



**HAL**  
open science

## Nonparametric relative error estimation of the regression function for censored data

Feriel Bouhadjera, Elias Ould Saïd, Mohamed Riad Remita

► **To cite this version:**

Feriel Bouhadjera, Elias Ould Saïd, Mohamed Riad Remita. Nonparametric relative error estimation of the regression function for censored data. 2019. hal-01994512

**HAL Id: hal-01994512**

**<https://hal.science/hal-01994512>**

Preprint submitted on 25 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nonparametric relative error estimation of the regression function for censored data

**BOUHADJERA Ferial.** <sup>1,2</sup>

<sup>1</sup> *Université Badji-Mokhtar, Lab. de Probabilités et Statistique. BP 12, 23000 Annaba, Algérie.*

**OULD SAÏD Elias.** Corresponding author.

<sup>2</sup> *Université du Littoral Côte d'Opale. Lab. de Math. Pures et Appliquées. IUT de Calais. 19, rue Louis David. Calais, 62228, France.*

**REMITA Mohamed Riad.**

*Université Badji-Mokhtar, Lab. de Probabilités et Statistique. BP 12, 23000 Annaba, Algérie.*

**ABSTRACT.** Let  $(T_i)_i$  be a sequence of independent identically distributed (i.i.d.) random variables (r.v.) of interest distributed as  $T$  and  $(X_i)_i$  be a corresponding vector of covariates taking values on  $\mathbb{R}^d$ . In censorship models the r.v.  $T$  is subject to random censoring by another r.v.  $C$ . In this paper we built a new kernel estimator based on the so-called synthetic data of the mean squared relative error for the regression function. We establish the uniform almost sure convergence with rate over a compact set and its asymptotic normality. The asymptotic variance is explicitly given and as product we give a confidence bands. A simulation study has been conducted to comfort our theoretical results.

*Keywords:* Asymptotic normality. Censored data. Kernel estimate. Relative regression error. Uniform almost sure convergence. V-C classes.

# 1 Introduction

Let  $(X_i, T_i)_i$  be a  $\mathbb{R}^d \times \mathbb{R}_+^*$ , ( $d \geq 1$ ) valued sequence of random vectors that we assume drawn from the pair  $(X, T)$  which is defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . The purpose of this work is to study the effect of a random covariable  $X$  on a r.v.  $T$  which is subject to right censoring by another r.v.  $C$ . This relation of regression is modeled by:

$$T = r(X) + \epsilon, \tag{1.1}$$

where  $r(\cdot)$  is the regression function and  $\epsilon$  a sequence of error independent to  $X$ . Usually,  $r(\cdot) = \mathbb{E}[T|X]$  is estimated by minimizing the mean squared loss function  $\mathbb{E}[(T - r(X))^2|X]$ . However, this loss function is based on some restrictive conditions that is the variance of the residual is the same for all the observations, which is inadequate when the data contains some outliers. Therefore, in order to overcome this drawback we consider an alternative approach allow to construct an efficient predictor even if the data is affected by the presence of outliers. So, in this paper the limitations of the classical regression are counteracted by estimating the regression function with respect to the minimization of the following mean squared relative error, for  $T > 0$ ,

$$\mathbb{E} \left[ \left( \frac{T - r(X)}{T} \right)^2 | X \right]. \tag{1.2}$$

The latter is a more meaningful measure of performance of a predictor than the usual error in the presence of outliers. It is easy to see that the solution of the minimization problem of (1.2) is given by

$$r(X) = \frac{\mathbb{E}[T^{-1}|X]}{\mathbb{E}[T^{-2}|X]}. \tag{1.3}$$

Park and Stefanski (1998) have shown that the solution given by (1.3) satisfies

$$\frac{\mathbb{E}[T^{-1}|X]}{\mathbb{E}[T^{-2}|X]} \leq \mathbb{E}[T|X] \quad \text{a.s.} \tag{1.4}$$

provided that the first two conditional inverse moments are finite. The authors consider parametric approaches to estimate the regression function  $r(\cdot)$  which focused on estimating the mean and

variance functions modeling methods (Carroll and Ruppert, 1988) of the inverse response  $T^{-1}$  as function of  $X$ . Without claiming to be exhaustive, we can quote Narula and Wellington (1977) who studied an estimation method for minimizing the sum of absolute relative residuals. Farum (1990) developed an estimation method designed to reduce absolute relative error. Khoshgoftaar *et al.* (1992) studied the asymptotic properties of the estimators minimizing the sum of squared relative errors. In this contribution, we focus on nonparametric approach. To the best of our knowledge, only the paper of Park *et al.* (2008) study the nonparametric regression using the relative error as loss function. They studied the asymptotic properties of an estimator minimizing the sum of squared relative errors by applying local linear approach.

In many estimation problems, it is not always possible, to make complete measurements when the available sample data is incomplete in the sense that measures are not available for all members of a random sample. For example, in medical follow-up studies, it often happens for various reasons, that the duration of interest can not be observed. This may be due to the loss of view of the patient at the beginning or end of the study period. These values are censored. The censored values, although unknown, must be taken into account to obtain a correct estimate and precise conclusions. For such practical observations, conventional statistical procedures are no longer valid and more elaborate techniques are used to model such observations.

One of the classical cases for incomplete data is the right-censored data. In this case, we observe another r.v.  $C$  with continuous distribution function (d.f.)  $G$ , we can only observe a sample  $(X_i, Y_i, \delta_i)$  where  $Y_i = T_i \wedge C_i$  and  $\delta_i = \mathbb{1}_{\{T_i \leq C_i\}}$ , for  $i = 1, \dots, n$ , with  $\wedge$  denotes the minimum and  $\mathbb{1}_A$  is the indicator function of the event  $A$ .

When we talk about censored data, several authors like Carbonez *et al.* (1995), Kohler *et al.* (2002), Delecroix *et al.* (2008) and Guessoum and Ould Said (2008) uses the synthetic data that take into account the effect of censorship on the distribution. For that we consider the sample  $(Y_i, \delta_i)_i$ , for

$1 \leq i \leq n$  and we put :

$$T_i^{*, -\ell} = \frac{\delta_i Y_i^{-\ell}}{\overline{G}(Y_i)}, \quad \ell = 1, 2 \quad (1.5)$$

where  $\overline{G}$  is the survival function of the censoring rv  $C$ .

All along this paper, we suppose that:

$$(T_i, X_i)_i \text{ and } (C_i)_i \text{ are independent for } 1 \leq i \leq n. \quad (1.6)$$

Then from the equation (1.5) and the hypothesis (1.6), we get,

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

This paper offers then an alternative approach to traditional estimation models by considering the minimization of the least relative error for regressions models when the data are randomly right censored. We establish the strong and uniform consistencie (with rate) of the constructed estimator and then the asymptotic normality has been shown. At the best of our acknowledge there is no result concerning the nonparametric regression function for censoring data using the relative error. The rest of the paper is organized as follows: Section 2 is devoted to the presentation of the new estimator of the mean squared relative error of the regression function. The assumptions and main results are given in Section 3. Simulations are drawn in Section 4. Finally, the proofs are relegated to Section 5 with some auxiliary results.

## 2 Definition of the new estimator

Let  $(T_i)_{1 \leq i \leq n}$  be an i.i.d.  $n$ -sample of r.v. of interest with commun unknown continuous d.f.  $F$  and let  $(X_i)_{1 \leq i \leq n}$  be a corresponding vector of covariates with joint density function  $f(\cdot)$ . As mentioned

before the solution of (1.2) is given by

$$r(x) = \frac{\mathbb{E}[T^{-1}|X = x]}{\mathbb{E}[T^{-2}|X = x]} := \frac{\bar{r}_1(x)}{\bar{r}_2(x)}, \quad (2.1)$$

with  $r_\ell(\cdot) = \frac{\bar{r}_\ell(\cdot)}{f(\cdot)}$  where  $\bar{r}_\ell(x) = \int_{\mathbb{R}_+^*} t^{-\ell} f_{T,X}(t, x) dt$  for  $\ell = 1, 2$ .

Recall that, in the case of complete data, a well-known Nadaraya Watson (N-W) estimator of  $r(\cdot)$  is given by

$$r^*(x) = \sum_{i=1}^n T_i W_{i,n}(x)$$

with

$$W_{i,n}(x) = \begin{cases} \frac{K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}, & \text{if } \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right) \neq 0; \\ 1/n, & \text{otherwise,} \end{cases}$$

where  $h_n$  is a sequence of positive real numbers (bandwidth) that decreases to zero when  $n$  goes to infinity and  $K$  is a kernel function defined in  $\mathbb{R}^d$ . Thus, a natural estimator of (2.1) is given by

$$r_n(x) = \frac{\sum_{i=1}^n T_i^{-1} K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^n T_i^{-2} K\left(\frac{x-X_i}{h_n}\right)}, \quad (2.2)$$

this is the analogous N-W estimator which is nothing other than a special case of the censored case.

As mentioned before, when the r.v.  $T$  is subject to right censoring by another r.v.  $C$ , we define

$\tilde{r}_n(x)$  as a "pseudo-estimator" of  $r(x)$  that is, for any  $x \in \mathbb{R}$ , we have,

$$\tilde{r}_n(x) = \frac{\sum_{i=1}^n \frac{\delta_i Y_i^{-1}}{\bar{G}(Y_i)} K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^n \frac{\delta_i Y_i^{-2}}{\bar{G}(Y_i)} K\left(\frac{x-X_i}{h_n}\right)} =: \frac{\bar{\tilde{r}}_{1,n}(x)}{\bar{\tilde{r}}_{2,n}(x)}. \quad (2.3)$$

The latter can not be calculated as  $\bar{G}$  is unknown. Then to define a genuine estimator of  $r(\cdot)$ , we replace  $\bar{G}$  by its Kaplan-Meier (1958) estimator which is defined by

$$\bar{G}_n(t) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1 - \delta_i}{n - i + 1}\right)^{\mathbb{1}_{\{Y_i \leq t\}}} & \text{if } t < Y_{(n)}, \\ 0 & \text{otherwise,} \end{cases} \quad (2.4)$$

where  $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)}$  are the order statistics of the  $Y_i$  and  $\delta_i$  is the indicator of non-censoring. The properties of  $\bar{G}_n(t)$  have been studied by many authors. So a calculable estimator of  $r(\cdot)$  is given by

$$r_n(x) = \frac{\sum_{i=1}^n \frac{\delta_i Y_i^{-1}}{\bar{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n \frac{\delta_i Y_i^{-2}}{\bar{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right)} =: \frac{r_{1,n}(x)}{r_{2,n}(x)} \quad (2.5)$$

where

$$r_{\ell,n}(x) = \frac{\bar{r}_{\ell,n}(x)}{f_n(x)} = \frac{\sum_{i=1}^n \frac{\delta_i Y_i^{-\ell}}{\bar{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)}, \quad (2.6)$$

for  $\ell = 1, 2$  and  $f_n(\cdot)$  is the well-known kernel estimator of the joint density function  $f(\cdot)$ .

### 3 Hypotheses and main results

In order to state our results, we introduce some notations. For any d.f.  $L$ , let  $\tau_L = \sup\{y, L(y) < 1\}$  be a upper endpoint of  $L$ . Assume that  $\tau_F < \infty$ ,  $\bar{G}(\tau_F) > 0$ . All along the paper, when no confusion is possible, we denote by  $M$  any generic strictly positive constant such that  $M \geq T^{-\ell}$  and by  $r_\ell(\cdot) = \mathbb{E}[T^{-\ell}|X = \cdot]$  the conditional  $\ell$ -inverse moments of  $T$  given  $X$  and  $\ell = 1, 2$ . Furthermore  $\bar{r}_\ell(\cdot) = r_\ell(\cdot)f(\cdot)$ , with  $f$  is the density of  $X$ . On the other hand,  $\log_2(\cdot) = \log \log(\cdot)$  denotes the iterated logarithm function. Finally denote  $\mathcal{C}_0 = \{x \in \mathbb{R}/f(x) > 0\}$  the open set and  $\mathcal{C}$  be a compact subset of  $\mathcal{C}_0$ .

We will make use of the following hypotheses.

**H.** The bandwidth  $h_n$  satisfies:

$$i) \lim_{n \rightarrow \infty} h_n = 0, \quad \lim_{n \rightarrow \infty} nh_n = +\infty, \quad \lim_{n \rightarrow \infty} \frac{\log n}{nh_n} = 0$$

$$ii) h_n \log_2 n = o(1).$$

iii)  $\lim_{n \rightarrow \infty} nh_n^5 = 0$ .

**K.** The kernel  $K(\cdot)$  is:

i) Continuously differentiable compactly supported density function.

ii)  $\int_{\mathbb{R}} |t|K(t)dt = 0$  and  $\int_{\mathbb{R}} t^2K(t)dt < \infty$  holds.

iii)  $\int_{\mathbb{R}} tK^2(t)dt < \infty$

**D.** i) The function  $\bar{r}_\ell(\cdot)$ , for  $\ell = 1, 2$ , is twice continuously differentiable and

$$\sup_{x \in \mathcal{C}} |\bar{r}_\ell''(x)| < +\infty.$$

ii) The function  $\Upsilon_\ell(\cdot)$ , for  $\ell = 2, 3, 4$ , is continuously differentiable and

$$\sup_{x \in \mathcal{C}} \Upsilon_\ell'(x) < +\infty.$$

iii) There exists  $\Gamma > 0$  such that  $\bar{r}_2(x) > \Gamma$  for all  $x \in \mathcal{C}$ .

### 3.1 Discussions on the hypotheses

1. *The independence assumption between  $(C_i)_i$  and  $(T_i, X_i)_i$  may seem to be strong and one can think of replacing it by a classical conditional independence assumption between  $(C_i)_i$  and  $(T_i)_i$  given  $(X_i)_i$ . However in the conditionally hypothesis we propose the following estimator for the regression function  $r(x)$  where  $\bar{r}_{\ell,n}$  for  $\ell = 1, 2$  are given by*

$$\bar{r}_{\ell,n}(x) = \sum_{i=1}^n \frac{\delta_i Y_i^{-\ell}}{\bar{G}_n(Y_i|X_i)} K\left(\frac{x - X_i}{h_n}\right) \quad (3.1)$$

*where  $\bar{G}_n(Y_i|X_i)$  is Beran's estimator of the survival conditional distribution of the censored r.v.  $C$  given  $X$ . Then we get an analogous estimator as in (2.5) using (3.1). As mentioned before and as far as we know there is no rate of convergence for this estimate as in the unconditional case (see Deheuvels and Einmahl, 2000). We think that this issue has to be addressed if we aim to get rates of convergence. Moreover our framework is classical and*

was considered by Carbonez et al. (1995) and Kohler et al. (2002) among others. Note finally that this assumption implies the independence between  $(C_i)_i$  and  $(T_i)_i$  which ensures the identifiability of the model.

2. The hypothesis  $\tau_F < \tau_G$  is classical for asymptotic normality results in the censorship framework. It implies that  $\overline{G}(T) \geq \overline{G}(\tau_F) > 0$  a.s.
3. The Hypotheses **H** i) and **K** concern the smoothing parameter  $h_n$  and the kernel  $K(\cdot)$  and are standard in nonparametric regression estimation for complete or incomplete data. Moreover, **D** i) is needed to study the bias term. On another side, hypothesis **D** iii) is used to state the uniform consistency of the constructed estimator. Finally, hypotheses **H** ii), iii) and **D** ii) are needed for get asymptotic normality.

## 3.2 Results

We can now present our results. The proofs of these are established in Section 5. We first state a uniform consistency result with rate for  $r_n(\cdot)$ .

**Theorem 1.** *Under hypotheses **H** i), **K** and **D** i), iii), we have:*

$$\sup_{x \in \mathcal{C}} |r_n(x) - r(x)| = O_{a.s.} \left\{ \max \left( \left( \frac{\log n}{nh_n^2} \right)^{1/2}, h_n^2 \right) \right\} \quad \text{as } n \rightarrow \infty.$$

**Remark 1.** *It is clear that we can give the same result in  $\mathbb{R}^d$ ,  $d > 1$  without difficulties. The proofs are analogous. Therefore, the Theorem 1 becomes:*

$$\sup_{x \in \mathcal{C}} |r_n(x) - r(x)| = O_{a.s.} \left\{ \max \left( \left( \frac{\log n}{nh_n^{d+1}} \right)^{1/2}, h_n^2 \right) \right\} \quad \text{as } n \rightarrow \infty.$$

In what follows we will state the asymptotic normality result. For this, let

$$\Sigma(x) = \begin{pmatrix} \Upsilon_2(x) & \Upsilon_3(x) \\ \Upsilon_3(x) & \Upsilon_4(x) \end{pmatrix}$$

be the covariance matrix, with

$$\Upsilon_{2\ell}(x) = \int \frac{t^{-2\ell}}{\overline{G}(t)} f_{T,X}(t, x) dt \quad \text{and} \quad \Upsilon_3(x) = \int \frac{t^{-3}}{\overline{G}(t)} f_{T,X}(t, x) dt, \quad \text{for } \ell = 1, 2.$$

Now we are in position to give our asymptotic normality result.

**Theorem 2.** *Suppose that hypotheses **H**, **K** and **D**i), ii) holds. Let  $\mathcal{A} = \{x \in \mathcal{C} \text{ and } r_\ell(x) \neq 0, \ell = 1, 2 \text{ and } \Upsilon_j(x) \neq 0, j = 2, 3, 4\}$ , we have*

$$\sqrt{nh_n}(r_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)) \quad \text{as } n \rightarrow \infty$$

where

$$\sigma^2(x) = \kappa \frac{\Upsilon_2(x)\bar{r}_2^2(x) - 2\Upsilon_3(x)\bar{r}_1(x)\bar{r}_2(x) + \Upsilon_4(x)\bar{r}_1^2(x)}{\bar{r}_2^4(x)}$$

for  $\kappa = \int K^2(t)dt$  and  $\xrightarrow{\mathcal{D}}$  denotes the convergence in distribution.

### 3.3 Confidence interval

The determination of confidence interval requires the estimation of the unknown quantity  $\sigma_n(x)$ . A plug-in estimate and using the following estimate of  $\Upsilon_{2\ell}(x)$ , for  $\ell = 1, 2$ , and  $\Upsilon_3(x)$  given by

$$\widehat{\Upsilon}_{2\ell}(x) = \frac{1}{nh_n} \sum_{i=1}^n \frac{Y_i^{-2\ell}}{\overline{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right) \quad \text{and} \quad \widehat{\Upsilon}_3(x) = \frac{1}{nh_n} \sum_{i=1}^n \frac{Y_i^{-3}}{\overline{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right)$$

respectively and (2.5) we get a consistent estimate of  $\sigma^2(x)$ . This yields a confidence interval of asymptotic level  $1 - \zeta$  for  $r(x)$  given by

$$\left[ r_n(x) - t_{1-\zeta/2} \frac{\sigma_n(x)}{\sqrt{nh_n}} ; r_n(x) + t_{1-\zeta/2} \frac{\sigma_n(x)}{\sqrt{nh_n}} \right]$$

where  $t_{1-\zeta/2}$  denotes the  $1 - \zeta/2$  quantile of the standard normal distribution.

### 3.4 Comeback to complete data

At the best of our knowledge there are no analogous results for the complete. The analogous results can be state by putting  $C = +\infty$  and therefore  $\overline{G}(\cdot) = 1$ .

To give an overview of the performance of our estimator, we graph it in the next section.

## 4 Simulation Study

The main objective of this part is to evaluate the good behavior of our estimator for different censoring rates and sample sizes and to show the efficiency of this approach compared to the classical one.

### 4.1 Consistency

#### 4.1.1 Simulations settings

For this purpose, simulation data are generated from model (1.1) where covariates  $X$  have normal distribution on  $\mathcal{N}(5, 2)$  and random effect  $\epsilon$  have standard normal distribution. For the rest, we proceed in the following way:

- Generate the censoring variable  $C$  according to the normal law with  $(\mu = 11, \sigma = 1)$ .
- Calculate the response variable  $T = \alpha X + \beta + c\epsilon$  with  $(\alpha = 2, \beta = 1$  and  $c = 0.2)$ .
- The censored data are calculated as  $Y = T \wedge C$  and  $\delta = \mathbb{1}_{\{T \leq C\}}$ . The observed data therefore becomes  $(X, Y, \delta)$ .
- The Kaplan-Meier estimator is calculated for the distribution function of censorship variable  $C$  in (2.4).
- The choice of  $K$  is not decisive, we choose then the standard Gaussian kernel ( $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2})$ ). In contrast, the choice of bandwidth is crucial that's why we take the optimal one  $h_n = 0.55 \left(\frac{\log(n)}{n}\right)^{0.33}$ .
- Finally, we calculate the expression of our estimator obtained from (2.5) for a compact set  $\mathcal{C} = [1, 4]$ .

Under each simulation setting, 100,300 and 500 replications are conducted.

### 4.1.2 Simulation results

**Effect of sample size with fixed censorship rate.** From Figure 1, we can see that the quality of fit increases with  $n$  when censoring rate (CR) and bandwidth kept unchanged.

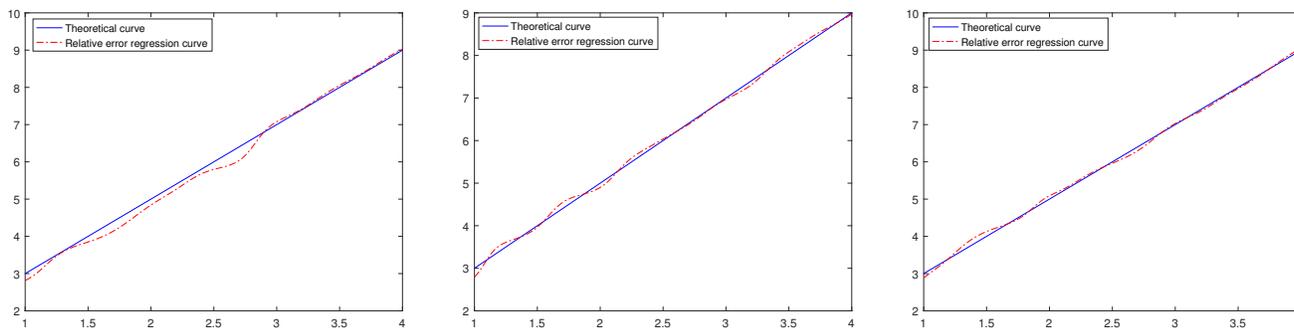


Figure 1:  $r(x) = 2x + 1$ ,  $CR \approx 50\%$  and  $n = 100, 300$  and  $500$ , respectively.

**Effect of censoring rate (CR) with a fixed sample size.** Figure 2 is obtained by varying the censoring rate for a fixed sample size ( $n = 300$ ) and for that we push the variable of interest on the right by increasing the average of the normal distribution to observe more censorship variable (the number of complete observation decreases). It can be seen that the forecasting quality decreases when the CR rises in particular in the border.

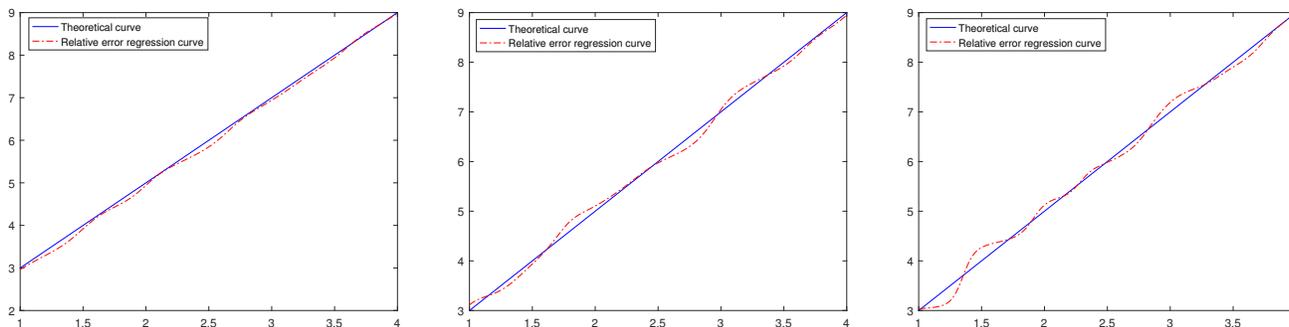


Figure 2:  $r(x) = 2x + 1$ ,  $n = 300$  and  $CR \approx 15\%, 50\%$  and  $80\%$ , respectively.

**Nonlinear functions** We consider the case of nonlinear regression by choosing this three kinds of model:

- (1) Parabolic  $T = x^2 + 1 + \varepsilon$ ,
- (2) Sinusoidal  $T = \sin\left(\frac{1}{2}x\right)^2 + 1 + \varepsilon$ ,
- (3) Exponential  $T = \exp\left(\frac{1}{2}x\right) + \varepsilon$ .

The curves are shown in Figure 3. Note that the quality of fit deteriorates when the period is very small.

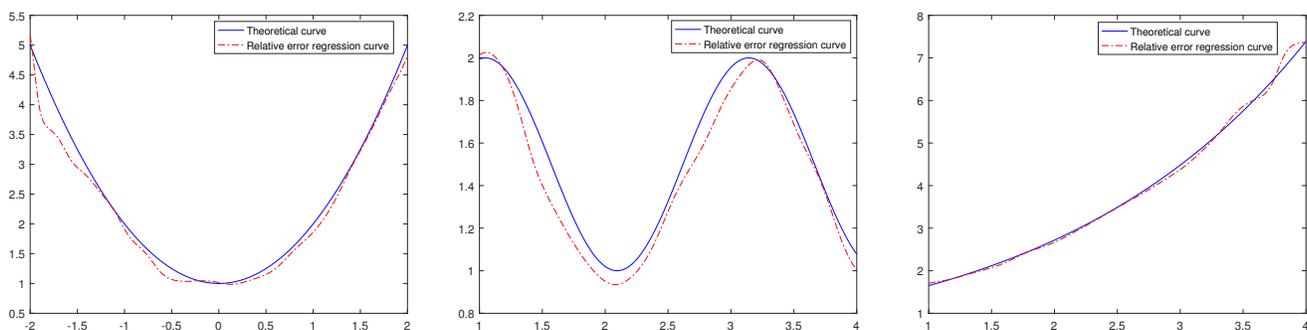


Figure 3:  $CR \approx 50\%$ ,  $n = 500$  for parabolic, sinus and exponential, respectively.

### Classical regression versus Relative error regression with respect to the censorship rate

In order to highlight the efficiency of relative error estimation, we draw up a comparative study. For that, we simulated the classical regression estimator for randomly right censored data defined in Guessoum and Ould Saïd (2008) by

$$\hat{r}(x) = \frac{\sum_{i=1}^n \frac{\delta_i Y_i}{\bar{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)},$$

for the same parameters listed below. From Figure 4 below, it is clear that the classical regression estimator deteriorates when the censorship rate increases considerably.

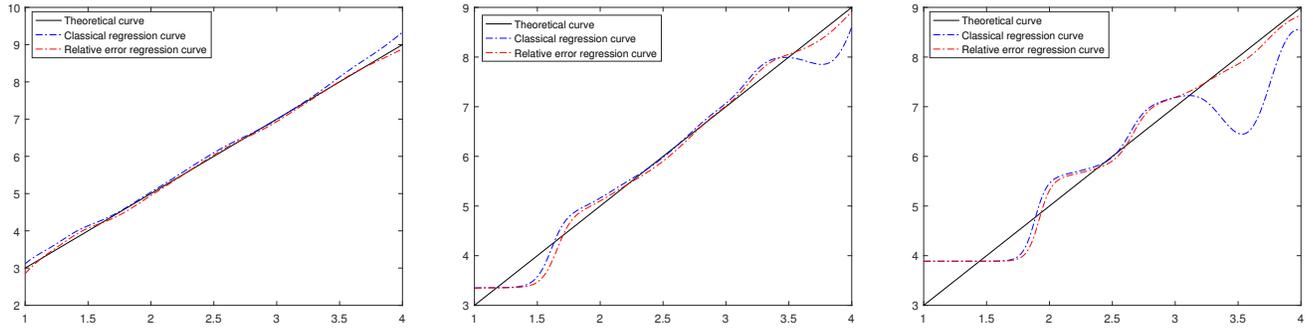


Figure 4:  $r(x) = 2x + 1$ ,  $n = 100$  and  $CR \approx 15\%$ ,  $50\%$  and  $75\%$ , respectively.

**Effect of outliers for the two methods with a fixed sample size and censorship rate.**

To show the robustness of our approach, we generate the case where the data contains outliers. For that we set both that sample size and censorship rate ( $n = 500$  and  $TC \approx 50$ ). To create this outlier effect, 20 values of this sample are multiplied by a factor called  $MF$ . From Figure 5, we can see that our estimator is very close to the theoretical curve with respect to the classical one. Then, it is very clear that our approach is widely better than the classical one in the presence of outliers.

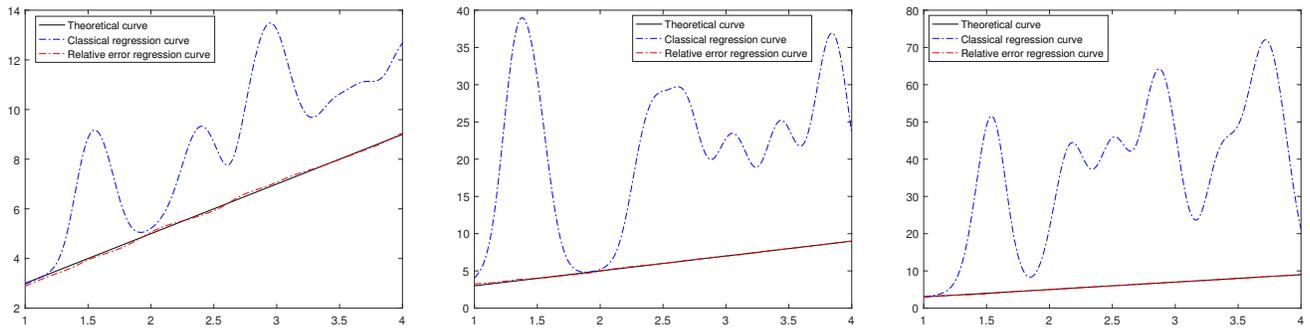


Figure 5:  $r(x) = 2x + 1$ ,  $n = 500$ ,  $CR \approx 50\%$  and  $MF = 10, 50$  and  $100$  respectively.

**Table of mean squared error.** To close this part, we compare the values of the mean squared error ( $MSE$ ) for the classical regression ( $MSE^1$ ) and the relative error regression ( $MSE^2$ ) methods represented by

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

for three sample sizes and censoring level.

It can be seen from Table 1 that the variability of the mean squared error ( $MSE$ ) of the two methods for a low censoring rate is not significantly considerable, i.e. the performance is the same for both methods. However, when the data is affected by the presence of censoring the  $MSE$  of relative error regression becomes smaller than the classical regression. It means that the relative error regression model is more stable than the classical regression in the presence of censorship.

Sample size	CR ( $\approx$ %)	$MSE^1$ (Classical Regression)	$MSE^2$ (Relative Error Regression)
n=100	20	0.0150	0.0027
	50	0.0802	0.0339
	80	0.4009	0.0885
n=300	20	0.0103	0.0057
	50	0.0275	0.0138
	80	0.1284	0.0425
n=500	20	0.0078	0.0008
	50	0.0138	0.0032
	80	0.1359	0.0136

Table 1: The  $MSE$  errors according to the censoring rates.

## 4.2 Asymptotic normality

The purpose of this part is to highlight the theoretical results obtained in Theorem 2, by studying by simulation the asymptotic normality. To do this, we compare the shape of the estimated density to that of the standard normal density in the case of a linear regression model:

$$T = 2X + 1 + \epsilon,$$

we reproduce the same steps as in the previous subsection for  $X \rightsquigarrow \exp(1.5)$  and  $C \rightsquigarrow \exp(3)$ . Throughout this subsection, we fix  $x = 0$  and replicate  $m$  independent  $n$ -sample size. Then, we calculate the asymptotic variance. For that we replace the  $\Upsilon_k(\cdot)$  for  $k = 2, 3, 4$  by their estimators in (3.3) and  $\bar{r}_{\ell,n}(\cdot)$  for  $\ell = 1, 2$  by their estimators in (2.6). A calculable estimator of the normalized deviation is given by:

$$\sigma_n^2(0) = \kappa \frac{\hat{\Upsilon}_2(0)\bar{r}_{2,n}^2(0) - 2\hat{\Upsilon}_3(0)\bar{r}_{1,n}(0)\bar{r}_{2,n}(0) + \hat{\Upsilon}_4(0)\bar{r}_{1,n}^2(0)}{\bar{r}_{2,n}^4(0)},$$

we consider now the sequence:

$$A_j = \left( \frac{nh_n}{\sigma_{n,j}^2(0)} \right)^{1/2} (r_{n,j}(0) - 1),$$

which under Theorem 2,  $A_j$  follows asymptotically to  $\mathcal{N}(0, 1)$ . Then, we build a kernel density estimator for the  $A_k$  that we compare with the standard normal law for different values of  $n$  and  $h^* = c \left( \frac{\log(m)}{m} \right)^{0.2}$  where the constant  $c$  is chosen appropriately. Finally, for a sample size  $m = 200$  and a censorship rate ( $CR \approx 66\%$ ), we conduct  $n = 100, 300$  and  $500$  replications. The figure 6 show the quality of goodness of fit.

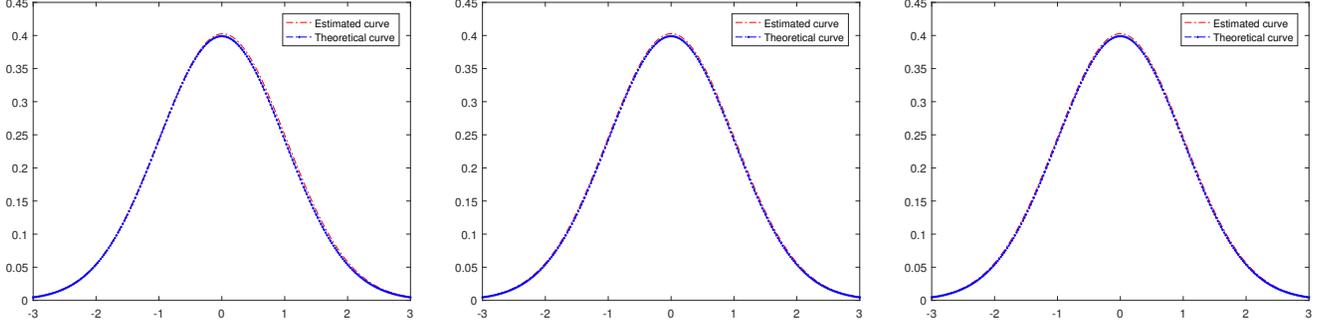


Figure 6:  $CR \approx 66\%$  and  $n = 100, 300$  and  $500$ , respectively.

## 5 Auxiliary results and proofs

*Proof.* Using (2.1), (2.3), (2.5), (2.6) and for  $x \in \mathbb{R}$ , we consider the following decomposition :

$$r_n(x) - r(x) = \frac{1}{\bar{r}_{2,n}(x)} \left\{ \left[ (\bar{r}_{1,n}(x) - \tilde{r}_{1,n}(x)) + (\tilde{r}_{1,n}(x) - \mathbb{E}[\tilde{r}_{1,n}(x)]) + (\mathbb{E}[\tilde{r}_{1,n}(x)] - \bar{r}_1(x)) \right] \right. \\ \left. + r(x) \left[ (\tilde{r}_{2,n}(x) - \bar{r}_{2,n}(x)) + (\mathbb{E}[\tilde{r}_{2,n}(x)] - \tilde{r}_{2,n}(x)) + (\bar{r}_2(x) - \mathbb{E}[\tilde{r}_{2,n}(x)]) \right] \right\}$$

which by triangle inequality, we have

$$\sup_{x \in \mathcal{C}} |r_n(x) - r(x)| \\ \leq \frac{1}{\inf_{x \in \mathcal{C}} |\bar{r}_{2,n}(x)|} \left\{ \sup_{x \in \mathcal{C}} \{ |\bar{r}_{1,n}(x) - \tilde{r}_{1,n}(x)| + |\tilde{r}_{1,n}(x) - \mathbb{E}[\tilde{r}_{1,n}(x)]| + |\mathbb{E}[\tilde{r}_{1,n}(x)] - \bar{r}_1(x)| \} \right. \\ \left. + \sup_{x \in \mathcal{C}} |r(x)| \{ |\tilde{r}_{2,n}(x) - \bar{r}_{2,n}(x)| + |\mathbb{E}[\tilde{r}_{2,n}(x)] - \tilde{r}_{2,n}(x)| + |\bar{r}_2(x) - \mathbb{E}[\tilde{r}_{2,n}(x)]| \} \right\}.$$

In the sequel, we give a sequence of lemmas that are helpful in proving our results.

**Lemma 1.** *Under hypotheses  $\mathbf{H}i)$  and  $\mathbf{K}$ , we have, for  $\ell = 1, 2$ , that*

$$\sup_{x \in \mathcal{C}} |\bar{r}_{\ell,n}(x) - \tilde{r}_{\ell,n}(x)| = O_{a.s.} \left\{ \left( \frac{\log_2 n}{n} \right)^{1/2} \right\} \quad \text{as } n \rightarrow \infty. \quad (5.1)$$

*Proof.* For  $\ell = 1, 2$ , we have

$$\begin{aligned}
|\bar{r}_{\ell,n}(x) - \tilde{r}_{\ell,n}(x)| &= \frac{1}{nh_n} \left| \sum_{i=1}^n \frac{\delta_i Y_i^{-\ell}}{\bar{G}_n(Y_i)} K\left(\frac{x - X_i}{h_n}\right) - \frac{\delta_i Y_i^{-\ell}}{\bar{G}(Y_i)} K\left(\frac{x - X_i}{h_n}\right) \right| \\
&= \frac{1}{nh_n} \left| \sum_{i=1}^n \frac{T_i^{-\ell}}{\bar{G}_n(T_i)} \mathbb{1}_{\{T_i \leq C_i\}} K\left(\frac{x - X_i}{h_n}\right) - \frac{T_i^{-\ell}}{\bar{G}_n(T_i)} \mathbb{1}_{\{T_i \leq C_i\}} K\left(\frac{x - X_i}{h_n}\right) \right| \\
&\leq \frac{1}{nh_n} \sum_{i=1}^n \left| T_i^{-\ell} K\left(\frac{x - X_i}{h_n}\right) \right| \left| \frac{1}{\bar{G}_n(T_i)} - \frac{1}{\bar{G}(T_i)} \right| \\
&\leq \frac{1}{nh_n} \sum_{i=1}^n \left| T_i^{-\ell} K\left(\frac{x - X_i}{h_n}\right) \right| \left| \frac{\bar{G}_n(T_i) - \bar{G}(T_i)}{\bar{G}_n(T_i)\bar{G}(T_i)} \right| \\
&\leq \frac{\sup_{t \leq \tau_F} |\bar{G}_n(t) - \bar{G}(t)|}{\bar{G}_n(\tau_F)\bar{G}(\tau_F)} \frac{M}{nh_n} \sum_{i=1}^n \left| K\left(\frac{x - X_i}{h_n}\right) \right|,
\end{aligned}$$

then by using the strong law of large numbers (SLLN) and law of iterated logarithm (LIL) on the censoring law (see formula (4.28) in Deheuvels and Einmahl, 2000), we get,

$$\sup_{x \in \mathcal{C}} |\bar{r}_{\ell,n}(x) - \tilde{r}_{\ell,n}(x)| \leq \frac{M}{\bar{G}^2(\tau_F)} \mathbb{E} \left( \left| \frac{1}{h_n} K\left(\frac{x - X_1}{h_n}\right) \right| \right) \sqrt{\frac{\log_2 n}{n}}.$$

Then hypotheses **H***i*) and **K** complete the proof of the lemma.  $\square$

**Lemma 2.** *Under hypotheses **H***i*), **K** and **D***i*), we have, for  $\ell = 1, 2$ , that:*

$$\sup_{x \in \mathcal{C}} |\mathbb{E}[\tilde{r}_{\ell,n}(x)] - \bar{r}_{\ell}(x)| = O(h_n^2) \quad \text{as} \quad n \rightarrow \infty. \quad (5.2)$$

*Proof.* Using the conditional expectation properties, we get,

$$\begin{aligned}
\mathbb{E}[\tilde{r}_{\ell,n}(x)] &= \frac{1}{h_n} \mathbb{E} \left[ \frac{\delta_1 Y_1^{-\ell}}{\bar{G}(Y_1)} K\left(\frac{x - X_1}{h_n}\right) \right] \\
&= \frac{1}{h_n} \mathbb{E} \left[ K\left(\frac{x - X_1}{h_n}\right) \mathbb{E} \left( \frac{\delta_1 Y_1^{-\ell}}{\bar{G}(Y_1)} \middle| X_1 \right) \right] \\
&= \frac{1}{h_n} \int K\left(\frac{x - u}{h_n}\right) \mathbb{E} \left[ \frac{\delta_1 Y_1^{-\ell}}{\bar{G}(Y_1)} \middle| X_1 = u \right] f(u) du,
\end{aligned}$$

and as  $\mathbb{1}_{\{T_1 \leq C_1\}} Y_1^{-\ell} = \mathbb{1}_{\{T_1 \leq C_1\}} T_1^{-\ell}$ , we get

$$\mathbb{E} \left[ \frac{T^{-\ell}}{\bar{G}(T_1)} \mathbb{E} [\mathbb{1}_{\{T_1 \leq C_1\}} | T_1] \middle| X_1 = u \right] = r_{\ell}(u). \quad (5.3)$$

By a change of variable and using  $\bar{r}_\ell(\cdot) = r_\ell(\cdot)f(\cdot)$ , we get

$$\begin{aligned} \left| \mathbb{E} \left( \frac{1}{h_n} \frac{\delta_1 Y^{-\ell}}{\bar{G}(Y)} K \left( \frac{x - X}{h_n} \right) \right) - \bar{r}_\ell(x) \right| &= \left| \int \frac{1}{h_n} K \left( \frac{x - u}{h_n} \right) r_\ell(u) f(u) du - \bar{r}_\ell(x) \right| \\ &= \left| \int \frac{1}{h_n} K \left( \frac{x - u}{h_n} \right) \bar{r}_\ell(u) du - \bar{r}_\ell(x) \right| \\ &= \left| \int K(t) [\bar{r}_\ell(x - h_n t) - \bar{r}_\ell(x)] dt \right|. \end{aligned}$$

We use a Taylor expansion to  $\bar{r}_\ell(\cdot)$  for  $\zeta \in ]x - h_n t, x[$ , we get

$$\begin{aligned} \sup_{x \in \mathcal{C}} |(\mathbb{E}[\bar{r}_{\ell,n}(x)] - \bar{r}_\ell(x))| &= \sup_{x \in \mathcal{C}} \left| \int K(t) [-h_n t \bar{r}'_\ell(x) - \frac{h_n^2 t^2}{2} \bar{r}''_\ell(\zeta)] dt \right| \\ &\leq h_n \sup_{x \in \mathcal{C}} \left| \int t K(t) \bar{r}'_\ell(x) dt \right| + h_n^2 \sup_{x \in \mathcal{C}} \left| \int \frac{t^2}{2} K(t) \bar{r}''_\ell(\zeta) dt \right| \\ &\leq h_n \sup_{x \in \mathcal{C}} \left| \int t K(t) \bar{r}'_\ell(x) dt \right| + \frac{h_n^2}{2} \sup_{x \in \mathcal{C}} \left| \int \bar{r}''_\ell(\zeta) t^2 K(t) dt \right|. \end{aligned}$$

Under Hypotheses **H***i*) and **K***ii*), the first term is equal to zero. The second term goes to zero for  $n \rightarrow \infty$  from hypotheses **D***i*) and **K***ii*). The last result complete the proof of the lemma.  $\square$

**Lemma 3.** *Under hypotheses **H***i*) and **K***i*), we have, for  $\ell = 1, 2$*

$$\sup_{x \in \mathcal{C}} | \bar{r}_{\ell,n}(x) - \mathbb{E}[\bar{r}_{\ell,n}(x)] | = O_{a.s.} \left( \sqrt{\frac{\log n}{nh_n}} \right) \quad \text{as } n \rightarrow \infty. \quad (5.4)$$

*Proof.* Let us consider the i.i.d sequence  $(X_1, Y_1, \delta_1), \dots, (X_n, Y_n, \delta_n)$  and define

$$\Phi_n = \left\{ \theta_x : \mathbb{R} \times \mathbb{R}_+^* \times \{0, 1\} \rightarrow \mathbb{R}^+ / \theta_x(u, y, \delta) = \frac{\delta y^{-\ell}}{nh_n \bar{G}(y)} K \left( \frac{x - u}{h_n} \right), \quad x \in \mathbb{R} \right\}.$$

By Lemma (3b) in Giné and Guillou (1999),  $\Phi_n$  is Vapnik-Cervonenkis (V-C) class of no-negative measurable functions. These are uniformly bounded with respective envelopes  $\Theta = \frac{M \|K\|_\infty}{nh_n \bar{G}(\tau_F)}$ .

Moreover,

$$\sup_{x \in \mathcal{C}} \mathbb{E} [\theta_x(X_1, Y_1, \delta_1)] \leq \frac{M \|K\|_\infty}{nh_n \bar{G}(\tau_F)} =: U_n$$

In the same way, we get

$$\begin{aligned} \sup_{x \in \mathcal{C}} \text{Var} [\theta_x(X_1, Y_1, \delta_1)] &\leq \sup_{x \in \mathcal{C}} \mathbb{E} [\theta_x^2(X_1, Y_1, \delta_1)] \\ &\leq \frac{M^2 \|K\|_2^2 \|f\|_\infty}{n^2 h_n \bar{G}^2(\tau_F)} =: \sigma_n^2 \end{aligned}$$

with  $\sigma_n \leq U_n$  for  $n$  large enough.

Now applying Talagrand's inequality [see Proposition 2.2 in Giné and Guillou (2001)], with  $t \geq A\sqrt{\frac{\log n}{nh_n}}$ , there exist two positives constants  $L$  and  $B$  such that

$$\begin{aligned} & \mathbb{P} \left[ \sup_{\theta_x \in \Phi_n} \left| \sum_{i=1}^n (\theta_x(X_i, Y_i, \delta_i) - \mathbb{E}[\theta_x(X_1, Y_1, \delta_1)]) \right| > A\sqrt{\frac{\log n}{nh_n}} \right] \\ & \leq L \exp \left( - \frac{A\sqrt{\frac{\log n}{nh_n}}}{L \frac{M\|K\|_\infty}{nh_n \bar{G}(\tau_F)}} \log \left[ 1 + \frac{A \frac{\sqrt{\log n} \|K\|_\infty}{(nh_n)^{3/2} \bar{G}(\tau_F)}}{L \left( \frac{M\|K\|_2 \sqrt{\|f\|_\infty}}{\sqrt{nh_n \bar{G}(\tau_F)}} + \frac{M\|K\|_\infty}{nh_n \bar{G}(\tau_F)} \sqrt{\log B \frac{\|K\|_\infty}{\sqrt{h_n \|K\|_2^2 \sqrt{\|f\|_\infty}}}} \right)^2} \right] \right), \end{aligned}$$

and using  $\log(1+x) \approx x$  (for  $x \rightarrow 0$ ), the right-hand of the last equation becomes an order of

$$L \exp \left( - \frac{A \bar{G}(\tau_F) \sqrt{\frac{\log n}{nh_n}}}{LM\|K\|_\infty} nh_n \frac{A \frac{\sqrt{\log n} \|K\|_\infty}{(nh_n)^{3/2} \bar{G}(\tau_F)}}{L \left( \frac{M\|K\|_2 \sqrt{\|f\|_\infty}}{\sqrt{nh_n \bar{G}(\tau_F)}} \right)^2} \right) = Ln^{-\frac{\bar{G}^2(\tau_F)}{M^2 \|K\|_2^2 \|f\|_\infty} \left(\frac{A}{L}\right)^2},$$

which by an appropriate choice of the constant  $A$ , can be made  $O(n^{-3/2})$ . The latter being a general term of summable series and by Borel-Cantelli's lemma we conclude the proof.  $\square$

Then Lemma 1- Lemma 3 permit to conclude the proof of the Theorem 1.  $\square$

Next we proceed to the proof of the Theorem 2.

*Proof.* Our goal is to show

$$\sqrt{nh_n}(r_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)) \quad \text{as } n \rightarrow \infty.$$

Note that, for  $\ell = 1, 2$ ,

$$\begin{aligned} \sqrt{nh_n}(\bar{r}_{\ell,n}(x) - \bar{r}_\ell(x)) &= \sqrt{nh_n}(\bar{r}_{\ell,n}(x) - \tilde{r}_{\ell,n}(x)) + \sqrt{nh_n}(\tilde{r}_{\ell,n}(x) - \mathbb{E}[\tilde{r}_{\ell,n}(x)]) \\ &+ \sqrt{nh_n}(\mathbb{E}[\tilde{r}_{\ell,n}(x)] - \bar{r}_\ell(x)) \\ &=: \Lambda_{\ell,n}(x) + \Gamma_{\ell,n}(x) + \Sigma_{\ell,n}(x). \end{aligned}$$

First, we consider the negligible terms  $\Lambda_{\ell,n}$  and  $\Sigma_{\ell,n}$ .

**Lemma 4.** Under **H ii), iii)** and by Lemma 1, Lemma 2, both  $\sqrt{nh_n}\Lambda_{\ell,n}(x)$  and  $\sqrt{nh_n}\Sigma_{\ell,n}(x)$  are  $o(1)$  as  $n \rightarrow \infty$ .

*Proof.* From Lemma 1, under **H ii)**, we get,

$$\Lambda_{\ell,n}(x) = \sqrt{nh_n}(\bar{r}_{\ell,n}(x) - \tilde{r}_{\ell,n}(x)) = O_{a.s.}\left(\sqrt{h_n \log_2 n}\right) = o_{a.s.}(1). \quad (5.5)$$

In the same way, from Lemma 2, under **H iii)**, we have,

$$\Sigma_{\ell,n}(x) = \sqrt{nh_n}(\mathbb{E}[\tilde{r}_{\ell,n}(x)] - \bar{r}_{\ell}(x)) = O\left(\sqrt{nh_n^5}\right) = o(1). \quad (5.6)$$

□

Now we consider the dominant terms  $\Gamma_{\ell,n}(x)$  for  $\ell \in \{1, 2\}$  and prove Lemma 5.

**Lemma 5.** Under hypotheses **H i), K** and **D i), ii)**, we have

$$(\Gamma_{1,n}(x), \Gamma_{2,n}(x))^T \xrightarrow{\mathcal{D}} \mathcal{N}(0, \kappa\Sigma(x)) \quad \text{as } n \rightarrow \infty.$$

*Proof.* We first estimate the asymptotic variance, for  $\ell = 1, 2$ , we get

$$\begin{aligned} \text{Var} [\Gamma_{\ell,n}(x)] &= nh_n \left\{ \mathbb{E} [\tilde{r}_{\ell,n}^2(x)] - \mathbb{E}^2 [\tilde{r}_{\ell,n}(x)] \right\} \\ &= h_n^{-1} \left\{ \mathbb{E} \left[ \frac{\delta_1 Y_1^{-2\ell}}{\bar{G}^2(Y_1)} K^2 \left( \frac{x - X_1}{h_n} \right) \right] - \mathbb{E}^2 \left[ \frac{\delta_1 Y_1^{-\ell}}{\bar{G}(Y_1)} K \left( \frac{x - X_1}{h_n} \right) \right] \right\} \\ &= \mathcal{V}_1 - \mathcal{V}_2. \end{aligned}$$

For  $\mathcal{V}_2$  proceeding as in Lemma 2 and under **H i), K** and **D i)**, we have

$$\mathcal{V}_2 = h_n^{-1} \left[ h_n \int K(s) \bar{r}_{\ell}(x - h_n s) ds \right]^2 = o(1). \quad (5.7)$$

Furthermore for  $\mathcal{V}_1$ , we get,

$$\begin{aligned}
\mathcal{V}_1 &= h_n^{-1} \mathbb{E} \left[ \frac{\delta_1 Y_1^{-2\ell}}{\overline{G}^2(Y_1)} K^2 \left( \frac{x - X_1}{h_n} \right) \right] = h_n^{-1} \mathbb{E} \left[ K^2 \left( \frac{x - X_1}{h_n} \right) \mathbb{E} \left[ \frac{T_1^{-2\ell}}{\overline{G}(T_1)} | X_1 \right] \right] \\
&= h_n^{-1} \int K^2 \left( \frac{x - u}{h_n} \right) \int \frac{t^{-2\ell}}{\overline{G}(t)} f_{T_1|X_1}(t|u) dt f(u) du \\
&= h_n^{-1} \int K^2 \left( \frac{x - u}{h_n} \right) \int \frac{t^{-2\ell}}{\overline{G}(t)} f_{T_1, X_1}(t, u) dt du \\
&= h_n^{-1} \int K^2 \left( \frac{x - u}{h_n} \right) \Upsilon_{2\ell}(u) du
\end{aligned}$$

by a change of variable and Taylor expansion, by **K iii)**, we get,

$$\begin{aligned}
&= h_n^{-1} \int K^2(t) \Upsilon_{2\ell}(x - h_n t) h_n dt \\
&= \Upsilon_{2\ell}(x) \int K^2(t) dt + o(h_n) \\
&= \Upsilon_{2\ell}(x) \kappa + o(h_n)
\end{aligned} \tag{5.8}$$

which together with (5.7) and (5.8) gives under **D ii)**, for  $\ell = 1, 2$

$$\text{Var} [\Gamma_{\ell, n}(x)] \longrightarrow \Upsilon_{2\ell}(x) \int K^2(t) dt \quad \text{as } n \rightarrow \infty.$$

In addition under **D ii)** and **K iii)**, we get easily

$$\begin{aligned}
\text{Cov}(\Gamma_{1, n}(x), \Gamma_{2, n}(x)) &= \mathbb{E}[\Gamma_{1, n}(x)\Gamma_{2, n}(x)] - \mathbb{E}[\Gamma_{1, n}(x)]\mathbb{E}[\Gamma_{2, n}(x)] \\
&= h_n^{-1} \left\{ \mathbb{E} \left[ \frac{\delta_1 Y_1^{-3}}{\overline{G}^2(Y_1)} K^2 \left( \frac{x - X_1}{h_n} \right) \right] - \mathbb{E} \left[ \frac{\delta_1 Y_1^{-1}}{\overline{G}(Y_1)} K \left( \frac{x - X_1}{h_n} \right) \right] \mathbb{E} \left[ \frac{\delta_1 Y_1^{-2}}{\overline{G}(Y_1)} K \left( \frac{x - X_1}{h_n} \right) \right] \right\} \\
&= \Upsilon_3(x) \kappa + o(1).
\end{aligned}$$

Next, we will show that any linear combinations are asymptotically gaussian. For let  $(z_1, z_2)^t$  be a real numbers, we put

$$\Delta_n(x) = \sum_{\ell=1}^2 z_\ell \Gamma_{\ell, n}(x) =: \sum_{i=1}^n \left( z_1 \Delta_{i, n}^1(x) + z_2 \Delta_{i, n}^2(x) \right) \tag{5.9}$$

where

$$\Delta_{i, n}^\ell(x) := (nh_n)^{-1/2} \left\{ \frac{\delta_i Y_i^{-\ell}}{\overline{G}(Y_i)} K \left( \frac{x - X_i}{h_n} \right) - \mathbb{E} \left[ \frac{\delta_i Y_i^{-\ell}}{\overline{G}(Y_i)} K \left( \frac{x - X_i}{h_n} \right) \right] \right\} \text{ for } \ell = 1, 2.$$

Now in order to show that (5.9) is asymptotically normal we verify the Berry-Essèen condition (Chow and Teicher (1997), p. 322). For that we need to prove :

$$\rho_n^3 := \sum_{i=1}^n \mathbb{E} \left[ |\Delta_{i,n}^\ell(x)|^3 \right] \longrightarrow 0 \quad (5.10)$$

with

$$\mathbb{E} [|\Delta_{i,n}^\ell(x)|^3] = \left(\frac{h_n}{n}\right)^{3/2} \mathbb{E} \left[ \left| \frac{\delta_i Y_i^{-\ell}}{h_n \overline{G}(Y_i)} K \left( \frac{x - X_i}{h_n} \right) - \mathbb{E} \left[ \frac{\delta_i Y_i^{-\ell}}{h_n \overline{G}(Y_i)} K \left( \frac{x - X_i}{h_n} \right) \right] \right|^3 \right].$$

Applying the  $C_r$  inequality (see Loève (1963), p. 155), we get

$$\begin{aligned} \mathbb{E} [|\Delta_{i,n}^\ell(x)|^3] &\leq 4 \left(\frac{h_n}{n}\right)^{3/2} \left\{ \mathbb{E} \left[ \frac{|\delta_i Y_i|^{-3\ell}}{h_n^3 \overline{G}^3(Y_i)} K^3 \left( \frac{x - X_i}{h_n} \right) \right] + \left| \mathbb{E} \left[ \frac{\delta_i Y_i^{-\ell}}{h_n \overline{G}(Y_i)} K \left( \frac{x - X_i}{h_n} \right) \right] \right|^3 \right\} \\ &\leq 4 \left(\frac{h_n}{n}\right)^{3/2} \{ \mathcal{M}_1 + \mathcal{M}_2 \}. \end{aligned}$$

and as  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are bounded under  $\mathbf{K}$  which gives that  $\rho_n^3 = O\left(\left(\frac{h_n^3}{n}\right)^{1/2}\right) = o(1)$ . Now, under **H*i***) the property (5.10) is satisfied which proves the asymptotic normality of  $\Gamma_{\ell,n}(x)$  and, together with (5.5) and (5.6), complete the proof of Lemma 5. □

Now to complete the proof of Theorem 2, consider the mapping  $\theta$  from  $\mathbb{R} \times \mathbb{R}_+^*$  to  $\mathbb{R}$  defined by  $\theta(x, y) = x/y$ . We deduce from Mann-Wald's Theorem (see Rao 1965, p. 321) that:

$$\sqrt{nh_n}(r_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \kappa \nabla \theta^T \Sigma(x) \nabla \theta)$$

where the gradient  $\nabla \theta^T = \left( \frac{\partial \theta}{\partial x}, \frac{\partial \theta}{\partial y} \right)$  is evaluated at  $(\bar{r}_1(x), \bar{r}_2(x))$ . Simple algebra gives then the variance

$$\sigma^2(x) = \kappa \frac{\Upsilon_2(x) \bar{r}_2^2(x) - 2\Upsilon_3(x) \bar{r}_1(x) \bar{r}_2(x) + \Upsilon_4(x) \bar{r}_1^2(x)}{\bar{r}_2^4(x)},$$

which completes the proof. □

## References

- [1] Carbonez A., Györfi L., Van Der Meulen E.C. (1995). Partitioning estimates of a regression function under random censoring. *Statist. and Decisions*. **76**, 1335-1344.
- [2] Carroll R.J., Ruppert, D. Transformation and Weighting in Regression. *Chapman and Hall, London*. 1988.
- [3] Chow Y.S., Teicher H. *Probability theory. Independence, interchangeability, martingales*. Springer, New York. 1977.
- [4] Deheuvels P., Einmahl J.H.J. (2000). Functional limit laws for the increments of Kaplan-Meier product limit processes and applications. *Ann Probab.*, **28**, 1301-1335.
- [5] Delecroix M., Lopez O., Patilea V. (2008). Nonlinear censored regression using synthetic data. *Scand. J. Stat.*, **35**, 248-256.
- [6] Farum, N.R. (1990). Improving the relative error of estimation. *The Amer. Stat.*, **44**, 288-289.
- [7] Foldès A., Rejtő L. (1981). A LIL type result for the product limit estimator. *Probab. Theory and Related Fields*, **56**, 75-86.
- [8] Giné E., Guillou A. (1999). Law of the iterated logarithm for censored data. *Ann. of Probab.*, **27**, 2042- 2067.
- [9] Giné E., Guillou A. (2002). Rates of strong uniform consistency for multivariate kernel density estimators. *Ann. I. H. Poincaré*. **38**, 907- 921.
- [10] Guessoum Z., Ould Saïd E. (2008). On nonparametric estimation of the regression function under random censorship model. *Stat. and Decisions*, **26**, 1001-1020.

- [11] Kaplan, E.L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Stat. Assoc.*, **53**, 458-481.
- [12] Khoshgoftaar, T.M., Bhattacharyya, B.B., Richardson, G.D. (1992). Prediction software errors, during development, using nonlinear regression models: comparative study. *IEEE Trans. Reliab.*, **41**, 390-395.
- [13] Köhler M., Máthè K., Pintër M. (2002). Prediction from randomly right censored data. *J. Multivar. Anal.*, **80**, 73-100.
- [14] Loève M., *Probability theory*. Springer-Verlag, New York, 1963.
- [15] Narula S.C., Wellington, J.F. (1977). Prediction, linear regression and the minimum sum of relative errors. *Technometrics*, **19**, 185-190.
- [16] Park H., Shin K.I., Jones M.C., Vines S.K. (2008). Relative error prediction via kernel regression smoothers. *Journal of Stat. Plann. and Infer.*, 138, 2887-2898.
- [17] Park H., Stefanski L.A. (1998). Relative error prediction. *Stat. & Probab. Lett.*, **40**, 227-236.
- [18] Rao C.R. *A linear statistical inference and its applications*, New-York; Wiley. 1965.