac{p}{2}} \Big| J_n(p) - I_n(p) \Big| = o_p(1)$$

□

Acknowledgment

We would like to thank the referees for contributing to the final edition of the manuscript with their insightful comments.

References

Ajami, M., Fakoor, V., Jomhoori, S. (2013). Some asymptotic results of kernel density estimator in length-biased sampling. *J. Sci.* 24(1):55–62.

Bhattacharyya, B.B., Franklin, L.A., Richardson, G.D. (1988). A comparison of nonparametric unweighted and length-biased density estimation of fibres. *Commun. Stat. A* 17:3629–3644.

Bickel, P.J., Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Stat.* 6:1071–1095.
Blumenthal, S. (1967). Proportional sampling in life length studies. *Technometrics.* 9:205–218.
Cox, D.R. (1969). Some sampling problems in technology. In: Johnson, N.L., Smith, H. eds. *New Developments in Survey Sampling*. New York: Wiley.
Cristóbal, J.A., Alcalá, J.T. (2001). An overview of nonparametric contributions to the problem of functional estimation from biased data. *Test.* 10(2):309–332.
Csörgő, M., Horváth, L. (1988). Central limit theorems for L_p-norm of density estimators. *Probab. Theory Relat. Fields* 80:269–291.
Csörgő, M., Gombay, E., Horváth, L. (1991). Central limit theorems for L_p distances of kernel estimators of densities under random censorship. *Ann. Stat.* 19(4):1813–1831.
Devroye, L., Györfi, L. (1985). *Nonparametric Density Estimation: The L_1 View*. New York: Wiley.
Durot, C. (2007). On the L_p-error of monotonicity constrained estimators. *Ann. Stat.* 35:1080–1104.
Efromovich, S. (2004). Density estimation for biased data. *Ann. Stat.* 32(3):1137–1161.
Gill, R.D.; Vardi, Y.; Wellner, J.A. (1988). Large sample theory of empirical distributions in biased sampling models. *Ann. Stat.* 16:1069–1112.
Groeneboom, P. (1985). Estimating a monotone density. In: Le Cam, L.M., Olshen, R.A., eds. *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (vol. 2, pp. 539–555). Belmont, CA: Wadsworth.
Hall, P. (1982). Limit theorems for stochastic measures of the accuracy of density estimators. *Stoch. Process Appl.* 13:11–25.
Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J. Multivariate Anal.* 14:1–16.
Horváth, L. (1985). Estimation from a length-biased distribution. *Stat. Decis.* 3:91–113.
Horváth, L. (1991). On L_p-norm of multivariate density estimators. *Ann. Stat.* 19(4):1933–1949.
James, B.R. (1975). A functional law of the iterated logarithm for weighted empirical distributions. *Ann. Probab.* 3:762–772.
Jones, M.C. (1991). Kernel density estimation for length biased data. *Biometrika* 78:511–519.
Komlós, J., Major, P., Tusnády, G. (1975). An approximation of partial sums of independent r.v.'s and the sample d.f. *I.Z. Wahrscheinlichkeitstheorie Verw. Geb.* 32:111–131.
McFadden, J.A. (1962). On the lengths of intervals in a stationary point process. *J. Royal Soc. Ser. B* 24:364–382.
Mojirsheibani, M. (2009). Some approximations to L_p-statistics of kernel density estimators. *Statistics* 41(3):203–220.
Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Stat.* 27:832–837.
Steele, J.M. (1978). Invalidity of average squared error criterion in density estimation. *Can. J. Stat.* 6:193–200.
Vardi, Y. (1982). Nonparametric estimation in renewal processes. *Ann. Stat.* 10:772–785.
Vardi, Y. (1985). Empirical distributions in selection bias models. *Ann. Stat.* 13:178–205.
Vardi, Y. (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density. *Biometrika* 76:751–761.
Wegman, E.J. (1972). Nonparametric probability density estimation: A comparison of density estimation methods. *J. Stat. Comput. Simul.* 1:225–245.
Wicksell, S.D. (1925). The corpuscle problem. A mathematical study of a biometric problem. *Biometrika* 17(1/2):84–99.

Appendix

In order to make the proofs easier, we need some auxiliary results and notations.

Lemma 1. *Let* $\Gamma_n^{(1)}(x) = \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) dW(y)$. *Assume* $\mathbf{K(1)}$–$\mathbf{K(2)}$, $\mathbf{C(1)}$, *and* $\mathbf{F(1)}$ *hold. If* $\lim_{n\to\infty} h_n = 0$, *then, as* $n \to \infty$, *we have*

$$\left((h_n)^{p+1}\sigma^2(p)\right)^{-\frac{1}{2}} \left\{\int_0^T \left|P(x)\Gamma_n^{(1)}(x)\right|^p d\mu(x) - h_n^{\frac{p}{2}} m(p)\right\} \xrightarrow{\mathcal{D}} N(0,1)$$

Proof. The proof of Lemma 1 goes along the lines of the proof of Lemmas 1 and 2 in Csörgő and Horváth (1988). □

Lemma 2. *Let* $\Gamma_n^{(2)}(x) = \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) P(y)dW(y)$. *Assume* **K(1)**–**K(3)**, **C(1)**, *and* **F(2)** *hold. If* $\lim_{n\to\infty} h_n = 0$, *then, as* $n \to \infty$, *we have*

$$(h_n^{p+1}\sigma^2(p))^{-\frac{1}{2}} \left\{ \int_0^T |\Gamma_n^{(2)}(x)|^p d\mu(x) - h_n^{\frac{p}{2}} m(p) \right\} \xrightarrow{\mathcal{D}} N(0,1)$$

Proof.

$$\begin{aligned}
\Gamma_n^{(2)}(x) &= P(x)\Gamma_n^{(1)}(x) + \int_{-1}^{1} (W(x-yh_n) - W(x))(P(x-yh_n) - P(x))dK(y) \\
&\quad + \int_{-1}^{1} \left(W(x-yh_n) - W(x)\right) K(y)dP(x-yh_n) \\
&:= P(x)\Gamma_n^{(1)}(x) + A_n^{(1)}(x) + A_n^{(2)}(x)
\end{aligned}$$

Using now continuity of Wiener process, the mean value theorem, and **F(2)**, we get

$$\begin{aligned}
\sup_{0<x\le T} |A_n^{(1)}(x)| &\le \sup_{0<t\le T} \sup_{0<s\le h_n} |W(t+s) - W(t)| \\
&\quad \times \int_{\mathbb{R}} \left|(P(x-yh_n) - P(x))\right| \left|dK(y)\right| \\
&\le Ch_n \sup_{0<t\le T} \sup_{0<s\le h_n} |W(t+s) - W(t)| \\
&\quad \times \sup_{0<x\le T+h_n} \left| \frac{f'(x)}{x^{\frac{1}{2}} f^{\frac{1}{2}}(x)} - \frac{f^{\frac{1}{2}}(x)}{x^{\frac{3}{2}}} \right| = o_p(h_n)
\end{aligned} \tag{A.1}$$

A similar argument gives

$$\sup_{0<x\le T} |A_n^{(2)}(x)| = o_p(h_n) \tag{A.2}$$

Applying (9), Lemma 1, (A.1), and (A.2), we obtain

$$\begin{aligned}
&\left| \int_0^T |\Gamma_n^{(2)}(x)|^p d\mu(x) - \int_0^T |P(x)\Gamma_n^{(1)}(x)|^p d\mu(x) \right| \\
&\le p2^{p-1} \int_0^T \left|A_n^{(1)}(x) + A_n^{(2)}(x)\right|^p d\mu(x) \\
&\quad + p2^{p-1} \left(\int_0^T \left|P(x)\Gamma_n^{(1)}(x)\right|^p d\mu(x) \right)^{1-\frac{1}{p}} \\
&\quad \times \left(\int_0^T \left|A_n^{(1)}(x) + A_n^{(2)}(x)\right|^p d\mu(x) \right)^{\frac{1}{p}} \\
&= o_p(h_n^p) + o_p\left(h_n^{\frac{(p+1)}{2}} \right)
\end{aligned} \tag{A.3}$$

(A.3) and Lemma 1 complete the proof. □

In the following, we use the strong approximation for the empirical process $\alpha_n(t) = \sqrt{n}[F_n(t) - F(t)]$, which has been proven by Horváth (1985). He defined a mean zero Gaussian process

$$\Gamma(t, n) = \mu \int_0^t y^{-1} dB(y, n) - \mu F(t) \int_0^\infty y^{-1} dB(y, n) \tag{A.4}$$

with covariance function

$$E[\Gamma(x, n)\Gamma(y, m)] = (mn)^{-\frac{1}{2}}(m \wedge n)\big[\sigma(x \wedge y) - F(x)\sigma(y) - F(y)\sigma(x) + F(x)F(y)\sigma\big] \tag{A.5}$$

such that it approximates the empirical process $\alpha_n(t)$. In (A.4), $B(t, n)$ is a two-parameter Gaussian process with zero mean and covariance function

$$E[B(x, n)B(y, m)] = (mn)^{-\frac{1}{2}}(m \wedge n)[G(x \wedge y) - G(x)G(y)] \ (a \wedge b = \min(a, b))$$

which approximate the empirical process

$$\beta_n(t) = \sqrt{n}[G_n(t) - G(t)], \quad t \geq 0$$

as obtained by Komlós et al. (1975). Let $\{W(u, v),\ u, v \geq 0\}$ denote a two-parameter Wiener process. By (A.5), the following representation holds:

$$\left\{n^{\frac{1}{2}}\Gamma(t, n), t \geq 0, n \geq 1\right\} \stackrel{D}{=} \{W(\sigma(t), n) - F(t)W(\sigma, n), t \geq 0, n \geq 1\} \tag{A.6}$$

where $\stackrel{D}{=}$ denotes an equal in distribution.

Lemma 3. *Let* $\Gamma_{n,2}(x) = \int_{\mathbb{R}} K(\frac{x-y}{h_n})d\Gamma(y, n)$. *Assume* $\mathbf{K(1)}$–$\mathbf{K(3)}$, $\mathbf{C(1)}$, *and* $\mathbf{F(2)}$ *hold. If* $\lim_{n\to\infty} h_n = 0$, *then, as* $n \to \infty$, *we can write*

$$\left(h_n^{p+1}\sigma^2(p)\right)^{-\frac{1}{2}} \left\{\int_0^T |\Gamma_{n,2}(x)|^p d\mu(x) - h_n^{\frac{p}{2}} m(p)\right\} \xrightarrow{\mathcal{D}} N(0, 1)$$

Proof. By using (A.6),

$$\Gamma_{n,2}(x) = \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\Gamma(y, n) \stackrel{D}{=} \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\left(\frac{W(\sigma(y), n)}{\sqrt{n}} - F(y)\frac{W(\sigma, n)}{\sqrt{n}}\right)$$

Since

$$\left\{\frac{W(x, n)}{\sqrt{n}}; 0 \leq x < \infty, n \geq 1\right\} \stackrel{D}{=} \{W_n(x); \quad 0 \leq x < \infty, n \geq 1\} \tag{A.7}$$

where $W_n(x)$ is a sequence of standard Wiener processes and $W_n(x) \stackrel{D}{=} W(t)$ for each n, it is enough to show that

$$\left(h_n^{p+1}\sigma^2(p)\right)^{-\frac{1}{2}} \left\{\int_0^T |\Gamma'_{n,2}(x)|^p d\mu(x) - h_n^{\frac{p}{2}} m(p)\right\} \xrightarrow{D} N(0, 1)$$

where $\Gamma'_{n,2}(x) = \int_{\mathbb{R}} K(\frac{x-y}{h_n})d(W(\sigma(y)) - F(y)W(\sigma))$. At first note that

$$\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) W(\sigma)dF(y)$$

is normally distributed with mean 0 and variance

$$\sum_{n,x} = \sigma\left(\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) dF(y)\right)^2$$

Therefore, for each x,

$$\left|\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) W(\sigma)dF(y)\right| \stackrel{D}{=} |N|\left|\sigma^{1/2}\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) dF(y)\right|$$

By **F(1)**

$$\begin{aligned}\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) f(y)dy &= h_n\int_{\mathbb{R}} K(u)f(x-uh_n)du\\ &\leq Mh_n\int_{\mathbb{R}} K(u)du\\ &= O(h_n)\end{aligned}$$

where $M=\sup_{0\leq x\leq\tau} f(t)$. Hence

$$\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) W(\sigma)dF(y) = O_p(h_n) \tag{A.8}$$

Since for each n,

$$\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\big(W(\sigma(y))\big) \stackrel{D}{=} {\Gamma_n}^{(2)}(x) \tag{A.9}$$

(9), (A.8), and Lemma 2 imply that

$$\begin{aligned}&\left|\int_0^T\left|\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\big(W(\sigma(y))\big)\right|^p d\mu(x) - \int_0^T |\Gamma'_{n,2}(x)|^p d\mu(x)\right|\\ &\leq p2^{p-1}\int_0^T\left|\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) W(\sigma)dF(y)\right|^p d\mu(x)\\ &+p2^{p-1}\left(\int_0^T\left|\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\big(W(\sigma(y))\big)\right|^p d\mu(x)\right)^{1-\frac{1}{p}}\\ &\times\left(\int_0^T\left|\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) W(\sigma)dF(y)\right|^p d\mu(x)\right)^{\frac{1}{p}}\\ &= o_p(h_n^p) + o_p\left(h_n^{\frac{p}{2}(1-\frac{1}{p})}\right) o_p(h_n)\end{aligned}$$

By (A.9) and Lemma 2 the proof of Lemma 3 is complete. □

Let

$$\hat{I}_n(p) = \int_0^T \left|f_n(t) - f_{(n)}(t)\right|^p d\mu(t)$$

where $f_{(n)}(t)$ is given by (11).

Lemma 4. *Assume that* **K(1)**–**K(3)**, **F(2)**, *and* **C(1)** *hold. If as* $n\to\infty$,

$$h_n\to 0,\quad h_n^{-1}n^{-B}\to 0,$$

(for any $0 < B < \frac{1}{2} - \frac{1}{r}$, $(r > 2)$*), then*

$$(h_n\sigma^2(p))^{-\frac{1}{2}}\left\{(nh_n)^{\frac{p}{2}}\hat{I}_n(p) - m(p)\right\} \xrightarrow{\mathcal{D}} N(0,1)$$

Proof. Using (9), we obtain

$$\begin{aligned}
&\int_0^T \left|\left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\alpha_n(x)\right|^p - \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Gamma(x,n)\right|^p\right| d\mu(t) \\
&\le p2^{p-1}\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d(\alpha_n(x) - \Gamma(x,n))\right|^p d\mu(t) \\
&\quad + p2^{p-1}\left(\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Gamma(x,n)\right|^p d\mu(t)\right)^{1-\frac{1}{p}} \\
&\quad \times \left(\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Big(\Gamma(x,n) - \alpha_n(x)\Big)\right|^p d\mu(t)\right)^{\frac{1}{p}} \\
&:= A_n^{(1)} + A_n^{(2)}
\end{aligned} \tag{A.10}$$

It follows from (A.16) that

$$A_n^{(1)} = O_p\left(n^{-Bp}\right) \tag{A.11}$$

Also, Lemma 3 and (A.16) imply

$$A_n^{(2)} = n^{-B}O_p\left(h_n^{\frac{p-1}{2}}\right) \tag{A.12}$$

Hence, by (A.10), (A.11), and (A.12), we get

$$\begin{aligned}
\hat{I}_n(p) &= \int_0^T \left|\frac{1}{h_n}\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Big(F_n(x) - F(x)\Big)\right|^p d\mu(t) \\
&= h_n^{-p}n^{-\frac{p}{2}}\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\alpha_n(x)\right|^p d\mu(t) \\
&= h_n^{-p}n^{-\frac{p}{2}}\int_0^T \left(\left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\alpha_n(x)\right|^p - \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Gamma(x,n)\right|^p\right) d\mu(t) \\
&\quad + h_n^{-p}n^{-\frac{p}{2}}\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Gamma(x,n)\right|^p d\mu(t) \\
&= h_n^{-p}n^{-\frac{p}{2}}\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Gamma(x,n)\right|^p d\mu(t) \\
&\quad + h_n^{-p}n^{-\frac{p}{2}}O_p(n^{-Bp}) + h_n^{-p}n^{-\frac{p}{2}}O_p\left(n^{-B}h_n^{\frac{p-1}{2}}\right)
\end{aligned}$$

which immediately gives

$$\begin{aligned}
&(h_n\sigma^2(p))^{-\frac{1}{2}}\left\{(nh_n)^{\frac{p}{2}}\hat{I}_n(p) - m(p)\right\} \\
&= h_n^{\frac{-(p+1)}{2}}O_p\left(n^{-Bp} \vee n^{-B}h_n^{\frac{p-1}{2}}\right) \\
&\quad + (h_n^{p+1}\sigma^2(p))^{\frac{-1}{2}}\left\{\int_0^T \left|\int_{\mathbb{R}} K\left(\frac{t-x}{h_n}\right) d\Gamma(x,n)\right|^p d\mu(t) - h_n^{\frac{p}{2}}m(p)\right\}
\end{aligned}$$

Now Lemma 4 follows from Lemma 3. $\square$

Lemma 5. *Suppose the Assumption* **G(1)** *holds. Then, we have*

$$\sup_{0<x<\tau} |F_n(x) - F(x)| = O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right) \quad a.s.$$

Proof.

$$\begin{aligned}
\sup_{0<x<\tau} \left|F_n(x) - F(x)\right| &\leq \left(\sup_{0<x<\tau} \int_0^x y^{-1} dG_n(x)\right) \left|\upsilon_n^{-1} - \upsilon^{-1}\right| \\
&\quad + \upsilon^{-1} \sup_{0<x<\tau} \left|\int_0^x y^{-1} d\big(G_n(y) - G(y)\big)\right| \\
&\leq \upsilon^{-1}\left|\upsilon - \upsilon_n\right| + \upsilon^{-1} \sup_{0<x<\tau} \left|\int_0^x y^{-1} d\big(G_n(y) - G(y)\big)\right| \\
&=: D_n^{(1)} + D_n^{(2)} \qquad\qquad (A.13)
\end{aligned}$$

where

$$\upsilon_n := {\mu_n}^{-1} = \int_0^\tau y^{-1} dG_n(y)$$

and

$$\upsilon := \mu^{-1} = \int_0^\tau y^{-1} dG(y)$$

Using theorem of James (1975) and **G(1)**, we obtain for any $0 < \delta < \frac{1}{2} - \frac{1}{r}$

$$\begin{aligned}
D_n^{(1)} &= \upsilon^{-1} \left|\int_0^\tau y^{-1} d\big(G_n(y) - G(y)\big)\right| \\
&= \upsilon^{-1} n^{-\frac{1}{2}} \left|\int_0^\tau y^{-1} d\beta_n(y)\right| \\
&= \upsilon^{-1} n^{-\frac{1}{2}} \left|\int_0^\tau y^{-2} \beta_n(y) dy\right| \\
&\leq \upsilon^{-1} n^{-\frac{1}{2}} \sup_{0<y<\tau} (G(y))^{\delta-\frac{1}{2}} \left|\beta_n(y)\right| \left|\int_0^\tau y^{-2} (G(y))^{\frac{1}{2}-\delta} dy\right| \\
&= O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right) \quad a.s. \qquad\qquad (A.14)
\end{aligned}$$

A similar argument gives

$$\begin{aligned}
D_n^{(2)} &= \upsilon^{-1} n^{-\frac{1}{2}} \sup_{0<x<\tau} \left|\int_0^x y^{-1} d\beta_n(y)\right| \\
&\leq \upsilon^{-1} n^{-\frac{1}{2}} \sup_{0<x<\tau} \left|\frac{\beta_n(x)}{x}\right| + \upsilon^{-1} n^{-\frac{1}{2}} \sup_{0<x<\tau} \left|\int_0^x y^{-2} \beta_n(y)\right| \\
&= \upsilon^{-1} n^{-\frac{1}{2}} \sup_{0<x<\tau} \left((G(x))^{\delta-\frac{1}{2}} \left|\beta_n(x)\right| \frac{(G(x))^{\frac{1}{2}-\delta}}{x}\right) \\
&\quad + \upsilon^{-1} n^{-\frac{1}{2}} \sup_{0<x<\tau} \left|\int_0^x y^{-2} \beta_n(y)\right|
\end{aligned}$$

$$
\begin{aligned}
&\leq \upsilon^{-1}n^{-\frac{1}{2}}\sup_{0<x<\tau}\left((G(x))^{\delta-\frac{1}{2}}\left|\beta_n(x)\right|\right)\left(\sup_{0<x<\tau}\frac{(G(x))^{\frac{1}{2}-\delta}x}{x^2}\right)\\
&\quad+\upsilon^{-1}n^{-\frac{1}{2}}\sup_{0<x<\tau}\left|\int_0^x y^{-2}\beta_n(y)dy\right|\\
&\leq \upsilon^{-1}n^{-\frac{1}{2}}\sup_{0<x<\tau}\left((G(x))^{\delta-\frac{1}{2}}\left|\beta_n(x)\right|\right)\left(\tau\sup_{0<x<\tau}\frac{(G(x))^{\frac{1}{2}-\delta}}{x^2}\right)\\
&\quad+\upsilon^{-1}n^{-\frac{1}{2}}\sup_{0<x<\tau}\left(\left|\beta_n(y)\right|(G(y))^{\delta-\frac{1}{2}}\right)\left(\int_0^\tau y^{-2}(G(y))^{\frac{1}{2}-\delta}dy\right)\\
&=O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)+O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)\\
&=O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right) \qquad \text{(A.15)}
\end{aligned}
$$

Collecting together (A.13), (A.14), and (A.15), we obtain the result. □

Lemma 6. *Let $p > 1$ be an even integer and define $I_n(p)$ and $J_n(p)$ as in (7) and (8), respectively. Suppose that* $\mathbf{K(1)}$–$\mathbf{K(2)}$, $\mathbf{K(5)}$–$\mathbf{K(6)}$, $\mathbf{F(2)}$–$\mathbf{F(3)}$, $\mathbf{C(1)}$, *and* $\mathbf{G(1)}$ *hold. If as $n \to \infty$, $h_n \to 0$, and $h_n^{-1}n^{-\frac{1}{2}} \to 0$, then for any $0 < B < \frac{1}{2} - \frac{1}{r}$, $(r > 2)$, one has*

$$
|J_n(p)-I_n(p)| = O_p(h_n^{2p}) + O_p(n^{-p(B+1/2)}h_n{}^{-p}) + O_p\left(n^{-\frac{p}{2}}h_n{}^{-\frac{(p+2)}{2}}\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)
$$

Proof.

$$
\begin{aligned}
f_n(x) &= h_n^{-1}\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)f(y)dy + h_n^{-1}n^{-\frac{1}{2}}\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)d(\Gamma(y,n))\\
&\quad+h_n^{-1}n^{-\frac{1}{2}}\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)d(\alpha_n(y)-\Gamma(y,n))
\end{aligned}
$$

Denote the last term on the right-hand side of the above expression by $r_n(x)$ and observe that

$$
r_n(x) = h_n^{-1}n^{-\frac{1}{2}}\int_{\mathbb{R}}[\alpha_n(x-uh_n)-\Gamma(x-uh_n,n)]dK(u)
$$

Consequently

$$
\begin{aligned}
|r_n(x)| &\leq h_n^{-1}n^{-\frac{1}{2}}\sup_{0<t<\infty}|\alpha_n(t)-\Gamma(t,n)|\int_{\mathbb{R}}\left|dK(u)\right|\\
&= h_n^{-1}n^{-\frac{1}{2}}O(n^{-B}) \quad a.s. \qquad \text{(A.16)}
\end{aligned}
$$

for any $0 < B < \frac{1}{2} - \frac{1}{r}$. Last line results by Theorem 4.2 of Horváth (1985). $\mathbf{F(3)}$, $\mathbf{K(6)}$, and $\mathbf{K(5)}$ in conjunction with the two-term Taylor expansion

$$
f(x-uh_n)-f(x) = -uh_nf'(x) + (uh_n)^2\frac{f''(x)}{2} - (uh_n^3)\frac{f'''(x^*)}{6}
$$

where $x^* \in (x\wedge(x-uh_n), x\vee(x-uh_n))$, immediately imply that for some constant $0 < C_1 < \infty$

$$
\begin{aligned}
f_n(x)-f(x) &= \int_{\mathbb{R}}\Big(f(x-uh_n)-f(x)\Big)K(u)du\\
&\quad+n^{-\frac{1}{2}}h_n^{-1}\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)d(\Gamma(y,n)) + h_n^{-1}n^{-\frac{1}{2}}O(n^{-B}) \quad a.s.
\end{aligned}
$$

$$= C_1 h_n^2 f''(x) - h_n^3 6^{-1} \int_{\mathbb{R}} u^3 f'''(x^*) K(u) du$$
$$+ n^{-\frac{1}{2}} h_n^{-1} \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d(\Gamma(y,n)) + h_n^{-1} n^{-\frac{1}{2}} O(n^{-B}) \quad a.s.$$

where the term involving f''' satisfies

$$\left| 6^{-1} \int_{\mathbb{R}} |u|^3 f'''(x^*) K(u) du \right| < some\ M < \infty$$

We have

$$\begin{aligned} \left| J_n(p) - I_n(p) \right| &\le \left| \upsilon^{\frac{p}{2}} - \upsilon_n^{\frac{p}{2}} \right| \int_0^T \left| f_n(x) - f(x) \right|^p x^{\frac{p}{2}} dF(x) \\ &\quad + \mu_n^{-\frac{p}{2}} \left| \int_0^T \left| f_n(x) - f(x) \right|^p x^{\frac{p}{2}} d\Big(F_n(x) - F(x)\Big) \right| \\ &:= C_1 + C_2 \end{aligned} \tag{A.17}$$

Using the facts that

$$\left| f_n(x) - f(x) \right|^p \le C_1^p h_n^{2p} |f''(x)|^p + \left| n^{-\frac{1}{2}} h_n^{-1} \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\Gamma(y,n) \right|^p$$
$$+ h_n^{3p} M^p + h_n^{-p} n^{-\frac{p}{2}} n^{-Bp} O(1) \quad a.s.$$

and

$$| f_n(x) - f(x) |^p \ge -C_1^p h_n^{2p} |f''(x)|^p + \left| n^{-\frac{1}{2}} h_n^{-1} \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\Gamma(y,n) \right|^p$$
$$- h_n^{3p} M^p - h_n^{-p} n^{-\frac{p}{2}} n^{-Bp} O(1) \quad a.s.$$

it is not difficult to show that

$$\begin{aligned} C_1 \le \left| \upsilon^{\frac{p}{2}} - \upsilon_n^{\frac{p}{2}} \right| \Big[& C_1^p h_n^{2p} T^{\frac{p}{2}} \int_0^T \left| f''(x) \right|^p dF(x) \\ & + \mu^{-\frac{p}{2}} n^{-\frac{p}{2}} h_n^{-p} \int_0^T \left| \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\Gamma(y,n) \right|^p \frac{x^{\frac{p}{2}}}{\mu^{\frac{p}{2}}} dF \\ & + h_n^{3p} M^p T^{\frac{p}{2}} \int_0^T dF(x) + h_n^{-p} n^{-\frac{p}{2}} n^{-Bp} O_p(1) T^{\frac{p}{2}} \int_0^T dF(x) \Big] \end{aligned}$$

and

$$\begin{aligned} C_2 \le \mu_n^{-\frac{p}{2}} \Big[& C_1^p h_n^{2p} T^{\frac{p}{2}} \int_0^T |f''(x)|^p d(F_n(x) + F(x)) \\ & + \left| n^{-\frac{p}{2}} h_n^{-p} \int_0^T \left| \int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d(\Gamma(y,n)) \right|^p x^{\frac{p}{2}} d\big(F_n(x) - F(x)\big) \right| \\ & + \Big(O(h_n^{3p}) + h_n^{-p} n^{-\frac{p}{2}} n^{-Bp} O(1) \Big) \int_0^T d(F_n(x) + F(x)) \Big] \quad a.s. \\ := I_n + |II_n| + O(h_n^{3p}) + III_n \quad a.s. & \end{aligned} \tag{A.18}$$

Note that by Lemma 3

$$n^{-\frac{p}{2}}h_n^{-p}\int_0^T\Big|\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)d\Gamma(y,n)\Big|^p\frac{x^{\frac{p}{2}}}{\mu^{\frac{p}{2}}}dF(x)=O_p\Big(n^{-\frac{p}{2}}h_n^{-\frac{p}{2}}\Big) \tag{A.19}$$

Using the inequality

$$\Big||a(x)|^p-|b(x)|^p\Big|\leq p2^{p-1}\big|a(x)-b(x)\big|^p+p2^{p-1}|b(x)|^{p-1}\big|a(x)-b(x)\big|$$

(for $p\geq 1$), and (A.14), we may write

$$\begin{aligned}\big|v_n{}^{\frac{p}{2}}-v^{\frac{p}{2}}\big| &\leq \frac{p}{2}2^{\frac{p}{2}-1}\big|v_n-v\big|^{\frac{p}{2}}+\frac{p}{2}2^{\frac{p}{2}-1}|v|^{\frac{p}{2}-1}\big|v_n-v\big|\\ &=O_p\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)\end{aligned} \tag{A.20}$$

(A.19) and (A.20) conclude that

$$\begin{aligned}C_1=&\,h_n^{2p}O_p\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)+(nh_n)^{-\frac{p}{2}}O_p\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)\\ &+h_n^{-p}n^{-p\left(B+\frac{1}{2}\right)}O_p\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right)\end{aligned} \tag{A.21}$$

On the other hand, by using Lemma 5 and $\mathbf{F(3)}$ for some $0<M<\infty$, we get

$$\begin{aligned}\int_0^T|f''(x)|^p d\Big(F_n(x)+F(x)\Big)&\leq M\Big(F_n(T)+F(T)\Big)\\ &\leq M|F_n(T)-F(T)|+2MF(T)=O_p(1)\end{aligned}$$

Hence,

$$I_n=O_p(h_n^{2p}) \tag{A.22}$$

Next, to deal with II_n, observe that when $p>1$ is an even integer one has

$$\begin{aligned}II_n=&\,\frac{p}{2}n^{-\frac{p}{2}}h_n^{-p}\int_0^T\Big(F_n(x)-F(x)\Big)x^{\frac{p}{2}-1}\Big[\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)d\Gamma(y,n)\Big]^p dx\\ &+pn^{-\frac{p}{2}}h_n^{-(p+1)}\int_0^T x^{\frac{p}{2}}\Big(F_n(x)-F(x)\Big)\Big[\int_{\mathbb{R}}\Gamma(x-uh_n,n)dK(u)\Big]^{p-1}\\ &\times\Big[\int_{\mathbb{R}}\Gamma(x-uh_n,n)d\psi(u)\Big]dx\\ :=&\,K_n^{(1)}+K_n^{(2)}\end{aligned} \tag{A.23}$$

where $\psi(u)=K'(u)$. But

$$\begin{aligned}K_n^{(1)}\leq&\,\frac{p}{2}n^{-\frac{p}{2}}h_n^{-p}\left(\int_0^T x^{2[\frac{p}{2}-1]}\Big(F_n(x)-F(x)\Big)^2dx\right)^{\frac{1}{2}}\\ &\times\left(\int_0^T\left(\int_{\mathbb{R}}K\left(\frac{x-y}{h_n}\right)d\Gamma(y,n)\right)^{2p}dx\right)^{\frac{1}{2}}\end{aligned}$$

$$\leq \frac{p}{2} n^{-\frac{p}{2}} h_n^{-p} T^{\frac{p}{2}-1} \sup_{0<x<\tau} \left| F_n(x) - F(x) \right|$$
$$\times \left(\int_0^T \left(\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\Gamma(y,n) \right)^{2p} dx \right)^{\frac{1}{2}} \tag{A.24}$$

Furthermore, since for each n,

$$\left| \Gamma_n^{(1)}(x) \right| \stackrel{D}{=} |N| h_n^{\frac{1}{2}} \left(\int_{\mathbb{R}} K^2(u) du \right)^{\frac{1}{2}}$$

(A.25) and a quick look at proofs of Lemma 3 and Lemma 2 imply that

$$\int_{\mathbb{R}} K\left(\frac{x-y}{h_n}\right) d\Gamma(y,n) = O_p\left(h_n^{\frac{1}{2}}\right) \tag{A.25}$$

Now, by (A.24), (A.25), and Lemma 5, we get

$$K_n^{(1)} = O_p\left(n^{-\frac{p}{2}} h_n^{-\frac{p}{2}} \left(\frac{\log\log n}{n} \right)^{\frac{1}{2}} \right) \tag{A.26}$$

Also,

$$K_n^2 \leq pn^{-\frac{p}{2}} h_n^{-(p+1)} \left(\int_0^T (F_n(x) - F(x))^2 x^p \left[\int_{\mathbb{R}} \Gamma(x - uh_n, n) dK(u) \right]^{2(p-1)} dx \right)^{\frac{1}{2}}$$
$$\times \left(\int_0^T \left[\int_{\mathbb{R}} \Gamma(x - uh_n, n) d\psi(u) \right]^2 dx \right)^{\frac{1}{2}}$$
$$\leq pn^{-\frac{p}{2}} h_n^{-(p+1)} T^p \sup_{x>0} |F_n(x) - F(x)|$$
$$\times \left\{ \int_0^T \left[\int_{\mathbb{R}} \Gamma(x - uh_n, n) dK(u) \right]^{2(p-1)} dx \right\}^{\frac{1}{2}}$$
$$\times \left\{ \int_0^T \left[\int_{\mathbb{R}} \Gamma(x - uh_n, n) d\psi(u) \right]^2 dx \right\}^{\frac{1}{2}}$$
$$:= pn^{-\frac{p}{2}} h_n^{-(p+1)} T^p \times R_{n,1} \times R_{n,2} \times R_{n,3} \tag{A.27}$$

By Lemma 5,

$$R_{n,1} = O\left(\left(\frac{\log\log n}{n} \right)^{\frac{1}{2}} \right) \quad a.s. \tag{A.28}$$

and (A.25) implies

$$R_{n,2} = O_p\left(h_n^{\frac{p-1}{2}} \right) \tag{A.29}$$

To deal with the term $R_{n,3}$, put

$$\Gamma_n(x) := \int_{\mathbb{R}} \psi\left(\frac{x-y}{h_n}\right) d(W(\sigma(y)) - F(y)W(\sigma))$$

where $W(\cdot)$ is a Wiener process. Clearly for each x, $\Gamma_n(x)$ is normally distributed with mean 0 and variance

$$\begin{aligned}\sum_{n,x} = &\int_{\mathbb{R}} \psi^2\left(\frac{x-y}{h_n}\right) d(\sigma(y)) + \sigma\Big(\int_{\mathbb{R}} \psi\left(\frac{x-y}{h_n}\right) dF(y)\Big)^2 \\ &-2\Big(\int_{\mathbb{R}} \psi\left(\frac{x-y}{h_n}\right) dF(y)\Big)\int_{\mathbb{R}} \psi\left(\frac{x-y}{h_n}\right) d\sigma(y) \\ := &\, g_1(x,n) + g_2(x,n) - g_3(x,n)\end{aligned}$$

Therefore, for each x,

$$\left|h_n^{-\frac{1}{2}}\Gamma_n(x)\right|^2 \overset{D}{=} |N|^2.\left|h_n^{-1}g_1(x,n) + h_n^{-1}g_2(x,n) - h_n^{-1}g_3(x,n)\right|$$

Since

$$\begin{aligned}h_n^{-1}g_1(x,n) &= h_n^{-1}\mu\int_{\mathbb{R}} \psi^2\left(\frac{x-y}{h_n}\right)\frac{f(y)}{y}dy \\ &= \mu\int_{\mathbb{R}} \psi^2(u)\frac{f(x-uh_n)}{(x-uh_n)}du \\ &\longrightarrow \frac{f(x)}{x}\int_{\mathbb{R}} \psi^2(u)du \quad as\ h_n \to 0\end{aligned}$$

and

$$h_n^{-1}g_2(x,n) = h_n\sigma\left(\int_{\mathbb{R}} \psi(u)f(x-uh_n)du\right)^2$$

one concludes that, as $h_n \to 0$, by the Cauchy–Schwarz inequality and dominated convergence theorem for some $M < \infty$,

$$\begin{aligned}&\int_{\mathbb{R}}\left|h_n^{-1}g_1(x,n) + h_n^{-1}g_2(x,n) - h_n^{-1}g_3(x,n)\right|dx \\ &\le \int_{\mathbb{R}} h_n^{-1}g_1(x,n)dx + \int_{\mathbb{R}} h_n^{-1}g_2(x,n)dx, \\ &\le \int_{\mathbb{R}} h_n^{-1}g_1(x,n)dx + MT\sigma\left(\int_{\mathbb{R}}|\psi(u)|^2du\right) \\ &\longrightarrow \left(\int_{\mathbb{R}}\psi^2(u)du\right)\left(\int_{\mathbb{R}}\frac{f(x)}{x}dx\right) + O(1) \\ &= O(1) + O(1) = O(1) \quad a.s.\end{aligned}$$

Hence

$$\left(\int_0^T \Gamma_n^2(x)dx\right)^{\frac{1}{2}} = O_p\left(h_n^{\frac{1}{2}}\right)$$

Furthermore, since $R_{n,3} \overset{D}{=} \left(\int_0^T \Gamma_n^2(x)dx\right)^{\frac{1}{2}}$ for each $n \ge 1$, (A.6) and (A.7) imply that

$$R_{n,3} = O_p\left(h_n^{\frac{1}{2}}\right) \tag{A.30}$$

Putting (A.28), (A.29), and (A.30) in (A.27)

$$K_n^{(2)} = O_p\left(n^{-\frac{p}{2}} h_n^{\frac{-(p+2)}{2}} \left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right) \tag{A.31}$$

By (A.23), (A.26), and (A.31), we can write

$$|II_n| = O_p\left(n^{-\frac{p}{2}} h_n^{-\frac{(p+2)}{2}} \left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right) \tag{A.32}$$

Also Lemma 5 implies that

$$\begin{aligned}\left|\int_0^T d\left(F_n(x) + F(x)\right)\right| &\leq |F_n(T) - F(T)| + 2F(T)\\ &= O_p(1)\end{aligned}$$

Hence

$$III_n = O_p\big(h_n^{3p}\big) + h_n^{-p} n^{-p[B+\frac{1}{2}]} O_p(1) \tag{A.33}$$

By (A.18), (A.22), (A.32), and (A.33), we obtain

$$C_2 = O_p\left(h_n^{2p}\right) + O_p\left(n^{-\frac{p}{2}} h_n^{-\frac{(p+2)}{2}} \left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right) + O_p\left(h_n^{-p} n^{-p[\frac{1}{2}+B]}\right) \tag{A.34}$$

Combining (A.17), (A.21), and (A.34) completes the proof. □