Theory of Stochastic Processes
Vol. 28 (44), no. 1, 2024, pp. 34–49

S. KHARDANI, W. NEFZI, AND C. THABET

RELATIVE ERROR PREDICTION FROM CENSORED DATA UNDER $\alpha-$MIXING CONDITION

In this paper, we address the case of a randomly right-censored model when the data exhibit some kind of dependency. We build and study a new nonparametric regression estimator by using the mean squared relative error as a loss function. Under classical conditions, we establish the uniform consistency with rate and asymptotic normality of the estimator suitably normalized.

1. Introduction

The problem of censored data arises in several applied fields, such as medicine, biology, public health, economics and demography.

In this paper, we are interested in estimating the relative regression function under right-censored data under an $\alpha-$mixing condition. The main motivations of this work are to propose a new estimator able to improve the performance of the other competitors such as the one introduced by Nadaraya [26] and the one proposed by Guessoum and Ould Said [15]. Relative error estimation has been recently used in regression analysis as an alternative to the restrictions imposed by the classical regression approach.

Let $\{Z_i = (X_i, T_i), 1 \leqslant i \leqslant n\}$ be n stationary random processes, identically distributed as the random pair $Z = (X, T)$ with values in $\mathbb{R}^d \times \mathbb{R}, (d \geqslant 1)$. A common problem in nonparametric statistics is the need to predict T given X. The ordinary way to study the relationship between X and T is to suppose that

$$T = r(X) + \epsilon, \tag{1}$$

where ϵ is a random error variable independent to X and r is a function obtained by minimizing the expected squared loss function

$$\mathbb{E}\left[(T - r(x))^2 |X = x\right].$$

In nonparametric forecasting, we often use the least squares and the least absolute deviation as criteria to construct the predictors.

The nonparametric estimation of the operator r is one of the most important tools to predict the relationship between T and X. There exist several nonparametric procedures allowing to estimate this operator. A popular one is the functional version of the Nadaraya-Watson estimator. The resulting estimator enjoys some important optimalities, such as simplicity, flexibility, and consistency.

However, for many practical situations in time series analysis the mean squared relative error is more appropriate as a measure of performance than the two previous criteria (see, for instance, Khoshgoftaar *et al.* [1] for some models in software engineering, Chatfield [8] for some examples in medicine or Chen [9] for some financial applications). Indeed the previous loss function as a measure of prediction performance may not be suitable in some situations. In particular, the presence of outliers can lead to unreasonable results since all variables have the same weight.

2000 *Mathematics Subject Classification.* 62G05, 62G08, 62G20, 62H12.

Key words and phrases. Asymptotic normality, Censored data, Consistency, Relative regression, Mean squared relative error.

In this paper, we construct a new estimator of the regression operator r. This estimator is obtained by combining the idea of the relative error regression (MSRE) and when the data are subject to random right censoring under an $\alpha-$mixing condition. We extend the work of Khardani and Slaoui [21] to the dependent data. Our alternative solution is to consider the estimation of the regression function via the minimization of the modified loss function

$$\mathbb{E}\left[\left(\frac{T-r\left(x\right)}{T}\right)^{2}|X=x\right], \quad \text{for} \quad T>0. \tag{2}$$

This criterium has been widely studied for parametric models, we refer to Chen *et al.* [9] for a discussion of the previous works and Hirose and Masuda [16] for a real example of electricity consumption. When the first two conditional inverse moments of T given X are finite, Park and Stefanski [28] showed that the solution of (2), for any fixed x, is given by the following ratio

$$r(x)=\frac{\mathbb{E}[T^{-1}|X=x]}{\mathbb{E}[T^{-2}|X=x]}=:\frac{m_1(x)}{m_2(x)}.$$

where $m_\ell(x)=r_\ell(x)/f(x)$ and $r_\ell(x)=\int_{\mathbb{R}}t^{-\ell}f_{X,T}(x,t)dt$ for $\ell=1,2$, where $f_{X,T}(\cdot,\cdot)$ and $f_X(\cdot)$ are the joint and marginal density of the couple (T,X) and X respectively. Several works have studied the relative regression problem. For example, Jones *et al.* [17] studied nonparametric prediction via relative error regression. Khardani [20] considered the problem of non-parametric relative regression for twice censored data. Mechab and Laksaci [24] studied this regression model when the observations are weakly dependent. Recently, Khardani and Slaoui [21] studied the asymptotic properties of a consistent estimator of this model by using the kernel method for twice-censored data. Moreover, Hu [14] established the consistency and the asymptotic normality of the regression function based on the least product relative error. Bouhadjara and Ould Said [2] studied a nonparametric local linear estimation of the relative error regression function for the censorship model. Narula and Wellington [27] studied an estimation method for minimizing the sum of absolute relative residuals. Farnum [13] developed an estimation method designed to reduce absolute relative error. Park and Stefanski [28] studied prediction for situations in which relative prediction error is more important than the usual prediction error and derived the form of the best mean squared relative error predictor.

The main purpose of the present work is to consider a general framework and the characterization of the asymptotic properties of the kernel relative regression estimators based on censored and dependent data, this generalization of the work of Jones *et al.* [17] and Khardani [20] is far from being trivial and harder to control the estimator of Kaplan-Meier and the mixing condition, which form a basically unsolved open problem in the literature. We aim to fill this gap in the literature by combining Jones *et al.* [17] results with techniques handling the Kaplan Meier estimate. However, as will be seen later, the problem requires much more than "simply' combining ideas from the existing results.

The paper is organized as follows: we present our model and the general idea of the mean squared relative error function in Section 2. Assumptions and theoretical results are given in Section 3. In Section 4, we examine the performances of our estimator with a simulation study. Finally, the proofs are given in Section 5.

2. Description of the model and estimator

We consider a regression model in which the response variable is subject to random right censoring. Censored data are present in many practical applications in a wide variety of fields, including economics, medicine, biology, and biostatistics. For example, let $(T_i)_{1\leqslant i\leqslant n}$ be the survival (or failure) time of individuals who are involved in a clinical study or a possible monotone transformation of it. We consider $(X_i)_{1\leqslant i\leqslant n}$ a random covariate, such as the age, the dose of a drug or the cholesterol level. As often occurs in practice, the response $(T_i)_{1\leqslant i\leqslant n}$ is subject to random right censoring. In other words, instead of observing $(T_i)_{1\leqslant i\leqslant n}$, one observes the n pairs of variables (Y_i, δ_i), where

$$Y_i = \min(T_i, C_i), \delta_i = 1_{\{T_i \leq C_i\}},$$

and C_i represents the censoring time with common continuous distribution function (df) G, which is assumed to be independent of T_i.

For any distribution function L let $\tau_L = \sup\{t, L(t) < 1\}$ be its support's right endpoint. Further, we denote by $F(\cdot)$ (resp. $G(\cdot)$) the df of T (resp. of C) and by τ_F (resp. τ_G) the upper endpoints of the survival function $\overline{F}$ (resp. of $\overline{G}$). In what follows, we assume that $\tau_F < \infty$, $\overline{G}(\tau_F) > 0$, $\tau_H < \min(\tau_F, \tau_G)$ and C is independent to (X, T).

In this kind of model, it is well known that the empirical distribution is not a consistent estimator for the distribution function G. Therefore, Kaplan and Meier (1958) proposed a consistent estimator $\bar{G}_n(\cdot)$, for the survival function $\bar{G} = 1 - G$. which is defined as

$$\bar{G}_n(t) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1-\delta_{(i)}}{n-i+1}\right)^{1\{Y_{(i)} \leq t\}} & \text{if} \quad t < Y_{(n)}, \\ 0 & \text{otherwise}, \end{cases}$$

where $Y_{(1)} < Y_{(2)} < ... < Y_{(n)}$ are the order statistics of $(Y_i)_{1\leq i\leq n}$ and $\delta_{(i)}$ is the concomitant of $Y_{(i)}$.

First, we recall that, the model (1) suffers from censorship data, for that, we use the so-called "synthetic data" which allows us to take into account the censoring effect on the lifetime distribution (for more details, we refer to Carbonez et al. [7] and Kohler et al. [19]).

The extension of nonparametric estimation procedures to the censored framework requires replacing the unavailable data with a suitable construction of the observed data given by

$$\varphi(Y_i^*) = \frac{\delta_i \varphi(Y_i)}{\overline{G}(Y_i)}, \quad 1 \leq i \leq n \tag{3}$$

for any measurable function φ.
Assuming a sequence of covariates is given, we then observe the triplets $(Y_i, \delta_i, X_i)_{1\leqslant i\leqslant n}$. All along this paper, we suppose that:

$$(T_i, X_i)_i \quad \text{and} \quad (C_i)_i \quad \text{are independent.} \tag{4}$$

Then from (3) and (4) we get

$$\begin{aligned} \mathbb{E}\left[\varphi(Y_1^*)|X_1\right] &= \mathbb{E}\left[\frac{\delta_1\varphi(Y_1)}{\overline{G}(Y_1)}|X_1\right] = \mathbb{E}\left\{\mathbb{E}\left[\frac{\delta_1\varphi(Y_1)}{\overline{G}(Y_1)}|T_1, X_1\right]|X_1\right\} \\ &= \mathbb{E}\left\{\mathbb{E}\left[\frac{\delta_1\varphi(T_1)}{\overline{G}(T_1)}|T_1, X_1\right]|X_1\right\} = \mathbb{E}\left\{\frac{\varphi(T_1)}{\overline{G}(T_1)}\mathbb{E}\left[1_{\{T_1\leq C_1\}}|T_1\right]|X_1\right\} \\ &= \mathbb{E}\left(\varphi(T_1)|X_1\right). \end{aligned} \tag{5}$$

In the current work, we propose to use the MSRE rather than the MSE criterion. Thus, our proposed estimators are: Firstly, we consider a "pseudo-estimator" given by

$$\widetilde{r}_n(x) = \frac{\widetilde{m}_{1,n}(x)}{\widetilde{m}_{2,n}(x)},$$

where

$$\widetilde{m}_{\ell,n}(x) = \frac{1}{nh_n^d}\sum_{i=1}^n \frac{\delta_i Y_i^{-\ell}}{\overline{G}(Y_i)} K\left(\frac{X_i - x}{h_n}\right) \quad \text{for} \quad \ell \in \{1, 2\}.$$

Second, we consider the estimator:

$$r_n(x) = \frac{\widetilde{m}_1(x)}{\widetilde{m}_2(x)}, \tag{6}$$

where

$$\widetilde{m}_\ell(x) = \frac{1}{nh_n^d}\sum_{i=1}^{n}\frac{\delta_i Y_i^{-\ell}}{\overline{G}_n(Y_i)}K\left(\frac{X_i - x}{h_n}\right) \quad \text{for} \quad \ell \in \{1,2\}.$$

In what follows, we suppose that $X_1, \ldots, X_n$ satisfy the α-mixing dependency property, whose definition is given below.

Let $Z = (Z_i)_{i\geqslant 1}$ be a sequence of random variables. Given a positive integer n, set

$$\alpha(n) = \sup_k \sup\left\{|\,\mathbb{P}(A\cap B) - \mathbb{P}(A)\,\mathbb{P}(B)\,|,\ A \in \mathcal{F}_1^k(Z) \text{ and } B \in \mathcal{F}_{k+n}^\infty(Z)\right\},$$

where $\mathcal{F}_i^k(Z)$ is the σ-field of events generated by $\{Z_j, i \leqslant j \leqslant k.\}$
The sequence is said to be α-mixing if the mixing coefficient $\alpha(n) \longrightarrow 0$ as $n \longrightarrow \infty$. The α-mixing condition, also called strong mixing is the weakest among mixing conditions known in the literature. Many stochastic processes satisfy the α-mixed condition: the ARMA processes are geometrically strongly mixing, *i.e.* there exists $\rho \in (0,1)$ such that, $\alpha(n) = O(\rho^n)$. The threshold models, the EXPAR model, the simple ARCH models, their GARCH extension and the bilinear Markovian models are geometrically and strongly mixing, under some general ergodicity conditions.

The α/ρ-mixing has many practical applications (see, *e.g.*, Cai, (1998, 2001), Bradley, (2007) and Dedecker *et al.*, (2007) for more details). As an example (see Bosq (1999)): A linear process is defined by

$$X_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j} \quad t \in \mathbb{Z},$$

where $a_j = O(e^{-rj}), r > 0$ and ϵ are independent with zero-mean real random variables with a common density and a finite second moment. Then the series above converge in quadratic mean, and (X_t) is ρ–mixing and therefore α–mixing with coefficients that decrease to zero at an exponential rate. For more examples of α-mixing models (see Doukhan (1994)).

3. Assumptions and main results

Throughout the paper, we denote by $\mathbb{E}[T^{-\ell}|X = \cdot]$ the conditional ℓ-inverse moments of T given X and by $m_\ell(\cdot) = \mathbb{E}[T^{-\ell}|X = \cdot]f(\cdot)$, with f the density of X; $\ell \in \{1,2,3,4\}$ and let $\gamma_\ell(\cdot) = \mathbb{E}[\frac{T^{-\ell}}{\bar{G}(T)}|X = \cdot]f(\cdot)$. When no confusion is possible, we denote by M and/or M' any generic positive constant. In this section, we aim to establish the uniform convergence ($a.s.$) of $r_n(x)$ to $r(x)$ over a compact S.

Finally, let us consider the following assumptions about the process (X_n, T_n).

M1 $(\xi_n = (X_n, T_n)_{n\in\mathbb{N}})$ is a strong mixing process such that the mixing coefficient satisfies $\alpha(n) = O(n^{-\nu})$ for some $\nu > 4$.

(D1) For each $k \neq k'$, one has :

$$C := \sup_{|k-k'|\geq 1}\left\|f_{(X_k,X_{k'})|(Y_k,Y_{k'})}(s|u,t|v) - f_{X_k|Y_k}(s|u)\, f_{X_{k'}|Y_{k'}}(t|v)\right\|_\infty < \infty$$

D2: The joint density $f^*_{k,k'}$ of $((X_k,T_k),(X_{k'},T_{k'}))$ exists and satisfies

$$\sup_{(\mathbb{R}^d\times\mathbb{R})^2}\left|f^*_{k,k'}(\cdot,\cdot) - f_k(\cdot,\cdot)f_{k'}(\cdot,\cdot)\right| \leq \mathcal{C} < \infty, \quad \text{for any } |k-k'| \geq 1.$$

D3: The joint density $f(\cdot,\cdot)$ is bounded and differentiable up to order 3. Moreover, all the partial derivatives are uniformly bounded up to order 3.

(K1) The kernel K is bounded, symmetric and has a compact support. It is also Holder of order γ and $\int_{\mathbb{R}^d} u_\ell^i K^j(u)du < \infty$ with $u = (u_1, \cdots, u_d)^T$ for $\ell = 1, \cdots, d$, $i \in \{0,1,2\}$ and $j \in \{1,2\}$.
(K2) The function $m_\ell(x)$ is twice differentiable and $\sup_{x\in S} |m_\ell''(x)| < +\infty, \quad \ell \in \{1,2\}$.
(K3) $m_2(x) > M, \forall \quad x \in S$, the inverse moments of the response variable satisfy: $\forall k \geqslant 1, \quad \overline{r}_k(x) = \mathbb{E}\left[\frac{T^{-k}}{\overline{G}(T)}|X = x\right]$ exists and are of class C^2 in S.
(H1) For $d \geqslant 1$, the bandwidth (h_n) satisfies:
(i) $\lim_{n\to+\infty} \frac{\log n}{nh_n^d} = 0$,
$(ii) \quad nh_n^{\frac{d}{4}(\nu-4)(\nu+4)} \to 0, \quad nh_n^{d\left(\frac{a+1-\nu}{\nu}\right)} \to \infty$, for a such that $1 < a < \nu - 1$, where ν is as in **M1**.

Discussion of the assumptions. Assumption **M1** specify the model and the rate of mixing coefficients. **D1–D3** are mild regularity assumptions that are usually required to obtain convergence rates and the asymptotic normality in regression kernel estimation. More precisely, **D1-D2** allow to get the same rates as for the iid case. Regarding Assumptions **K1–K3**, we refer to Silverman [30] for the univariate case. Finally, **H1** gives a condition for the bandwidth which allows the estimation of the covariance term.

The independence assumption in (4) between $(C_n)_n$ and $(X_n, T_n)_n$ may seem to be strong, and one can think of replacing it with a classical conditional independence assumption between $(C_n)_n$ and $(T_n)_n$ given $(X_n)_n$.

Theorem 3.1. *Under Assumptions **(M)**, **(D)**, **(K)** and **(H)**, we have*

$$\sup_{x\in S} |r_n(x) - r(x)| = O\left(\max\left\{\sqrt{\frac{\log n}{nh_n^d}}, h_n^2\right\}\right) \qquad a.s.$$

For $h_n = \left(\left(\frac{\log n}{n}\right)^{1/(d+4)}\right)$, we have

$$\sup_{x\in S} |r_n(x) - r(x)| = O\left(\left(\frac{\log n}{n}\right)^{2/(d+4)}\right) \qquad \text{a.s.}$$

Theorem 3.2. *If the assumptions **(M)**, **(D)**, **(K)** and **(H)**, hold, then for all* $\ell = (1,2)$

$$Var(\widetilde{m}_{1,n}(x)) = \frac{1}{nh_n^d}\mathbb{E}\left[\frac{T_i^{-2}}{\overline{G}(T_i)}|X_i = x\right] f(x) \int_{\mathbb{R}^d} K^2(z)dz, \ a.s. \tag{7}$$

$$Var(\widetilde{m}_{2,n}(x)) = \frac{1}{nh_n^d}\mathbb{E}\left[\frac{T_i^{-4}}{\overline{G}(T_i)}|X_i = x\right] f(x) \int_{\mathbb{R}^d} K^2(z)dz, \ a.s. \tag{8}$$

$$nh_n^d Cov[\widetilde{m}_{1,n}(x), \widetilde{m}_{2,n}(x)] \to \mathbb{E}\left[\frac{T^{-3}}{\overline{G}(T)}|X = x\right] f(x) \int_{\mathbb{R}^d} K^2(z)dz, \ a.s. \tag{9}$$

Theorem 3.3. *Under assumptions **(M)**, **(D)**, **(K)** and **(H)**, then for any $x \in A$, we obtain*

$$\left(\frac{nh_n^d}{\sigma^2(x)}\right)^{1/2} (r_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty,$$

where

$$\sigma^2(x) = \frac{\{\gamma_2(x) - 2r(x)\gamma_3(x) + r^2(x)\gamma_4(x)\}}{m_2^2(x)} \int_{\mathbb{R}^d} K^2(x)dx.$$

and

$$A = \left\{ x \in S, \frac{\{\gamma_2(x) - 2r(x)\gamma_3(x) + r^2(x)\gamma_4(x)\}}{m_2^2(x)} \neq 0 \right\},$$

$\xrightarrow{\mathcal{D}}$ *denotes convergence in distribution,* $\mathcal{N}$ *the Gaussian distribution.*

Remark 3.1. **(Comeback to complete data).** In absence of censoring ($\overline{G}(\cdot) = 1$), the asymptotic variance becomes

$$\sigma^2(x) = \frac{\{m_2(x) - 2r(x)m_3(x) + r^2(x)m_4(x)\}}{m_2^2(x)} \int_{\mathbb{R}^d} K^2(x)dx.$$

4. Simulations results

In this section, we discuss the feasibility and the performance of the relative nonparametric method through an empirical study. Precisely, our main purpose is to compare the efficiency, in the sense of the MSE, between relative and classical kernel regression estimation in the time series case. For this aim, we consider a sequence of m-dependent variables

$$X_i = \sum_{i}^{i+m} Z_i^2$$

where $(Z_i)_i$ are independent standard normal variables. Obviously, these random variables form an α–mixing process satisfying Assumption (M_1). Concerning, the response variables, we use the following regression model to generate them

$$T_i = 2\cos(\frac{3}{2}X_i) + \epsilon_i, \qquad i = 1, \dots n.$$

For this simulation study, we consider two types of $(\epsilon_i)_i$ that are generated from the following normal mixture distributions (see, Marron. and Wand [23] for more details on the normal mixture distributions):

Law	Distribution function
Standard normal distribution (S.N.D.)	$\mathcal{N}(0,1)$
Skewed bimodal distribution (S.B.D.)	$\frac{3}{4}\mathcal{N}(0,1) + \frac{1}{4}\mathcal{N}(\frac{3}{2}, \frac{1}{9})$

Recall that, our main goal of this empirical study is to show the applicability of our procedure in practice, with special attention on the influence of the three fundamental parameters involved in this approach, namely the mixing condition, the percentage of censoring and the number of outliers data. While the first (mixing) is controlled by m, we control the effect of the censoring model by considering censoring variables C distributed by a normal distribution $N(0,\sigma)$ and the percentage of the outliers data. So, the behavior of both estimators is evaluated over several parameters, such as the dimension of the regressors, the sample size n, the percentage of censoring τ (controlled by σ), the correlation of the data (controlled by m) and the number of the outliers data (see Table 1). For the sake of shortness, we consider only the unidimensional case. We fix the sample size $n = 200$ and consider three censoring types ($\tau = 8$, $\tau = 37$ and $\tau = 65$) and three dependence cases ($m = 2, 6, 10$).

We use the Gaussian kernel and we consider the well-known smoothing parameter defined by $h_n = \sigma_n^2 n^{-1/5}$ where

$$\sigma_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad \text{and} \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

We compare the values of the mean squared error (MSE) for the classical regression (MSE^1) (see Guessoum and Ould Said (2010) and the relative error regression (MSE^2) methods represented by

$$MSE^1 = \frac{1}{n}\sum_{i=1}^{n}(\hat{r}_n(X_i) - Y_i)^2 \quad \text{and} \quad MSE^2 = \frac{1}{n}\sum_{i=1}^{n}(r_n(X_i) - Y_i)^2.$$

Cond. dis.	m	τ	MSE^1	MSE^2
S.N.D	2	8	0.12	0.15
		37	0.38	0.36
		65	0.87	0.65
	6	8	0.34	0.42
		37	0.76	0.74
		65	1.24	1.17
	10	8	0.66	0.62
		37	1.22	1.09
		65	1.64	1.56
S.B.D	2	8	0.86	1.15
		37	1.57	1.42
		65	2.19	2.02
	6	8	1.03	1.20
		37	1.94	1.57
		65	2.34	2.12
	10	8	1.31	1.42
		37	2.18	1.95
		65	2.74	2.87

Table 1 The MSE errors according to the censoring rates and dependence degree.

It appears that the performance of both estimators is less affected by the percentage of censoring and the degree of dependence. However, the quality of estimation is dramatically destroyed in cases of strong dependence and strong censoring.

To show the robustness of our approach and the effect of outliers data, we generate the case where the data contains outliers (see Table 2). For this purpose, we set the sample size and the censorship rate (n = 400, $\tau = 40\%$ and m=6). To create this outlier effect, the number of values of this sample is multiplied by a factor called *MF*.

Outliers	1	5	10	15	20	30	40
MSE^1	0.02	0.525	7.06	164.1753	658.1290	5.4 e+10	9.478e+12
MSE^2	0.012115	0.015677	0.012498	0.012450	0.012450	0.012450	0.012450

Table 2. The MSE errors according to the number of outlier data with $\tau = 40\%$ and $m = 6$.

Conclusion

In this paper, we have investigated the asymptotic properties of a nonparametric estimator of the relative error regression given a dependent explanatory variable, in the case of a scalar censored response. We have used the mean squared relative error as a loss function to construct a nonparametric estimator of the regression operator of these censored functional data. We have established the almost surely convergence and asymptotic normality of the proposed estimator. Additionally, a simulation study was performed to support the theoretical results and to compare the quality of predictive performances of the relative error regression estimator with those obtained from standard kernel regression estimates.

5. Auxiliary results and proofs

To prove our results, we state the following lemmas, which will be proven in the Appendix subsection.

Lemma 5.1. *Under assumptions of Theorem 3.1 and for $\ell \in \{1,2\}$, we have*

$$\sup_{x\in S}|\widetilde{m}_\ell(x) - \widetilde{m}_{\ell,n}(x)| = O\left(\sqrt{\frac{\log\log n}{n}}\right).$$

Lemma 5.2. *Under assumptions of Theorem 3.1 and for $\ell \in \{1,2\}$, we have*

$$\sup_{x\in S}|\mathbb{E}(\widetilde{m}_{\ell,n}(x)) - m_{\ell,}(x)| = O(h_n^2).$$

Lemma 5.3. *Under assumptions of Theorem 3.1, we obtain*

$$\sup_{x\in S}|\widetilde{m}_{\ell,,n}(x) - \mathbb{E}(\widetilde{m}_{\ell,,n}(x))| = O\left(\sqrt{\frac{\log n}{nh_n^d}}\right).$$

Proof of Theorem 3.1. We consider the following decomposition:

$$\begin{aligned}\sup_{x\in S}|r_n(x) - r(x)| \leq \frac{1}{\inf_{x\in S}\widetilde{m}_2(x)}\Bigg\{&\sup_{x\in S}|\widetilde{m}_1(x) - \widetilde{m}_{1,n}(x)| + \sup_{x\in S}|\widetilde{m}_{1,n}(x) - \mathbb{E}(\widetilde{m}_{1,n}(x))| \\ &+ \sup_{x\in S}|\mathbb{E}(\widetilde{m}_{1,n}(x)) - m_1(x)| + \sup_{x\in S}(|m_1(x)|)\Bigg\{\sup_{x\in S}|\widetilde{m}_2(x) - \widetilde{m}_{2,n}(x)| \\ &+ \sup_{x\in S}|\widetilde{m}_{2,n}(x) - \mathbb{E}(\widetilde{m}_{2,n}(x))| + \sup_{x\in S}|\mathbb{E}(\widetilde{m}_{2,n}(x)) - m_2(x)|\Bigg\}\Bigg\}.\end{aligned}$$

Thus, the proof of Theorem 3.1 is a direct consequence of Lemmas 5.1-5.3. □

Proof of Theorem 3.3. From (6), we have the following decomposition:

$$r_n(x) - r(x) = r_n(x) - \widetilde{r}_n(x) + \widetilde{r}_n(x) - r(x) =: \mathcal{I}_{1n}(x) + \mathcal{I}_{2n}(x)$$

where

$$\mathcal{I}_{1n}(x) =: r_n(x) - \widetilde{r}_n(x) \quad \text{and} \quad \mathcal{I}_{2n}(x) =: \widetilde{r}_n(x) - r(x).$$

The proof is derived by showing first that $\mathcal{I}_{1n}(x)$ is negligible, whereas $\mathcal{I}_{2n}(x)$ is asymptotically normal distributed.

First, we can write that

$$\mathcal{I}_{2n}(x) \quad = \quad \frac{1}{\widetilde{m}_2(x)}[B_n + V_n(\widetilde{m}_2(x) - \mathbb{E}[\widetilde{m}_2(x)])] + V_n,$$

where

$$\begin{aligned}V_n &= \frac{1}{\mathbb{E}[\widetilde{m}_{2,n}(x)]\, m_2(x)}\{\mathbb{E}[\widetilde{m}_{1,n}(x)]\, m_2(x) - [\mathbb{E}[\widetilde{m}_{2,n}(x)]]m_1(x)\} \\ B_n &= \frac{1}{m_2(x)}[[\widetilde{m}_{1,n}(x) - \mathbb{E}[\widetilde{m}_{1,n}(x)]]m_2(x) + [\mathbb{E}[\widetilde{m}_{2,n}(x)] - \widetilde{m}_{2,n}(x)]m_1(x)].\end{aligned}$$

Then, we have,

$$\begin{aligned}\widetilde{r}_n(x) - r(x) - V_n &= \frac{1}{\widetilde{m}_{2,n}(x)}[B_n + V_n(\widetilde{m}_{2,n}(x) - \mathbb{E}[\widetilde{m}_{2,n}(x)])] \\ &=: \frac{B_n + V_n\mathcal{J}_{2n}(x)}{\widetilde{m}_{2,n}(x)},\end{aligned}$$

where

$$\mathcal{J}_{jn}\left(x\right)=\widetilde{m}_{j,n}(x)-\mathbb{E}\left[\widetilde{m}_{j,n}(x)\right],\quad \text{for}\quad j=1,2.$$

Concerning $\mathcal{I}_{2n}(x)$, on the one hand, the denominator $\widetilde{m}_{2,n}(x)$ converges to $m_2(x)$ (see Lemma 5.5), while for the numerator, both

$$\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2}V_n\quad\text{and}\quad\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2}V_n\mathcal{J}_{2n}(x)$$

are negligible (see the following Lemma 5.5). In contrast, B_n is asymptotically normal (see Lemma 5.4 below). For the latter, we first evaluate its asymptotic variance before dealing with its asymptotic normality. Finally, Theorem 3.3 is a consequence of Lemmas 5.1–5.5. □

Lemma 5.4. *Under assumptions of Theorem 3.3, we obtain*

$$\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2}\left(B_n-\mathbb{E}[B_n]\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1).$$

Lemma 5.5. *Under assumptions of Theorem 3.3, we obtain*

$$\widetilde{m}_{2,n}(x)\xrightarrow{\mathbb{P}}m_2(x),$$

$$\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2}V_n\to 0,$$

and

$$\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2}V_n\left(\widetilde{m}_{2,n}(x)-\mathbb{E}\left[\widetilde{m}_{2,n}(x)\right]\right)\xrightarrow{\mathbb{P}}0,$$

where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability.

Proof of Theorem 3.2. **Proof of (7)**: Using equation (5) and Lemma 1, we have

$$\begin{aligned}Var(\widetilde{m}_{1,n}(x)) &= \sum_{i=1}^{n}Var\left[\nabla_{i,n}^1(x)\right]+2\sum_{1\le i<j\le n}Cov\left[\nabla_{i,n}^1(x)\nabla_{j,n}^1(x)\right].\\ &:= \mathcal{I}_{1,n}+\mathcal{I}_{2,n}.\end{aligned}$$

Note that for any function φ, we have $\delta_i\varphi(Y_i)=\delta_i\varphi(T_i)$. Then, since $G(\cdot)$ is continuous, we have:

$$\begin{aligned}\mathcal{I}_{1,n}=\sum_{i=1}^{n}\text{Var}\left[\nabla_{i,n}^1(x)\right] &= \frac{1}{nh_n^{2d}}\mathbb{E}\left[\frac{T_1^{-2}}{\overline{G}^2(T_1)}K^2\left(\frac{X_1-x}{h_n}\right)\mathbb{E}\left[1_{\{T_i\le C_i\}}|X_i,T_i\right]\right]\\ &\quad-\frac{1}{nh_n^{2d}}\mathbb{E}^2\left[\frac{T_1^{-1}}{\overline{G}^1(T_1)}K\left(\frac{X_1-x}{h_n}\right)\mathbb{E}\left[1_{\{T_i\le C_i\}}|X_i,T_i\right]\right]\\ &=: \mathcal{K}_{1n}(x)+\mathcal{K}_{2n}(x).\end{aligned}$$

For $\mathcal{K}_{2n}(x)$, we once again use the conditional expectation properties (see equation (5)) to get, under **K1**, **K3** and **H1**,

$$\mathcal{K}_{2n}(x)=\frac{1}{nh_n^{2d}}\mathbb{E}^2\left[T_1^{-1}K\left(\frac{X_1-x}{h_n}\right)\right]=o(1). \tag{10}$$

For $\mathcal{K}_{1n}(x)$, once again using the conditional expectation properties (see equation (5)), Taylor's expansion with integral remainder and assumptions **D3**, **K1** and **K3**, we get,

$$\text{(11)}\quad \mathcal{K}_{1n}(x) = \frac{1}{nh_n^{2d}}\mathbb{E}\left[\frac{T_1^{-2}}{\overline{G}^2(T_1)}K^2\left(\frac{X_1-x}{h_n}\right)\mathbb{E}\left[1_{\{T_i\le C_i\}}|X_i,T_i\right]\right]$$

$$= \frac{1}{nh_n^{2d}}\mathbb{E}\left[K^2\left(\frac{X_1-x}{h_n}\right)\mathbb{E}\left[\frac{T_1^{-2}}{\overline{G}^1(T_1)}\middle| X=v\right]\right] = \frac{1}{nh_n^d}\int_{\mathbb{R}^d}K^2\left(\frac{x-v}{h_n}\right)\gamma_2(v)dv$$

$$= \frac{1}{nh_n^d}\int_{\mathbb{R}^d}K^2(u)\left[\gamma_2(x) - \sum_{j=1}^{d}\frac{\partial\gamma_2(x)}{\partial x_j}(h_n u_j)\right.$$

$$\left. + \int_0^1 (1-s)\sum_{i,j=1}^{n}\frac{\partial^2\gamma_2(x-sh_n u)}{\partial x_i\partial x_j}(u_i h u_j h)ds\right]du$$

$$= \frac{1}{nh_n^d}\int_{\mathbb{R}^d}K^2(w)\left(\gamma_2(x)\right)dw + o\left(\frac{1}{nh_n^d}\right).$$

For what follows and up to the equation (17) below, we will show that the covariance quantity $\mathcal{I}_{2,n}$ is zero.

Hence, since $\nabla^{\ell}_{i,n}(x)$ is centered, we have

$$\mathcal{I}_{2,n} = 2\sum_{i=1}^{n}\sum_{i\neq j\ j=1}^{n} Cov\left[\nabla^1_{i,n}(x)\nabla^1_{j,n}(x)\right].$$

Notice that for any function φ, we have $\delta_i\varphi(Y_i) = \delta_i\varphi(T_i)$. Then, since $G(\cdot)$ is continuous, we have:

$$\text{(12)}\quad Cov\left[\nabla^1_{i,n}(x)\nabla^1_{j,n}(x)\right] = \frac{1}{(nh_n^d)^2}\int_{\mathbb{R}^d}\int_{\mathbb{R}}\int_{\mathbb{R}^d}\int_{\mathbb{R}}K\left(\frac{x-u}{h}\right)(rv)^{-1}K\left(\frac{x-s}{h}\right)$$

$$\times f_{1,j-i+1}(u,r,s,v)dudrdsdv - \frac{1}{(nh_n^d)^2}\left[\int_{\mathbb{R}^d}\int_{\mathbb{R}}K\left(\frac{x-u}{h}\right)v^{-1}f(u,v)dudv\right]^2$$

$$= O\left(\frac{1}{n^2}\right).$$

Now, following Masry [22], we define the sets

$$\mathcal{B}_1 = \{(i,j),\ \text{such that}\ 1\le |i-j|\le \eta_n\}\ \text{ and}$$
$$\mathcal{B}_2 = \{(i,j),\ \text{such that}\ \eta_n+1\le |i-j|\le n-1\},$$

where $\eta_n = o(n)$. Then, we have

$$\text{(13)}\quad \sum_{1\le i,j\le n}|Cov\left[\nabla_{i,n}(x)\nabla_{j,n}(x)\right]| = \sum_{(i,j)\in\mathcal{B}_1}|\mathbb{E}\left[\nabla_{i,n}(x)\nabla_{j,n}(x)\right]|$$

$$+ \sum_{(i,j)\in\mathcal{B}_2}|\mathbb{E}\left[\nabla_{i,n}(x)\nabla_{j,n}(x)\right]|.$$

By (12), we have

$$\text{(14)}\quad \sum_{(i,j)\in\mathcal{B}_1}|\mathbb{E}(\nabla_{i,n}(x)\nabla_{j,n}(x))| = \sum_{i=1}^{n-\eta_n}\sum_{j=i+1}^{\eta_n+i}|\mathbb{E}(\nabla_{i,n}(x)\nabla_{j,n}(x))| = O\left((\eta_n/n)\right).$$

To bound the sum over $\mathcal{B}_2$ in (13), we use the moment inequality in Rio [29] (page 10, Formula 1.12b). For that let a, b and c be real numbers greater than 1 such that

$1/a + 1/b + 1/c = 1$, then

$$\left|\mathbb{E}(\nabla^1_{i,n}(x,t)\nabla^1_{j,n}(x,t))\right| \leq \frac{1}{n^2h_n^{2d}} \frac{2^{1+\frac{1}{a}}(\alpha(|j-i|))^{1/a}}{\overline{G}^2(\tau)} \tag{15}$$

$$\times \left[\mathbb{E}\left|K_d\left(\frac{x-X_i}{h_n}\right)\left(\frac{\delta_i Y_i^{-1}}{\overline{G}(Y_i)}\right)\right|^b\right]^{1/b} \left[\mathbb{E}\left|K_d\left(\frac{x-X_j}{h_n}\right)\left(\frac{\delta_j (Y_j^{-1})}{\overline{G}(Y_j)}\right)\right|^c\right]^{1/c}.$$

Furthermore, observe that, under **K1**, **D3** and **M1**, we have for any (x,t)

$$\mathbb{E}\left|K_d\left(\frac{x-X_i}{h_n}\right)\left(\frac{\delta_i Y_i^{-1}}{\overline{G}(Y_i)}\right)\right|^b = h_n^d \int_{\mathbb{R}^d}\int_{\mathbb{R}} \left|K_d(s)r^{-1}\right|^b f(x-sh,r)dsdr = O\left(h_n^d\right),$$

which, when replaced in (15), yields (the (T_j, X_j)-term with exponent c is dealt with in the same manner)

$$\sum_{(i,j)\in\mathcal{B}_2} |\mathbb{E}(\nabla_{i,n}(x,t)\nabla_{j,n}(x,t))| \leq \mathcal{C}\frac{1}{n^2h_n^{2d}} h_n^{d(\frac{1}{b}+\frac{1}{c})} \sum_{(i,j)\in\mathcal{B}_2} (\alpha(|j-i|))^{1/a} \tag{16}$$

$$= \mathcal{C}\frac{1}{n^2h_n^{2d}} h_n^{(d(\frac{1}{d}+\frac{1}{c}))} \sum_{l=\eta_n+1}^{n-1} (n-l)(\alpha(l))^{1/a} = O\left(n^{-1}h_n^{-d(\frac{a+1}{a})}\eta_n^{1-\nu/a}\right).$$

Now, equalizing ((14)) and ((16)), we get

$$\eta_n \sim h_n^{-d\frac{(a+1)}{\nu}},$$

which replaced in (13) gives

$$\sum_{1\leq i,j\leq n} |\mathbb{E}(\nabla_{i,n}(x,t)\nabla_{j,n}(x,t))| = O\left(\frac{1}{nh_n^{\frac{a+1-\nu}{\nu}}}\right).$$

Under **H1**, we get

$$\mathcal{I}_{2,n} = o(1). \tag{17}$$

Then, (10),(11), (17) and a Taylor expansion ensures that

$$Var(\widetilde{m}_{1,n}(x)) = \frac{1}{nh_n^d}\mathbb{E}\left[\frac{T_i^{-2}}{\overline{G}(T_i)}|X_i = x\right] f(x)\int_{\mathbb{R}^d} K^2(z)dz.$$

Next, similar to the proof of (7) we find

$$Var(\widetilde{m}_{2,n}(x)) = \frac{1}{nh_n^d}\mathbb{E}\left[\frac{T_i^{-4}}{\overline{G}(T_i)}|X_i = x\right] f(x)\int_{\mathbb{R}^d} K^2(z)dz.$$

Proof of (9): Using the following decomposition:

$$\text{Cov}\left[\widetilde{m}_{1,n}(x), \widetilde{m}_{2,n}(x)\right] = \left[nh_n^d\right]^{-2}\left[\sum_{i=1}^n A_{ii} + \sum_{i=1}^n \sum_{i\neq j}\sum_{j=1}^n A_{ij}\right] := F_1 + F_2,$$

where for any s, t:

$$A_{s,t} := \text{Cov}\left[\frac{\delta_s Y_s^{-1}}{\overline{G}(Y_s)}K\left(\frac{x-X_s}{h_s}\right), \frac{\delta_t Y_t^{-2}}{\overline{G}(Y_t)}K\left(\frac{x-X_t}{h_t}\right)\right].$$

For F_2, choosing $c_n = o(n) \to +\infty$, we have:

$$F_2 \leq 2\left[nh_n^d\right]^{-2}\left[\sum_{k=1}^{c_n}\sum_{p=1}^n |A_{p,k+p}| + \sum_{k=c_{n+1}}^{n-1}\sum_{p=1}^n |A_{p,k+p}|\right] := F_{21} + F_{22}.$$

Next, we study the term F_{22}. Using the Davydov lemma, we obtain:

$$|A_{p,k+p}| \leq 8\,[\alpha(k)]^{\frac{1}{4}}\,\|K\|_\infty^2\,\mathrm{E}Y^{-3}.$$

Therefore, under **M1** and **K1** we obtain:

$$F_{22} \leq \ \mathrm{Cst}\ 2\,\|K\|_\infty^2\,\left[nh_n^d\right]^{-2}\sum_{k=c_n+1}^{n-1}\sum_{p=1}^{n-k} k^{-\frac{\nu_1}{4}}.$$

Consequently,

$$F_{22} = O\left[c_n^{\frac{-\nu}{4}}h_n^{-2d}\right]. \tag{18}$$

On the other hand, for F_{21}, we pose:

$$g_{k,k'}^{\star} := f_{(X_k,X_{k'})|(Y_k,Y_{k'})} - f_{X_k|Y_k}\otimes f_{X_{k'}|Y_{k'}}.$$

Under **D1**, we can write:

$$A_{p,k+p} \leq \quad h_n^{2d} cte\,\|K\|_\infty^2\,\sup_{k\geq 1}\left\|g_{k+p,p}^{\star}\right\|_\infty \mathrm{E}\left|Y^{-3}\right|.$$

and

$$F_{21} \leq \quad 2\left[nh_n^d\right]^{-2} nc_n h_n^{2d}\,\|K\|_\infty^2\,\sup_{k\geq 1}\left\|g_{k+p,p}^{\star}\right\|_\infty \mathrm{E}\left|Y^{-3}\right|.$$

We have

$$F_{21} = O\left(\frac{c_n}{n}\right) = o(1). \tag{19}$$

Choosing:

$$c_n := \left(\frac{n}{h_n^{2d}}\right)^{\frac{4}{4+\nu}}$$

Under (18) and (19), we have :

$$nh_n^d F_2 \to 0, \quad \text{as } n \to \infty.$$

For F_1, we have:

$$\begin{aligned} nh_n^d F_1 = h_n^{-d}\int_{\mathbb{R}^d} K^2\left(\frac{x-u}{h_n}\right)\mathbb{E}\left[\frac{T^{-3}}{\bar{G}(T)}|X=u\right]f(u)du \\ - h_n^d\int_{\mathbb{R}^d} K\left(\frac{x-u}{h_i}\right)\mathbb{E}\left[\frac{T^{-1}}{\bar{G}(T)}|X=u\right]f(u)du \\ \times\int_{\mathbb{R}^d} K\left(\frac{x-v}{h_i}\right)\mathbb{E}\left[\frac{T^{-2}}{\bar{G}(T)}|X=v\right]f(v)dv. \end{aligned}$$

A direct application of Bochner and Toplitz lemma allows us to show the convergence of the first term of the right-hand-side of $nh_n^d F_1$ to $\mathbb{E}\left[\frac{T^{-3}}{\bar{G}(T)}|X=x\right]f(x)\int_{\mathbb{R}^d} K^2(t)dt$ and the negligibility of the second term. Finally, we have

$$Cov(\widetilde{m}_{1,n}(x),\widetilde{m}_{2,n}(x)) = \frac{1}{nh_n^d}\mathbb{E}\left[\frac{T^{-3}}{\bar{G}(T)}|X=x\right]f(x)\int_{\mathbb{R}^d} K^2(z)dz + o\left(\frac{1}{nh_n^d}\right).$$

□

Appendix

Proof of Lemma 5.1. Using the fact that,

$$1_{\{T_1 \leq C_1\}}\varphi(Y_1) = 1_{\{T_1 \leq C_1\}}\varphi(T_1). \tag{20}$$

For all measurable function φ and for all $i = 1, \dots n$, we have

$$\begin{aligned}
|\widetilde{m}_l(x) - \widetilde{m}_{l,n}(x)| &= \frac{1}{nh_n^d}\left|\sum_{i=1}^n \frac{\delta_i Y_i^{-l}}{\overline{G}_n(Y_i)}K\left(\frac{x-X_i}{h_n}\right) - \frac{\delta_i Y_i^{-l}}{\overline{G}(Y_i)}K\left(\frac{x-X_i}{h_n}\right)\right| \\
&= \frac{1}{nh_n^d}\left|\sum_{i=1}^n \frac{1_{\{T_i<C_i\}}T_i^{-l}}{\overline{G}_n(T_i)}K\left(\frac{x-X_i}{h_n}\right) - \frac{1_{\{T_i<C_i\}}T_i^{-l}}{\overline{G}(T_i)}K\left(\frac{x-X_i}{h_n}\right)\right| \\
&\leq \frac{1}{nh_n^d}\sum_{i=1}^n \left|T_i^{-l}K\left(\frac{x-X_i}{h_n}\right)\frac{\overline{G}_n(T_i)-\overline{G}(T_i)}{\overline{G}_n(T_i)\overline{G}(T_i)}\right| \\
&\leq \frac{1}{\overline{G}_n(\tau_H)\overline{G}(\tau_H)}\sup_{t\leqslant\tau_H}\left(|\overline{G}_n(t)-\overline{G}(t)|\right)\frac{1}{nh_n^d}\sum_{i=1}^n\left|T_i^{-l}K\left(\frac{x-X_i}{h_n}\right)\right|.
\end{aligned}$$

Then, by using the strong law of large numbers (SLLN) and the law of iterated logarithm (LIL) on the censoring law (see formula (4.28) in Deheuvels and Einmahl [10]), we have

$$\sup_{x\in S}|\widetilde{m}_l(x) - \widetilde{m}_{l,n}(x)| \leq \frac{1}{\overline{G}^2(\tau_H)}\mathbb{E}\left(\left|\frac{1}{h_n^d}T_i^{-l}K\left(\frac{x-X_i}{h_n}\right)\right|\right)\sqrt{\frac{\log\log n}{n}}.$$

Then, assumptions (**D3**) and (**K1-K3**) ensures that

$$\sup_{x\in S}|\widetilde{m}_l(x) - \widetilde{m}_{l,n}(x)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) = o(1).$$

□

Proof of Lemma 5.2.

$$\begin{aligned}
|\mathbb{E}(\widetilde{m}_{l,n}(x)) - m_l(x)| &= \left|\mathbb{E}\left[\frac{1}{nh_n^d}\sum_{i=1}^n\frac{\delta_i Y_i^{-l}}{\overline{G}(Y_i)}K\left(\frac{x-X_i}{h_n}\right)\right] - m_l(x)\right| \\
&= \left|\mathbb{E}\left[\frac{1}{h_n^d}\frac{\delta_1 Y_1^{-l}}{\overline{G}(Y_1)}K\left(\frac{x-X_1}{h_n}\right)\right] - m_l(x)\right|.
\end{aligned}$$

Moreover, proceeding as in (5), using (20) and a Taylor expansion, we have

$$\begin{aligned}
\sup_{x\in S}|\mathbb{E}(\widetilde{m}_{l,n}(x)) - m_l(x)| &= \sup_{x\in S}\left|\int K(t)[m_l(x-h_nt) - m_l(x)]dt\right| \\
&= \sup_{x\in S}\left|\int K(t)[-h_ntm_l'(x) + \frac{h_n^2}{2}t^2m_l''(s)]dt\right| \\
&\leq h_n\sup_{x\in S}\left|\int tK(t)m_l'(x)dt\right| + h_n^2\sup_{x\in S}\left|\int\frac{t^2}{2}K(t)m_l''(s)dt\right|,
\end{aligned}$$

where s is between $x - h_nt$ and x. Assumptions (**K1** − **K2**) conclude the proof. □

Proof of Lemma 5.4. It is clear that: $\mathbb{E}[B_n] = 0$, so $Var[B_n] = \mathbb{E}[B_n^2]$. Then

$$\begin{aligned}
\mathbb{E}[B_n^2] &= \mathbb{E}\{[\widetilde{m}_{1,n}(x) - \mathbb{E}(\widetilde{m}_{1,n}(x))]^2\} + \left\{\frac{m_1(x)}{m_2(x)}\right\}^2 - 2\frac{m_1(x)}{m_2(x)}cov\left(\widetilde{m}_{1,n}(x), \widetilde{m}_{2,n}(x)\right) \\
&= Var(\widetilde{m}_{1,n}(x)) + r^2(x)Var(\widetilde{m}_{2,n}(x)) - 2r(x)Cov(\widetilde{m}_{1,n}(x), \widetilde{m}_{2,n}(x)).
\end{aligned}$$

$$Var[B_n] = \frac{f(x)}{nh_n^d}\left\{\mathbb{E}\left[\frac{T^{-2}}{\overline{G}(T)}|X=x\right] + r^2(x)\mathbb{E}\left[\frac{T^{-4}}{\overline{G}(T)}|X=x\right]\right.$$
$$\left. -2r(x)\mathbb{E}\left[\frac{T^{-3}}{\overline{G}(T)}|X=x\right]\right\}\int_{\mathbb{R}^d} K^2(z)dz.$$

Then, by the application of Theorem 3.2 it follows that

$$\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2}(B_n - \mathbb{E}[B_n]) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

□

Proof of Lemma 5.5. It follows from Lemma 5.1 that

$$\mathbb{E}[\widetilde{m}_{2,n}(x) - m_2(x)] \to 0.$$

We have

$$Var[\widetilde{m}_{2,n}(x)] = \frac{1}{n^2h_n^{2d}}\sum_{i=1}^n Var\left[\frac{\delta_i Y_i^{-2}}{\overline{G}(Y_i)}K\left(\frac{x-X_i}{h_n}\right)\right] = O\left(n^{-1}h_n^{-d}\right).$$

Hence $\widetilde{m}_{2,n}(x) \xrightarrow{\mathbb{P}} m_2(x)$.
Next, it is clear that the second limit of Lemma 5 is a consequence of the above convergence. Then, it suffices to treat the last result. For this, we use the fact that

$$Var[(\widetilde{m}_{2,n}(x) - \mathbb{E}[\widetilde{m}_{2,n}(x)])] = Var[\widetilde{m}_{2,n}(x)] \to 0.$$

Then, by applying Lemma 5.1, we obtain $V_n = O(h_n^2)$. Then, we deduce that

$$\left(\frac{nh_n^d}{\sigma^2(x)m_2^2(x)}\right)^{1/2} V_n(\widetilde{m}_{2,n}(x) - \mathbb{E}[\widetilde{m}_{2,n}(x)]) \xrightarrow{\mathbb{P}} 0.$$

□

Acknowledgment

The authors would like to thank the Editor and two anonymous reviewers for their valuable comments which substantially improved the quality of an earlier version of this paper.

References

1. T.M. Khoshgoftaar, B.B. Bhattacharyya and G.D. Richardson, *Predicting software errors, during development, using nonlinear regression models: a comparative study*, IEEE Transactions on Reliability **41** (1992), 390–395. https://doi.org/10.1109/24.159804
2. F. Bouhadjera and E. Ould Saïd, *Nonparametric local linear estimation of the relative error regression function for censorship model*, Journal of Nonpara. Statistics (2020).
3. D. Bosq, *Nonparametric statistics for stochastic processes: estimation and prediction*, Lecture Notes in Statistics **110**, Springer New York, NY, 1998. https://doi.org/10.1007/978-1-4612-1718-3
4. R. Bradley, *Introduction to Strong Mixing Conditions*, Kendrick Press, 2007.
5. Z. Cai, *Asymptotic properties of Kaplan-Meier estimator for censored dependent data*, J. Statistics and Prob. Letters **37** (1998), 381–389. https://doi.org/10.1016/S0167-7152(97)00141-7
6. Z. Cai, *Estimating a distribution function for censored time series data*, J. Multivariate Anal **78** (2001), 299–318. https://doi.org/10.1006/jmva.2000.1953
7. A. Carbonez, L. Györfi, E.C. van der Meulen, *Partitioning estimates of a regression function under random censoring*, Statist. Decis. **13** (1995), 21–37. https://doi.org/10.1524/strm.1995.13.1.21
8. C. Chatfield, *The joys of consulting*, Significance **4** (2007), 33–36. https://doi.org/10.1111/j.1740-9713.2007.00219.x

9. K. Chen, S. Guo, Y. Lin, Z. Ying, *Least absolute relative error estimation*, J. Am. Stat. Assoc. **105** (2010), 1104–1112. https://doi.org/10.1198/jasa.2010.tm09307
10. P. Deheuvels and J.H.J. Einmahl, *Functional limit laws for the increments of Kaplan-Meier product-limit processes and applications*, Ann. Probab. **28** (2000), 1301–1335. https://doi.org/10.1214/aop/1019160336
11. J. Dedecker, P. Doukhan, G. Lang, J.R. Leon, S. Louhichi, C. Prieur, *Weak Dependence with Examples and Applications*, Lecture Notes in Statistics **190**, Springer-Verlag, 2007. https://doi.org/10.1007/978-0-387-69952-3
12. P. Doukhan, *Mixing: properties and Examples*, Lectures Notes in Statistics **85**, Springer New York, NY, 1994. https://doi.org/10.1007/978-1-4612-2642-0
13. N.R. Farnum, *Improving the relative error of estimation*, Amer. Statist. **44** (1990), 288–289. https://doi.org/10.1080/00031305.1990.10475741
14. D. Hu, *Local least product relative error estimation for varying coefficient multiplicative regression model*, Acta. Math. Appl. Sinica. **35** (2019), 274–286. https://doi.org/10.1007/s10255-018-0794-2
15. Z. Guessoum and E. Ould Saïd, *Kernel regression uniform rate estimation for censored data under α-mixing condition*, Electron. J. Stat. **4** (2010), 117–132. https://doi.org/10.1214/08-EJS195
16. K. Hirose and H. Masuda, *Robust relative error estimation*, Entropy **20** (2018), Paper No. 632, 24 pp. https://doi.org/10.3390/e20090632
17. M.C. Jones, H. Park, K.L. Shin, S.K. Vines, S.O. Jeong, *Relative error prediction via kernel regression smoothers*, J. Stat. Plan. Inference **138** (2008), 2887–2898. https://doi.org/10.1016/j.jspi.2007.11.001
18. E.L. Kaplan and P. Meier, *Nonparametric estimation from incomplete observations*, J. Amer. Statist. Assoc. **53** (1958), 457–481. https://doi.org/10.1080/01621459.1958.10501452
19. M. Kohler, K. Máthé, M. Pinter, *Prediction from randomly Right Censored Data*, J. Multivariate Anal **80** (2002), 73–100. https://doi.org/10.1006/jmva.2000.1973
20. S. Khardani, *Relative error prediction for twice censored data*, Math. Methods Statist. **28** (2019), 291–306. https://doi.org/10.3103/S1066530719040045
21. S. Khardani and Y. Slaoui, *Nonparametric relative regression under random censorship model*, Statistics and Probability Letters **151** (2019), 116–122. https://doi.org/10.1016/j.spl.2019.03.019
22. E. Masry, *Recursive probability density estimation for weakly dependent stationary process*, IEEE Trans. Inform. Theory **32** (1986), 254–267. https://doi.org/10.1109/TIT.1986.1057163
23. J.S. Marron and M.P. Wand, *Exact mean integrated squared error*, The Annals of Statistics **20** (1992), 712–736. https://doi.org/10.1214/aos/1176348653
24. W. Mechab and A. Lacksaci, *Nonparametric relative regression for associated random variables*, Metron **74** (2016), 75–97. https://doi.org/10.1007/s40300-016-0084-9
25. A. Mokkadem and M. Pelletier, *The law of the iterated logarithm for the multivariate kernel mode estimator*, ESAIM: Probability and Statistics **7** 2003, 1–21. https://doi.org/10.1051/ps:2003004
26. E.A. Nadaraya, *On non-parametric estimates of density functions and regression curves*, Theory Probab. Appl. **10** (1965), 186–190. https://doi.org/10.1137/1110024
27. S.C. Narula and J.F. Wellington, *Prediction, linear regression and the minimum sum of relative errors*, Technometrics **19** (1977), 185–190. https://doi.org/10.1080/00401706.1977.10489526
28. H. Park and L.A. Stefanski, *Relative-error prediction*, Statist. Probab. Lett. **40** (1998), 227–236. https://doi.org/10.1016/S0167-7152(98)00088-1
29. E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants*, ESAIM, Springer **31**, 2000.
30. B. Silverman, *Weak and strong uniform consistency of the kernel estimate of a density and its derivatives*, Ann. Statist. **6** (1978), 177–184. https://doi.org/10.1214/aos/1176344076

Université El-Manar, Laboratoire de Modélisation mathématique, Statistique Fiabilité et Analyse stochastique, Tunisia
E-mail address: khardani_salah@yahoo.fr

Université El-Manar, Laboratoire de Modélisation mathématique, Statistique Fiabilité et Analyse stochastique, Tunisia.
E-mail address: nefziwiem24@gmail.com

Université El-Manar, Laboratoire de Modélisation mathématique, Statistique Fiabilité et Analyse stochastique, Tunisia
E-mail address: salah.khardani@fst.utm.tn