# Nonparametric estimation of the conditional extreme-value index with random covariates and censoring

Pathé Ndao, Aliou Diop, Jean-François Dupuy

# Nonparametric estimation of the conditional extreme-value index with random covariates and censoring

Pathé NDAO[a], Aliou DIOP[a], Jean-François DUPUY[b]

[a] *LERSTAD, Université Gaston Berger, Saint Louis, Sénégal.*
[b] *IRMAR-Institut National des Sciences Appliquées de Rennes, France.*

**Abstract**

Estimation of the extreme-value index of a heavy-tailed distribution is addressed when some random covariate information is available and the data are randomly right-censored. An inverse-probability-of-censoring-weighted kernel version of Hill's estimator of the extreme-value index is proposed and its asymptotic normality is established. Based on this, a Weissman-type estimator of conditional extreme quantiles is also constructed. A simulation study is conducted to assess the finite-sample behaviour of the proposed estimators.

*Keywords:* Conditional extreme-value index, conditional extreme quantile, conditional Kaplan-Meier estimator, kernel estimator, simulations.

## 1. Introduction

Estimation of extreme quantiles has become a crucial issue in many fields, such as hydrology, insurance, medicine and reliability. Let $Y$ be a random variable with cumulative distribution function $F$ and let $Y_1, \ldots, Y_n$ be independent copies of $Y$. Extreme quantiles of $F$ are defined as quantities of the form

$$F^{\leftarrow}(1-\alpha) = \inf\{y : F(y) \geq 1-\alpha\},$$

where $\alpha$ is so small that this quantile falls beyond the range of the observed $Y_1, \ldots, Y_n$. This problem is closely related to estimation of the extreme-value index of $Y$. The extreme-value index drives the behavior of $F$ in its right tail and thus plays a central role in the analysis of extremes. Recent monographs on extreme value theory and in particular on estimation of the extreme-value index and extreme quantiles include Embrechts *et al.* (1997), Beirlant *et al.* (2004), Reiss and Thomas (2007) and Novak (2012).

When some covariate information $X$ is available and the distribution of $Y$ depends on $X$, the problem is to estimate the conditional extreme-value index and conditional extreme quantiles $F^{\leftarrow}(1-\alpha|x) = \inf\{y : F(y|x) \geq 1-\alpha\}$ of the distribution $F(\cdot|x)$ of $Y$ given $X = x$. Motivating examples include estimation of extreme rainfalls given the geographical location

*Email addresses:* ndao.pathe@yahoo.fr (Pathé NDAO), aliou.diop@ugb.edu.sn (Aliou DIOP), Jean-Francois.Dupuy@insa-rennes.fr (Jean-François DUPUY)

(Gardes and Girard, 2010), analysis of extreme temperatures given topological parameters (Ferrez *et al.*, 2011) and the study of extreme earthquakes given the location (Pisarenko and Sornette, 2003). Estimation of the conditional extreme-value index and conditional extreme quantiles with fixed (or non-random) covariates has been investigated rather extensively in the recent extreme value literature. We refer to Beirlant *et al.* (2004), Gardes and Girard (2008), Gardes *et al.* (2010), Stupfler (2013) and the references therein for an overview of the available methodology, including the case where the covariate is functional (Gardes and Girard, 2012). To date, less attention has been paid to the random covariate case, despite its practical interest. Gardes and Stupfler (2014) and Goegebeur *et al.* (2014b) adapt Hill's estimator of the extreme-value index of a heavy-tailed distribution to the presence of a random covariate. The moment estimator introduced by Dekkers *et al.* (1989) is adapted to the presence of random covariates by Goegebeur *et al.* (2014a). Daouia *et al.* (2011) proposed a kernel-based estimator of conditional extreme quantiles with random covariates.

In this paper, we address estimation of the conditional extreme-value index and conditional extreme quantiles with random covariates when moreover, the observations $Y_1, \ldots, Y_n$ are randomly right-censored. Censoring commonly occurs in the analysis of event time data. For example, $Y$ may represent the duration until the occurrence of some event of interest (such as death of a patient, ruin of a company...). If censoring is present, the observations consist of triplets $(Z_i, \delta_i, X_i)$, $i = 1, \ldots, n$, where $Z_i = \min(Y_i, C_i)$, $\delta_i = 1_{\{Y_i \leq C_i\}}$, $1_{\{\cdot\}}$ is the indicator function and $C_i$ is a random censoring time which provides a lower bound on $Y_i$ if $\delta_i = 0$. When there is no covariate information, estimation of the extreme-value index from censored data is considered by Delafosse and Guillou (2002), Gomes and Oliveira (2003), Beirlant *et al.* (2010), Gomes and Neves (2011), Brahimi *et al.* (2013) and Worms and Worms (2014). Matthys *et al.* (2004), Beirlant *et al.* (2007) and Einmahl *et al.* (2008) additionally address estimation of extreme quantiles. Ndao *et al.* (2014) address estimation of the conditional extreme-value index and conditional extreme quantiles with fixed covariates and censoring. The authors combine a moving-window technique (for tackling the covariate information) and the inverse-probability-of-censoring weighting method.

To our knowledge, estimation of the conditional extreme-value index and extreme quantiles with random covariates and censoring has not yet been addressed. This is the topic for the present paper. We first construct an estimator of the conditional extreme-value index and we establish its asymptotic normality. Our proposal combines a kernel version of Hill's estimator of the extreme-value index (such as developed in Goegebeur *et al.* (2014b) in the uncensored case) with the inverse-probability-of-censoring weighting principle (such as used in Einmahl *et al.* (2008) and Brahimi *et al.* (2013) for estimating the unconditional extreme-value index with censoring). Then, we propose a Weissman-type estimator of conditional extreme quantiles under censoring. The finite-sample performance of the proposed estimators are assessed via simulations and compared with several alternative estimators.

The remainder of this paper is organized as follows. In Section 2, we construct our estimator of the conditional extreme-value index and we establish its asymptotic normality. In Section 3, we propose an estimator of conditional extreme quantiles under censoring. In Section 4, we assess via simulations the finite sample behaviour of our estimators. Some

perspectives are given in Section 5. All proofs are deferred to an appendix.

## 2. Construction of the estimator and asymptotic properties

*2.1. The proposed estimator*

Let $(X_i, Y_i), i = 1, \ldots, n$, be independent copies of the random pair $(X, Y)$ where $Y$ is a non-negative random variable and $X \in \mathcal{X}$ (with $\mathcal{X}$ some bounded set of $\mathbb{R}^p$) is a $p$-dimensional covariate with probability density function $g$. We assume that $Y$ can be right-censored by a non-negative random variable $C$. Thus we really observe independent triplets $(X_i, \delta_i, Z_i), i = 1, \ldots, n$, where $Z_i = \min(Y_i, C_i)$, $\delta_i = 1_{\{Y_i \leq C_i\}}$ and $1_{\{A\}}$ is the indicator function of the event $A$. The random variable $C$ is defined on the same probability space $(\Omega, \mathcal{C}, \mathbb{P})$ as $Y$. We assume that $C_1, \ldots, C_n$ are independent of each other and that $Y$ and $C$ are independent given $X$. Let $F(\cdot|x)$ and $G(\cdot|x)$ denote the conditional cumulative distribution functions of $Y$ and $C$ given $X = x$, respectively. Let also $\bar{F}(\cdot|x) = 1 - F(\cdot|x)$ and $\bar{G}(\cdot|x) = 1 - G(\cdot|x)$ be the conditional survival functions of $Y$ and $C$ given $X = x$.

In this paper, we focus on heavy tails. Precisely, we assume that the conditional survival functions of $Y$ and $C$ given $X = x$ satisfy

**(C1)** $\bar{F}(u|x) = u^{-1/\gamma_1(x)} L_1(u|x) \quad$ and $\quad \bar{G}(u|x) = u^{-1/\gamma_2(x)} L_2(u|x)$,

where $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are unknown positive continuous functions of the covariate $x$ and for $x$ fixed, $L_1(\cdot|x)$ and $L_2(\cdot|x)$ are slowly varying functions at infinity, that is, for all $\lambda > 0$,

$$\lim_{u \to \infty} \frac{L_i(\lambda u|x)}{L_i(u|x)} = 1, \qquad i = 1, 2.$$

This amounts to saying that $\bar{F}(\cdot|x)$ and $\bar{G}(\cdot|x)$ are regularly varying functions at infinity with index $-1/\gamma_1(x)$ and $-1/\gamma_2(x)$ respectively. Condition (C1) also amounts to assuming that the conditional distributions of $Y$ and $C$ given $X = x$ are in the Fréchet maximum domain of attraction. In what follows, the functions $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are referred to as conditional extreme-value index functions.

**Remark 1.** By conditional independence of $Y$ and $C$, the conditional cumulative distribution function $H(\cdot|x)$ of $Z$ given $X = x$ is also heavy-tailed, with conditional extreme-value index $\gamma(x) = \gamma_1(x)\gamma_2(x)/(\gamma_1(x) + \gamma_2(x))$. To see this, note that for every $u$ and $x$,

$$\begin{aligned} \bar{H}(u|x) := 1 - H(u|x) &= \bar{F}(u|x)\bar{G}(u|x) \\ &= u^{-1/\gamma_1(x)} L_1(u|x) u^{-1/\gamma_2(x)} L_2(u|x) \\ &= u^{-1/\gamma(x)} L(u|x), \end{aligned}$$

where $\gamma(x)$ is as above and $L(u|x) = L_1(u|x)L_2(u|x)$. Moreover,

$$\lim_{u \to \infty} \frac{L(\lambda u|x)}{L(u|x)} = \lim_{u \to \infty} \frac{L_1(\lambda u|x)}{L_1(u|x)} \frac{L_2(\lambda u|x)}{L_2(u|x)} = 1.$$

□

If the $Z_i$ were uncensored (that is, $Z_i = Y_i$ for all $i$), Goegebeur *et al.* (2014b) propose to estimate $\gamma_1(x)$ by the following kernel version of Hill's estimator:

$$\widehat{\gamma}^H_{t_n}(x) = \sum_{i=1}^{n} K_h(x - X_i)(\log Z_i - \log t_n)1_{\{Z_i > t_n\}} \Big/ \sum_{i=1}^{n} K_h(x - X_i)1_{\{Z_i > t_n\}}, \tag{2.1}$$

where $K_h(x) := h^{-p}K(x/h)$, $K$ is a probability density function on $\mathbb{R}^p$, $h := h_n$ is a positive non-random bandwidth sequence such that $h \to 0$ as $n \to \infty$ and $t_n$ is a positive non-random threshold sequence with $t_n \to \infty$ for $n \to \infty$.

The estimator (2.1) is not consistent for $\gamma_1(x)$ if it is directly applied to the censored sample $(X_i, \delta_i, Z_i), i = 1, \ldots, n$. Indeed, under appropriate regularity assumptions, estimator (2.1) will converge to the extreme-value index $\gamma(x)$ of the conditional distribution of $Z$ given $X = x$ (Goegebeur *et al.*, 2014b). To accommodate censoring, we suggest to divide (2.1) by the proportion $\widehat{p}_{t_n}(x)$ of uncensored observations among the $\{Z_i, i = 1, \ldots, n\}$ that are larger than $t_n$, in a neighborhood of $x$:

$$\widehat{p}_{t_n}(x) = \frac{\bar{H}^1_n(t_n|x)}{\bar{H}_n(t_n|x)}, \tag{2.2}$$

where $\bar{H}_n(t_n|x) = \sum_{i=1}^n B_i(x)1_{\{Z_i > t_n\}}$, $\bar{H}^1_n(t_n|x) = \sum_{i=1}^n B_i(x)1_{\{Z_i > t_n, \delta_i = 1\}}$ and the weights $B_i(x)$ are defined as:

$$B_i(x) = K\left(\frac{x - X_i}{h}\right) \Big/ \sum_{j=1}^{n} K\left(\frac{x - X_j}{h}\right).$$

Therefore, we propose to estimate $\gamma_1(x)$ by

$$\widehat{\gamma}^{(c,H)}_{t_n}(x) := \frac{\widehat{\gamma}^H_{t_n}(x)}{\widehat{p}_{t_n}(x)}. \tag{2.3}$$

This estimator depends on both the bandwidth $h$ and threshold $t_n$. The choice of $h$ and $t_n$ will be discussed in the simulation study of Section 4.
The limiting distribution of $\widehat{\gamma}^{(c,H)}_{t_n}(x)$ is investigated in the next section.

*2.2. Asymptotic results*

Some regularity conditions are needed for proving our results (these conditions are adapted from Daouia *et al.* (2011) and Goegebeur *et al.* (2014b)). First, we require that some Lipschitz conditions hold (in these conditions, $d$ denotes the Euclidean distance on $\mathbb{R}^p$). For all $(x, x') \in \mathcal{X} \times \mathcal{X}$, we assume that:

**(C2)** There exists $c_\gamma > 0$ such that $\left|\frac{1}{\gamma(x)} - \frac{1}{\gamma(x')}\right| \leq c_\gamma d(x, x')$.

**(C3)** There exist $c_L > 0$ and $u_0 > 1$ such that $\sup_{u \geq u_0} \left|\frac{\log L(u|x)}{\log u} - \frac{\log L(u|x')}{\log u}\right| \leq c_L d(x, x')$.

**(C4)** There exists $c_g > 0$ such that $|g(x) - g(x')| \leq c_g d(x, x')$.

The next condition is classical in extreme value theory where it is called a second-order condition.

**(C5)** There exists a function $\rho(x) < 0$ and a regularly varying function $b(\cdot|x)$ with index $\rho(x)$ such that for all $u > 0$,

$$\lim_{t\to\infty} \frac{H^{\leftarrow}\left(1 - \frac{1}{tu} \mid x\right) / H^{\leftarrow}\left(1 - \frac{1}{t} \mid x\right) - u^{\gamma(x)}}{b(t|x)} = u^{\gamma(x)} \frac{u^{\rho(x)} - 1}{\rho(x)}.$$

Finally, we impose the following condition on the kernel function $K$:

**(C6)** $K$ is a bounded density function on $\mathbb{R}^p$, with support $S$ included in the unit ball of $\mathbb{R}^p$.

We are now in position to state our results. We first investigate the asymptotic distribution of $\widehat{p}_{t_n}(x)$, which is interesting on its own (thereafter, $\xrightarrow{\mathcal{D}}$ will denote the convergence in distribution). The proof is given in Appendix A.

**Proposition 2.1.** *Suppose (C1)-(C6) hold. Let $t_n$ be such that $t_n \to \infty$, $nh^p\bar{H}(t_n|x) \to \infty$ and $nh^{p+2}\bar{H}(t_n|x)(\log t_n)^2 \to 0$ as $n \to \infty$. Then, for all $x \in \mathcal{X}$ such that $g(x) > 0$,*

$$\sqrt{nh^p\bar{H}(t_n|x)}(\widehat{p}_{t_n}(x) - p_x) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{p_x(1-p_x)\|K\|_2^2}{g(x)}\right) \qquad \text{as } n \to \infty,$$

*where $p_x = \gamma_2(x)/(\gamma_1(x) + \gamma_2(x))$ and $\|K\|_2^2 = \int K^2(u)du$.*

We now derive the limiting distribution of the proposed estimator (2.3) of $\gamma_1(x)$. The proof is given in Appendix A.

**Theorem 2.2.** *Suppose that conditions (C1)-(C6) hold. Let $t_n$ be such that $t_n \to \infty$, $nh^p\bar{H}(t_n|x) \to \infty$, $nh^{p+2}\bar{H}(t_n|x)(\log t_n)^2 \to 0$ and $\sqrt{nh^p\bar{H}(t_n|x)}b(t_n|x) \to 0$ as $n \to \infty$. Then, for all $x \in \mathcal{X}$ such that $g(x) > 0$,*

$$\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{\gamma}_{t_n}^{(c,H)}(x) - \gamma_1(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\gamma_1^3(x)}{\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right) \qquad \text{as } n \to \infty. \tag{2.4}$$

We note that the asymptotic variance in (2.4) is the same (up to the scale factor $\|K\|_2^2/g(x)$) as the asymptotic variance obtained by Beirlant *et al.* (2007) for the adapted Hill estimator under the Hall model when censoring occurs and there is no covariate. We also note that when there is no censoring, the asymptotic variance in (2.4) reduces to the asymptotic variance of Goegebeur *et al.* (2014b)'s estimator (2.1).

## 3. Estimation of conditional extreme quantiles

In this section, we address the problem of estimating conditional extreme quantiles $q(\alpha_n|x)$ of order $1-\alpha_n$ of the conditional distribution $F(\cdot|x)$ of $Y$ given $X=x$. Such quantiles verify $\bar{F}(q(\alpha_n|x)|x)=\alpha_n$ where $\alpha_n \to 0$ as $n \to +\infty$.

First, we consider the kernel conditional Kaplan-Meier estimator of $\bar{F}(\cdot|x)$ (see Dabrowska (1992)), defined for all $(x,t) \in \mathcal{X} \times [0,\infty)$ by

$$\widehat{\bar{F}}_n(t|x) = \prod_{i=1}^{n} \left[ 1 - \frac{B_i(x)}{\sum_{j=1}^{n} 1_{\{Z_j \geq Z_i\}} B_j(x)} \right]^{1_{\{Z_i \leq t, \delta_i = 1\}}} \qquad \text{if } t \leq Z_{(n)} \tag{3.5}$$

and $\widehat{\bar{F}}_n(t|x)=0$ if $t > Z_{(n)}$, where $Z_{(1)} \leq ... \leq Z_{(n)}$ denote the order statistics of $Z_1, \ldots, Z_n$. Based on this, an estimator of the conditional quantile $q(\alpha|x)$ (for fixed $\alpha \in (0,1)$) can be defined via the generalized inverse of $\widehat{\bar{F}}_n(\cdot|x)$, as:

$$\widehat{q}_n^c(\alpha|x) := \widehat{\bar{F}}_n^{\leftarrow}(\alpha|x) = \inf\{u, \widehat{\bar{F}}_n(u|x) \leq \alpha\}. \tag{3.6}$$

When $\alpha_n \to 0$ as $n \to \infty$, we propose to estimate the conditional extreme quantile $q(\alpha_n|x)$ by the following Weissman-type (Weissman, 1978) estimator:

$$\widehat{q}_n^{(c,W)}(\alpha_n|x) := \widehat{q}_n^c\left(\widehat{\bar{F}}_n(Z_{(n-k)}|x)|x\right)\left(\widehat{\bar{F}}_n(Z_{(n-k)}|x)/\alpha_n\right)^{\widehat{\gamma}_{Z_{(n-k)}}^{(c,H)}(x)}. \tag{3.7}$$

Here, $\widehat{q}_n^c(\widehat{\bar{F}}_n(Z_{(n-k)}|x)|x)$ is the kernel quantile estimator (3.6), $(\widehat{\bar{F}}_n(Z_{(n-k)}|x)/\alpha_n)^{\widehat{\gamma}_{Z_{(n-k)}}^{(c,H)}(x)}$ is an extrapolation factor and $\widehat{\gamma}_{Z_{(n-k)}}^{(c,H)}(x)$ is the estimator (2.3) with $t_n$ replaced by $Z_{(n-k)}$. The choice of the threshold $Z_{(n-k)}$ (or equivalently, of the indice $k$) is discussed in the simulation study.

**Remark 2.** When there is no censoring, $\widehat{\bar{F}}_n(t|x)$ reduces to

$$\widetilde{\bar{F}}_n(t|x) := \sum_{i=1}^{n} K_h(x-X_i) 1_{\{Z_i > t\}} / \sum_{i=1}^{n} K_h(x-X_i) \tag{3.8}$$

and $\widehat{q}_n^c(\alpha|x)$ becomes

$$\widehat{q}_n(\alpha|x) := \inf\{u, \widetilde{\bar{F}}_n(u|x) \leq \alpha\}. \tag{3.9}$$

Several authors have considered the estimator (3.9) when $\alpha \in (0,1)$ is fixed. In particular, the consistency of (3.9) was established by Stone (1977) and Gannoun (1990). Asymptotic normality was proved by Stute (1986), Samanta (1989) and Berlinet *et al.* (2001). The asymptotic normality of (3.8) was established by Daouia *et al.* (2011) when $t := t_n \to \infty$ (*i.e.*, when (3.8) is used to estimate small tail probabilities). Daouia *et al.* (2011) also proved asymptotic normality of (3.9) when estimating extreme quantiles (*i.e.*, when $\alpha := \alpha_n \to 0$ as $n \to \infty$). □

In the present paper, we assess the properties of the proposed estimator (3.7) via simulations.

## 4. Simulation study

In this section, we assess the finite-sample performance of the proposed estimators (2.3) and (3.7).

### *4.1. The study design*

The simulation design is as follows. We simulate $N = 100$ samples of size $n$ ($n = 200, 400, 600, 800$) of independent replicates $(Z_i, \delta_i, X_i)$, where $Z_i = \min(Y_i, C_i)$ and $X_i \in \mathbb{R}$ is uniformly distributed on $[0, 1]$. The conditional distribution of $Y_i$ given $X_i = x$ is Pareto with parameter $\gamma_1(x) = 0.5\left(0.1 + \sin(\pi x) \times \left(1.1 - 0.5\exp\left(-64\,(x - 0.5)^2\right)\right)\right)$. The distribution of $C_i$ is Pareto and its parameter $\gamma_2$ is chosen to yield the desired censoring percentage $c$ ($c = 10\%, 25\%, 40\%$). The pattern of $\gamma_1(\cdot)$ is given in Figure 1.

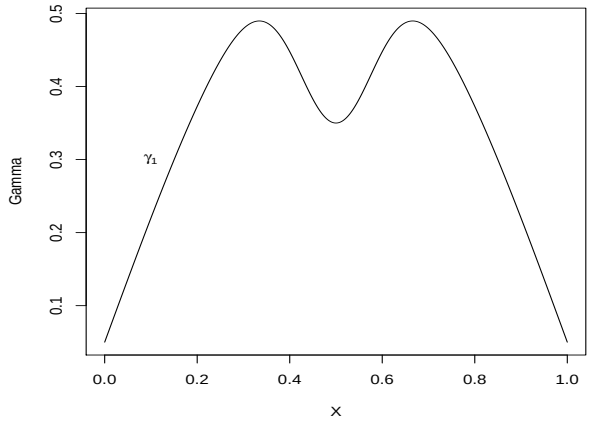


Figure 1: Pattern of $\gamma_1(\cdot)$ on $[0, 1]$.

For each of the $N$ simulated samples, we estimate $\gamma_1(\cdot)$ at $x = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$ using the estimator (2.3) with a biquadratic kernel $K(x) = \frac{15}{16}(1 - x^2)^2 1_{\{-1 \le x \le 1\}}$. In order to calculate our estimator, we need to choose the bandwidth $h$ and threshold $t_n$. As mentioned in Section 3, we take $t_n$ as the $(n - k)$th order statistic $Z_{(n-k)}$, as is classical in extreme value statistics.

Now, we propose an algorithm for choosing $(h, k)$. This algorithm is adapted from Goegebeur *et al.* (2014b). First, we select the bandwidth $h$ using the following cross-validation criterion, introduced by Yao (1999) and also implemented by Daouia *et al.* (2011):

$$h^* := \arg\min_{h \in \mathbb{H}} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(1_{\{Z_i > Z_j\}} - \widehat{\overline{F}}_{n,-i}(Z_j | x_i)\right)^2,$$

where $\widehat{\overline{F}}_{n,-i}$ is the kernel conditional Kaplan-Meier estimator (3.5) (depending on $h$) calculated on the sample $\{(X_j, \delta_j, Z_j), j = 1, \ldots, n, j \neq i\}$ and $\mathbb{H}$ is a grid of $h$ (we take

$\mathbb{H} = \{h_1 < h_2 < \ldots < h_M\}$, where the points $h_2, \ldots, h_{M-1}$ are regularly distributed between $h_1 = 1/(5 \log n)$ and $h_M = 0.5$, with $M = 60$). Once the bandwidth is determined, we choose $k$ as follows:

1. we compute the estimate $\widehat{\gamma}^{(c,H)}_{Z_{(n-k)}}(x)$ with $k = 1, \ldots, n-1$,
2. we form several successive "blocks" of estimates $\widehat{\gamma}^{(c,H)}_{Z_{(n-k)}}(x)$ (one block for $k \in \{1, \ldots, 15\}$, a second block for $k \in \{16, \ldots, 30\}$ and so on),
3. we calculate the standard deviation of the estimates within each block,
4. we determine the $k$-value to be used (thereafter denoted by $k^*$) from the block with minimal standard deviation. Precisely, we take the middle value of the $k$-values in the block.

Finally, we estimate $\gamma_1(x)$ from (2.3) with $(h, k) = (h^*, k^*)$.

*4.2. Results*

For each configuration `sample size`×`censoring percentage`×`covariate value` of the simulation design parameters, we calculate the averaged estimate of $\gamma_1(x)$ and the empirical root mean square error (RMSE) and mean absolute error (MAE) over the $N$ estimates. We also obtain asymptotic 95%-level confidence intervals for $\gamma_1(x)$ (the lower and upper bounds are averaged over the $N$ samples) and the averaged amplitude of the $N$ intervals. Variance estimates are obtained from a plug-in approach. The results are given in Table 1.

To assess the value of our estimator, we also provide a comparison with two simple alternative estimation strategies. These alternatives will provide a useful benchmark for evaluating the gain obtained by using our estimator. The first alternative is a complete-case procedure ("CC" for short): we remove all censored observations from the simulated samples and we calculate Goegebeur *et al.* (2014b)'s kernel version of Hill's estimator (2.1) on the resulting datasets. In the second alternative, we treat each observation $Z_i$ as if it were uncensored and we calculate the estimator (2.1) on the whole simulated datasets. We refer this alternative to as "CI" (for Censoring-Ignored). We obtain the averaged (over the $N$ simulated datasets) estimates of $\gamma_1(x)$ for both alternative strategies. We also provide the empirical RMSE and MAE. The results are reported in Table 3 (for CC) and Table 4 (for CI) in Appendix B (in order to save space and since they do not add much to the comparison, the confidence intervals and averaged amplitudes are omitted).

From Table 1, the proposed estimator (2.3) of $\gamma_1(x)$ performs quite well in every simulation scenario. As expected, its quality deteriorates as the censoring percentage increases and the sample size decreases. However, the numerical results indicate that the estimator still behaves reasonably when censoring is large or the sample size is moderate. Indeed in both cases, the bias and variability stay limited. Overall, the results are quite robust to censoring. Figure 2 shows the boxplots of the $N$ realizations of estimator (2.3) for each $x$ ($x = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$). From this figure, it appears that the estimator (2.3) captures quite well the shape of the conditional extreme-value index function $x \mapsto \gamma_1(x)$. From Table 3 and Table 4, we observe that the CC and CI estimators of $\gamma_1(x)$ are generally biased (sometimes strongly) even when censoring is moderate. The proposed estimator (2.3) clearly outperforms these simple alternative strategies.

*4.3. Estimation of conditional extreme quantiles*

In this paragraph, we assess performance of the estimator (3.7) of the conditional extreme quantile $q(1/1000|x)$ of order $1-1/1000$ of the conditional distribution of $Y$ given $X = x$, for $x = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$. The simulation setting is the same as above. For each configuration of the simulation design parameters, we calculate the estimate (3.7) with $(h^*, k^*)$ obtained in Section 4.1. Based on the $N$ simulated samples, we obtain the averaged estimate of $q(1/1000|x)$ and the RMSE and MAE. The results are given in Table 2. Similarly as above, we apply the CC and CI strategies to estimate $q(1/1000|x)$ from the simulated samples. The results are given in Table 5 (for CC) and Table 6 (for CI) in Appendix B.

From Table 2, we observe, as expected, that the performance of the proposed estimator deteriorates as censoring increases and the sample size decreases. When the sample size is moderate ($n = 200$, say) and the censoring percentage is moderate to high ($c \geq 25\%$, say), the bias can be quite large. For moderate sample size ($n = 200$), the bias stays limited however, when censoring does not exceed 10%. When the sample size is sufficiently large ($n \geq 400$, say), the bias of the proposed estimator stays limited in almost all simulation scenarios. Finally, Figure 3 shows the boxplots of the $N$ realizations of estimator (3.7) for every $x$ in $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. Similarly as above, we observe that estimator (3.7) captures quite well the shape of the conditional extreme quantile function $x \mapsto q(1/1000|x)$. From Table 5 and Table 6, the CC and CI estimators of $q(1/1000|x)$ are strongly biased (the bias is particularly noticeable when censoring exceeds 10%). These estimators generally underestimate the true quantile, which could be expected since these methods ignore the residual survival experience of the censored patients.

## 5. Conclusion and perspectives

In this paper, we address estimation of the extreme-value index and extreme quantiles of a heavy-tailed distribution when some random covariate information is available and the data are randomly right-censored. We constructed an inverse-probability-of-censoring-weighted kernel version of Hill's estimator of the extreme-value index and we established its asymptotic normality. We also proposed a Weissman-type estimator of conditional extreme quantiles. We assessed the finite-sample performance of these estimators via simulations. From these simulations, the proposed estimators perform well provided that the sample size is reasonably large, even when the censoring percentage is high. When the sample size is moderate, the proposed estimators still perform well provided that censoring stays limited. Moreover, the proposed estimators appear to correctly capture the shape of the unknown conditional extreme-value index function and conditional extreme quantile function. Finally, the proposed estimators clearly outperform simple existing alternatives, such as the classical complete-case approach.

Now, several issues deserve attention. In particular, a rigorous derivation of asymptotic properties of the proposed estimator (3.7) of conditional extreme quantiles is needed. This is a topic for our future research. Extending our convergence results to uniform (over $x$) convergence results is also of interest. Uniform convergence results would allow construc-

tion of simultaneous confidence bands for the conditional extreme-value index function and conditional extreme quantile function.

| | $\gamma_1(0.1) = 0.2199$ | | | $\gamma_1(0.2) = 0.3728$ | | | $\gamma_1(0.3) = 0.4793$ | | | $\gamma_1(0.4) = 0.4477$ | | | $\gamma_1(0.5) = 0.35$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 0.2255 | 0.2258 | 0.2225 | 0.3595 | 0.3679 | 0.3623 | 0.4673 | 0.4724 | 0.4695 | 0.4250 | 0.4221 | 0.4295 | 0.3620 | 0.3652 | 0.3631 |
| | (0.2082) | (0.2165) | (0.2191) | (0.2444) | (0.2690) | (0.2759) | (0.2764) | (0.2947) | (0.3227) | (0.2598) | (0.2623) | (0.2930) | (0.2412) | (0.2633) | (0.2682) |
| | [0.0032] | [0.0033] | [0.0039] | [0.0050] | [0.0081] | [0.0086] | [0.0090] | [0.0124] | [0.0200] | [0.0072] | [0.0074] | [0.0128] | [0.0056] | [0.0078] | [0.0082] |
| | [0.1394, | [0.1342, | [0.1222, | [0.2150, | [0.2078, | [0.1839, | [0.2800, | [0.2622, | [0.2261, | [0.2486, | [0.2256, | [0.2022, | [0.2244, | [0.2146, | [0.1964, |
| | 0.3115] | 0.3174] | 0.3228] | 0.5041] | 0.5280] | 0.5408] | 0.6547] | 0.6827] | 0.7130] | 0.6013] | 0.6185] | 0.6568] | 0.4996] | 0.5158] | 0.5299] |
| | 0.1721† | 0.1832† | 0.2006† | 0.2891† | 0.3202† | 0.3568† | 0.3747† | 0.4205† | 0.4869† | 0.3527† | 0.3929† | 0.4546† | 0.2752† | 0.3012† | 0.3335† |
| 400 | 0.2220 | 0.2195 | 0.2141 | 0.3666 | 0.3680 | 0.3664 | 0.4579 | 0.4621 | 0.4615 | 0.4368 | 0.4419 | 0.4409 | 0.3784 | 0.3830 | 0.3809 |
| | (0.1534) | (0.1557) | (0.1609) | (0.1945) | (0.2053) | (0.2130) | (0.2433) | (0.2458) | (0.2551) | (0.2305) | (0.2560) | (0.2741) | (0.2118) | (0.2272) | (0.2298) |
| | [0.0011] | [0.0013] | [0.0017] | [0.0025] | [0.0027] | [0.0032] | [0.0051] | [0.0061] | [0.0069] | [0.0044] | [0.0063] | [0.0086] | [0.0034] | [0.0045] | [0.0048] |
| | [0.1624, | [0.1562, | [0.1454, | [0.2660, | [0.2581, | [0.2412, | [0.3257, | [0.3145, | [0.2926, | [0.3144, | [0.3047, | [0.2828, | [0.2833, | [0.2784, | [0.2620, |
| | 0.2815] | 0.2828] | 0.2827] | 0.4671] | 0.4780] | 0.4915] | 0.5901] | 0.6096] | 0.6304] | 0.5591] | 0.5792] | 0.5991] | 0.4735] | 0.4876] | 0.4999] |
| | 0.1192† | 0.1266† | 0.1373† | 0.2011† | 0.2199† | 0.2503† | 0.2644† | 0.2951† | 0.3378† | 0.2446† | 0.2746† | 0.3163† | 0.1903† | 0.2092† | 0.2380† |
| 600 | 0.2220 | 0.2192 | 0.2153 | 0.3718 | 0.3689 | 0.3638 | 0.4562 | 0.4532 | 0.4612 | 0.4368 | 0.4362 | 0.4396 | 0.3727 | 0.3770 | 0.376 |
| | (0.1491) | (0.1523) | (0.1604) | (0.1906) | (0.1956) | (0.2121) | (0.2228) | (0.2323) | (0.2394) | (0.1996) | (0.2148) | (0.2184) | (0.1928) | (0.2006) | (0.2045) |
| | [0.0008] | [0.0010] | [0.0012] | [0.0022] | [0.0024] | [0.0030] | [0.0036] | [0.0043] | [0.0051] | [0.0022] | [0.0031] | [0.0039] | [0.0019] | [0.0026] | [0.0032] |
| | [0.1732, | [0.1680, | [0.1592, | [0.2892, | [0.2779, | [0.2618, | [0.3487, | [0.3339, | [0.3222, | [0.3371, | [0.3262, | [0.3122, | [0.3060, | [0.3041, | [0.2939, |
| | 0.2709] | 0.2704] | 0.2714] | 0.4544] | 0.4598] | 0.4658] | 0.5638] | 0.5725] | 0.6001] | 0.5366] | 0.5463] | 0.5671] | 0.4395] | 0.4500] | 0.4593] |
| | 0.0977† | 0.1024† | 0.1121† | 0.1651† | 0.1819† | 0.2040† | 0.2150† | 0.2385† | 0.2779† | 0.1995† | 0.2201† | 0.2549† | 0.1549† | 0.1681† | 0.1898† |
| 800 | 0.2238 | 0.2226 | 0.2205 | 0.3646 | 0.3627 | 0.3626 | 0.4619 | 0.4628 | 0.4637 | 0.4431 | 0.4385 | 0.4350 | 0.3734 | 0.3747 | 0.3767 |
| | (0.1372) | (0.1420) | (0.1510) | (0.1713) | (0.1825) | (0.2000) | (0.1944) | (0.2179) | (0.2354) | (0.1874) | (0.1960) | (0.2097) | (0.1870) | (0.1956) | (0.2042) |
| | [0.0005] | [0.0006] | [0.0008] | [0.0014] | [0.0016] | [0.0024] | [0.0024] | [0.0036] | [0.0049] | [0.0021] | [0.0023] | [0.0031] | [0.0018] | [0.0024] | [0.0031] |
| | [0.1827, | [0.1787, | [0.1729, | [0.2932, | [0.2844, | [0.2751, | [0.3705, | [0.3600, | [0.3450, | [0.3580, | [0.3430, | [0.3250 | [0.2960, | [0.2906, | [0.2817, |
| | 0.2649] | 0.2665] | 0.2681] | 0.4359] | 0.4411] | 0.4502] | 0.5534] | 0.5656] | 0.5823] | 0.5283] | 0.5340] | 0.5449] | 0.4395] | 0.4500] | 0.4593] |
| | 0.0822† | 0.0878† | 0.0952† | 0.1427† | 0.1567† | 0.1751† | 0.1829† | 0.2056† | 0.2373† | 0.1703† | 0.1910† | 0.2199† | 0.1335† | 0.1459† | 0.1654† |

**Table 1. Simulation results for $\gamma_1(x)$.** For each configuration of the simulation parameters $(n, c, x)$, the first line gives the averaged value of the $N = 100$ estimates of $\gamma_1(x)$. $(\cdot)$: empirical RMSE. $[\cdot]$: empirical MAE. $[\cdot, \cdot]$: 95%-level asymptotic confidence interval for $\gamma_1(x)$. †: averaged amplitude of the confidence intervals.

| | $\gamma_1(0.6) = 0.4477$ | | | $\gamma_1(0.7) = 0.4793$ | | | $\gamma_1(0.8) = 0.3728$ | | | $\gamma_1(0.9) = 0.2199$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 0.4423 | 0.4439 | 0.4496 | 0.4650 | 0.4657 | 0.4919 | 0.3759 | 0.3818 | 0.3869 | 0.2236 | 0.2201 | 0.2185 |
| | (0.2789) | (0.2910) | (0.3020) | (0.2857) | (0.2961) | (0.3617) | (0.2596) | (0.2806) | (0.2950) | (0.1912) | (0.1962) | (0.2057) |
| | [0.0094] | [0.0112] | [0.0135] | [0.0125] | [0.0143] | [0.0362] | [0.0070] | [0.0104] | [0.0121] | [0.0022] | [0.0026] | [0.0030] |
| | [0.2668, | [0.2476, | [0.2276, | [0.2766, | [0.2548, | [0.2468, | [0.2314, | [0.2239, | [0.2075, | [0.1370, | [0.1288, | [0.1184, |
| | 0.6178] | 0.6401] | 0.6716] | [0.6535] | [0.6766] | [0.7369] | 0.5203] | 0.5397] | 0.5662] | 0.3103] | 0.3114] | 0.3185] |
| | 0.3510† | 0.3925† | 0.4441† | 0.3770† | 0.4219† | 0.4900† | 0.2889† | 0.3159† | 0.3587† | 0.1732† | 0.1826† | 0.2001† |
| 400 | 0.4338 | 0.4390 | 0.4359 | 0.4680 | 0.4689 | 0.4782 | 0.3665 | 0.3681 | 0.3669 | 0.2202 | 0.2183 | 0.2163 |
| | (0.2188) | (0.2322) | (0.2379) | (0.2484) | (0.2577) | (0.2932) | (0.2148) | (0.2148) | (0.2384) | (0.1755) | (0.1811) | (0.1857) |
| | [0.0038] | [0.0055] | [0.0056] | [0.0056] | [0.0074] | [0.0127] | [0.0032] | [0.0034] | [0.0052] | [0.0015] | [0.0017] | [0.0019] |
| | [0.3124, | [0.3028, | [0.2810, | [0.3368, | [0.3209, | [0.3078, | [0.2644, | [0.2570, | [0.2399, | [0.1608, | [0.1554, | [0.1486, |
| | 0.5552] | 0.5753] | 0.5909] | 0.5991] | 0.6170] | 0.6487] | 0.4685] | 0.4791] | 0.4938] | 0.2796] | 0.2812] | 0.2840] |
| | 0.2428† | 0.2725† | 0.3099† | 0.2622† | 0.2961† | 0.3409† | 0.2040† | 0.2220† | 0.2539† | 0.1188† | 0.1257† | 0.1354† |
| 600 | 0.4353 | 0.4363 | 0.4411 | 0.4655 | 0.4575 | 0.4583 | 0.3701 | 0.3676 | 0.3662 | 0.2178 | 0.2141 | 0.2092 |
| | (0.2060) | (0.2200) | (0.2271) | (0.2161) | (0.2195) | (0.2434) | (0.1880) | (0.1881) | (0.1898) | (0.1450) | (0.1503) | (0.1622) |
| | [0.0029] | [0.0035] | [0.0042] | [0.0031] | [0.0036] | [0.0052] | [0.0014] | [0.0016] | [0.0019] | [0.0007] | [0.0008] | [0.0010] |
| | [0.3362, | [0.3254, | [0.3131, | [0.3592, | [0.3384, | [0.3193, | [0.2875, | [0.2762, | [0.2630, | [0.1694, | [0.1629, | [0.1530 |
| | 0.5344] | 0.5473] | 0.5690] | 0.5718] | 0.5765] | 0.5974] | 0.4527] | 0.4591] | 0.4694] | 0.2663] | 0.2652] | 0.2654] |
| | 0.1983† | 0.2219† | 0.2560† | 0.2125† | 0.2381† | 0.2782† | 0.1651† | 0.1828† | 0.2064† | 0.0969† | 0.1023† | 0.1124† |
| 800 | 0.4397 | 0.4401 | 0.4408 | 0.4623 | 0.4604 | 0.4614 | 0.3697 | 0.3691 | 0.3648 | 0.2211 | 0.2186 | 0.2165 |
| | (0.1940) | (0.1957) | (0.2057) | (0.2135) | (0.2214) | (0.2360) | (0.1759) | (0.1823) | (0.1939) | (0.1327) | (0.1340) | (0.1434) |
| | [0.0022] | [0.0022] | [0.0028] | [0.0032] | [0.0041] | [0.0050] | [0.0015] | [0.0017] | [0.0021] | [0.0003] | [0.0005] | [0.0006] |
| | [0.3544, | [0.3453, | [0.3301, | [0.3696, | [0.3560, | [0.3419, | [0.2990, | [0.2920, | [0.2772, | [0.1801, | [0.1751, | [0.1687, |
| | 0.5250] | 0.5349] | 0.5514] | 0.5550] | 0.5649] | 0.5809] | 0.4405] | 0.4462] | 0.4525] | 0.2621] | 0.2622] | 0.2642] |
| | 0.1706† | 0.1895† | 0.2213† | 0.1854† | 0.2089† | 0.2391† | 0.1415† | 0.1543† | 0.1753† | 0.0820† | 0.0871† | 0.0954† |

**Table 1 (continued).**

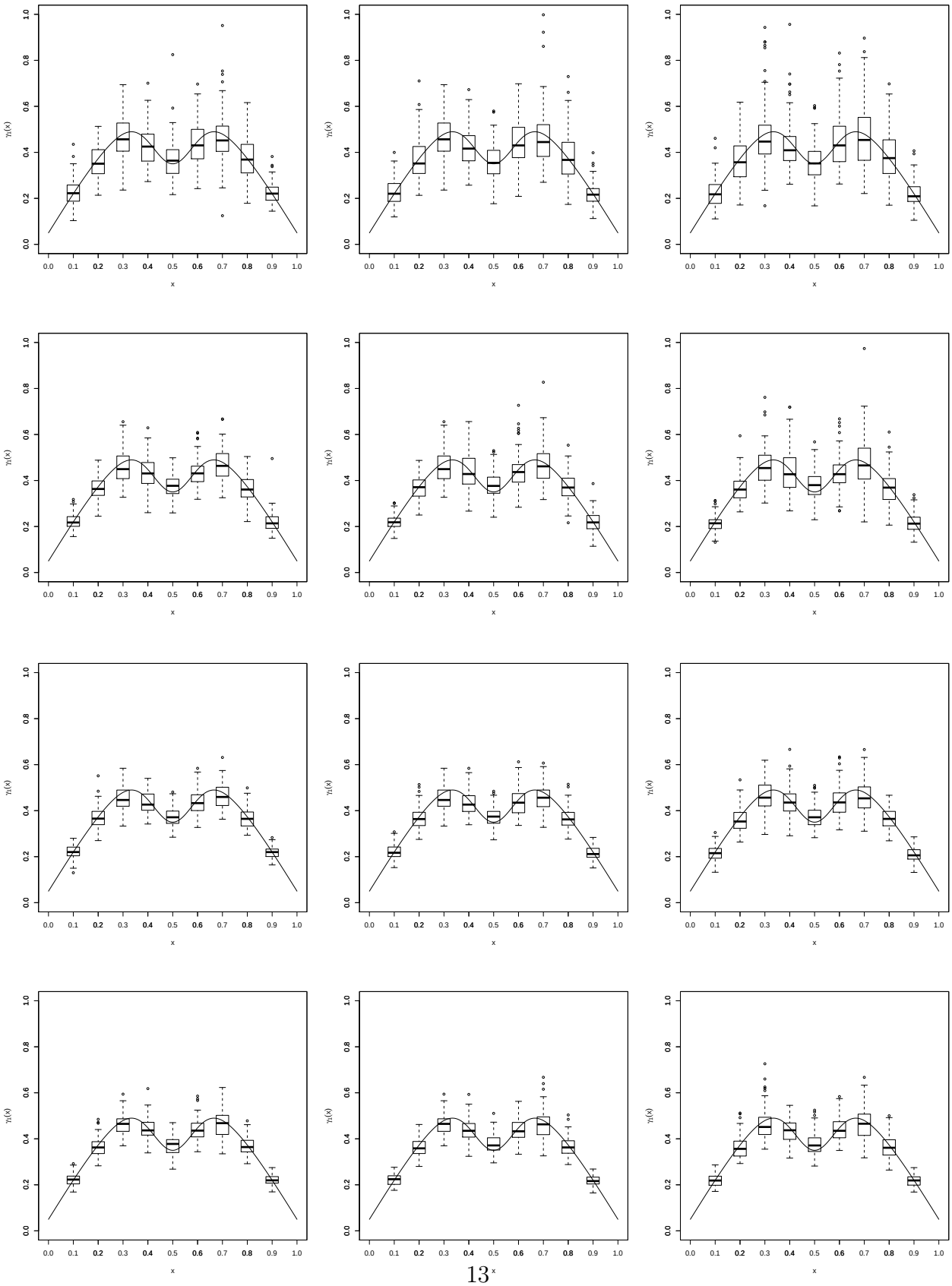




Figure 2: Boxplots of the $N$ estimates of $\gamma_1(x)$ for $n = 200$ (1st line), $n = 400$ (2nd line), $n = 600$ (3rd line), $n = 800$ (4th line). Left: $c = 10\%$, center: $c = 25\%$, right: $c = 40\%$.

| | $q(1/1000, 0.1) = 4.5695$ | | | $q(1/1000, 0.2) = 13.1358$ | | | $q(1/1000, 0.3) = 27.4140$ | | | $q(1/1000, 0.4) = 22.0358$ | | | $q(1/1000, 0.5) = 11.2201$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 4.9051 | 4.9217 | 4.9775 | 12.8627 | 15.1259 | 14.6754 | 29.7765 | 34.2158 | 47.4512 | 21.1591 | 20.8619 | 30.0911 | 15.1480 | 17.4580 | 19.7782 |
| | (2.0984) | (2.1909) | (2.8898) | (6.5390) | (14.6945) | (15.1103) | (22.1035) | (33.9952) | (91.4621) | (13.9045) | (14.5388) | (65.2830) | (11.1821) | (17.2417) | (19.2146) |
| | [1.4987] | [1.5564] | [1.7181] | [5.2294] | [7.7024] | [8.4554] | [15.1600] | [19.8775] | [35.1628] | [9.9591] | [10.0064] | [20.0666] | [6.5352] | [8.8535] | [11.2048] |
| 400 | 4.5509 | 4.4909 | 4.3758 | 12.7265 | 13.0571 | 13.2056 | 25.4270 | 27.4535 | 28.4287 | 21.8468 | 23.7294 | 25.7412 | 13.6515 | 13.5535 | 14.1533 |
| | (1.0149) | (1.0231) | (1.0838) | (4.1748) | (4.6262) | (6.0845) | (14.6121) | (19.7593) | (23.8713) | (11.1816) | (14.4708) | (23.1045) | (5.8417) | (6.8835) | (7.4824) |
| | [0.7783] | [0.7846] | [0.8026] | [3.3625] | [3.6549] | [4.2554] | [10.8706] | [12.0761] | [13.1128] | [8.2532] | [9.6850] | [13.5354] | [3.8271] | [4.2458] | [5.0220] |
| 600 | 4.5291 | 4.4845 | 4.3823 | 13.6544 | 13.4414 | 13.1955 | 24.0469 | 24.1875 | 26.2753 | 20.5668 | 21.2501 | 22.4150 | 13.7379 | 13.9236 | 14.0653 |
| | (0.8798) | (0.9537) | (0.9731) | (4.3521) | (4.5436) | (4.6639) | (10.3738) | (11.3229) | (13.2170) | (6.8417) | (9.0705) | (12.0648) | (5.5173) | (6.0139) | (6.7546) |
| | [0.7116] | [0.7504] | [0.7849] | [3.1794] | [3.3703] | [3.7097] | [8.7212] | [9.2147] | [10.3688] | [5.8623] | [7.0867] | [7.5981] | [3.7719] | [4.1005] | [4.2414] |
| 800 | 4.6094 | 4.5768 | 4.5582 | 12.3614 | 12.2954 | 12.6640 | 24.7095 | 25.5801 | 27.4010 | 21.7130 | 21.0711 | 20.9701 | 13.0676 | 13.6430 | 13.8643 |
| | (0.7210) | (0.7395) | (0.8746) | (3.6101) | (3.6328) | (3.7755) | (8.6029) | (10.9137) | (12.2523) | (6.7012) | (7.4548) | (7.8365) | (3.9094) | (4.8973) | (6.0034) |
| | [0.5762] | [0.6098] | [0.7028] | [2.6401] | [3.0161] | [3.6113] | [6.8036] | [8.3678] | [10.2965] | [5.4887] | [5.7443] | [6.1833] | [3.0581] | [3.5205] | [3.8947] |

| | $q(1/1000, 0.6) = 22.0358$ | | | $q(1/1000, 0.7) = 27.4140$ | | | $q(1/1000, 0.8) = 13.1358$ | | | $q(1/1000, 0.9) = 4.5695$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 24.8973 | 26.8478 | 31.6122 | 35.4726 | 37.9542 | 38.4741 | 15.0307 | 17.6032 | 18.7812 | 4.7138 | 4.6911 | 4.7025 |
| | (18.6830) | (23.4396) | (40.3262) | (42.2866) | (49.5125) | (56.0340) | (9.5039) | (17.6637) | (18.1308) | (1.6713) | (1.8586) | (2.0443) |
| | [12.7777] | [15.0691] | [19.6097] | [20.8515] | [24.0656] | [25.0722] | [6.5695] | [9.3951] | [10.5987] | [1.1836] | [1.2528] | [1.4079] |
| 400 | 21.0084 | 23.0927 | 22.8012 | 24.7805 | 27.1309 | 32.7301 | 12.8802 | 13.2370 | 13.7795 | 4.6228 | 4.5499 | 4.5391 |
| | (9.9573) | (16.5794) | (17.8647) | (12.4526) | (16.3368) | (33.0097) | (5.0678) | (5.4929) | (7.9506) | (1.0399) | (1.4602) | (1.4660) |
| | [7.1595] | [9.1380] | [9.1503] | [9.6803] | [11.6037] | [16.9023] | [4.0083] | [4.0705] | [5.2544] | [1.0226] | [1.0585] | [1.0976] |
| 600 | 20.7236 | 21.3903 | 22.7184 | 25.5869 | 24.5530 | 25.8020 | 12.8633 | 12.7666 | 12.6675 | 4.3906 | 4.3147 | 4.1995 |
| | (8.0589) | (9.5612) | (12.2271) | (10.3777) | (10.6816) | (14.2859) | (3.9845) | (4.5512) | (4.9016) | (0.7658) | (0.8289) | (0.9169) |
| | [6.1070] | [7.2747] | [8.3499] | [8.3564] | [8.5976] | [10.6484] | [3.1846] | [3.2996] | [3.4189] | [0.6193] | [0.6974] | [0.7860] |
| 800 | 21.2026 | 21.3579 | 21.7652 | 24.9908 | 25.1628 | 26.8321 | 12.8467 | 12.9886 | 12.7717 | 4.5190 | 4.4598 | 4.4162 |
| | (7.5896) | (7.6956) | (8.6432) | (9.3165) | (10.0739) | (12.0559) | (3.5802) | (4.1186) | (4.4042) | (0.6725) | (0.6914) | (0.7651) |
| | [5.6381] | [5.7519] | [6.4576] | [6.9324] | [8.0160] | [8.6939] | [2.8018] | [3.0938] | [3.2062] | [0.5373] | [0.5651] | [0.6219] |

**Table 2. Simulation results for** $q(1/1000|x)$. For each configuration of the simulation parameters $(n, c, x)$, the first line gives the averaged value of the $N = 100$ estimates of $q(1/1000|x)$. $(\cdot)$: empirical RMSE. $[\cdot]$: empirical MAE.

## 6. Appendix A: proofs

We first provide some additional notations and intermediate technical lemmas. Proofs of Proposition 2.1 and Theorem 2.2 are given in paragraph 6.2.

*6.1. Preliminary results*

Note first that $\bar{H}_n(t|x)$ and $\bar{H}_n^1(t|x)$ in (2.2) can be rewritten as $\bar{H}_n(t|x) = \hat{\varphi}_n(t,x)/\hat{g}_n(x)$ and $\bar{H}_n^1(t|x) = \hat{\psi}_n(t,x)/\hat{g}_n(x)$ respectively, where $\hat{\varphi}_n(t,x) = \frac{1}{n}\sum_{i=1}^n K_h(x-X_i)1_{\{Z_i>t\}}$, $\hat{\psi}_n(t,x) = \frac{1}{n}\sum_{i=1}^n K_h(x-X_i)1_{\{Z_i>t,\delta_i=1\}}$ and $\hat{g}_n(x) = \frac{1}{n}\sum_{i=1}^n K_h(x-X_i)$ is the classical kernel estimator of the density function $g$.
Recall that $\bar{H}(t|x) = \mathbb{P}(Z>t|X=x)$ is the conditional survival function of $Z$ given $X=x$ and let $\bar{H}^1(t|x) = \mathbb{P}(Z>t, \delta=1|X=x)$ be the sub-distribution conditional survival function of $Z$. We further define $\varphi(t,x) = \bar{H}(t|x)g(x)$, $\psi(t,x) = \bar{H}^1(t|x)g(x)$,

$$\mathbb{W}_{n,\varphi}(x) = \sqrt{nh^p\varphi(t_n,x)}\left(\frac{\hat{\varphi}_n(t_n,x) - \mathbb{E}\left[\hat{\varphi}_n(t_n,x)\right]}{\varphi(t_n,x)}\right),$$

and

$$\mathbb{W}_{n,\psi}(x) = \sqrt{nh^p\psi(t_n,x)}\left(\frac{\hat{\psi}_n(t_n,x) - \mathbb{E}\left[\hat{\psi}_n(t_n,x)\right]}{\psi(t_n,x)}\right).$$

Finally in what follows, $\top$ will denote the transpose. Lemma 6.1 investigates the limit of $\psi(t,x)/\varphi(t,x)$ as $t\to\infty$. A similar lemma is obtained by Brahimi *et al.* (2013) in the unconditional case.

**Lemma 6.1.** *Suppose that (C1) holds and let $x\in\mathcal{X}$. Then $\lim_{t\to\infty}\psi(t,x)/\varphi(t,x) = p_x$.*

**Proof.** Note first that $\psi(t,x)/\varphi(t,x) = \bar{H}^1(t|x)/\bar{H}(t|x)$. A straightforward calculation yields $\bar{H}^1(t|x) = \int_t^\infty \bar{G}(u|x)dF(u|x)$. The change of variable $\bar{F}(u|x) = 1/v$ yields

$$\begin{aligned}\bar{H}^1(t|x) &= \int_{1/\bar{F}(t|x)}^\infty v^{-2}\bar{G}\left(F^{\leftarrow}(1-1/v|x)\,|x\right)dv\\ &:= \int_{1/\bar{F}(t|x)}^\infty v^{-1}\bar{R}(v|x)dv,\end{aligned}$$

where $v\mapsto\bar{R}(v|x)$ is a regularly varying function with index $-(1+\gamma_1(x)/\gamma_2(x))$. By Theorem 1.2.2 of de Haan and Ferreira (2006),

$$\bar{H}^1\left(t|x\right) = \int_{1/\bar{F}(t|x)}^\infty v^{-1}\bar{R}(v|x)dv \sim \left(1+\frac{\gamma_1(x)}{\gamma_2(x)}\right)^{-1}\bar{R}(1/\bar{F}(t|x)|x)$$

when $t$ is large. Now, $\bar{R}(1/\bar{F}(t|x)|x) = \bar{F}(t|x)\bar{G}(t|x) = \bar{H}(t|x)$ thus

$$\bar{H}^1(t|x) \sim \left(1+\frac{\gamma_1(x)}{\gamma_2(x)}\right)^{-1}\bar{H}\left(t|x\right)$$

when $t$ is large, which implies that $\lim_{t\to\infty}\bar{H}^1\left(t|x\right)/\bar{H}\left(t|x\right) = \gamma_2(x)/(\gamma_1(x)+\gamma_2(x))$, that is, $\lim_{t\to\infty}\psi(t,x)/\varphi(t,x) = p_x$. □

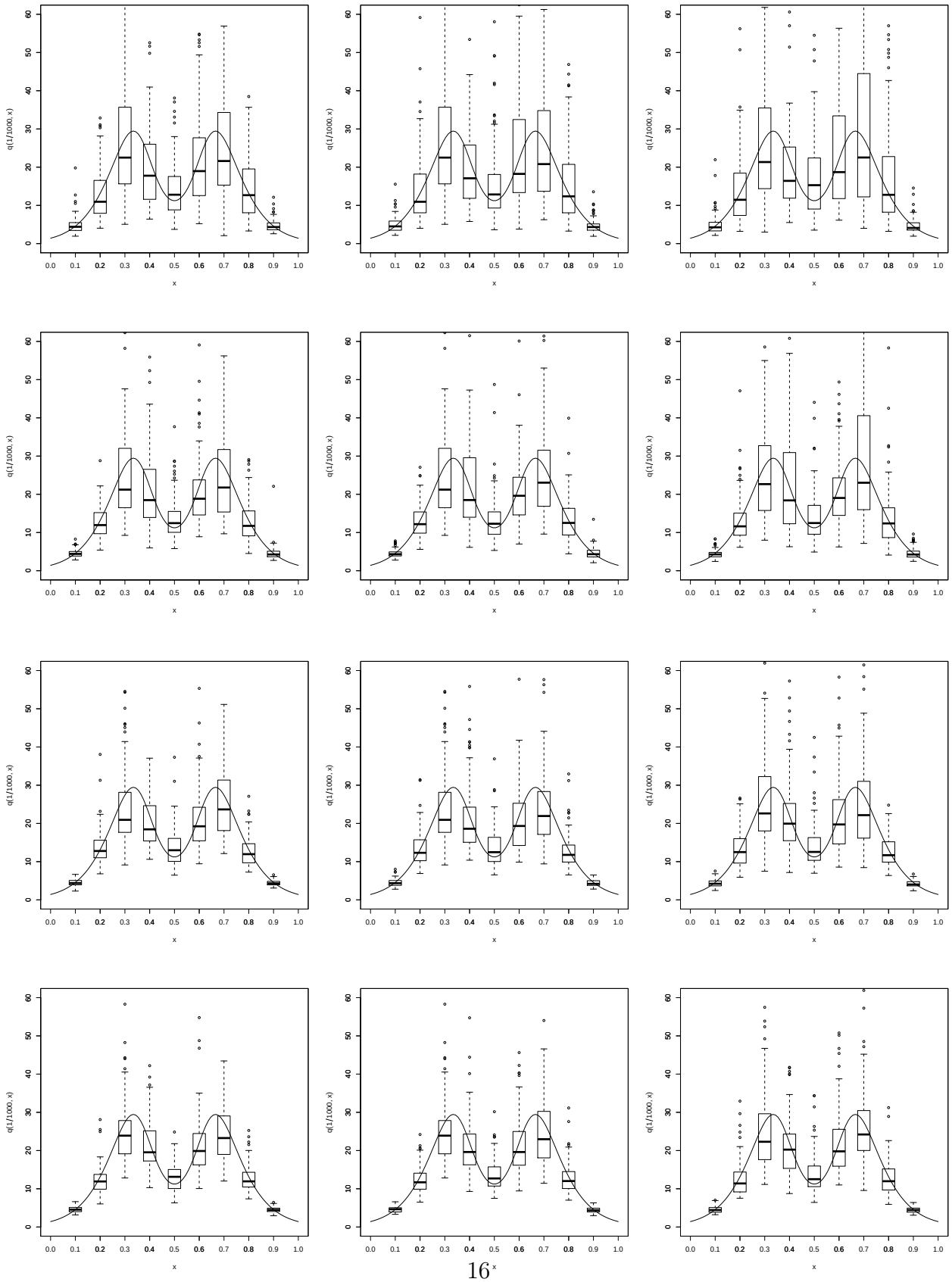




Figure 3: Boxplots of the $N$ estimates of $q(1/1000|x)$ for $n = 200$ (1st line), $n = 400$ (2nd line), $n = 600$ (3rd line), $n = 800$ (4th line). Left: $c = 10\%$, center: $c = 25\%$, right: $c = 40\%$.

**Lemma 6.2.** *Suppose (C1), (C2) and (C3) hold. If $t_n \to \infty$ and $h \log t_n \to 0$ as $n \to \infty$, then*

$$\sup_{d(x,x')\leq h} \left| \frac{\bar{H}(t_n|x)}{\bar{H}(t_n|x')} - 1 \right| = O(h \log t_n).$$

**Proof.** The proof is similar to proof of Lemma 1 in Daouia *et al.* (2011) and is therefore omitted. □

The next lemma is dedicated to the asymptotic properties of $\hat{\varphi}_n(t,x)$ and $\hat{\psi}_n(t,x)$ .

**Lemma 6.3.** *Suppose (C1)-(C6) hold. Let $t_n$ be such that $t_n \to \infty$, $h \log t_n \to 0$ and $nh^p \bar{H}(t_n|x) \to \infty$ as $n \to \infty$. Then, for all $x \in \mathcal{X}$ such that $g(x) > 0$,*

1. $\mathbb{E}\left[\hat{\varphi}_n(t_n,x)\right] = \varphi(t_n,x)(1+O(h\log t_n))$ *and* $\mathbb{E}[\hat{\psi}_n(t_n,x)] = \psi(t_n,x)(1+O(h\log t_n))$.
2. *As $n \to \infty$, $\mathbb{W}_{n,\varphi}(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \|K\|_2^2)$ and $\mathbb{W}_{n,\psi}(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \|K\|_2^2)$. Moreover as $n \to \infty$, $\mathbb{W}_n(x) := (\mathbb{W}_{n,\varphi}(x), \mathbb{W}_{n,\psi}(x))^\top$ converges in distribution to a bivariate Gaussian vector $\mathcal{N}(0,M)$, where*

$$M := \begin{pmatrix} \|K\|_2^2 & \|K\|_2^2\sqrt{p_x} \\ \|K\|_2^2\sqrt{p_x} & \|K\|_2^2 \end{pmatrix}.$$

**Proof.**

1. We prove that $\mathbb{E}\left[\hat{\varphi}_n(t_n,x)\right] = \varphi(t_n,x)(1+O(h\log t_n))$ (the proof is similar for $\hat{\psi}_n(t_n,x)$ and is thus omitted). The observations $(Z_i, X_i, \delta_i)$, $i = 1, ..., n$ are identically distributed, thus

$$\begin{aligned} \mathbb{E}\left[\hat{\varphi}_n(t_n,x)\right] &= \int_{\mathbb{R}^p} K_h(x-s)\bar{H}(t_n|s)g(s)ds \\ &= \int_S K(u)\bar{H}(t_n|x-hu)g(x-hu)du \end{aligned}$$

under (C6). It follows that

$$\begin{aligned} \left|\mathbb{E}\left[\hat{\varphi}_n(t_n,x)\right] - \varphi(t_n,x)\right| &\leq \bar{H}(t_n|x)\int_S K(u)\left|g(x-hu)-g(x)\right|du \\ &\quad + \bar{H}(t_n|x)\int_S K(u)\left|\frac{\bar{H}(t_n|x-hu)}{\bar{H}(t_n|x)} - 1\right| g(x-hu)du \\ &:= A_{1,x} + A_{2,x}. \end{aligned}$$

Under $(C4)$ and since $g(x) > 0$, we have

$$A_{1,x} \leq \bar{H}(t_n|x)c_g h \int_S d(u,0)K(u)du = \varphi(t_n,x)O(h). \tag{6.10}$$

By Lemma 6.2, we have

$$\sup_{u\in S}\left|\frac{\bar{H}(t_n|x-hu)}{\bar{H}(t_n|x)}-1\right|=O(h\log t_n),$$

thus

$$\begin{aligned}A_{2,x} &= \bar{H}(t_n|x)O(h\log t_n)\int_S K(u)g(x-hu)du\\ &= \varphi(t_n,x)O(h\log t_n).\end{aligned}\tag{6.11}$$

Combining (6.10) and (6.11) concludes the proof.

2. We first prove asymptotic normality of $\mathbb{W}_{n,\varphi}(x)$ (the proof for $\mathbb{W}_{n,\psi}(x)$ is similar and is thus omitted). Let

$$\mathbb{W}_{n,\varphi}(x):=\frac{\hat{\varphi}_n(t_n,x)-\mathbb{E}\left[\hat{\varphi}_n(t_n,x)\right]}{\varphi(t_n,x)\Lambda_n(x)}$$

where $\Lambda_n(x)=(nh^p\varphi(t_n,x))^{-1/2}$. Straightforward calculations allow to rewrite $\mathbb{W}_{n,\varphi}(x)$ as:

$$\begin{aligned}\mathbb{W}_{n,\varphi}(x) &= \frac{1}{\varphi(t_n,x)\Lambda_n(x)}\frac{1}{n}\sum_{i=1}^n\left\{K_h(x-X_i)1_{\{Z_i>t_n\}}-\mathbb{E}\left[K_h(x-X)1_{\{Z>t_n\}}\right]\right\}\\ &:= \frac{1}{\varphi(t_n,x)\Lambda_n(x)n}\sum_{i=1}^n T_{i,n},\end{aligned}$$

that is, as the row sum of the triangular array of random variables $(T_{i,n}, i=1,\ldots,n)$, $n\geq 1$, where for fixed $n$, $T_{i,n}$, $i=1,\ldots,n$ are independent and identically distributed zero-mean random variables. Thus, to establish the asymptotic normality of $\mathbb{W}_{n,\varphi}(x)$, we verify Lyapunov's condition for triangular arrays of random variables. As a preliminary step, we first calculate the variance of $\mathbb{W}_{n,\varphi}(x)$. We have:

$$\begin{aligned}\mathrm{var}(T_{i,n}) &= \mathrm{var}\left(K_h(x-X_i)1_{\{Z_i>t_n\}}\right)\\ &= \mathbb{E}\left[K_h^2(x-X_i)1_{\{Z_i>t_n\}}\right]-\mathbb{E}\left[K_h(x-X_i)1_{\{Z_i>t_n\}}\right]^2.\end{aligned}$$

From Lemma 2 of Goegebeur *et al.* (2014b), $\mathbb{E}\left[K_h^2(x-X_i)1_{\{Z_i>t_n\}}\right]=\varphi(t_n,x)h^{-p}\|K\|_2^2(1+O(h\log t_n))$ and from statement 1 in Lemma 6.3, $\mathbb{E}\left[K_h(x-X_i)1_{\{Z_i>t_n\}}\right]=\varphi(t_n,x)(1+O(h\log t_n))$. It follows that

$$\mathrm{var}(T_{i,n})=\varphi(t_n,x)h^{-p}\|K\|_2^2(1+O(h\log t_n))-\varphi^2(t_n,x)(1+O(h\log t_n))^2\tag{6.12}$$

and thus,

$$\begin{aligned}\mathrm{var}(\mathbb{W}_{n,\varphi}(x)) &= \frac{1}{\varphi^2(t_n,x)\Lambda_n^2(x)n^2}n\mathrm{var}(T_{i,n})\\ &= \|K\|_2^2(1+O(h\log t_n))-h^p\varphi(t_n,x)(1+O(h\log t_n))^2.\end{aligned}$$

From this, we deduce that $\mathrm{var}(\mathbb{W}_{n,\varphi}(x)) \to \|K\|_2^2$ as $n \to \infty$. Now, we prove that Lyapunov's condition for triangular arrays is satisfied, namely we show that $\sum_{i=1}^n \mathbb{E}|\widetilde{T}_{i,n}|^3 \to 0$ as $n \to \infty$, where $\widetilde{T}_{i,n} := T_{i,n}/\left\{\varphi(t_n,x)\Lambda_n(x)n\right\}$. It is straightforward to prove that

$$|\widetilde{T}_{i,n}| \leq \frac{2\|K\|_\infty}{\sqrt{nh^p\varphi(t_n,x)}},$$

where $\|\cdot\|_\infty$ denotes the supremum norm and thus $|\widetilde{T}_{i,n}|^3 \leq \frac{2\|K\|_\infty}{\sqrt{nh^p\varphi(t_n,x)}}|\widetilde{T}_{i,n}|^2$. Taking expectation on both sides of this inequality yields

$$\mathbb{E}|\widetilde{T}_{i,n}|^3 \leq \frac{2\|K\|_\infty h^{2p}}{(nh^p\varphi(t_n,x))^{3/2}}\mathrm{var}(T_{i,n}),$$

and using (6.12), we obtain that

$$\sum_{i=1}^n \mathbb{E}|\widetilde{T}_{i,n}|^3 \leq \frac{2\|K\|_\infty\|K\|_2^2}{\sqrt{nh^p\varphi(t_n,x)}}(1+o(1)) - \frac{2\|K\|_\infty h^p\varphi(t_n,x)}{\sqrt{nh^p\varphi(t_n,x)}}(1+o(1)) \to 0$$

as $n \to \infty$. Lyapunov's condition is verified, thus $\mathbb{W}_{n,\varphi}(x)/\sqrt{\mathrm{var}(\mathbb{W}_{n,\varphi}(x))} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$. Finally, $\mathbb{W}_{n,\varphi}(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\|K\|_2^2)$.
We now prove that $\mathbb{W}_n(x) := (\mathbb{W}_{n,\varphi}(x), \mathbb{W}_{n,\psi}(x))^\top$ converges in distribution to $\mathcal{N}(0,M)$. According to Cramér-Wold device (*e.g.*, van der Vaart (1998)), it is sufficient to prove that $\ell^\top\mathbb{W}_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ell^\top M\ell)$ for all $\ell = (\ell_1,\ell_2)^\top \in \mathbb{R}^2$, $\ell \neq 0$. Some simple algebra yields:

$$\ell^\top\mathbb{W}_n(x) := \frac{1}{\varphi(t_n,x)\Lambda_n(x)n}\sum_{i=1}^n T_{i,n}^*,$$

where

$$\begin{aligned} T_{i,n}^* = K_h\left(x-X_i\right)&\left\{\ell_1 1_{\{Z_i>t_n\}} + \ell_2 1_{\{Z_i>t_n,\delta_i=1\}}\sqrt{\frac{\varphi(t_n,x)}{\psi(t_n,x)}}\right\} \\ &-\mathbb{E}\left[K_h\left(x-X\right)\left\{\ell_1 1_{\{Z>t_n\}} + \ell_2 1_{\{Z>t_n,\delta=1\}}\sqrt{\frac{\varphi(t_n,x)}{\psi(t_n,x)}}\right\}\right]. \end{aligned}$$

Similar calculations as for $\mathrm{var}(T_{i,n})$ yield

$$\begin{aligned} \mathrm{var}(T_{i,n}^*) = l_1^2&\left(\varphi(t_n,x)h^{-p}\|K\|_2^2(1+O(h\log t_n)) - \varphi^2(t_n,x)(1+O(h\log t_n))^2\right) \\ &+l_2^2\left(p_x^{-1}+o(1)\right)\left(\psi(t_n,x)h^{-p}\|K\|_2^2(1+O(h\log t_n))\right. \\ &\qquad\left.-\psi^2(t_n,x)(1+O(h\log t_n))^2\right) \\ &+2l_1l_2\left(p_x^{-1/2}+o(1)\right)\left(\psi(t_n,x)h^{-p}\|K\|_2^2(1+O(h\log t_n))\right. \\ &\qquad\left.-\varphi(t_n,x)\psi(t_n,x)(1+O(h\log t_n))^2\right), \end{aligned}$$

and thus

$$\begin{aligned}\lim_{n\to\infty} \mathrm{var}(\ell^\top \mathbb{W}_n(x)) &= \lim_{n\to\infty} \frac{1}{h^{-p}\varphi(t_n,x)} \mathrm{var}(T^*_{i,n}) \\ &= \ell_1^2 \|K\|_2^2 + \ell_2^2 \|K\|_2^2 + 2\ell_1\ell_2 \|K\|_2^2 \sqrt{p_x} \\ &= \ell^\top M \ell.\end{aligned}$$

Asymptotic normality of $\ell^\top \mathbb{W}_n(x)$ follows from Lyapunov's condition: $\lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}|T^*_{i,n}/\{\varphi(t_n,x)\Lambda_n(x)n\}|^3 = 0$ (calculations are similar as above and are omitted for conciseness). Thus, for all $\ell \in \mathbb{R}^2$, $\ell \neq 0$, $\ell^\top \mathbb{W}_n(x)$ converges in distribution to the univariate normal distribution $\mathcal{N}(0, \ell^\top M \ell)$. Cramér-Wold device finally implies that $\mathbb{W}_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, M)$. □

*6.2. Proofs of main results*

**Proof of Proposition 2.1.** We decompose

$$\begin{aligned}\sqrt{nh^p \bar{H}(t_n|x)}(\widehat{p}_{t_n}(x) - p_x) = \sqrt{nh^p \bar{H}(t_n|x)} \left(\widehat{p}_{t_n}(x) - \frac{\psi(t_n,x)}{\varphi(t_n,x)}\right) \\ + \sqrt{nh^p \bar{H}(t_n|x)} \left(\frac{\psi(t_n,x)}{\varphi(t_n,x)} - p_x\right).\end{aligned}$$

We first prove that the first term in this sum is asymptotically normal. This follows from an application of the Delta method. We have:

$$\begin{aligned}\sqrt{nh^p \bar{H}(t_n|x)} \begin{pmatrix} \frac{\hat{\psi}_n(t_n,x)}{\varphi(t_n,x)} - \frac{\psi(t_n,x)}{\varphi(t_n,x)} \\ \frac{\hat{\varphi}_n(t_n,x)}{\varphi(t_n,x)} - 1 \end{pmatrix} &= \begin{pmatrix} \frac{1}{\sqrt{g(x)}} \sqrt{\frac{\psi(t_n,x)}{\varphi(t_n,x)}} \mathbb{W}_{n,\psi}(x) \\ \frac{1}{\sqrt{g(x)}} \mathbb{W}_{n,\varphi}(x) \end{pmatrix} \\ &+ \begin{pmatrix} \frac{1}{\sqrt{g(x)}} \sqrt{\frac{\psi(t_n,x)}{\varphi(t_n,x)}} \sqrt{nh^p \psi(t_n,x)} \left(\frac{\mathbb{E}[\hat{\psi}_n(t_n,x)]}{\psi(t_n,x)} - 1\right) \\ \frac{1}{\sqrt{g(x)}} \sqrt{nh^p \varphi(t_n,x)} \left(\frac{\mathbb{E}[\hat{\varphi}_n(t_n,x)]}{\varphi(t_n,x)} - 1\right) \end{pmatrix}.\end{aligned}$$

As $n \to \infty$, $\psi(t_n,x)/\varphi(t_n,x) \to p_x$ by Lemma 6.1. Moreover, by statement 1 in Lemma 6.3, $\mathbb{E}[\hat{\psi}_n(t_n,x)]/\psi(t_n,x) - 1 = O(h \log t_n)$ and $\mathbb{E}[\hat{\varphi}_n(t_n,x)]/\varphi(t_n,x) - 1 = O(h \log t_n)$ (note that under the conditions of Proposition 2.1, the condition $h \log t_n \to 0$ of Lemma 6.3 is satisfied). Finally by assumption, $nh^{p+2}\bar{H}(t_n|x)(\log t_n)^2 \to 0$ as $n \to \infty$. If follows that

$$\sqrt{nh^p \bar{H}(t_n|x)} \begin{pmatrix} \frac{\hat{\psi}_n(t_n,x)}{\varphi(t_n,x)} - \frac{\psi(t_n,x)}{\varphi(t_n,x)} \\ \frac{\hat{\varphi}_n(t_n,x)}{\varphi(t_n,x)} - 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{g(x)}} \sqrt{\frac{\psi(t_n,x)}{\varphi(t_n,x)}} \mathbb{W}_{n,\psi}(x) \\ \frac{1}{\sqrt{g(x)}} \mathbb{W}_{n,\varphi}(x) \end{pmatrix} + o(1).$$

By statement 2 in Lemma 6.3, this converges in distribution to a bivariate Gaussian vector $\mathcal{N}(0, \Sigma)$, with

$$\Sigma := \begin{pmatrix} \frac{\|K\|_2^2 p_x}{g(x)} & \frac{\|K\|_2^2 p_x}{g(x)} \\ \frac{\|K\|_2^2 p_x}{g(x)} & \frac{\|K\|_2^2}{g(x)} \end{pmatrix}.$$

Consider the function $\phi(x,y) = \frac{x}{y}$. The map $\phi$ is differentiable at $(x,y)$ (if $y \neq 0$), with derivative $\phi'(x,y) = (\frac{1}{y}, -\frac{x}{y^2})$. It follows from the Delta method (van der Vaart, 1998) that

$$\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{p}_{t_n}(x) - \frac{\psi(t_n,x)}{\varphi(t_n,x)}\right) = \sqrt{nh^p\bar{H}(t_n|x)}\left(\phi\left(\frac{\hat{\psi}_n(t_n,x)}{\varphi(t_n,x)}, \frac{\hat{\varphi}_n(t_n,x)}{\varphi(t_n,x)}\right) - \phi\left(\frac{\psi(t_n,x)}{\varphi(t_n,x)}, 1\right)\right)$$

converges in distribution to $\mathcal{N}(0, \phi'(p_x,1)\Sigma(\phi'(p_x,1))^\top)$, that is, to the normal distribution $\mathcal{N}\left(0, p_x(1-p_x)\|K\|_2^2/g(x)\right)$. Using similar arguments as Beirlant *et al.* (2007), one shows that $\sqrt{nh^p\bar{H}(t_n|x)}\left(\frac{\psi(t_n,x)}{\varphi(t_n,x)} - p_x\right) = o(1)$, which concludes the proof. □

**Proof of Theorem 2.2.** First, we decompose $\sqrt{nh^p\bar{H}(t_n|x)}(\widehat{\gamma}_{t_n}^{(c,H)}(x) - \gamma_1(x))$ as:

$$\begin{aligned}
\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{\gamma}_{t_n}^{(c,H)}(x) - \gamma_1(x)\right) &= \frac{1}{\widehat{p}_{t_n}(x)}\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{\gamma}_{t_n}^{H}(x) - \gamma(x)\right) \\
&\quad - \frac{\gamma_1(x)}{\widehat{p}_{t_n}(x)}\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{p}_{t_n}(x) - p_x\right) \\
&= \frac{1}{p_x}\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{\gamma}_{t_n}^{H}(x) - \gamma(x)\right) \\
&\quad - \frac{\gamma_1(x)}{p_x}\sqrt{nh^p\bar{H}(t_n|x)}\left(\widehat{p}_{t_n}(x) - p_x\right) + o_{\mathbb{P}}(1) \\
&:= \frac{1}{p_x}P_{n,x} - \frac{\gamma_1(x)}{p_x}R_{n,x} + o_{\mathbb{P}}(1). \qquad (6.13)
\end{aligned}$$

Then, some simple but tedious algebra yields

$$\begin{aligned}
P_{n,x} &= \sqrt{\frac{nh^p}{g(x)\varphi(t_n,x)}}\left(\frac{1}{n}\sum_{i=1}^n K_h(x-X_i)\log\left(\frac{Z_i}{t_n}\right)1_{\{Z_i>t_n\}} - \int_{t_n}^\infty \frac{\varphi(z,x)}{z}dz\right) \\
&\quad -\gamma(x)\sqrt{\frac{nh^p}{g(x)\varphi(t_n,x)}}\left(\hat{\varphi}_n(t_n,x) - \varphi(t_n,x)\right) + o_{\mathbb{P}}(1), \\
R_{n,x} &= \sqrt{\frac{nh^p}{g(x)\varphi(t_n,x)}}\left(\hat{\psi}_n(t_n,x) - \psi(t_n,x)\right) - p_x\sqrt{\frac{nh^p}{g(x)\varphi(t_n,x)}}\left(\hat{\varphi}_n(t_n,x) - \varphi(t_n,x)\right) + o_{\mathbb{P}}(1).
\end{aligned}$$

Therefore, asymptotic normality of $\widehat{\gamma}_{t_n}^{(c,H)}(x)$ will be proved if we can establish the asymptotic normality of the random vector

$$\mathbb{X}_n(x) = \sqrt{\frac{nh^p}{g(x)\varphi(t_n,x)}}\begin{pmatrix} \hat{\varphi}_n(t_n,x) - \varphi(t_n,x) \\ \hat{\psi}_n(t_n,x) - \psi(t_n,x) \\ \frac{1}{n}\sum_{i=1}^n K_h(x-X_i)\log\left(\frac{Z_i}{t_n}\right)1_{\{Z_i>t_n\}} - \int_{t_n}^\infty \frac{\varphi(z,x)}{z}dz \end{pmatrix},$$

whereafter the result will follow by a straightforward application of the Delta method. The asymptotic normality of $\mathbb{X}_n(x)$ is proved by using Cramér-Wold device. Let $\ell = (\ell_1, \ell_2, \ell_3)^\top \in \mathbb{R}^3$, $\ell \neq 0$. Then

$$\ell^\top \mathbb{X}_n(x) := \sqrt{\frac{nh^p}{g(x)\varphi(t_n, x)}} \frac{1}{n} \sum_{i=1}^{n} T_{i,n}^{\dagger},$$

where

$$\begin{aligned} T_{i,n}^{\dagger} &= l_1 K_h(x - X_i) 1_{\{Z_i > t_n\}} + l_2 K_h(x - X_i) 1_{\{Z_i > t_n, \delta_i = 1\}} + l_3 K_h(x - X_i) \log\left(\frac{Z_i}{t_n}\right) 1_{\{Z_i > t_n\}} \\ &\quad - \left( l_1 \varphi(t_n, x) + l_2 \psi(t_n, x) + l_3 \int_{t_n}^{\infty} \frac{\varphi(z, x)}{z} dz \right). \end{aligned}$$

From Lemma 2 of Goegebeur *et al.* (2014b), it holds that:

$$\begin{aligned} \mathbb{E}\left[K_h(x - X_i) \log(Z_i/t_n) 1_{\{Z_i > t_n\}}\right] &= \gamma(x)\varphi(t_n, x)(1 + O(h \log t_n)), \\ \mathbb{E}\left[K_h^2(x - X_i) \log(Z_i/t_n) 1_{\{Z_i > t_n\}}\right] &= \gamma(x)\varphi(t_n, x) h^{-p} \|K\|_2^2 (1 + O(h \log t_n)) \end{aligned}$$

and

$$\mathbb{E}\left[K_h^2(x - X_i) \log^2(Z_i/t_n) 1_{\{Z_i > t_n\}}\right] = 2\gamma^2(x)\varphi(t_n, x) h^{-p} \|K\|_2^2 (1 + O(h \log t_n)).$$

Using this, lengthy but simple calculations yield

$$\begin{aligned} \mathrm{var}(\ell^\top \mathbb{X}_n(x)) &= \frac{h^p}{g(x)\varphi(t_n, x)} \mathrm{var}(T_{i,n}^{\dagger}) \\ &= \ell_1^2 \frac{\|K\|_2^2}{g(x)} + \ell_2^2 \frac{p_x \|K\|_2^2}{g(x)} + \ell_3^2 \frac{2\gamma^2(x)\|K\|_2^2}{g(x)} + 2\ell_1\ell_2 \frac{p_x \|K\|_2^2}{g(x)} \\ &\quad + 2\ell_1\ell_3 \frac{\gamma(x)\|K\|_2^2}{g(x)} + 2\ell_2\ell_3 \frac{\gamma(x) p_x \|K\|_2^2}{g(x)} + o(1) \\ &= \ell^\top \Gamma \ell + o(1), \end{aligned}$$

where

$$\Gamma := \begin{pmatrix} \frac{\|K\|_2^2}{g(x)} & \frac{\|K\|_2^2 p_x}{g(x)} & \frac{\|K\|_2^2 \gamma(x)}{g(x)} \\ \frac{\|K\|_2^2 p_x}{g(x)} & \frac{\|K\|_2^2 p_x}{g(x)} & \frac{\|K\|_2^2 p_x \gamma(x)}{g(x)} \\ \frac{\|K\|_2^2 \gamma(x)}{g(x)} & \frac{\|K\|_2^2 p_x \gamma(x)}{g(x)} & \frac{2\|K\|_2^2 \gamma^2(x)}{g(x)} \end{pmatrix}.$$

Now, to establish the asymptotic normality of $\ell^\top \mathbb{X}_n(x)$, we verify Lyapounov's criterion for triangular arrays of random variables. In the present context, this consists in proving that $\lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}|T_{i,n}^{\dagger} / \{ng(x)h^{-p}\varphi(t_n, x)\}^{1/2}|^3 = 0$ (calculations are similar as in proof of Lemma 6.3 and are thus omitted). If follows that for all $\ell = (\ell_1, \ell_2, \ell_3)^\top \in \mathbb{R}^3$, $\ell \neq 0$, $\ell^\top \mathbb{X}_n(x)$ converges in distribution to $\mathcal{N}(0, \ell^\top \Gamma \ell)$. By Cramér-Wold device, $\mathbb{X}_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Gamma)$. Finally, based on the decomposition (6.13), on the expressions of $P_{n,x}$ and $R_{n,x}$ and on the asymptotic normality of $\mathbb{X}_n(x)$, a straightforward application of the Delta method completes the proof of Theorem 2.2. □

## 7. Appendix B: simulation results for the complete-case and censoring-ignored strategies

| | $\gamma_1(0.1) = 0.2199$ | | | $\gamma_1(0.2) = 0.3728$ | | | $\gamma_1(0.3) = 0.4793$ | | | $\gamma_1(0.4) = 0.4477$ | | | $\gamma_1(0.5) = 0.35$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 0.2014 | 0.1719 | 0.1556 | 0.3181 | 0.2720 | 0.2099 | 0.3799 | 0.3107 | 0.2453 | 0.4112 | 0.2953 | 0.2229 | 0.3689 | 0.2828 | 0.2208 |
| | (0.2137) | (0.2344) | (0.2993) | (0.2939) | (0.3251) | (0.4053) | (0.3262) | (0.4122) | (0.4830) | (0.3130) | (0.4716) | (0.4741) | (0.2556) | (0.2756) | (0.3625) |
| | [0.0034] | [0.0041] | [0.0069] | [0.0104] | [0.0141] | [0.0309] | [0.0153] | [0.0390] | [0.0613] | [0.0134] | [0.0447] | [0.0548] | [0.0060] | [0.0080] | [0.0203] |
| 400 | 0.2060 | 0.1817 | 0.1434 | 0.3220 | 0.2680 | 0.2149 | 0.3859 | 0.3107 | 0.2507 | 0.3717 | 0.2996 | 0.2314 | 0.3234 | 0.2672 | 0.2146 |
| | (0.1648) | (0.2107) | (0.2767) | (0.2518) | (0.3248) | (0.3974) | (0.3085) | (0.4106) | (0.4782) | (0.2848) | (0.3850) | (0.4651) | (0.2275) | (0.2905) | (0.3579) |
| | [0.0012] | [0.0025] | [0.0064] | [0.0054] | [0.0131] | [0.0265] | [0.0118] | [0.0309] | [0.0543] | [0.0088] | [0.0243] | [0.0487] | [0.0026] | [0.0078] | [0.0199] |
| 600 | 0.2008 | 0.1836 | 0.1477 | 0.3209 | 0.2661 | 0.2041 | 0.3981 | 0.3082 | 0.2485 | 0.3817 | 0.3044 | 0.2342 | 0.3231 | 0.2702 | 0.2124 |
| | (0.1623) | (0.1973) | (0.2689) | (0.2340) | (0.3237) | (0.3907) | (0.2884) | (0.4037) | (0.4605) | (0.2656) | (0.3786) | (0.4621) | (0.2013) | (0.2842) | (0.3510) |
| | [0.0011] | [0.0020] | [0.0057] | [0.0038] | [0.0126] | [0.0253] | [0.0089] | [0.0280] | [0.0510] | [0.0065] | [0.0226] | [0.0465] | [0.0023] | [0.0075] | [0.0198] |
| 800 | 0.2039 | 0.1753 | 0.1446 | 0.3186 | 0.2690 | 0.2086 | 0.3863 | 0.3218 | 0.2379 | 0.3735 | 0.3091 | 0.2338 | 0.3267 | 0.2697 | 0.2099 |
| | (0.1564) | (0.1870) | (0.2545) | (0.2338) | (0.3223) | (0.3752) | (0.2759) | (0.3969) | (0.4514) | (0.2639) | (0.3723) | (0.4525) | (0.1967) | (0.2839) | (0.3443) |
| | [0.0008] | [0.0018] | [0.0052] | [0.0037] | [0.0116] | [0.0228] | [0.0073] | [0.0261] | [0.0491] | [0.0062] | [0.0205] | [0.0447] | [0.0021] | [0.0072] | [0.0202] |

| | $\gamma_1(0.6) = 0.4477$ | | | $\gamma_1(0.7) = 0.4793$ | | | $\gamma_1(0.8) = 0.3728$ | | | $\gamma_1(0.9) = 0.2199$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 0.3698 | 0.2973 | 0.2272 | 0.3902 | 0.3245 | 0.2370 | 0.3126 | 0.2594 | 0.2144 | 0.2091 | 0.1806 | 0.1492 |
| | (0.3029) | (0.3900) | (0.4707) | (0.3265) | (0.4066) | (0.4922) | (0.2681) | (0.3394) | (0.3994) | (0.2111) | (0.2194) | (0.2735) |
| | [0.0121] | [0.0284] | [0.0534] | [0.0169] | [0.0309] | [0.0636] | [0.0072] | [0.0161] | [0.0289] | [0.0031] | [0.0034] | [0.0071] |
| 400 | 0.3685 | 0.3127 | 0.2289 | 0.3801 | 0.3153 | 0.2408 | 0.3129 | 0.2569 | 0.2072 | 0.1988 | 0.1754 | 0.1507 |
| | (0.2936) | (0.3688) | (0.4678) | (0.3189) | (0.4059) | (0.4884) | (0.2593) | (0.3305) | (0.3969) | (0.1897) | (0.2124) | (0.2631) |
| | [0.0092] | [0.0259] | [0.0497] | [0.0129] | [0.0304] | [0.0594] | [0.0061] | [0.0149] | [0.0282] | [0.0023] | [0.0029] | [0.0064] |
| 600 | 0.3734 | 0.3045 | 0.2260 | 0.3916 | 0.3127 | 0.2417 | 0.3153 | 0.2701 | 0.2084 | 0.2024 | 0.1745 | 0.1464 |
| | (0.2753) | (0.3584) | (0.4608) | (0.2982) | (0.4042) | (0.4875) | (0.2446) | (0.3206) | (0.3954) | (0.1547) | (0.2051) | (0.2613) |
| | [0.0074] | [0.0220] | [0.0491] | [0.0096] | [0.0292] | [0.0579] | [0.0046] | [0.0116] | [0.0277] | [0.0018] | [0.0027] | [0.0059] |
| 800 | 0.3749 | 0.3010 | 0.2313 | 0.3894 | 0.3167 | 0.2412 | 0.3231 | 0.2684 | 0.2128 | 0.2023 | 0.1742 | 0.1487 |
| | (0.2723) | (0.3430) | (0.4552) | (0.2905) | (0.4033) | (0.4863) | (0.2358) | (0.3132) | (0.3814) | (0.1453) | (0.1958) | (0.2579) |
| | [0.0070] | [0.0186] | [0.0483] | [0.0095] | [0.0280] | [0.0576] | [0.0035] | [0.0112] | [0.0263] | [0.0008] | [0.0025] | [0.0055] |

**Table 3. Simulation results for $\gamma_1(x)$: complete-case procedure**. For each configuration of the simulation parameters $(n, c, x)$, the first line gives the averaged value of the $N = 100$ CC-estimates of $\gamma_1(x)$. $(\cdot)$: empirical RMSE. $[\cdot]$: empirical MAE.

| | $\gamma_1(0.1) = 0.2199$ | | | $\gamma_1(0.2) = 0.3728$ | | | $\gamma_1(0.3) = 0.4793$ | | | $\gamma_1(0.4) = 0.4477$ | | | $\gamma_1(0.5) = 0.35$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 0.2096 | 0.1794 | 0.1510 | 0.3094 | 0.2632 | 0.2106 | 0.3800 | 0.3097 | 0.2372 | 0.3765 | 0.3063 | 0.2355 | 0.3236 | 0.2664 | 0.2054 |
| | (0.2118) | (0.2192) | (0.2644) | (0.2686) | (0.3311) | (0.4128) | (0.3433) | (0.4321) | (0.5021) | (0.2863) | (0.3760) | (0.4647) | (0.2455) | (0.2978) | (0.3803) |
| | [0.0041] | [0.0050] | [0.0058] | [0.0071] | [0.0141] | [0.0288] | [0.0151] | [0.0421] | [0.0640] | [0.0093] | [0.0240] | [0.0476] | [0.0046] | [0.0098] | [0.0227] |
| 400 | 0.2071 | 0.1828 | 0.1527 | 0.3199 | 0.2665 | 0.2078 | 0.3727 | 0.3000 | 0.2302 | 0.3718 | 0.3032 | 0.2330 | 0.3269 | 0.2724 | 0.2175 |
| | (0.1563) | (0.2092) | (0.2594) | (0.2431) | (0.3261) | (0.4063) | (0.3278) | (0.4235) | (0.4991) | (0.2845) | (0.3702) | (0.4634) | (0.2182) | (0.2807) | (0.3640) |
| | [0.0015] | [0.0026] | [0.0050] | [0.0046] | [0.0125] | [0.0279] | [0.0140] | [0.0339] | [0.0633] | [0.0088] | [0.0231] | [0.0473] | [0.0042] | [0.0073] | [0.0183] |
| 600 | 0.2038 | 0.1788 | 0.1480 | 0.3181 | 0.2642 | 0.2057 | 0.3988 | 0.3201 | 0.2425 | 0.3623 | 0.2974 | 0.2287 | 0.3217 | 0.2705 | 0.2107 |
| | (0.1517) | (0.2067) | (0.2483) | (0.2377) | (0.3196) | (0.4048) | (0.2873) | (0.3990) | (0.4866) | (0.2740) | (0.3678) | (0.3686) | (0.1965) | (0.2719) | (0.3533) |
| | [0.0010] | [0.0023] | [0.0045] | [0.0040] | [0.0122] | [0.0274] | [0.0082] | [0.0265] | [0.0568] | [0.0078] | [0.0227] | [0.0386] | [0.0026] | [0.0070] | [0.0179] |
| 800 | 0.2019 | 0.1790 | 0.1504 | 0.3192 | 0.2676 | 0.2098 | 0.3911 | 0.3184 | 0.2430 | 0.3734 | 0.3033 | 0.2342 | 0.3245 | 0.2699 | 0.2127 |
| | (0.1431) | (0.2033) | (0.2137) | (0.2356) | (0.3144) | (0.4037) | (0.2783) | (0.3012) | (0.3861) | (0.2728) | (0.3400) | (0.3621) | (0.1963) | (0.2630) | (0.3406) |
| | [0.0008] | [0.0021] | [0.0031] | [0.0039] | [0.0118] | [0.0270] | [0.0065] | [0.0240] | [0.0564] | [0.0066] | [0.0217] | [0.0360] | [0.0021] | [0.0063] | [0.0162] |

| | $\gamma_1(0.6) = 0.4477$ | | | $\gamma_1(0.7) = 0.4793$ | | | $\gamma_1(0.8) = 0.3728$ | | | $\gamma_1(0.9) = 0.2199$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 0.3689 | 0.2971 | 0.2291 | 0.3946 | 0.3165 | 0.2432 | 0.3253 | 0.2763 | 0.2133 | 0.2031 | 0.1784 | 0.1529 |
| | (0.2958) | (0.3881) | (0.4676) | (0.3180) | (0.4070) | (0.4959) | (0.2616) | (0.3198) | (0.4094) | (0.1920) | (0.2157) | (0.2590) |
| | [0.0105] | [0.0259] | [0.0497] | [0.0141] | [0.0316] | [0.0589] | [0.0083] | [0.0128] | [0.0293] | [0.0024] | [0.0028] | [0.0053] |
| 400 | 0.3691 | 0.3038 | 0.2352 | 0.3882 | 0.3149 | 0.2370 | 0.3184 | 0.2647 | 0.2068 | 0.1994 | 0.1760 | 0.1470 |
| | (0.2868) | (0.3793) | (0.4609) | (0.3105) | (0.4055) | (0.4923) | (0.2423) | (0.3288) | (0.4074) | (0.1776) | (0.2129) | (0.2202) |
| | [0.0089] | [0.0228] | [0.0465] | [0.0121] | [0.0293] | [0.0598] | [0.0049] | [0.0132] | [0.0285] | [0.0017] | [0.0026] | [0.0057] |
| 600 | 0.3702 | 0.2998 | 0.2298 | 0.3908 | 0.3116 | 0.2366 | 0.3187 | 0.2649 | 0.2089 | 0.2028 | 0.1781 | 0.1499 |
| | (0.2808) | (0.3746) | (0.4568) | (0.2980) | (0.4035) | (0.4827) | (0.2387) | (0.3285) | (0.4049) | (0.1590) | (0.2045) | (0.2147) |
| | [0.0079] | [0.0211] | [0.0443] | [0.0096] | [0.0291] | [0.0597] | [0.0042] | [0.0124] | [0.0274] | [0.0009] | [0.0021] | [0.0052] |
| 800 | 0.3742 | 0.3057 | 0.2322 | 0.3888 | 0.3164 | 0.2405 | 0.3213 | 0.2669 | 0.2104 | 0.2053 | 0.1804 | 0.1503 |
| | (0.2712) | (0.3668) | (0.4542) | (0.2204) | (0.4017) | (0.4187) | (0.2285) | (0.3255) | (0.4030) | (0.1406) | (0.1997) | (0.2040) |
| | [0.0067] | [0.0210] | [0.0412] | [0.0092] | [0.0273] | [0.0575] | [0.0034] | [0.0117] | [0.0267] | [0.0006] | [0.0018] | [0.0051] |

**Table 4. Simulation results for $\gamma_1(x)$: censoring-ignored procedure**. For each configuration of the simulation parameters $(n, c, x)$, the first line gives the averaged value of the $N = 100$ CI-estimates of $\gamma_1(x)$. $(\cdot)$: empirical RMSE. $[\cdot]$: empirical MAE.

| | $q(1/1000, 0.1) = 4.5695$ | | | $q(1/1000, 0.2) = 13.1358$ | | | $q(1/1000, 0.3) = 27.4140$ | | | $q(1/1000, 0.4) = 22.0358$ | | | $q(1/1000, 0.5) = 11.2201$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 4.1080 | 3.2491 | 2.8883 | 9.8051 | 9.6386 | 13.4212 | 13.6245 | 8.6236 | 5.5780 | 13.5252 | 7.7951 | 4.5972 | 11.0114 | 6.9258 | 4.3742 |
| | (3.0417) | (3.6299) | (3.9964) | (8.9486) | (15.2268) | (15.5815) | (25.4495) | (39.4626) | (93.1253) | (14.3937) | (15.8764) | (67.5979) | (13.5663) | (21.2262) | (22.9915) |
| | [2.2933] | [2.4723] | [2.8388] | [6.8961] | [7.8926] | [10.0760] | [16.3037] | [21.8665] | [37.8360] | [11.1897] | [15.2898] | [27.4387] | [7.8146] | [11.1363] | [13.8459] |
| 400 | 4.0718 | 3.4227 | 2.5847 | 9.2295 | 6.2094 | 4.2650 | 13.9239 | 8.3275 | 5.4401 | 12.6516 | 7.5983 | 4.7965 | 9.0064 | 6.1299 | 4.2915 |
| | (2.0162) | (2.3928) | (3.0292) | (5.2196) | (7.2136) | (8.9502) | (15.4528) | (20.2842) | (25.0422) | (10.5159) | (14.7007) | (27.3086) | (6.5536) | (8.3824) | (9.6400) |
| | [0.8698] | [1.2616] | [1.9848] | [4.7442] | [6.9404] | [8.8708] | [15.5384] | [18.0865] | [20.9739] | [9.7117] | [14.4417] | [17.2394] | [5.0531] | [7.1533] | [9.5287] |
| 600 | 3.9357 | 3.4370 | 2.7061 | 8.7731 | 6.0104 | 4.9871 | 15.2458 | 8.1524 | 5.4168 | 13.1873 | 7.9479 | 4.8380 | 9.1764 | 6.1949 | 4.1538 |
| | (0.9810) | (1.2672) | (1.9026) | (4.8776) | (7.1786) | (8.1874) | (13.2615) | (19.4176) | (20.0551) | (9.7162) | (14.3171) | (17.2845) | (6.4945) | (8.2429) | (9.1162) |
| | [0.8568] | [1.1605] | [1.8634] | [4.5176] | [6.1254] | [8.1487] | [12.3215] | [17.2616] | [19.9972] | [8.9936] | [14.0880] | [17.1978] | [4.9579] | [7.0686] | [9.0664] |
| 800 | 3.9981 | 3.2871 | 2.6670 | 8.6456 | 6.1839 | 4.0785 | 14.0891 | 8.8163 | 4.9508 | 12.6915 | 8.1450 | 4.8126 | 9.2001 | 6.2326 | 4.1118 |
| | (0.8457) | (1.2627) | (1.8285) | (4.7951) | (7.0478) | (8.0909) | (12.9882) | (18.7219) | (19.4846) | (8.9834) | (14.0463) | (17.2540) | (4.8982) | (5.1481) | (7.1425) |
| | [0.7563] | [1.0877] | [1.5025] | [4.5147] | [5.9519] | [8.0573] | [11.3990] | [16.5977] | [19.4632] | [8.3843] | [13.8909] | [16.2232] | [4.4539] | [6.9889] | [8.1084] |

| | $q(1/1000, 0.6) = 22.0358$ | | | $q(1/1000, 0.7) = 27.4140$ | | | $q(1/1000, 0.8) = 13.1358$ | | | $q(1/1000, 0.9) = 4.5695$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 12.7008 | 7.8237 | 4.7413 | 15.9442 | 9.7660 | 5.1182 | 8.7584 | 5.9670 | 4.2613 | 4.1983 | 3.4545 | 2.7492 |
| | (22.1353) | (24.8535) | (41.5405) | (44.3813) | (51.9383) | (58.4641) | (12.7927) | (18.5359) | (19.0613) | (2.4890) | (2.5222) | (3.0214) |
| | [12.9255] | [15.2122] | [22.3945] | [21.5389] | [27.2877] | [29.2957] | [7.2900] | [11.2382] | [13.8745] | [2.3626] | [2.5279] | [2.8785] |
| 400 | 12.7121 | 8.4974 | 4.7249 | 13.5699 | 8.9238 | 5.1427 | 8.6218 | 5.7283 | 3.9875 | 3.8840 | 3.2568 | 2.7641 |
| | (11.9283) | (17.9500) | (19.3630) | (14.8161) | (17.9546) | (36.3483) | (5.5306) | (7.5235) | (9.0031) | (1.9958) | (2.4203) | (2.8869) |
| | [10.2847] | [13.6622] | [17.3110] | [14.1996] | [17.5721] | [22.2713] | [5.0160] | [7.4075] | [9.1483] | [1.8479] | (2.3574) | [2.8754] |
| 600 | 12.9843 | 7.8050 | 4.5757 | 14.4691 | 8.3937 | 5.0858 | 8.5311 | 6.2330 | 4.0638 | 3.9726 | 3.2510 | 2.6818 |
| | (9.9377) | (13.3955) | (17.3035) | (13.6632) | (17.1534) | (21.3662) | (5.1311) | (7.0562) | (8.8028) | (0.8761) | (1.3547) | (1.8280) |
| | [9.2349] | [13.2308] | [17.2601] | [12.9716] | [16.0203] | [20.9282] | [4.7467] | [6.9028] | [9.0720] | [0.7527] | [1.3308] | [1.8577] |
| 800 | 13.0067 | 7.7419 | 4.8322 | 14.1145 | 8.6326 | 5.1386 | 9.0117 | 6.1506 | 4.1788 | 3.9813 | 3.2770 | 2.7368 |
| | (9.7905) | (12.4117) | (17.2368) | (12.8020) | (16.9134) | (20.3035) | (4.5578) | (7.0047) | (7.9912) | (0.8668) | (1.3321) | (1.7768) |
| | [9.0793] | [12.2940] | [17.2036] | [12.3011] | [15.7814] | [20.2754] | [4.1976] | [6.0852] | [8.9570] | [0.7406] | [1.3160] | [1.8327] |

**Table 5. Simulation results for $q(1/1000|x)$: complete-case procedure**. For each configuration of the simulation parameters $(n, c, x)$, the first line gives the averaged value of the $N = 100$ CC-estimates of $q(1/1000|x)$. $(\cdot)$: empirical RMSE. $[\cdot]$: empirical MAE.

| | $q(1/1000, 0.1) = 4.5695$ | | | $q(1/1000, 0.2) = 13.1358$ | | | $q(1/1000, 0.3) = 27.4140$ | | | $q(1/1000, 0.4) = 22.0358$ | | | $q(1/1000, 0.5) = 11.2201$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 4.2897 | 3.4537 | 2.8371 | 8.7684 | 6.2767 | 4.3502 | 15.0531 | 9.0258 | 5.3598 | 14.3938 | 8.6377 | 5.2354 | 9.7966 | 6.5747 | 4.2460 |
| | (2.3776) | (2.4435) | (2.8471) | (8.5582) | (15.1524) | (17.8639) | (24.5730) | (38.8192) | (95.7441) | (15.3041) | (18.8140) | (68.8797) | (13.2965) | (15.4608) | (27.1106) |
| | [2.1389] | [2.4090] | [2.7546] | [6.9917] | [9.8591] | [10.7856] | [17.2359] | [22.4204] | [42.0542] | [11.2257] | [15.8982] | [26.8004] | [8.0146] | [9.0510] | [13.9742] |
| 400 | 4.0967 | 3.4640 | 2.8344 | 9.1486 | 6.2933 | 4.1896 | 13.5288 | 8.1083 | 4.9534 | 13.5414 | 8.3314 | 5.0511 | 9.5372 | 6.5826 | 4.4676 |
| | (1.8949) | (2.2388) | (2.7832) | (4.8504) | (7.0134) | (8.5750) | (15.7748) | (20.4761) | (25.4959) | (12.0645) | (16.9884) | (26.0269) | (6.2990) | (6.9478) | (9.9912) |
| | [1.7477] | [2.1237] | [2.7351] | [4.4303] | [6.8425] | [8.7462] | [13.9894] | [19.3057] | [22.4606] | [9.1570] | [13.7044] | [15.9848] | [4.8359] | [6.6975] | [7.7526] |
| 600 | 4.0016 | 3.3798 | 2.7494 | 8.8928 | 6.1247 | 4.1215 | 15.8611 | 9.1643 | 5.3512 | 12.2564 | 7.8021 | 4.8465 | 9.1106 | 6.3891 | 4.2628 |
| | (0.8834) | (1.2082) | (1.7530) | (4.7068) | (7.0079) | (8.0362) | (12.5660) | (18.3917) | (21.0892) | (9.9940) | (14.3552) | (15.2116) | (5.8370) | (6.9393) | (6.9889) |
| | [0.7296] | [1.2090] | [1.7201] | [4.3913] | [6.0111] | [8.0143] | [11.9221] | [18.2497] | [21.0628] | [8.9485] | [13.2338] | [15.1894] | [4.4512] | [5.8311] | [5.9573] |
| 800 | 3.9709 | 3.3922 | 2.7954 | 9.0251 | 6.3347 | 4.2389 | 15.0829 | 9.0157 | 5.3787 | 13.1601 | 8.1165 | 5.0220 | 9.3226 | 6.4006 | 4.3235 |
| | (0.8421) | (1.1631) | (1.7023) | (4.6250) | (6.9185) | (7.9195) | (11.1630) | (18.1203) | (20.0558) | (9.3573) | (14.0095) | (15.0295) | (4.7082) | (5.7484) | (5.9203) |
| | [0.6318] | [1.1845] | [1.6741] | [4.2471] | [5.8011] | [7.8969] | [10.4731] | [17.3983] | [20.0353] | [8.8834] | [12.9193] | [15.0138] | [3.2976] | [4.8196] | [5.8967] |

| | $q(1/1000, 0.6) = 22.0358$ | | | $q(1/1000, 0.7) = 27.4140$ | | | $q(1/1000, 0.8) = 13.1358$ | | | $q(1/1000, 0.9) = 4.5695$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% | 10% | 25% | 40% |
| 200 | 13.5598 | 8.1817 | 5.0081 | 17.6563 | 9.9514 | 5.7560 | 9.9470 | 7.0556 | 4.4594 | 4.0331 | 3.4000 | 2.8543 |
| | (20.7934) | (24.2654) | (47.1019) | (46.2637) | (52.5574) | (58.8686) | (11.6445) | (18.8980) | (19.7960) | (1.9243) | (2.0226) | (2.4991) |
| | [15.7558] | [19.8598] | [23.0277] | [23.8800] | [28.1125] | [31.6580] | [7.0836] | [9.4752] | [10.6764] | [1.6999] | [1.9085] | [2.3158] |
| 400 | 13.1138 | 8.3195 | 5.1619 | 15.4384 | 9.0572 | 5.1732 | 9.1203 | 6.2692 | 4.2023 | 3.9124 | 3.3383 | 2.7317 |
| | (14.2812) | (17.9705) | (20.9315) | (17.0349) | (22.6133) | (36.2739) | (6.8832) | (8.2716) | (9.9819) | (1.6807) | (1.9761) | (2.4784) |
| | [13.3751] | [17.7164] | [20.8739] | [15.0190] | [19.3568] | [26.2408] | [6.3280] | [7.8666] | [9.9335] | [1.5269] | [1.8638] | [2.2378] |
| 600 | 13.0675 | 7.9925 | 4.8972 | 14.9991 | 8.6069 | 5.1462 | 8.9504 | 6.1782 | 4.2053 | 3.9653 | 3.3457 | 2.7773 |
| | (13.8580) | (17.1865) | (19.1670) | (15.1700) | (18.9371) | (22.2930) | (6.1652) | (7.8561) | (8.9552) | (0.8865) | (1.2941) | (1.8213) |
| | [13.1027] | [17.0433] | [19.1387] | [14.4417] | [17.8071] | [23.2178] | [6.1672] | [7.2576] | [9.5305] | [0.7655] | [1.2238] | [1.7922] |
| 800 | 13.3315 | 8.2681 | 4.9618 | 14.6017 | 8.8815 | 5.2610 | 9.1107 | 6.2694 | 4.2371 | 4.0549 | 3.4273 | 2.7940 |
| | (13.2773) | (16.8695) | (19.0896) | (14.1858) | (17.9119) | (22.1671) | (6.0907) | (7.1299) | (8.9132) | (0.7301) | (1.2121) | (1.7983) |
| | [12.7043] | [16.7677] | [19.0740] | [13.8123] | [17.5325] | [23.1530] | [6.0647] | [7.1664] | [8.4987] | [0.6367] | [1.1474] | [1.7755] |

**Table 6. Simulation results for $q(1/1000|x)$: censoring-ignored procedure**. For each configuration of the simulation parameters $(n, c, x)$, the first line gives the averaged value of the $N = 100$ CI-estimates of $q(1/1000|x)$. $(\cdot)$: empirical RMSE. $[\cdot]$: empirical MAE.

**Acknowledgements**

The authors acknowledge financial support from the "Agence Universitaire de la Francophonie" and AIRES-Sud (a program from the French Ministry of Foreign and European Affairs implemented by the "Institut de Recherche pour le Développement", IRD-DSF, grant number: 7168).

**References**

Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., 2004. Statistics of Extremes: Theory and Applications. John Wiley & Sons, Ltd.

Beirlant, J., Guillou, A., Dierckx, G., Fils-Villetard, A., 2007. Estimation of the extreme value index and extreme quantiles under random censoring. Extremes 10(3), 151-174.

Beirlant, J., Guillou, A., Toulemonde, G., 2010. Peaks-Over-Threshold modeling under random censoring. Communications in Statistics - Theory and Methods 39, 1158-1179.

Berlinet, A., Gannoun, A., Matzner-Løber, E., 2012. Asymptotic normality of convergent estimates of conditional quantiles. Statistics 35(2), 139-169.

Brahimi, B., Meraghni, D., Necir, A., 2013. On the asymptotic normality of Hill's estimator of the tail index under random censoring. Preprint: arXiv-1302.1666.

Dabrowska, D.M., 1992. Variable bandwidth conditional Kaplan-Meier estimate. Scandinavian Journal of Statistics 19, 351-361.

Daouia, A., Gardes, L., Girard, S., Lekina, A., 2011. Kernel estimators of extreme level curves. Test 20, 311-333.

de Haan, L., Ferreira, A., 2006. Extreme Value Theory. Springer, New York.

Dekkers, A.L.M., Einmahl, J.H.J, de Haan, L., 1989. A moment estimator for the index of an extreme-value distribution. Annals of Statistics 17(4), 1833-1855.

Delafosse, E., Guillou, A., 2002. Almost sure convergence of a tail index estimator in the presence of censoring. Comptes Rendus Mathématique. Académie des Sciences. Paris 335(4), 375-380.

Einmahl, J.H.J., Fils-Villetard, A., Guillou, A., 2008. Statistics of extremes under random censoring. Bernoulli 14(1), 207-227.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events For Insurance and Finance. Springer, Berlin.

Ferrez, J., Davison, A.C., Rebetez, M., 2011. Extreme temperature analysis under forest cover compared to an open field. Agricultural and Forest Meteorology 151, 992-1001.

Gannoun, A., 1990. Estimation non paramétrique de la médiane conditionnelle, médianogramme et méthode du noyau. Publications de l'Institut de Statistique de l'Université de Paris XXXXVI, 11-22.

Gardes, L., Girard, S., 2008. A moving window approach for nonparametric estimation of the conditional tail index. Journal of Multivariate Analysis 99(10), 2368-2388.

Gardes, L., Girard, S., 2010. Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels. Extremes 13(2), 177-204.

Gardes, L., Girard, S., 2012. Functional kernel estimators of large conditional quantiles. Electronic Journal of Statistics 6, 1715-1744.

Gardes, L., Girard, S., Lekina, A., 2010. Functional nonparametric estimation of conditional extreme quantiles. Journal of Multivariate Analysis 101(2), 419-433.

Gardes, L., Stupfler, G., 2014. Estimation of the conditional tail index using a smoothed local Hill estimator. Extremes 17(1), 45-75.

Goegebeur, Y., Guillou, A., Osmann, M., 2014a. A local moment type estimator for the extreme value index in regression with random covariates. Canadian Journal of Statistics, to appear.

Goegebeur, Y., Guillou, A., Schorgen, A., 2014b. Nonparametric regression estimation of conditional tails: the random covariate case. Statistics 48(4), 732-755.

Gomes, M.I., Neves, M.M., 2011. Estimation of the extreme value index for randomly censored data. Biometrical Letters 48(1), 1-22.

Gomes, M.I., Oliveira, O., 2003. Censoring estimators of a positive tail index. Statistics & Probability Letters 65(3), 147-159.

Matthys, G., Delafosse, E., Guillou, A., Beirlant, J., 2004. Estimating catastrophic quantile levels for heavy-tailed distributions. Insurance: Mathematics & Economics 34(3), 517-537.

Ndao, P., Diop, A., Dupuy, J.-F., 2014. Nonparametric estimation of the conditional tail index and extreme quantiles under random censoring. Computational Statistics & Data Analysis 79, 63-79.

Novak, S.Y., 2012. Extreme value methods with applications to finance. CRC Press, Boca Raton, FL.

Pisarenko, V.F., Sornette, D., 2003. Characterization of the frequency of extreme earthquake events by the generalized Pareto distribution. Pure and Applied Geophysics 160, 2343-2364.

Reiss, R.-D., Thomas, M., 2007. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhäuser, Basel.

Samanta, M., 1989. Nonparametric estimation of conditional quantiles. Statistics & Probability Letters 7(5), 407-412.

Stone, C.J., 1977. Consistent nonparametric regression (with discussion). The Annals of Statistics 5(4), 595-645.

Stupfler, G., 2013. A moment estimator for the conditional extreme-value index. Electronic Journal of Statistics 7, 2298-2343.

Stute, W., 1986. Conditional empirical processes. The Annals of Statistics 14(2), 638-647.

van der Vaart, A.W., 1998. Asymptotic Statistics. Cambridge University Press, New York.

Weissman, I., 1978. Estimation of parameters and large quantiles based on the $k$ largest observations. Journal of the American Statistical Association 73, 812-815.

Worms, J., Worms, R., 2014. New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. Extremes 17(2), 337-358.

Yao, Q., 1999. Conditional predictive regions for stochastic processes. Technical report. Canterbury: Institute of Mathematics and Statistics, University of Kent.