Afrika Statistika

Afrika Statistika
Vol. 16 (2), 2021, pages 2647 - 2688.
DOI: http://dx.doi.org/10.16929/as/2021.2647.178

ISSN 2316-090X

Functional kernel estimation of the conditional extreme value index under random right censoring

Justin Ushize Rutikanga [(1)], **Aliou Diop** [2,*]

[1]Université d'Abomey-Calavi, Institut de Mathématiques et de Sciences Physiques (*IMSP-UAC*), Porto-Novo, Bénin.
[2]*LERSTAD*, Université Gaston Berger, Saint Louis, Sénégal.

Received on June 09, 2021; Accepted on July 26, 2021

Abstract. Estimation of the extreme-value index of a heavy-tailed distribution is investigated when some functional random covariate (i.e. valued in some infinite-dimensional space) information is available and the scalar response variable is right-censored. A weighted kernel version of Hill's estimator of the extreme-value index is proposed and its asymptotic normality is established under mild assumptions. A simulation study is conducted to assess the finite-sample behavior of the proposed estimator. An application to ambulatory blood pressure trajectories and clinical outcome in stroke patients is also provided.

Key words: censored data; conditional extreme value index; heavy-tailed distributions; functional data; kernel estimator.
AMS 2010 **Mathematics Subject Classification Objects :** 62G32; 62G30; 62N02.

---

*Corresponding author: Aliou Diop (aliou.diop@ugb.edu.sn)
Justin Ushize Rutikanga : ushizerj@gmail.com

Résumé. L'estimation de l'indice des valeurs extrêmes d'une distribution à queue lourde est étudiée lorsque des informations sur les covariables aléatoires fonctionnelles (appartenant à un espace de dimension infinie) sont disponibles et que la variable réponse scalaire est censurée à droite. Une version à noyau pondéré de l'estimateur de Hill de l'indice des valeurs extrêmes est proposée et sa normalité asymptotique est établie sous des hypothèses légères. Une étude de simulation est menée pour évaluer le comportement de l'estimateur proposé à distance finie. Une application aux mesures de la pression artérielle ambulatoire et aux résultats cliniques des patients victimes d'un accident vasculaire cérébral est également fournie.

The authors.
Justin Ushize Rutikanga, M.Sc., is preparing a Ph.D, under the supervision of the second author at Abomey-Calavi University, (Physics and Mathematics Institute) [Institut de Mathématiques et de Sciences Physiques (**IMSP-UAC**), Porto-Novo Bénin].
Aliou Diop, Ph.D., is full professor at the Department of Applied Mathematics of Université Gaston Berger, Saint Louis, Sénégal.

1. Introduction

Let $(X_i, Y_i), \quad i = 1, \cdots, n$ be independent copies of random pair (X, Y), where Y is positive real random variable and $X \in \mathbf{E}$, where $\mathbf{E}$ is an infinite dimensional space associated to a semi-metric $d(.,.)$. We assume that the variable Y is randomly right censored by a positive random variable C defined on the same probability space $(\Omega, \mathcal{C}, \mathbb{P})$ as Y. Therefore, we really observe independent triplets (X_i, δ_i, Z_i), where $Z_i = \min(Y_i, C_i)$ and $\delta_i = \mathbf{1}_{\{Y_i \leqslant C_i\}}$ for $i = 1 \cdots, n$, where $\mathbf{1}_A$ is the indicator function of the event A.

We assume that Y and C are independent of given $X = x$, where $C_1, \cdots, C_n$ are independent each other.

Let $F(\cdot|x)$ and $G(\cdot|x)$ be the conditional cumulative distribution functions of random variables Y and C given $X = x$ respectively.

Let $\overline{F}(\cdot|x)$ and $\overline{G}(\cdot|x)$ be the conditional survival function of random variable Y and C given $X = x$ respectively.

In this paper, we focus on heavy tails. More specifically, we assume that the conditional survival functions satisfy the following assumption

(A1)

$$\overline{F}(t|x) = r_1(x) \exp\left\{ -\int_1^t \left(\frac{1}{\gamma_1(x)} - \varepsilon_1(\mu|x) \right) \frac{d\mu}{\mu} \right\} \tag{1}$$

and

$$\overline{G}(t|x) = r_2(x) \exp\left\{-\int_1^t \left(\frac{1}{\gamma_2(x)} - \varepsilon_2(\mu|x)\right)\frac{d\mu}{\mu}\right\}, \tag{2}$$

where $\gamma_1(x), \gamma_2(x)$ are positive unknown functions of the covariate x, r_1, r_2 are positive functions and $|\varepsilon_1(\mu|x)|$, $|\varepsilon_2(\mu|x)|$ are continuous and ultimately decreasing to zero. From (1) and (2), we can state that the conditional distribution functions of Y and C given $X = x$ are in Fréchet maximum domain of attraction. Thus, $\gamma_1(x)$ and $\gamma_2(x)$ are taken as the conditional extreme tail index functions. Therefore, for all $t > 0$, $\overline{F}(\cdot|x)$ and $\overline{G}(\cdot|x)$ are regularly varying functions at infinity with index $-1/\gamma_1(x)$ and $-1/\gamma_2(x)$ respectively. Thus,

$$\overline{F}(u|x) = u^{-\frac{1}{\gamma_1(x)}} L_1(u|x) \qquad \text{and} \qquad \overline{G}(u|x) = u^{-\frac{1}{\gamma_2(x)}} L_2(u|x)$$

where for x fixed, $L_1(.|x)$ and $L_2(.|x)$ are slowly varying functions at infinity, that is, for all $\lambda > 0$,

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

By conditional independence between Y and C, the conditional survival function $\overline{H}(\cdot|x)$ of Z given $X = x$ is also a regularly varying function at infinity with index $-\frac{1}{\gamma(x)}$ as expressed as follows:

$$\begin{aligned}\overline{H}(\mu|x) &= 1 - H(\mu|x) = \overline{F}(\mu|x)\overline{G}(\mu|x)\\ &= r(x)\exp\left\{-\int_1^z \left(\frac{1}{\gamma(x)} - (\varepsilon(\mu|x))\right)\frac{d\mu}{\mu}\right\},\end{aligned}$$

with $\gamma(x) = \gamma_1(x)p(x)$, where $p(x) = \gamma_2(x)/\gamma_1(x) + \gamma_2(x)$ is the ultimate proportion of uncensored observations among $Z_i, i = 1, \cdots, n$; (see Einmahl *et al.*(2008), Ndao (2015) for more details) and

$$r(x) = r_1(x)r_2(x), \varepsilon(\mu|x) = \varepsilon_1(\mu|x) + \varepsilon_2(\mu|x).$$

In the sequel, we further assume that $L_i(u|x)$, $i = 1, 2$ belong to the Hall class of slowly-varying functions defined as follows:

$$L_i(u|x) = C_i(x) + D_i(x)u^{-\beta_i(x)}(1 + o(1)) \quad \textit{as} \quad u \to \infty, \quad \textit{for} \quad i = 1, 2, \tag{3}$$

where $C_i(x)$ and $D_i(x)$ are functions in x with $C_i(x) > 0$ and $D_i(x) \in \mathbb{R}$ and $\beta_i(\cdot)$ is a positive function.

Therefore, the main contribution of this paper is to construct our estimator of the conditional extreme value index $\gamma_1(x)$ and establish its asymptotic normality from a sample of observations $(Z_1, \delta_1, x_1), \cdots, (Z_n, \delta_n, x_n)$.

More specifically, we are interested in nonparametric estimation of conditional tail index estimation of the conditional distribution of a randomly censored scalar response given a functional covariate. When some covariate is recorded simultaneously with the quantity of interest Y, the tail index can depend on the covariate and is referred in the sequel to as conditional tail index or conditional extreme value-index. This parameter drives the tail heaviness of the distribution of Y and thus plays a central role in the analysis of extremes, making its estimation a crucial issue. Then, we consider the situation where some covariate information X is available to the investigator, and the distribution of Y depends on X.

1.1. Literature review

In this paper, we deal with the problem of estimating a conditional extreme-value index of a heavy-tailed distribution when some functional covariate information $X \in \mathbf{E}$ is available, where $\mathbf{E}$ is an infinite dimensional space associated with a semi-metric $d(\cdot, \cdot)$.

The analysis of functional data has been extensively studied for example in Ramsay and Dalzell (1991). These authors have developed the fundamental theory around the functional data. Parametric model estimators have been developed in Ramsay and Silverman (2005), Bosq (2000) and nonparametric estimators have been proposed in Ferraty and Vieu (2004) by establishing the strong consistency of estimators related to regression function.

Despite the greater interest in practice, the study of the functional random covariate case have been initiated recently, we refer to the works in Gardes and Girard (2010) based on the estimation of extreme rainfalls given the geographical location. Ferrez *et al.*(2011) also studied extreme temperatures given topological parameters and Pisarenko and Sornette (2003) investigated the estimation of extreme earthquakes given the geographical location.

Normally, in real-life practical application, the variable of interest may be incomplete. In classical applications such as the analysis of lifetime data, a typical feature which appears is censorship. For example, in medical follow up, the response variable Y represents the time elapsed from the entry of a patient in, say, a follow-up study until death. If, at the time that the data collection is performed, the patient is still alive or has withdrawn from the study for some reason, the variable of interest Y will not be available. A convenient way to model this situation is the introduction of a random variable C (called a censoring random variable), independent of Y, such that only the pairs (Z_i, δ_i), $1 \leqslant i \leqslant n$ are observed where $Z_i = \min(Y_i, C_i)$ and $\delta_i = \mathbf{1}_{\{Y_i \leqslant C_i\}}$ and $\mathbf{1}_{\{A\}}$ is the indicator function of the event A.

Estimation of the extreme value-index or high quantiles with censored data when covariate information is not available was the purpose of many investigations.

In this framework, estimation of the extreme-value index from censored data is considered by Delafosse and Guillou (2002), Beirlant *et al.*(2010), Gomes and Neves (2011), Gomes and Neves (2010), Einmahl *et al.*(2008), Brahimi (2013) and Worms and Worms (2014). Beirlant *et al.*(2007) and Einmahl *et al.*(2008) additionally address estimation of extreme quantiles. However, considering Pareto-type distributions, Beirlant *et al.*(2018) dealt with bias reduced estimator for extreme value index and tail probability under random right censored. Ndao *et al.*(2016) and Ndao (2015) addressed estimation of the conditional extreme-value index and conditional extreme quantiles with fixed and random covariates. Beside, Stupfler (2016) investigated the conditional estimation of the extreme value index under random censoring for all domain of attraction. While Worms and Worms (2018) proposed estimator of the extreme value index by considering heavy tailed lifetime data under random censoring and competing risks, using the Aalen-Johansen estimator of the cumulative incidence function. Beirlant *et al.*(2019) introduced a new class of estimator which generalized the proposed estimator of Worms and Worms (2014) and Beirlant *et al.*(2018)

Some important literature is devoted to the estimation of the conditional quantile of a scalar response given a functional covariate. Gardes and Girard (2012) dealt with the estimation of conditional quantiles when the covariate is functional and when the order of the quantiles converges to one as the sample size increases. Chaouch and Khardani (2014) investigated the conditional quantile estimation of a randomly censored scalar response variable given a functional random covariate whenever a stationary ergodic data are considered.

However, based on the literature and our knowledge, estimation of the functional conditional extreme value-index of a heavy-tailed distribution under random right censoring has not yet been addressed, when covariate is a functional random variable. Our methodology 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 a weighting term whose role is to correct for censoring (such as in Einmahl *et al.*(2008) and Brahimi (2013), who estimated the unconditional extreme-value index with censoring). This idea is already used in Ndao *et al.*(2016). Following the same idea, we propose a functional Hill-type estimator depending on a semi-metric $d(.,.)$.

1.2. Construction of the estimator

Let (X_i, δ_i, Z_i), $i = 1, \cdots, n$, be independent realizations of the random vector (X, δ, Z) where $Z_i = \min(Y_i, C_i)$ and $\delta_i = \mathbf{1}_{\{Y_i \leqslant C_i\}}$ for $i = 1 \cdots, n$ and $(X, Z) \in \mathbf{E} \times \mathbb{R}_+^*$.

If Z_i were uncensored it means that $Z_i = Y_i$ for all i. In this situation, Goegebeur *et al.*(2014b) proposed a Hill's version of the conditional extreme value index when the covariate response is in $\mathbb{R}^p$. Following the same idea, we propose a functional Hill-type estimator depending on a semi-metric $d(.,.)$:

$$\hat{\gamma}_{y_n}^H(x) = \sum_{i=1}^n K(h^{-1}d(x,X_i))\,(\log Z_i - \log y_n)\,\mathbf{1}_{\{Z_i>y_n\}}/\sum_{i=1}^n K(h^{-1}d(x,X_i))\mathbf{1}_{\{Z_i>y_n\}}, \quad (4)$$

where $K(.)$ is a real-valued kernel function on $\mathbf{E}$, $h = h_n$ is a positive non-random bandwidth sequence such that $h \to 0$ as $n \to \infty$ and y_n is a local non-random threshold sequence for estimation with $y_n \to \infty$ as $n \to \infty$. Here, as stated in Goegebeur *et al.*(2014b), a local threshold means a threshold depending on the point x in the covariate space where the estimation takes place, though the threshold is constant in a neighborhood of x. The estimator (4) is not consistent for $\gamma_1(x)$ if it is directly applied to the censored sample $(X_i, \delta_i, Z_i), i = 1, \cdots, n$. Indeed, under appropriate regularity assumptions, estimator (4) will converge to the extreme-value index $\gamma(x)$ of the conditional distribution of Z given $X = x$. To accommodate censoring, we suggest, like in Ndao *et al.*(2016), to divide (4) by the proportion $\hat{p}_n(x)$ of uncensored observations among the $Z_i, i = 1, \cdots, n$ that are larger than y_n, in a neighborhood of x:

$$\hat{p}_n(x) = \frac{\overline{H}_n^1(y_n|x)}{\overline{H}_n(y_n|x)},$$

where

$$\overline{H}_n(y_n|x) = \sum_{i=1}^n B_i(x)\mathbf{1}_{\{Z_i>y_n\}},$$

$$\overline{H}_n^1(y_n|x) = \sum_{i=1}^n B_i(x)\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}$$

and $B_i(x)$'s are the well-known Nadaraya-Watson weights defined by

$$B_i(x) = \frac{K(h^{-1}d(x,X_i))}{\sum_{j=1}^n K(h^{-1}d(x,X_j))}.$$

The survival functions $\overline{H}_n(y_n|x)$ and $\overline{H}_n^1(y_n|x)$ can be rewritten as follows:

$$\overline{H}_n^1(y_n|x) = \hat{\psi}_n(y_n,x)/\hat{g}_n(x) \ and \ \overline{H}_n(y_n|x) = \hat{\zeta}_n(y_n,x)/\hat{g}_n(x),$$

respectively, where

$$\hat{\psi}_n(y_n,x) = \frac{1}{n(\mu_x^{(1)}(h))}\sum_{i=1}^{n} K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}};$$

$$\hat{\zeta}_n(y_n,x) = \frac{1}{n(\mu_x^{(1)}(h))}\sum_{i=1}^{n} K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}}$$

and

$$\hat{g}_n(x) = \frac{1}{n(\mu_x^{(1)}(h))}\sum_{i=1}^{n} K(\frac{d(x,X_i)}{h}).$$

Therefore we propose to estimate $\gamma_1(.)$ by

$$\hat{\gamma}_{y_n}^{c,H}(x) = \frac{\hat{\gamma}_n^H(x)}{\hat{p}_n(x)}. \tag{5}$$

This estimator depends on the bandwidth h, the threshold y_n and the semi-metric $d(\cdot,\cdot)$. The choice of the semi-metric is a crucial point in non-parametric functional data analysis (see Ferraty and Vieu (2006)). Once the semi-metric has been chosen, packages are available in the literature (see https://cran.r-project.org/web/packages/fda.usc/index.html) to evaluate approximates between functional data. In the simulation study in Section 3, we will discuss the impact of the degree of derivatives on the performance of our estimator when semi-metric based on derivatives are considered for smooth curves as covariates.

The remainder of this paper is organized as follows. Section 2 is devoted to assumptions and asymptotic normality of the proposed estimator. Section 3 illustrates the finite sample behavior of the proposed estimator via simulations and application to ambulatory blood pressure trajectories and clinical outcome in stroke patients is provided.

The conclusion and some perspectives are presented in Section 4. All proofs are presented in the Appendix in Section 5.

2. Assumptions and asymptotic results

In this section, we first give some regularity conditions for proving our results. Some of those assumptions have been adapted from different authors including Gardes and Girard (2008), Chaouch and Khardani (2014), Gardes and Girard (2012); Ndao *et al.*(2016).
Let us first introduce the following conditional expectation whose asymptotic expansion will enable us to study the asymptotic expansion of our proposed estimator (5)

$$\mathbb{S}_n(y_n, r; x) = \mathbb{E}\left[(\log Z - \log y_n)^r \mathbf{1}_{\{Z>y_n\}} | X = x\right]. \tag{6}$$

(A2) The kernel K is a bounded density function with support $[0,1]$, and there exist real numbers $0 < C < C' < \infty$ such that $C \leqslant K(t) \leqslant C'$ for all $t \in [0,1]$.

(A3) There exist $K_r > 0$, $K_\epsilon > 0$, $K_\gamma > 0$, $\mu_0 > 1$ such that for all $(x, x') \in \mathbf{E} \times \mathbf{E}$ and $\mu > \mu_0$,

$$\begin{aligned}
|\log r(x) - \log r(x')| &\leqslant K_r d(x, x'), \\
|\epsilon(\mu|x) - \epsilon(\mu|x')| &\leqslant K_\epsilon d(x, x'), \\
\left|\frac{1}{\gamma(x)} - \frac{1}{\gamma(x')}\right| &\leqslant K_\gamma d(x, x').
\end{aligned}$$

Under condition **(A2)**, we introduce the new notation that

$$\mu_x^{(j)}(h) = \mathbb{E}\left[K^j\left(h^{-1}d(x, X)\right)\right] \; for j = 1, 2.$$

Let $B(t, r)$ be the ball of center t and radius r. Then we define the function

$$\varphi_x(h) = \mathbb{P}(X \in B(x, h)),$$

which is the small ball probability. By condition **(A2)** and Lemma 2 in the Appendix, it is shown that for $j > 0$, $\mu_x^{(j)}(h)$ has the same asymptotic order as $\varphi_x(h)$.

(A4) Suppose that there exist a function $\rho(x) < 0$ and a regularly varying function $A(\cdot|x)$ which not changing the sign eventually at infinity with index $\rho(x)$ and with $\lim_{y_n \to \infty} A(y_n|x) = 0$ such that for all $u > 0$,

$$\lim_{t \to \infty} \frac{H^{\leftarrow}\left(1 - \frac{1}{ut}|x\right) / H^{\leftarrow}\left(1 - \frac{1}{t}|x\right) - u^{\gamma(x)}}{A(t|x)} = u^{\gamma(x)} \frac{u^{\rho(x)} - 1}{\rho(x)},$$

where $H^{\leftarrow}(\cdot|x)$ is the generalized inverse function defined by

$$H^{\leftarrow}(q|x) = \inf\{y, H(y|x) \geq q\}, \; 0 < q < 1.$$

The aforementioned condition is called second-order condition in classical extreme value theory. Considering the condition **(A4)** which controls the behavior of $H^{\leftarrow}(\cdot|x)$, we need an additional condition to control the oscillation of $H^{\leftarrow}(u|x)$ when considered as a function of the covariate x. That condition is expressed in terms of conditional expectation $\mathbb{S}_n(y_n, r; x)$ given in (6) as defined in

Goegebeur *et al.*(2014a).

(A5) The conditional expectation $\mathbb{S}_n(y_n, r; x)$ satisfies that, for $y_n \to \infty, h \to 0$ as $n \to \infty$ and $(x', x) \in \mathbf{E} \times \mathbf{E}$

$$\Phi_n(y_n, h; x) := \sup_{r\in\{1,2\}} \sup_{x'\in B(x,h)} \left[\left|\frac{\mathbb{S}_n(y_n, r; x')}{\mathbb{S}_n(y_n, r; x)} - 1\right| \mathbf{1}_{\{d(x',x)<h\}}\right] \to 0 \text{ if } n \to \infty.$$

Remark 1. For sake of simplicity, we choose to deal with the naive kernel estimator. It is easy to construct a smooth version of this naive estimator. To do so, it suffices to change the basic indicator function into a smooth cumulative distribution function. For more details, see Ferraty and Vieu (2006). However, the theoretical properties stated in Theorem 1 can not be considered as valid for the smooth version of our estimator.

In the following, we first investigate the asymptotic distribution of $\hat{p}_n(x)$, which plays an important role in the definition of the conditional extreme-value index. Then, we derive the limiting distribution of the proposed estimator (5) of $\gamma_1(x)$. The proofs are given in the Appendix.

Proposition 1. *Suppose* ***(A1)*** *–* ***(A4)*** *hold. If*

$$y_n \to \infty,\ \sigma_n^{-1}(x)\left(h \log y_n \vee \varphi_x(h) \vee A\left(\frac{1}{\overline{H}(y_n|x)}|x\right)\right) \to 0,$$

then for all $x \in \mathbf{E}$*, such that* $\varphi_x(h) > 0$*,*

$$\sigma_n^{-1}(x)\left(\hat{p}_n(x) - p(x)\right) \xrightarrow{D} \mathcal{N}\left(0, p(x)(1-p(x))\right).$$

Theorem 1. *Suppose* ***(A1)–(A5)*** *hold. Let* y_n *be non-random threshold sequence and* h *be a sequence of bandwidth, such that* $y_n \to \infty$ *and* $h \to 0$*,* $n\overline{H}(y_n|x) \to \infty$*,*

$$\sigma_n^{-1}(x)\left(A\left(\frac{1}{\overline{H}(y_n|x)}|x\right) \vee (h \log y_n) \vee \varphi_x(h)\right) \to 0 \text{ as } n \to \infty.$$

For all $x \in \mathbf{E}$*, such that* $\varphi_x(h) > 0$*,*

$$\sigma_n^{-1}(x)\left(\hat{\gamma}_{y_n}^{c,H}(x) - \gamma_1(x)\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{\gamma_1^3(x)}{\gamma(x)}\right),$$

where

$$\sigma_n(x) = \left(n\overline{H}(y_n|x)\frac{\left(\mu_x^{(1)}(h)\right)^2}{\mu_x^{(2)}(h)}\right)^{-1/2}. \square$$

The condition

$$\sigma_n^{-1}(x)\left(A\left(\frac{1}{\overline{H}(y_n|x)}|x\right)\vee(h\log y_n)\vee\varphi_x(h)\right)\to 0$$

in Theorem 1 is a combination of three conditions. The first one

$$\sigma_n^{-1}(x)A\left(\frac{1}{\overline{H}(y_n|x)}|x\right)\to 0$$

is similar to the one in Ndao *et al.*(2016) and Goegebeur *et al.*(2014b) and is classical in extreme value theory. Since

$$\frac{\left(\mu_x^{(1)}(h)\right)^2}{\mu_x^{(2)}(h)}$$

has the same asymptotic order as $\varphi_x(h)$ according to Lemma 2 below,

$$\sigma_n^{-1}(x)\left((h\log y_n)\vee\varphi_x(h)\right)\to 0$$

is equivalent to

$$n\overline{H}(y_n|x)\varphi_x(h)\left((h\log y_n)\vee\varphi_x(h)\right)^2\to 0.$$

It imposes to $h\log y_n$ and $\varphi_x(h)$ to be negligible compared to the standard deviation $\sigma_n(x)$ of the estimator. For practical implementation, we have to choose the bandwidth h and the non-random threshold sequence such that $y_n = Z_{(n-k)}$ is in the ball $B(x,h)$, where $Z_{(n-k)}$ is the $(n-k)^{th}$ order statistic.

Remark 2. To prove Proposition 1, Lyapunov condition has been used. For more details of a general formulation of Lyapunov theorem and its proof, see Theorem 19 in Lo (2018) . Let us recall the Lyapunov Theorem. Suppose that $U_1, U_2, \cdots$ are independent centered random variables, with finite $n+\phi$ moment, $\phi>0$. For each $n>1$ we can define $T_n=\sum_{i=1}^n U_i$ and $s_n^2=\sum_{i=1}^n v\,\mathrm{Var}(U_i)$. Suppose that

$$\frac{1}{s_n^{2+\phi}}\sum_{i=1}^n \mathbb{E}|U_i|^{2+\phi}\to 0 \quad \textit{as} \quad n\to\infty.$$

Thus,

$$T_n/s_n \xrightarrow{D} \mathcal{N}(0,1).$$

In our context $\phi = 1$. Since Lyapunov condition implies the Linderberg condition which is less restrictive (requiring only square integrable variable), Linderberg condition can be used. See also Lo (2018) for more details.
The above Theorem allows one to obtain an asymptotic confidence interval in practice since all quantities can be estimated.

The above Theorem allows one to obtain an asymptotic confidence interval in practice since all quantities can be estimated.

Confidence intervals
Since the asymptotic conditional variance in the above result can be written in terms of $\gamma(x)$, $\gamma_1(x)$, $\bar{H}(.|x)$ and $\mu_x^{(i)}(h)$, $i = 1, 2$, it follows that one can construct asymptotic confidence intervals by replacing $\gamma(x)$ and $\gamma_1(x)$ respectively by the proposed estimators (4) and (5). The terms $\bar{H}(.|x)$ and $\mu_x^{(i)}(h)$, $i = 1, 2$ can be estimated by their empirical counterparts. This procedure is implemented in the simulation section. To avoid the calculation of the conditional variance, resampling techniques, such as bootstrap can be used and are implemented in the real-data application.

3. Simulations studies and applications

In this section, we assess, via simulations, finite-sample performance of our estimator. We also provide comparisons with two simple estimation strategies of the tail index of a heavy-tailed distribution under random censoring.

3.1. Simulation design

The main goal of this section is to demonstrate the efficiency of the proposed estimator in terms of consistency. For this purpose, we consider the simulation of $N = 500$ replications of a sample of size n $(n = 200, n = 500)$ of random triple (X_i, δ_i, Z_i) where $Z_i = \min(C_i, Y_i)$ and X_i is a functional covariate $X \in \mathbf{E}$ which is defined by

$$X(t) = \Omega(2 - cos(\pi W t)) + (1 - \Omega)cos(\pi W t), \quad \text{for all} \quad t \in [0, 1],$$

where W is normally distributed with mean zero and unit variance and Ω is a random variable Bernoulli distributed with parameter $p = 1/2$ as in Chaouch and Khardani (2014).

Figure 1 below illustrates some realizations of random curves of the given functional random variable $X(\cdot)$.

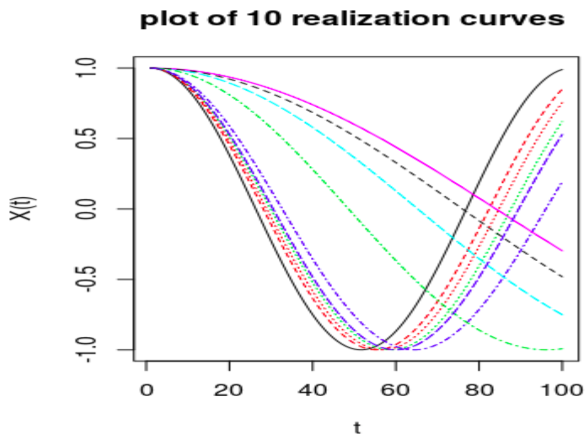


Fig. 1. Ten realizations of the random curves of function $X(\cdot)$.

The conditional distribution of Y given $X = x$ is a Burr distribution with parameter $\tau(x) = 2, \lambda_1(x) = 2/(8 \parallel X \parallel_2^2 -3)$, which implies that $\gamma_1(x) = 1/(\tau(x)\lambda_1(x))$, with

$$\parallel X \parallel_2^2 = \int_0^1 X^2(t)dt = 4\Omega^2 - 4\Omega(2\Omega-1)\frac{\sin(\pi W)}{\pi W} + (2\Omega-1)^2\left[\frac{1}{2} + \frac{\sin(2\pi W)}{4\pi W}\right].$$

The conditional distribution of C given $X = x$ is also Burr distribution with parameter $\gamma_2(x) = 1/(\tau(x)\lambda_2(x))$, where the parameter $\gamma_2(x)$ is chosen to yield various values for the overall censoring percentage c ($c = 10\%, 20\%, 30\%, 40\%$). Since $\gamma(x) = \gamma_1(x)p(x)$ with $p(x) = \gamma_2(x)/(\gamma_1(x) + \gamma_2(x)) = \lambda_1(x)/(\lambda_1(x) + \lambda_2(x))$ is the ultimate proportion of uncensored observations among Z_i for $i = 1, \cdots, n$. We choose $\gamma_2(x)$ such that $1 - p(x)$ is approximately to $(10\%, 20\%, 30\%, 40\%)$ as censoring percentage.

In this simulation, we are interested to calculate the conditional tail index estimator presented in Equation (5). In practice there is some parameters to be fixed. Let K be an asymmetric linear kernel defined as $K(u) = (1.9 - 1.8u)1_{[0,1]}(u)$. Our simulation show that the choice of kernel has no impact. The estimator $\hat{\gamma}_{y_n}^{c,H}(x)$ depends on the bandwidth parameter h which is chosen using the cross-validation implemented in Gardes and Girard (2012).

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

with $\widehat{\widehat{F}}_{n-i}$ is the kernel conditional Kaplan-Meier estimator used in Ndao *et al.*(2016) which depends on parameter h. The aforementioned estimator is calculated on the sample (X_j, δ_j, Z_j), for $j = 1, \cdots, n$ and $i \neq j$.

In case the bandwidth has already been selected then, next step is to determine the non random threshold y_n. We take y_n as the $(n-k)^{th}$ order statistic $Z_{(n-k)}$, as is classical in extreme value statistics. This data-driven method for estimating local non-random threshold is not new in extreme value theory. See for instance, Goegebeur *et al.*(2014a), for Pareto-type models and Goegebeur *et al.*(2014b), for estimating the extreme value index in regression with random covariates.

Several methods have been mentioned in literature and in this paper we adopted the method used by Ndao *et al.*(2016) as described in the following lines:

1. we compute the estimate $\hat{\gamma}^{c,H}_{Z_{(n-k)}}(x)$ with $k = 1, \cdots, n-1$,
2. we form several successive "blocks" of estimates $\hat{\gamma}^{c,H}_{Z_{(n-k)}}(x)$ (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 (see Ndao *et al.*(2016), Goegebeur *et al.*(2014a)).

Finally, we estimate $\gamma_1(x)$ by $\hat{\gamma}^{c,H}_{Z_{(n-k)}}(x)$ with $(h, k) = (h^*, k^*)$.

Other thing to discuss is the selection of semi-metric whose choice will become an important stage for the behavior of nonparametric statistics for functional data. According to the literature the semi-metric based on the derivative are used in practice for the smoothing curve while semi-metric based on functional principal analysis and partial least squares are adapted for rough curves. For more details, we refer to Ferraty and Vieu (2006). Since the curves of $X(t)$ are smoothing curves according to Gardes and Girard (2012), the semi-metric distance based on the derivative will be used to determine the distance between two curves X_1 and X_2. We consider the semi-metric:

$$d_q^{derive}(X_1, X_2) = \sqrt{\int \left(X_1^{(q)}(t) - X_2^{(q)}(t)\right)^2 dt}, \tag{7}$$

where q is the degree of derivative and where $X^{(q)}$ denotes the q^{th} derivative of X. In the following, second, third and fourth derivatives are considered.

3.2. Results

The performance of our estimator $\hat{\gamma}^{c,H}_{Z_{(n-k^*)}}(x)$ defined in (5) is evaluated using Mean Squared Error (*MSE*) and Mean Absolute Error (*MAE*).
We also provide the averaged value (over the N samples) of the number of threshold excesses k^*.

The asymptotic confidence interval was computed using the asymptotic variance presented in Theorem 1.

The accuracy of our estimator depends on the censoring percentage and on the degree of derivative of the semi-metric $d_q^{derive}(.,.)$.

To illustrate the performance of our estimator, we make a comparison with two simple estimation strategies. The first one is a complete-case procedure ("*CC*" for short): we remove all censored observations from the simulated samples. Then, we compute the estimator proposed in Goegebeur *et al.*(2014b) presented in Equation (4). While, the second strategy is the ignored case, where, we consider that δ_i for $i = 1, \cdots, n$ equally to one for all observations. We consider the observations $Z_i,\ i = 1, \ldots, n$ as if they were uncensored. That kind of strategy is called Ignored case ("*CI*" for Censoring-Ignored).

In Table 1, we give the different value of empirical *MSE* and empirical *MAE* of our estimators at different sample size with respect to the different censoring percentage and degree of derivative respectively.
We also obtain asymptotic 95%-level confidence intervals for $\gamma_1(x)$ (the lower and upper bounds are averaged over the N samples). The averaged widths of these intervals (over the N intervals) are provided. As expected, the quality of the estimator deteriorates as the censoring percentage increases and the sample size decreases. Our estimator performs well at high level of derivative of semi-metric distance. This is not surprising since semi-metric distance parameter is known to play a key role to guaranty the behavior of nonparametric statistic in functional data analysis specially when curves appear to be smooth. The interested reader is referred to Ferraty and Vieu (2006).

Considering the results in Table 1, *CI* and *CC* estimators of $\gamma_1(x)$ are quietly biased, even though when censoring is moderate. As result, our estimator proved a significant result regarding the issues of estimating the functional conditional extreme value index under censorship.

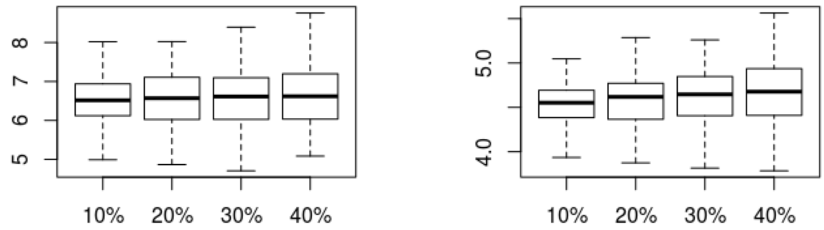
As illustrated from Table 1, the Hill's kernel version estimator under censorship of $\gamma_1(x)$ shows to be well performed in almost all simulation cases. As result, it performs quite better on low censoring percentage for large enough sample size and its quality becomes worst as the censored rate increases and sample size decreases. As by Ferraty and Vieu (2006) advice that in cases of the smoothing curve, the semi-metric based on the derivative is a better one. After going through we have observed that the high order derivative gives a better result.

Figure 2 shows the boxplots of the N realizations of our estimator for different values of the censoring percentage c (10%, 20%, 30%, 40%). This figure illustrates the distribution of the obtained estimates for $N = 500$ replications for sample size $n = 200$ and $n = 500$. One can observed that the proposed estimator performs quite well especially when the sample size is large enough and the censoring percentage decreases. This conclusion is expected, since it is confirmed by the results in Table 1.

Table 1. Table of *MSE* and *MAE* of the estimators value with sample size $n = 200$ and $n = 500$ for $N = 500$ replications and $\Omega = 1$ with 95%-level asymptotic confidence interval for $\gamma_1(x)$ and its empirical average width of confidence interval.

		$n = 200$					$n = 500$				
		MSE	*MAE*	*CI*	*AWCI*	$k^\star$	*MSE*	*MAE*	*CI*	*AWCI*	$k^\star$
		For censorship case									
order=4	10%	0.1082	0.2599	[3.2295, 3.2582]	0.0287	65.152	0.0423	0.1653	[3.2051, 3.2237]	0.0185	161.188
	20%	0.2277	0.4072	[2.7602,2.7964]	0.0361	72.096	0.2054	0.3945	[2.7342, 2.7581]	0.0238	185.388
	30%	0.6504	0.7666	[2.3695,2.4117]	0.0422	77.024	0.6503	0.7585	[2.3324, 2.3629]	0.0305	199.204
	40%	1.3434	1.1366	[1.9967,2.0440]	0.0473	81.700	1.3152	1.1357	[1.9818, 2.0183]	0.0365	207.872
order=3	10%	0.1097	0.2698	[3.2215, 3.2505]	0.0289	63.304	0.0449	0.1716	[3.2253, 3.2442]	0.0188	160.132
	20%	0.2392	0.4153	[2.7352, 2.7744]	0.0392	72.684	0.1903	0.4021	[2.7530, 2.7762]	0.0231	179.536
	30%	0.6790	0.7847	[2.3485, 2.3959]	0.0474	76.296	0.6542	0.7901	[2.3523, 2.3815]	0.0291	195.024
	40%	1.3555	1.1429	[1.9885, 2.0397]	0.0512	82.232	1.3478	1.1399	[1.9868, 2.0212]	0.0344	209.060
order=2	10%	0.1098	0.2714	[3.2184, 3.2476]	0.0291	64.032	0.0484	0.1770	[3.2253,3.2443]	0.0190	162.200
	20%	0.2474	0.4243	[2.7284, 2.7671]	0.0386	75.036	0.1968	0.4029	[2.7526, 2.7766]	0.0240	186.136
	30%	0.7083	0.8029	[2.3307, 2.3774]	0.0467	78.760	0.6677	0.8009	[2.3530, 2.3808]	0.0278	197.180
	40%	1.3692	1.1591	[1.9839, 2.0319]	0.0480	81.028	1.3401	1.1484	[1.9882,2.0199]	0.0316	207.652
		Censoring ignored-based estimation									
order=4	10%	0.1293	0.3010	[2.8960, 2.9318]	0.0358	67.280	0.0933	0.2657	[2.8946, 2.9173]	0.0227	160.836
	20%	0.3584	0.5448	[2.5966, 2.6362]	0.0395	67.756	0.3314	0.4586	[2.5852, 2.6114]	0.0261	166.248
	30%	0.7708	0.8481	[2.2871, 2.3308]	0.0436	71.116	0.7642	0.8320	[2.2797, 2.3103]	0.0305	171.396
	40%	1.3964	1.1642	[1.9687, 2.0169]	0.0481	74.812	1.3918	1.1521	[1.9670, 2.0027]	0.0357	189.216
order=3	10%	0.1351	0.3012	[2.8878, 2.9238]	0.0359	64.256	0.0937	0.2606	[2.8954, 2.9188]	0.0234	162.640
	20%	0.3863	0.5737	[2.5620, 2.6061]	0.0441	68.904	0.3352	0.5544	[2.5896, 2.6155]	0.0258	166.996
	30%	0.7904	0.8604	[2.2717, 2.3213]	0.0495	69.884	0.7790	0.8499	[2.2717, 2.3023]	0.0306	179.536
	40%	1.4184	1.1747	[1.9562, 2.0084]	0.0521	74.756	1.4050	1.1684	[1.9613, 1.9958]	0.0344	191.240
order=2	10%	0.1384	0.3020	[2.8813, 2.9180]	0.0367	66.076	0.0970	0.2638	[2.8938, 2.9174]	0.0236	165.192
	20%	0.3872	0.5763	[2.5712, 2.6133]	0.0421	70.024	0.3503	0.5601	[2.5756, 2.6022]	0.0266	172.144
	30%	0.8068	0.8704	[2.2625, 2.3105]	0.0479	69.996	0.7839	0.8702	[2.2691,2.2984]	0.0292	179.052
	40%	1.4223	1.1763	[1.9560, 2.0053]	0.0493	75.176	1.3912	1.1699	[1.9690, 2.0011]	0.0321	186.312
		Complete-case estimation									
order=4	10%	0.1452	0.3096	[2.9039, 2.9386]	0.0347	58.805	0.0973	0.2572	[2.8928, 2.9152]	0.0223	148.662
	20%	0.3941	0.5624	[2.5762, 2.6199]	0.0436	54.577	0.3498	0.4650	[2.5764, 2.6044]	0.0279	138.925
	30%	0.8326	0.8692	[2.2524, 2.3081]	0.0556	51.255	0.7988	0.8464	[2.2581, 2.2949]	0.0368	128.993
	40%	1.4637	1.1811	[1.9192, 1.9955]	0.0763	47.553	1.4191	1.1798	[1.9430, 1.9928]	0.0498	120.135
order=3	10%	0.1526	0.3190	[2.8784, 2.9142]	0.0358	60.603	0.0975	0.2573	[2.9039,2.9266]	0.0226	144.235
	20%	0.3999	0.5773	[2.5588, 2.6046]	0.0457	56.256	0.3472	0.5618	[2.5793, 2.6077]	0.0283	134.688
	30%	0.8524	0.8878	[2.2329, 2.2934]	0.0605	52.626	0.7915	0.8746	[2.2594, 2.2966]	0.0372	125.360
	40%	1.5323	1.2115	[1.8887, 1.9700]	0.0813	48.851	1.4318	1.1868	[1.9348, 1.9859]	0.0510	116.549
order=2	10%	0.1545	0.3200	[2.8756, 2.9111]	0.0354	58.339	0.0987	0.2652	[2.8978, 2.9207]	0.0228	144.933
	20%	0.4043	0.5805	[2.5549, 2.5995]	0.0446	54.242	0.3585	0.5631	[2.5783, 2.6069]	0.0285	135.297
	30%	0.8562	0.8899	[2.2311, 2.2900]	0.0589	50.807	0.8078	0.8818	[2.2519,2.2890]	0.0371	126.054
	40%	1.5907	1.2961	[1.9059, 1.9851]	0.0792	47.148	1.4424	1.1895	[1.9318, 1.9827]	0.0509	117.176

CI: is the average confidence interval.
AWCI: is average width of confidence interval.



(a) For sample size $n = 200$ (b) For sample size $n = 500$

Fig. 2. Boxplot of the $N = 500$ estimates of $\gamma_1(x)$ for censoring percentage $10\%, 20\%, 30\%, 40\%$ respectively for sample size $n = 200$ left and right $n = 500$.

3.3. Illustration on real data

In this section, we propose to illustrate the behaviour of our estimator on blood pressure and stroke recurrence data. The data contain 24-hr ambulatory blood pressure measurements and interesting clinical outcomes about 297 stroke patients. The primary endpoint is the time to the composite stroke recurrent event, including death, disability, or vascular events. The censoring rate is 40%. Each patient's *SBP* is measured every 15 min starting from $19 : 00$ for 24hr (see Fei *et al.*(2020) for more details). The covariate X_i is thus defined by $X_i = (x_{i,1}, \dots, x_{i,96})$ with $x_{i,j}$ the *SBP* for each patients for all $i = 1, \dots, 297$. The dataset can be found at `https://amstat.tandfonline.com/doi/suppl/10.1080/01621459.2019.1602047/suppl_file/uasa_a_1602047_sm0766.zip`. Figure 3 below illustrates some realizations of random curves of the given functional random variable $X(\cdot)$. The covariate X_i is in fact a discretized curve but the fineness of the grid spanning the discretization allows us to consider each subject as a continuous curve as stated in Gardes and Girard (2012). Hence, the covariate can be considered as belonging to an infinite dimensional space $\mathbf{E}$.

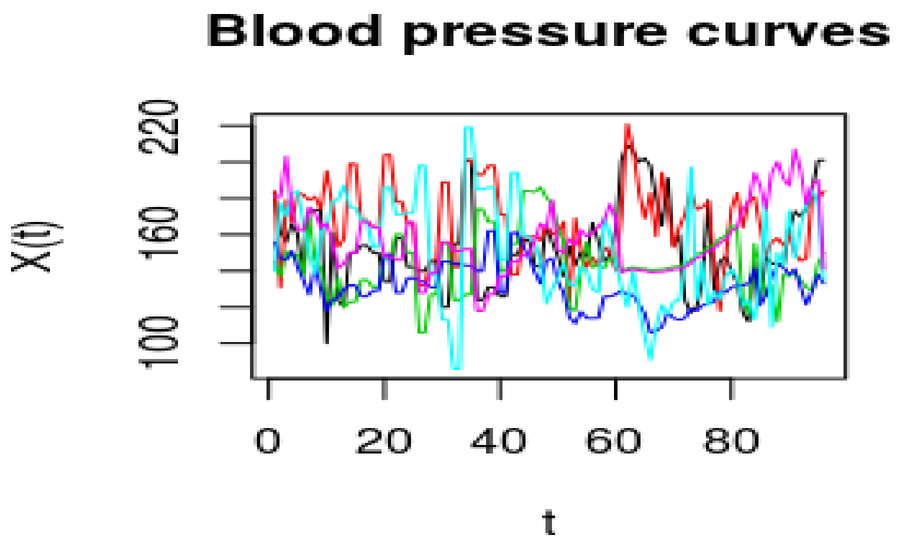


Fig. 3. Measurement of blood pressure

In this application, we choose the semi-metric based on derivative defined in (7) as advised in Ferraty and Vieu (2006). See also Gardes and Girard (2012).

First, we investigate the hypothesis that these data arise from a heavy-tailed distribution. The model described in Section 1 assumes that the cumulative distribution function of both survival and censoring times are heavy-tailed, which implies that the observed time Z is also heavy-tailed. In order to check these assumptions on the Stroke data, we apply a graphical tools.

We carry out, in Figure 4, visual checks of whether the heavy-tailed assumption makes sense for this sample of data. The boxplot, mean excess plot and histogram of the variable of interest show that the heavy tail assumption on the variable of interest is reasonable.

We therefore carry out our analysis of conditional tail index estimator using the methodology described in Section 3.1. The results, presented in Table 2, give an overview of the estimates of conditional extreme value index for different degree of derivative for semi-metric distance. In addition, the confidence interval is provided using resampling techniques which reveal that confidence interval becomes narrow as the degree of derivative increases. To get these empirical confidence intervals, we suggest a bootstrap methodology described as follows

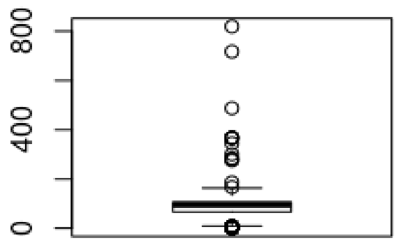
1. Draw $N = 500$ samples of the indexes of our dataset from $1, 2, \cdots, 297$ with replacement,

2. Generate $N = 500$ samples of $(Y_{i,1}, \delta_{i,1}, X_{i,1}) \cdots, (Y_{i,297}, \delta_{i,297}, X_{i,297})$ for corresponding indexes sampled in the first step,
3. Carry out on each of these N samples the estimation of the conditional extreme value index by $\hat{\gamma}^{c,H}_{Z_{(n-k^*)}}(x)$ using the procedure described in the simulation section (with the same (h^*, k^*)).
4. Take the interval bounded by the 2.5% and 97.5% quantile of the conditional extreme value index estimates as a confidence interval. Therefore, the average of low and upper bounds formed the 95%-level asymptotic empirical confidence interval presented on Table 2.

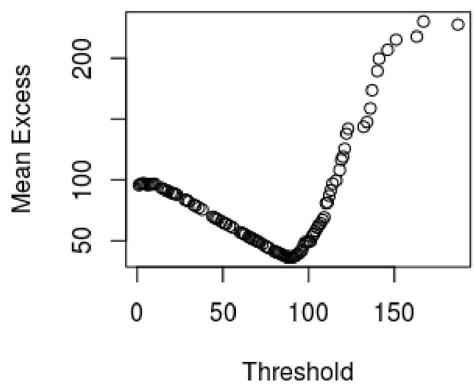
Table 2. Table of estimation result of $\gamma_1(x)$ for the stroke data, first line: estimator of $\gamma_1(x)$, $[\cdot]$ Bootstrap 95%-empirical confidence interval for $\gamma_1(x)$,$(\cdot)$ empirical width of the confidence interval. By considering that x is mean of curves X_i

4^{th} derivative	3^{rd} derivative	2^{nd} derivative
0.5086	0.5418393	0.4960085
$[0.1602, 0.8571]$	$[0.1478, 0.9358]$	$[0.0799, 0.9121]$
(0.6968)	(0.7880)	(0.8322)
$135^\star$	$153^\star$	$145^\star$

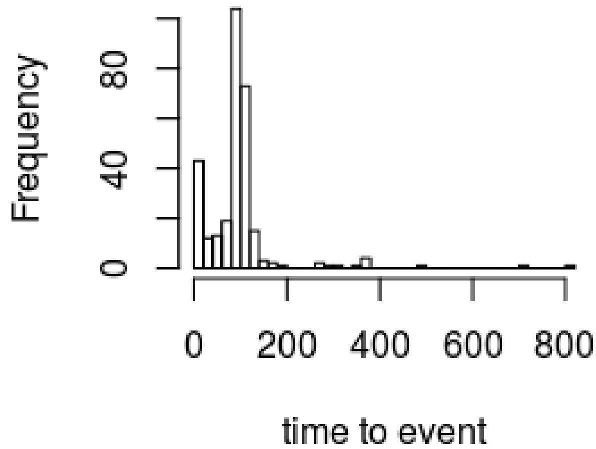
$\star$ is the threshold excesses.



(a) Boxplot for time to stroke



(b) Meplot of time to event



(c) Histogram

Fig. 4. Stroke patients data analysis. Top right panel: mean excess plot for Y_i. Top left panel: Boxplot of time to stroke appeared Y_i. Bottom panel: Histogram of the variable of interest Y_i.

4. Conclusion and further work

We considered the estimation of the conditional extreme value-index when some functional random covariate (i.e. valued in some infinite-dimensional space) information is available and the scalar response variable is right-censored. Its asymptotic properties were established. Its finite sample performance was illustrated on a simulation study and a real data analysis. Also a comparison with two simple estimation strategies has been provided. In a future research, we will focus on a functional conditional extreme value-index and establish its asymptotic behavior in the case where the distribution of the dependent variable Y is not heavy-tailed.

Acknowledgments

The authors acknowledge the Associate Editor and anonymous reviewers for their helpful comments and suggestions that led to an improved version of this paper.
J. Ushize Rutikanga was supported by *DAAD* bursary grants implemented by Institut de Mathématiques et de Sciences Physiques (*IMSP*)-Benin. Part of this work was made while J.Ushize Rutikanga was visiting *LERSTAD* of Gaston Berger University, Saint Louis, Senegal. The authors are grateful for helpful comments from El hadji Deme.

References

BEIRLANT, J., GUILLOU, A., DIERCKX, G., AND FILS-V ILLETARD , A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring. Extremes, 10 (3), 151-1174.

BEIRLANT, J., GUILLOU, A. AND TOULEMONDE, G. (2010). Peaks-over-threshold modeling under random censoring. Communications in Statistics-Theory and Methods,39(7), 1158-1179.

BEIRLANT, J., MARIBE, G., AND VERSTER, A. (2018). Penalized bias reduction in extreme value estimation for censored Pareto-type data, and long-tailed insurance applications. Insurance: Mathematics and Economics, 78, 114-122.

BEIRLANT, J., WORMS, J., AND WORMS, R. (2019). Estimation of the extreme value index in a censorship framework: Asymptotic and finite sample behavior. Journal of Statistical Planning and Inference, 202, 31-56.

BOSQ , D. (2000). Linear processes in function spaces: theory and applications. Journal of the Royal Statistical Society. Series B (Methodological).

BRAHIMI, B., MERAGHNI, D., AND NECIR, A. (2013). On the asymptotic normality of hill's estimator of the tail index under random censoring. Preprint: arXiv-1302.1666, 44.

CHAOUCH, M. AND KHARDANI, S. (2014). Randomly censored quantile regression estimation using functional stationary ergodic data. Journal of Nonparametric Statistics, 27 (1), 65-87.

DE HAAN, L. AND FERREIRA, A. (2007). Extreme value theory: An introduction. Springer Science & Business Media.

DELAFOSSE, E. AND GUILLOU, A. (2002). Almost sure convergence of a tail index estimator in the presence of censoring. Comptes Rendus Mathematique, 335 (4), 375-380.

DREES, H. (1998). On smooth statistical tail functionals. Scandinavian Journal of Statistics, 25 (1),187-210.

EINMAHL, J. H., FILS-VILLETARD , A. AND GUILLOU, A. (2008). Statistics of extremes under random censoring. Bernoulli, 14 (1), 207-227.

FEI, J., QING, C., GUOSHENG , Y., AND SHEN, H. (2020). Functional censored quantile regression. Journal of the American Statistical Association, 115 (530), 931-944. doi:10.1080/01621459.2019.1602047. URL: https:// doi.org/10.1080/01621459.2019.1602047

FERRATY, F. AND VIEU, P. (2004). Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination. Nonparametric Statistics, 16 (1-2),111125.

FERRATY, F. AND VIEU, P. (2006). Nonparametric Functional Data Analysis: Theory and Practice (Springer Series in Statistics). Springer-Verlag, Berlin, Heidelberg.

FERREZ, J., DAVISON, A., AND REBETEZ, M. (2011). Extreme temperature analysis under forest cover compared to an open field. Agricultural and Forest Meteorology, 151 (7), 992-1001.

GARDES, L. AND 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. AND 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. AND GIRARD, S. (2012). Functional kernel estimators of large conditional quantiles. Electronic Journal of Statistics, 6, 1715-1744.

GOEGEBEUR, Y., GUILLOU, A., AND OSMANN, M. (2014a). A local moment type estimator for the extreme value index in regression with random covariates.Canadian Journal of Statistics,42(3), 487-507.

GOEGEBEUR,Y., GUILLOU, A. AND SCHORGEN, A. (2014b). Nonparametric regression estimation of conditional tails: the random covariate case.Statistics,48(4), 732-755.

GOMES, M. I. AND NEVES, M. M. (2010). A note on statistics of extremes for censoring schemes on a heavy right tail.In Information Technology Interfaces (ITI), 2010 32nd International Conference on. IEEE, 539-544.

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

Lo, G. S. (2018). Mathematical foundations of probability theory. arXiv preprint arXiv:1808.01713

Ndao P. (2015). Modélisation de valeurs extrêmes conditionnelles en présence de censure. PhD thesis, Université Gaston Berger de Saint-Louis.

NDAO P, DIOP A, DUPUY JF. (2016). Nonparametric estimation of the conditional extreme-value index with random covariates and censoring. Journal of Statistical Planning and Inference, 168:20-37.

PISARENKO, V. AND SORNETTE, D. (2003). Characterization of the frequency of extreme earth-quake events by the generalized Pareto distribution.Pure and applied geophysics,160(12), 2343-2364.

RAMSAY, J. O. AND DALZELL, C. (1991). Some tools for functional data analysis.Journal of the Royal Statistical Society. Series B (Methodological), 539-572.

RAMSAY, J. O. AND SILVERMAN, B. W. (2005). Functional data analysis.Journal of the Royal Statistical Society. Series B (Methodological).

SCHERVISH, M. J. (2012). Theory of statistics. Springer Science & Business Media.

Stupfler G. (2016). Estimating the conditional extreme-value index under random right-censoring. Journal of Multivariate Analysis. 144:1-24.

VAN DERVAART, A. W. (2000).Asymptotic statistics, volume 3. Cambridge university press.

WORMS, J. AND WORMS, R.(2014). New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. Extremes, 17(2):337-358.
WORMS, J. AND WORMS, R. (2018). Extreme value statistics for censored data with heavy tails under competing risks. Metrika,81(7), 849-889.

5. Appendix

5.1. Preliminary

In order to prove our results, we now start by introducing the following notation:

$$\overline{H}(y|x) = \mathbb{P}(Z > y|X = x)$$

as the conditional survival function of Z given $X = x$ and

$$\overline{H}^1(y|x) = \mathbb{P}(Z > y, \delta = 1|X = x)$$

is a sub-distribution function of Z given $X = x$. We further define

$$\tilde{p}_x(y_n) = \frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)},$$

$$\mathbb{W}_{n\psi}(x) = \left(n \frac{\left(\mu_x^{(1)}(h)\right)^2 \overline{H}^1(y_n|x))}{\mu_x^{(2)}(h)} \right)^{1/2} \left(\frac{\hat{\psi}_n(y_n, x) - E\hat{\psi}_n(y_n, x)}{\overline{H}^1(y_n|x)} \right),$$

$$\mathbb{W}_{n\zeta}(x) = \left(n \frac{\left(\mu_x^{(1)}(h)\right)^2 \overline{H}(y_n|x)}{\mu_x^{(2)}(h)} \right)^{1/2} \left(\frac{\hat{\zeta}_n(y_n, x) - \mathrm{E}\,\hat{\zeta}_n(y_n, x)}{\overline{H}(y_n|x)} \right).$$

Lemma 1. *Suppose **(A1)** and **(A3)** hold. If $y_n \to \infty$ and $h \log y_n \to 0$ as $n \to \infty$, then*

$$\sup_{d(x,x')\leqslant h} \left| \frac{\overline{H}(y_n|x)}{\overline{H}(y_n|x')} - 1 \right| = \mathcal{O}(h \log y_n).$$

Proof of Lemma 1 From **(A1)**, we have

$$\log \frac{\overline{H}(y_n|x)}{\overline{H}(y_n|x')} = \log(r(x)) - \log(r(x')) + \int_1^{y_n} \left(\frac{1}{\gamma(x')} - \frac{1}{\gamma(x)} + \varepsilon(\mu|x') - \varepsilon(\mu|x) \right) \frac{du}{u}.$$

Under **(A3)**, we easily get

$$\begin{aligned}
\left|\log \frac{\overline{H}(y_n|x)}{\overline{H}(y_n|x')}\right| &\leqslant K_r d(x,x') + \int_1^{y_n} \{(K_\gamma + K_\varepsilon)d(x,x')\} \frac{du}{u} \\
&\leqslant K_r d(x,x') + \{(K_\gamma + K_\varepsilon)d(x,x')\} \int_1^{y_n} \frac{du}{u} \\
&\leqslant \{(K_r + (K_\gamma + K_\varepsilon)d(x,x')\} \log(y_n) \\
&\leqslant (K_r + (K_\gamma + K_\varepsilon)h \log(y_n).
\end{aligned}$$

Thus,

$$\sup_{d(x,x')\leqslant h} \left|\log \frac{\overline{H}(y_n|x)}{\overline{H}(y_n|x')}\right| = \mathcal{O}(h\log(y_n)) \to 0, as \quad n \to \infty$$

and considering that $\log(z+1) \sim z$ as $z \to 0$, we get the desired result.

Lemma 2. *Suppose **(A2)** holds and let $x \in \boldsymbol{E}$ such that $\varphi_x(h) > 0$, there exist constants $0 < \tilde{C}_1 < \tilde{C}_2 < \infty$ such that*

$$\tilde{C}_1 \varphi_x(h) \leqslant \frac{\left(\mu_x^{(1)}(h)\right)^2}{\mu_x^{(2)}(h)} \leqslant \tilde{C}_2 \varphi_x(h).$$

Proof of Lemma 2. Under Lemma 3 in Gardes and Girard (2012), we get,

$$\begin{aligned}
C_1 \mathbf{1}_{[0,1]}(\frac{d(x,X)}{h}) &\leqslant K(\frac{d(x,X)}{h}) \leqslant C_2 \mathbf{1}_{[0,1]}(\frac{d(x,X)}{h}) \\
C_1\varphi_x(h) &\leqslant \mathbb{E}\left[K(\frac{d(x,X)}{h})\right] \leqslant C_2\varphi_x(h) \\
(C_1\varphi_x(h))^2 &\leqslant \left(\mu_x^{(1)}(h)\right)^2 \leqslant (C_2\varphi_x(h))^2 .
\end{aligned} \tag{8}$$

Similarly, we can write

$$C_1^2\varphi_x(h) \leqslant \mu_x^{(2)}(h) \leqslant C_2^2\varphi_x(h).$$

Using (8) and (9), we prove the desired result.

Lemma 3. *Suppose **(A1)-(A3)** hold, let $y_n \to \infty$ and $h\log y_n \to 0$ as $n \to \infty$. Then, $\forall x \in \boldsymbol{E}$ such that $\varphi_x(h) > 0$,*

1. $\mathbb{E}[\hat{\zeta}_n(y_n,x)] = \overline{H}(y_n|x)\left[1 + \mathcal{O}(h\log y_n)\right]$

2. $\mathbb{E}[\hat{\psi}_n(y_n, x)] = \overline{H}^1(y_n|x)\,[1 + \mathcal{O}(h \log y_n)]$.

Proof of Lemma 3

1. Let us first recall that

$$\hat{\zeta}_n(y_n, x) = \frac{1}{n(\mu_x^{(1)}(h))} \sum_{i=1}^{n} K\left(\frac{d(x, X_i)}{h}\right) \mathbf{1}_{\{Z_i > y_n\}}.$$

The couple (X_i, Z_i), $i = 1, \ldots, n$ being identically distributed, using conditional expectation, we can write

$$\mathbb{E}[\hat{\zeta}_n(y_n, x)] = \frac{1}{\mu_x^{(1)}(h)} \mathbb{E}\left[K\left(h^{-1} d(x, X)\right) \overline{H}(y_n|X)\right].$$

Simple calculations yield

$$\begin{aligned}
\mathbb{E}[\hat{\zeta}_n(y_n, x)] - \overline{H}(y_n|x) &= \frac{1}{\mu_x^{(1)}(h)} \mathbb{E}\left[K\left(h^{-1} d(x, X)\right) \overline{H}(y_n|X)\right] - \overline{H}(y_n|x) \\
&= \frac{1}{\mu_x^{(1)}(h)} \mathbb{E}\left[K\left(h^{-1} d(x, X)\right) \left(\overline{H}(y_n|X) - \overline{H}(y_n|x)\right)\right] \\
|\mathbb{E}[\hat{\zeta}_n(y_n, x)] - \overline{H}(y_n|x)| &\leq \frac{\overline{H}(y_n|x)}{\mu_x^{(1)}(h)} \mathbb{E}\left[K\left(h^{-1} d(x, X)\right) \left|\frac{\overline{H}(y_n|X)}{\overline{H}(y_n|x)} - 1\right| \mathbf{1}_{\{d(x,X) < h\}}\right],
\end{aligned}$$

under Lemma 1

$$\left|\frac{\overline{H}(y_n|X)}{\overline{H}(y_n|x)} - 1\right| \mathbf{1}_{\{d(x,X) < h\}} \leqslant h \log y_n,$$

thus,

$$\mathbb{E}[\hat{\zeta}_n(y_n, x)] - \overline{H}(y_n|x) = \overline{H}(y_n|x)\mathcal{O}(h \log y_n). \tag{9}$$

2. Calculations for $\hat{\psi}_n(y_n, x)$ are similar as above and are omitted for conciseness.

Lemma 4. *Suppose **(A1)-(A3)** hold, let $y_n \to \infty$ and $h \log y_n \to 0$ as $n \to \infty$. Then, $\forall x \in$ **E** such that $\varphi_x(h) > 0$,*

$$\mathbb{W}_{n\psi}(x) \xrightarrow{D} \mathcal{N}(0, 1), \quad \mathbb{W}_{n\zeta}(x) \xrightarrow{D} \mathcal{N}(0, 1).$$

Therefore the random vector

$$\mathbb{W}_n(x) = (\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\zeta}(x))^T$$

converges in distribution to a bivariate Gaussian vector $\mathcal{N}(0, \mathcal{M})$ *where,*

$$\mathcal{M} := \begin{pmatrix} 1 & \sqrt{p(x)} \\ \sqrt{p(x)} & 1 \end{pmatrix}.$$

Proof of Lemma 4.

Let us first rewrite $\mathbb{W}_{n\psi}(x)$ and $\mathbb{W}_{n\zeta}(x)$ as follows:

$$\begin{aligned}\mathbb{W}_{n\zeta}(x) &= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right)^{1/2}\sum_{i=1}^{n} K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}} - EK(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}} \\ &:= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right)^{1/2}\sum_{i=1}^{n} Q_{i,n}\end{aligned}$$

and

$$\begin{aligned}\mathbb{W}_{n\psi}(x) &= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2}\sum_{i=1}^{n} K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} - EK(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} \\ &:= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2}\sum_{i=1}^{n} Q^{\star}_{i,n}\end{aligned}$$

where for all $i = 1, \cdots, n$, the random variables $Q^{\star}_{i,n}$ and $Q_{i,n}$ are defined as follows:

$$\begin{aligned}Q^{\star}_{i,n} &= K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} - \mathbb{E}K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} \\ Q_{i,n} &= K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}} - \mathbb{E}K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}}.\end{aligned}$$

Consider that $\{Q_{i,n}, Q^{\star}_{i,n}, \forall i = 1, \cdots, n\}$ are the set of centered, independent identically distributed random variables. Now let us focus on the calculation of their variances:

$$\begin{aligned}Var(Q_{i,n}) &= Var(K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}) \\ &= \mathbb{E}\left[K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}\right]^2 - \left(\mathbb{E}\left[K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}\right]\right)^2 \\ &= \mathbf{R}_1 - \mathbf{R}_2^2.\end{aligned} \tag{10}$$

By determining $\mathbf{R}_1$ and $\mathbf{R}_2$ separately, we get

$$\begin{aligned}\mathbf{R}_1 &= \mathbb{E}\left[K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}\right]^2 \\ &= \mathbb{E}\left[K^2\left(h^{-1}d(x,X)\right)\overline{H}(y_n|X)\right] \\ &= \overline{H}(y_n|x)\mathbb{E}\left[K^2\left(h^{-1}d(x,X)\right)\right] + \overline{H}(y_n|x)\mathbb{E}\left[K^2\left(h^{-1}d(x,X)\right)\left[\frac{\overline{H}(y_n|X)}{\overline{H}(y_n|x)} - 1\right]\right].\end{aligned}$$

Under Lipschitz conditions **(A3)** and Lemma 1,

$$\sup_{d(x,X)\leqslant h}\left|\frac{\overline{H}(y_n|X)}{\overline{H}(y_n|x)} - 1\right| = \mathcal{O}(h\log y_n),$$

then

$$\begin{aligned}\mathbf{R}_1 &= \overline{H}(y_n|x)E\left[K^2\left(h^{-1}d(x,X)\right)\right]\left[1+\mathcal{O}(h\log y_n)\right] \\ &= \overline{H}(y_n|x)\mu_x^{(2)}(h)\left[1+\mathcal{O}(h\log y_n)\right].\end{aligned} \tag{11}$$

Simple calculations yield

$$\begin{aligned}\mathbf{R}_2 &= \mathbb{E}\left[K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}\right] \\ &= \mathbb{E}\left[K\left(h^{-1}d(x,X)\right)\overline{H}(y_n|X)\right] \\ &= \overline{H}(y_n|x)\mathbb{E}\left[K\left(h^{-1}d(x,X)\right)\right] + \overline{H}(y_n|x)\mathbb{E}\left[K\left(h^{-1}d(x,X)\right)\left(\frac{\overline{H}(y_n|X)}{\overline{H}(y_n|x)} - 1\right)\right].\end{aligned}$$

Using the same arguments as above, we get

$$\mathbf{R}_2 = \overline{H}(y_n|x)\mu_x^{(1)}(h)\left[1+\mathcal{O}(h\log y_n)\right]. \tag{12}$$

By substituting the result of Equations (12) and (11) in Equation (10), we obtain

$$Var(Q_{i,n}) = \overline{H}(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right) - \left(\overline{H}(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2.$$

Then, under Lemma 2, the variance of $\mathbb{W}_{n\varsigma}(x)$ is given by

$$
\begin{aligned}
Var(\mathbb{W}_{n\varsigma}(x)) &= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right) nVar(Q_{i,n}) \\
&= \left(\frac{1}{\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right) \overline{H}(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right) \\
&- \left(\frac{1}{\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right) \left(\overline{H}(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2 \\
&= 1+\mathcal{O}(h\log y_n) - \left(\frac{1}{\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right) \left(\overline{H}(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2 \\
&\to 1 \quad \textit{as} \quad n \to \infty.
\end{aligned}
$$

The proof for $Var(\mathbb{W}_{n\psi}(x))$ is similar and is thus omitted. Thus, we have

$$
Var(\mathbb{W}_{n\psi}(x)) \to 1 \quad \textit{as} \quad n \to \infty. \tag{13}
$$

Let now determine the covariance of $Cov\left(\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\varsigma}(x)\right)$:

$$
\begin{aligned}
Cov\left(\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\varsigma}(x)\right) &= Cov\Bigg(\left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2}\sum_{i=1}^{n} Q_{i,n}, \\
&\qquad \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}(y_n|x)}\right)^{1/2}\sum_{l=1}^{n} Q^{\star}_{l,n}\Bigg) \\
&= \frac{1}{n\mu_x^{(2)}(h)\sqrt{\overline{H}^1(y_n|x)\overline{H}(y_n|x)}} Cov\left(\sum_{i=1}^{n} Q_{i,n}, \sum_{l=1}^{n} Q^{\star}_{l,n}\right) \\
&= \frac{\sqrt{\tilde{p}_x(y_n)}}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)} Cov\left(\sum_{i=1}^{n} Q_{i,n}, \sum_{l=1}^{n} Q^{\star}_{l,n}\right) \\
&= \frac{\sqrt{\tilde{p}_x(y_n)}}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)} \sum_{l=1}^{n}\sum_{i=1}^{n} Cov\left(Q_{i,n}, Q^{\star}_{l,n}\right) \\
&= \frac{\sqrt{\tilde{p}_x(y_n)}}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)} \sum_{i=1}^{n} Cov\left(Q_{i,n}, Q^{\star}_{i,n}\right)
\end{aligned}
$$

since $Cov\left(Q_{i,n}, Q^{\star}_{l,n}\right) = 0$ for $i \neq l$.
Now let us consider the covariance $Cov\left(Q_{1,n}, Q^{\star}_{1,n}\right)$ given by

$$\begin{aligned}
Cov\left(Q_{1,n}, Q^{\star}_{1,n}\right) &= Cov\left(K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n,\delta=1\}}, K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}\right) \\
&= \mathbb{E}\left(K^2(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n,\delta=1\}}\right) \\
&- \mathbb{E}\left(K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n,\delta=1\}}\right)\mathbb{E}\left(K(\frac{d(x,X)}{h})\mathbf{1}_{\{Z>y_n\}}\right) \\
&:= \mathbf{R}_3 - \mathbf{R}_4.
\end{aligned} \tag{14}$$

From the previous result, we obtain

$$\begin{aligned}
\mathbf{R}_3 &= \overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right) \\
\mathbf{R}_4 &= \overline{H}^1(y_n|x)\overline{H}(y_n|x)\left(\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2.
\end{aligned}$$

Thus, the covariance $Cov\left(Q_{1,n}, Q^{\star}_{1,n}\right)$ is given by

$$\begin{aligned}
Cov\left(Q_{1,n}, Q^{\star}_{1,n}\right) &= \overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right) \\
&- \overline{H}^1(y_n|x)\overline{H}(y_n|x)\left(\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2.
\end{aligned} \tag{15}$$

It follows from Equation (15) that

$$\begin{aligned}
Cov\left(\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\zeta}(x)\right) &= \frac{\sqrt{\tilde{p}_x(y_n)}}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\sum_{i=1}^{n} Cov\left(Q_{i,n}, Q^{\star}_{i,n}\right) \\
&= \left(\frac{\sqrt{\tilde{p}_x(y_n)}}{\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right) \\
&- \left(\frac{\sqrt{\tilde{p}_x(y_n)}}{\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)\overline{H}^1(y_n|x)\overline{H}(y_n|x)\left(\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2.
\end{aligned}$$

From Lemma 2, the covariance

$$Cov\left(\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\zeta}(x)\right) \to \sqrt{p(x)}$$

since

$$\lim_{y_n\to\infty} \tilde{p}_x(y_n) = p(x) as\ n \to \infty\ (see \text{Einmahl } et\ al. (2008),\ \text{Ndao (2015)}).$$

Now, we can prove that Lyapunov's condition for asymptotic normality of the sum of triangular arrays is verified, since we have deduce that $Var(\mathbb{W}_{n\psi}(x)) \to 1)$ as $n \to \infty$.

Let us consider that

$$\sum_{i=1}^{n}\mathbb{E}|Q^{\star\star}_{i,n}|^3 = n\mathbb{E}|Q^{\star\star}_{1,n}|^3 \to 0 \ as \ n\to\infty$$

by assuming that $Q^{\star\star}_{1,n}$ is bounded random variable with

$$Q^{\star\star}_{i,n} = \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2} Q^{\star}_{i,n} :$$

$$\begin{aligned}
|Q^{\star\star}_{1,n}| &\leqslant \frac{2}{\sqrt{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}}\\
|Q^{\star\star}_{1,n}|^3 &\leqslant \frac{2}{\sqrt{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}}|Q^{\star\star}_{1,n}|^2\\
&\leqslant \frac{2}{\sqrt{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}}\left|\left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2} Q^{\star}_{1,n}\right|^2\\
&\leqslant \frac{2}{\left(n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)\right)^{3/2}}\left|Q^{\star}_{1,n}\right|^2
\end{aligned}$$

$$\begin{aligned}
n\mathbb{E}|Q^{\star\star}_{1,n}|^3 &\leqslant \frac{2}{\left(n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)\right)^{3/2}} nVar(Q^{\star}_{1,n})\\
&\leqslant \frac{2}{\left(n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)\right)^{1/2}}(1+o(1))\\
&- \frac{2}{\left(n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)\right)^{1/2}}\overline{H}^1(y_n|x)\frac{(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}(1+o(1)).
\end{aligned}$$

Since

$$n\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\to\infty,$$

the Lyapunov condition is verified, therefore

$$\mathbb{W}_{n\psi}(x)/\sqrt{Var(\mathbb{W}_{n\psi}(x))} \xrightarrow{D} \mathcal{N}(0,1).$$

Now, we want to show that

$$\mathbb{W}_n(x) = (\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\zeta}(x))^T$$

converges in distribution to a Gaussian vector $\mathcal{N}(0, \mathcal{M})$ where,

$$\mathcal{M} := \begin{pmatrix} 1 & \sqrt{p(x)} \\ \sqrt{p(x)} & 1 \end{pmatrix}.$$

To prove that $\mathbb{W}_n(x) = (\mathbb{W}_{n\psi}(x), \mathbb{W}_{n\zeta}(x))^T$ converges in distribution to $\mathcal{N}(0, \mathcal{M})$, according to Cramér-Wold's device, it is sufficient to show that $e^T\mathbb{W}_n(x) \xrightarrow{D} \mathcal{N}(0, e^T\mathcal{M}e)$ for all $e = (e_1, e_2)^T$.

We can write that

$$\begin{aligned} e^T\mathbb{W}_n(x) &= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2} \sum_{i=1}^n \left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right) \\ &- \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2} \sum_{i=1}^n \mathbb{E}\left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right) \\ &= \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right)^{1/2} \sum_{i=1}^n Q_{i,n}^{\star\star\star} \end{aligned}$$

where

$$\begin{aligned} Q_{i,n}^{\star\star\star} &= \left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right) \\ &- \mathbb{E}\left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right). \end{aligned}$$

Hence the variance is given by

$$Var(e^T\mathbb{W}_n(x)) = \left(\frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)}\right) nVar(Q_{1,n}^{\star\star\star})$$

where

$$\begin{aligned} Var(Q_{1,n}^{\star\star\star}) &= Var\left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right) \\ &= \mathbb{E}\left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right)^2 \\ &- \left[\mathbb{E}\left(K(\frac{d(x,X_i)}{h}) \left[e_1 \mathbf{1}_{\{Z_i>y_n,\delta_i=1\}} + \sqrt{\tilde{p}_x(y_n)}e_2\mathbf{1}_{\{Z_i>y_n\}}\right]\right)\right]^2, \\ &=: \mathbf{R}_5 - \mathbf{R}_6. \end{aligned} \tag{16}$$

Now, some simple algebra yields

$$\begin{aligned}
\mathbf{R}_5 &= \mathbb{E}\left(K^2(\frac{d(x,X_i)}{h})\left[e_1^2\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}+\sqrt{\tilde{p}_x(y_n)}2e_1e_2\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}+\tilde{p}_x(y_n)e_2^2\mathbf{1}_{\{Z_i>y_n\}}\right]\right)\\
&= \mathbb{E}\left(K^2(\frac{d(x,X_i)}{h})e_1^2\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}\right)+2e_1e_2\sqrt{\tilde{p}_x(y_n)}\mathbb{E}\left(K^2(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}\right)\\
&+ \tilde{p}_x(y_n)e_2^2\mathbb{E}\left(K^2(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}}\right)\\
&= e_1^2\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right)+2e_1e_2\sqrt{\tilde{p}_x(y_n)}\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\\
&+ e_2^2\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\\
&= \left(e_1^2+2e_1e_2\sqrt{\tilde{p}_x(y_n)}+e_2^2\right)\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right). \qquad (17)
\end{aligned}$$

Similar calculations yield:

$$\begin{aligned}
\mathbf{R}_6 &= e_1^2\left(\mathbb{E}\left(K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}\right)\right)^2+\tilde{p}_x(y_n)e_2^2\left(\mathbb{E}\left(K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}}\right)\right)^2 \qquad (18)\\
&+ 2\sqrt{\tilde{p}_x(y_n)}e_1e_2\mathbb{E}\left(K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}\right)E\left(K(\frac{d(x,X_i)}{h})\mathbf{1}_{\{Z_i>y_n\}}\right)\\
&= e_1^2\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2+\tilde{p}_x(y_n)e_2^2\left(\overline{H}(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2\\
&+ 2\sqrt{\tilde{p}_x(y_n)}e_1e_2\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)\left(\overline{H}(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)\\
&= e_1^2\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2+\frac{1}{\tilde{p}_x(y_n)}e_2^2\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2\\
&+ 2\frac{1}{\sqrt{\tilde{p}_x(y_n)}}e_1e_2\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)\\
&= \left(e_1^2+2e_1e_2\frac{1}{\sqrt{\tilde{p}_x(y_n)}}+\frac{1}{p(x)}e_2^2\right)\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2.
\end{aligned}$$

By substituting the final result of Equation (17) and (18) in Equation (16) then, the variance of $Var(Q_{1,n}^{\star\star\star})$ is given by

$$\begin{aligned}
Var(Q_{1,n}^{\star\star\star}) &= \left(e_1^2+2e_1e_2\tilde{p}_x(y_n)+e_2^2\right)\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\\
&+ \left(e_1^2+2e_1e_2\frac{1}{\sqrt{\tilde{p}_x(y_n)}}+\frac{1}{\tilde{p}_x(y_n)}e_2^2\right)\left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1+\mathcal{O}(h\log y_n)\right)\right)^2.
\end{aligned}$$

Then,

$$
\begin{aligned}
Var(e^T \mathbb{W}_n(x)) &= \frac{1}{n\mu_x^{(2)}(h)\overline{H}^1(y_n|x)} nVar(Q_{1,n}^{\star\star\star}) \\
&= \frac{e_1^2 + 2e_1e_2\sqrt{\tilde{p}_x(y_n)} + e_2^2)}{\mu_x^{(2)}(h)\overline{H}^1(y_n|x)} \overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1 + \mathcal{O}(h\log y_n)\right) \\
&+ \frac{e_1^2 + 2e_1e_2\frac{1}{\sqrt{\tilde{p}_x(y_n)}} + \frac{1}{\tilde{p}_x(y_n)}e_2^2}{\mu_x^{(2)}(h)\overline{H}^1(y_n|x)} \left(\overline{H}^1(y_n|x)\mu_x^{(1)}(h)\left(1 + \mathcal{O}(h\log y_n)\right)\right)^2.
\end{aligned}
$$

Using again Lemma 2, we deduce that

$$
Var(e^T \mathbb{W}_n(x)) \to \left(e_1^2 + 2e_1e_2\sqrt{p(x)} + e_2^2\right) = e^T \mathcal{M} e \; as \; n \to \infty.
$$

Therefore, for all $e \in \mathbb{R}^2, e \neq 0$, $e^T \mathbb{W}_n(x)$ converges in distribution to the univariate normal distribution $\mathcal{N}(0, e^T \mathcal{M} e)$. The convergence

$$
\mathbb{W}_n(x) \xrightarrow{D} \mathcal{N}(0, \mathcal{M})
$$

follows from Cramer-Wold device.

5.2. Proof of main results

Proof of Proposition 1 Now, let us focus on the term $T_2(x)$ defined as follows

$$
T_2(x) := \sigma_n^{-1}(x)\left(\hat{p}_n(x) - p(x)\right).
$$

This term can be rewritten as

$$
T_2(x) = \sigma_n^{-1}(x)\left(\hat{p}_n(x) - \frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}\right) + \sigma_n^{-1}(x)\left(\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)} - p(x)\right).
$$

Let us deal with the first term of the above summation. Applying uniform Delta method and some algebra, we get

$$\sigma_n^{-1}(x)\begin{pmatrix}\frac{\hat\psi_n(y_n,x)}{\overline{H}(y_n|x)}-\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}\\ \frac{\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-1\end{pmatrix} = \left(n\frac{\overline{H}(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}\right)^{1/2}\begin{pmatrix}\frac{\hat\psi_n(y_n,x)-\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}\\ \frac{\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-1\end{pmatrix}$$

$$= \begin{pmatrix}\sqrt{\tilde p_x(y_n)}\sqrt{n\frac{\overline{H}^1(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\hat\psi_n(y_n,x)}{\overline{H}^1(y_n|x)}-1\right)\\ \sqrt{n\frac{\overline{H}(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-1\right)\end{pmatrix}$$

$$= \begin{pmatrix}\sqrt{\tilde p_x(y_n)}\sqrt{n\frac{\overline{H}^1(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\hat\psi_n(y_n,x)}{\overline{H}^1(y_n|x)}-\frac{\mathbb{E}\hat\psi_n(y_n,x)}{\overline{H}^1(y_n|x)}\right)\\ \sqrt{n\frac{\overline{H}(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-\frac{\mathbb{E}\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}\right)\end{pmatrix}$$

$$+ \begin{pmatrix}\sqrt{\tilde p_x(y_n)}\sqrt{n\frac{\overline{H}^1(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\mathbb{E}\hat\psi_n(y_n,x)}{\overline{H}^1(y_n|x)}-1\right)\\ \sqrt{n\frac{\overline{H}(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\mathbb{E}\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-1\right)\end{pmatrix}.$$

Since

$$\frac{\mathbb{E}\hat\zeta_n(y,x)}{\overline{H}(y_n|x)}-1=\mathcal{O}(h\log y_n)$$

and

$$\frac{\mathbb{E}\hat\psi_n(y,x)}{\overline{H}^1(y_n|x)}-1=\mathcal{O}(h\log y_n)$$

according to Lemma 3 and by assumption

$$\sigma_n^{-1}(x)\,(h\log y_n)\to 0,\ \ as\ n\to\infty,$$

it follows that

$$\sigma_n^{-1}(x)\begin{pmatrix}\frac{\hat\psi_n(y_n,x)}{\overline{H}(y_n|x)}-\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}\\ \frac{\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-1\end{pmatrix} = \begin{pmatrix}\sqrt{\tilde p_x(y_n)}\sqrt{n\frac{\overline{H}^1(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\hat\psi_n(y_n,x)}{\overline{H}^1(y_n|x)}-\frac{\mathbb{E}\hat\psi_n(y_n,x)}{\overline{H}^1(y_n|x)}\right)\\ \sqrt{n\frac{\overline{H}(y_n|x)(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)}}\left(\frac{\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}-\frac{\mathbb{E}\hat\zeta_n(y_n,x)}{\overline{H}(y_n|x)}\right)\end{pmatrix}$$

$$= \begin{pmatrix}\sqrt{\tilde p_x(y_n)}\mathbb{W}_{n\psi}(x)\\ \mathbb{W}_{n\zeta}(x)\end{pmatrix}+o(1). \tag{19}$$

Using the convergence of

$$\mathbb{W}_n(x)\xrightarrow{D}\mathcal{N}(0,\mathcal{M}),$$

Delta method and the fact that $\tilde{p}_x(y_n) \to p(x)$, the quantity in Equation (19) converges to bivariate Gaussian vector $\mathcal{N}(0, \Sigma)$ where

$$\Sigma := \begin{pmatrix} p(x) & p(x) \\ p(x) & 1 \end{pmatrix}.$$

Let consider a given function $\Phi(\omega_1, \omega_2) = \omega_1/\omega_2$ and apply again uniform Delta method (see Schervish (2012) ,Van Dervaat (2000)), we have

$$\sigma_n^{-1}(x)\left(\hat{p}_n(x) - \frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}\right) = \sigma_n^{-1}(x)\left(\Phi\left(\frac{\hat{\psi}_n(y_n,x)}{\overline{H}(y_n|x)}, \frac{\hat{\zeta}_n(y_n,x)}{\overline{H}(y_n|x)}\right) - \Phi\left(\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}, 1\right)\right)$$

converges in distribution to

$$\mathcal{N}(0, \Phi'(p(x), 1)\Sigma(\Phi'(p(x), 1))^T).$$

By adapting the similar idea in Beirlant *et al.*(2007), we want to show that

$$\sigma_n^{-1}(x)\left(\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)} - p(x)\right) = \sigma_n^{-1}(x)\left(\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)} - p(x)\right) = o_p(1).$$

Let us set $y_n \in [1, \infty[$

$$\begin{aligned}\frac{\int_{y_n}^{\infty} z^{-1/\gamma_1(x)-1/\gamma_2(x)-1} L(z)dz}{y_n^{-1/\gamma_1(x)-1/\gamma_2(x)} L(z)} - \gamma(x) &= \frac{\int_{y_n}^{\infty} \frac{\overline{H}(z|x)}{z} dz}{\overline{H}(y_n|x))} - \gamma(x) \\ &= \varepsilon(y_n|x).\end{aligned}$$

Let us consider $s = z/y_n$, we have

$$\lim_{y_n\to\infty} \frac{\int_{y_n}^{\infty} \frac{\overline{H}(z|x)}{z} dz}{\overline{H}(y_n|x)} = \lim_{y_n\to\infty} \int_1^{\infty} \frac{\overline{H}(y_n s|x)}{\overline{H}(y_n|x)} \frac{ds}{s}.$$

Since the function $\overline{H}(.|x)$ is regular varying,

$$\begin{aligned}\frac{\overline{H}(y_n s|x)}{\overline{H}(y_n|x)} &= \frac{(y_n s)^{-1/\gamma(x) L(y_n s|x)}}{y_n^{-1/\gamma(x)} L(y_n|x)} \\ &= s^{-1/\gamma(x)} \frac{L(y_n s|x)}{L(y_n|x)} \to s^{-1/\gamma(x)} \quad as \quad n \to \infty.\end{aligned}$$

Therefore

$$\lim_{y_n\to\infty}\frac{\int_{y_n}^{\infty}\frac{\overline{H}(z|x)}{z}dz}{\overline{H}(y_n|x)}=\gamma(x).$$

Hence,

$$\int_{y_n}^{\infty}\frac{\overline{H}(z|x)}{z}dz=\gamma(x)\overline{H}(y_n|x)\left(1+\mathcal{O}(\varepsilon(y_n|x))\right).$$

Remark that

$$\begin{aligned}\overline{H}^1(y_n|x) &= \mathbb{P}\left(\min(Y,C)>y_n,\delta=1|X=x\right)\\ &= \mathbb{P}\left(Y>y_n, Y-C<0|X=x\right).\end{aligned}$$

After some calculations, we can write

$$\begin{aligned}\overline{H}^1(y_n|x) &= \int_{y_n}^{\infty}\left(\int_{-\infty}^{0}f_Y(u|x)f_C(u-v|x)dv\right)du\\ &= \int_{y_n}^{\infty}f_Y(u|x)\left(\int_{u}^{\infty}f_C(w|x)dw\right)du\\ &= \int_{y_n}^{\infty}f_Y(u|x)\overline{G}(u|x)du.\end{aligned}$$

Thus,

$$\begin{aligned}\overline{H}^1(y_n|x) &= \int_{y_n}^{\infty}\overline{G}(u|x)dF(u|x)\\ &= \frac{1}{\gamma_1(x)}\int_{y_n}^{\infty}u^{-1/\gamma(x)-1}L_1(u|x)du-\int_{y_n}^{\infty}u^{-1/\gamma(x)}L_1'(u|x)du\end{aligned}$$

Set

$$\varepsilon_1(y_n|x)=\frac{y_nL_1'(y_n|x)}{L_1(y_n|x)}.$$

Under the class of Hall,

$$\varepsilon_1(y_n|x)=\mathcal{O}\left(\varepsilon(y_n|x)\right)$$

and

$$\varepsilon(y_n|x) = \mathcal{O}\left(A\left(\frac{1}{\overline{H}(y_n|x)}|x\right)\right).$$

Hence, we can rewrite as

$$\begin{aligned}
\overline{H}^1(y_n|x) &= \frac{1}{\gamma_1(x)}\int_{y_n}^{\infty}\frac{\overline{H}(u|x)}{u}du - \int_{y_n}^{\infty}\frac{\overline{H}(u|x)}{u}\varepsilon_1(u|x)du \\
&= \frac{1}{\gamma_1(x)}\gamma(x)\overline{H}(y_n|x)\left(1+\mathcal{O}(\varepsilon(y_n|x))\right) \\
&= p(x)\overline{H}(y_n|x)\left(1+\mathcal{O}(\varepsilon(y_n|x))\right).
\end{aligned}$$

Thus as result,

$$\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)} - p(x) = \mathcal{O}(\varepsilon(y_n|x)).$$

Consequently, we obtain

$$\sigma_n^{-1}(x)\left(\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)} - p(x)\right) = o_p(1).$$

Proof of Theorem 1. We decompose for any fixed $x \in \mathbf{E}$,

$$\begin{aligned}
\sigma_n^{-1}(x)\left(\hat{\gamma}_{y_n}^{c,H}(x) - \gamma_1(x)\right) &= \sigma_n^{-1}(x)\left(\frac{\hat{\gamma}_{y_n}^{H}(x)}{\hat{p}_n(x)} - \gamma_1(x)\right) \\
&= \frac{1}{\hat{p}_n(x)}\sigma_n^{-1}(x)\left(\hat{\gamma}_{y_n}^{H}(x) - \hat{p}_n(x)\gamma_1(x)\right) \\
&= \frac{1}{\hat{p}_n(x)}\sigma_n^{-1}(x)\left(\hat{\gamma}_{y_n}^{H}(x) - \gamma(x) + \gamma(x) - \hat{p}_n(x)\gamma_1(x)\right) \\
&= \frac{1}{p(x)}\sigma_n^{-1}(x)\left[\left(\hat{\gamma}_{y_n}^{H}(x) - \gamma(x)\right)\right] \\
&- \frac{\gamma_1(x)}{p(x)}\sigma_n^{-1}(x)\left[\left(\hat{p}_n(x) - p(x)\right)\right] + o_p(1) \\
&:= \frac{1}{p(x)}T_1(x) - \frac{\gamma_1(x)}{p(x)}T_2(x) + o_p(1).
\end{aligned}$$

From here, we do some decomposition

$$\begin{aligned}
T_1(x) &= \sigma_n^{-1}(x)\left[(\hat{\gamma}_{y_n}^H(x)-\gamma(x))\right] \\
&= \sigma_n^{-1}(x)\left[\sum_{i=1}^n K(h^{-1}d(x,X_i))\log(Z_i/y_n)\,\mathbf{1}_{\{Z_i>y_n\}}/\sum_{i=1}^n K(h^{-1}d(x,X_i))\mathbf{1}_{\{Z_i>y_n\}}-\gamma(x)\right] \\
&= \sigma_n^{-1}(x)\left[\frac{1}{n\mu_x^{(1)}(h)}\sum_{i=1}^n K(h^{-1}d(x,X_i))\log(Z_i/y_n)\,\mathbf{1}_{\{Z_i>y_n\}}/\hat{\zeta}_n(y_n,x)-\gamma(x)\right] \\
&= \sigma_n^{-1}(x)\left[\frac{1}{n\mu_x^{(1)}(h)}\sum_{i=1}^n K(h^{-1}d(x,X_i))\log(Z_i/y_n)\,\mathbf{1}_{\{Z_i>y_n\}}/\hat{\zeta}_n(y_n,x)-\gamma(x)\frac{\overline{H}(y_n|x)}{\hat{\zeta}_n(y_n,x)}\right] \\
&- \sigma_n^{-1}(x)\left[\gamma(x)-\gamma(x)\frac{\overline{H}(y_n|x)}{\hat{\zeta}_n(y_n,x)}\right].
\end{aligned}$$

Since

$$\overline{H}(y_n|x)/\hat{\zeta}_n(y_n,x)=1+o_p(1)$$

(from proof of Lemma 1) and using the fact that

$$\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}\to p(x)$$

under Lemma A1 in Ndao *et al.*(2016)

$$\begin{aligned}
T_1(x) &= \left(\frac{n\left(\mu_x^{(1)}(h)\right)^2}{\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\left[\frac{1}{n\mu_x^{(1)}(h)}\sum_{i=1}^n K(h^{-1}d(x,X_i))\log(Z_i/y_n)\,\mathbf{1}_{\{Z_i>y_n\}}-\gamma(x)\overline{H}(y_n|x)\right] \\
&- \gamma(x)\left(\frac{n\left(\mu_x^{(1)}(h)\right)^2}{\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\left[\hat{\zeta}_n(y_n,x)-\overline{H}(y_n|x)\right].
\end{aligned}$$

Similar calculations give

$$\begin{aligned}
T_2(x) &= \left(\frac{n\left(\mu_x^{(1)}(h)\right)^2}{\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\frac{\overline{H}(y_n|x)}{\hat{\zeta}_n(y_n,x)}\left[\left(\hat{\psi}_n(y_n,x)-\overline{H}^1(y_n|x)\right)-p(x)\left(\hat{\zeta}_n(y_n,x)-\overline{H}(y_n|x)\right)\right] \\
&+ \left(\frac{n\overline{H}(y_n|x)\left(\mu_x^{(1)}(h)\right)^2}{\mu_x^{(2)}(h)}\right)^{1/2}\frac{\overline{H}(y_n|x)}{\hat{\zeta}_n(y_n,x)}\left[\frac{\overline{H}^1(y_n|x)}{\overline{H}(y_n|x)}-p(x)\right] \\
&= \left(\frac{n\left(\mu_x^{(1)}(h)\right)^2}{\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\left[\left(\hat{\psi}_n(y_n,x)-\overline{H}^1(y_n|x)\right)-p(x)\left(\hat{\zeta}_n(y_n,x)-\overline{H}(y_n|x)\right)\right]+o_p(1).
\end{aligned}$$

Thus, to prove the asymptotic normality of our estimator, it is sufficient to establish asymptotic normality of the random vector

$$\Xi_n(x) := \left(\frac{n\left(\mu_x^{(1)}(h)\right)^2}{\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\left(\begin{array}{c}\hat{\psi}_n(y_n,x)-\overline{H}^1(y_n|x)\\ \frac{1}{n\mu_x^{(1)}(h)}\sum_{i=1}^n K(h^{-1}d(x,X_i))\log\left(Z_i/y_n\right)\mathbf{1}_{\{Z_i>y_n\}}-\gamma(x)\overline{H}(y_n|x)\end{array}\right).$$

To prove the asymptotic normality of $\Xi_n(x)$, we use Cramér-Wold's device. Let $e = (e_1, e_2)$ be a vector of real numbers, $e \neq 0$. Then,

$$e^T\Xi_n(x) = \left(\frac{1}{n\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\sum_{i=1}^n \chi_{i,n}(x),$$

where,

$$\begin{aligned}\chi_{i,n}(x) &= e_1\left(K(h^{-1}d(x,X_i))\mathbf{1}_{\{Z_i>y_n,\delta_i=1\}}\right)+e_2\left(K(h^{-1}d(x,X_i))\log\left(Z_i/y_n\right)\mathbf{1}_{\{Z_i>y_n\}}\right)\\ &\quad-\sqrt{n\left(\mu_x^{(1)}(h)\right)^2}\left[e_1\overline{H}^1(y_n|x)+e_2\gamma(x)\overline{H}(y_n|x)\right].\end{aligned}$$

From Remark 1.2.3 in De Haan and Ferreira (2007)

$$\lim_{y_n\to\infty}\mathbb{E}(\log Z-\log y_n|Z>y_n)=\gamma(x).$$

Let us introduce the new notation

$$\mathbb{S}_n(y_n,r;x)=\mathbb{E}\left[(\log Z-\log y_n)^r\mathbf{1}_{\{Z>y_n\}}|X=x\right]$$

and

$$\widetilde{\mathbb{S}}_n(y_n,r;x):=\mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\log(Z/y_n)^r\mathbf{1}_{\{Z>y_n\}}\right).$$

Consider a rate function $B(\cdot|x)$ with $B(t|x)\to 0$ for $t\to\infty$ of constant sign for large values of t, such that for all positive u,

$$\lim_{t\to\infty}\frac{\bar{H}(ut|x)/\bar{H}(t|x)-u^{-1/\gamma(x)}}{B(t|x)}=u^{-1/\gamma(x)}\frac{u^{\rho(x)/\gamma(x)}-1}{\gamma(x)\rho(x)}.$$

Let's define

$$\mathbb{T}_n(u;x):=\frac{\bar{H}(ut|x)/\bar{H}(t|x)-u^{-1/\gamma(x)}}{B(t|x)}-u^{-1/\gamma(x)}\frac{u^{\rho(x)/\gamma(x)}-1}{\gamma(x)\rho(x)}. \tag{20}$$

Some algebraic calculation yield here

$$\begin{aligned}\mathbb{S}_n(y_n, r; x) &= \int_{y_n}^{\infty} (\log z - \log y_n)^r d(H(z|x) \\ &= -\lim_{z\to\infty} (\log z - \log y_n)^r \bar{H}(z|x) + r\int_{y_n}^{\infty} (\log z - \log y_n)^{r-1} \frac{\bar{H}(z|x)}{z} dz.\end{aligned}$$

Considering that the first term equals to zero, then

$$\begin{aligned}\mathbb{S}_n(y_n, r; x) &= r\int_{y_n}^{\infty} (\log z - \log y_n)^{r-1} \frac{\bar{H}(z|x)}{z} dz \\ &= r\bar{H}(y_n|x)\int_{y_n}^{\infty} (\log z - \log y_n)^{r-1} \frac{\bar{H}(z|x)}{z\bar{H}(y_n|x)} dz \\ &= r\bar{H}(y_n|x)\int_{1}^{\infty} (\log w)^{r-1} \frac{\bar{H}(wy_n|x)}{w\bar{H}(y_n|x)} dw,\end{aligned}$$

with change of variable $z/y_n = w$.

Considering the ratio $\frac{\bar{H}(wy_n|x))}{\bar{H}(y_n|x)}$, from second order condition, $\mathbb{S}_n(y_n, r; x)$ can be rewritten as

$$\begin{aligned}\mathbb{S}_n(y_n, r; x) &= r\bar{H}(y_n|x)\int_{1}^{\infty} (\log w)^{r-1} w^{-1/\gamma(x)-1} dw \\ &+ r\bar{H}(y_n|x)B(y_n|x)\int_{1}^{\infty} (\log w)^{r-1} w^{-1/\gamma(x)-1} \frac{w^{\rho(x)/\gamma(x)} - 1}{\gamma(x)\rho(x)} dw \\ &+ r\bar{H}(y_n|x)B(y_n|x)\int_{1}^{\infty} (\log w)^{r-1} \mathbb{T}_n(w; x) \frac{dw}{w} \\ &= r\bar{H}(y_n|x)\left(\mathbb{L}_1 + B(y_n|x)(\mathbb{L}_2 + \mathbb{L}_3)\right)\end{aligned}$$

with,

$$\begin{aligned}\mathbb{L}_1 &= \int_{1}^{\infty} (\log w)^{r-1} w^{-1/\gamma(x)-1} dw \\ \mathbb{L}_2 &= \int_{1}^{\infty} (\log w)^{r-1} w^{-1/\gamma(x)-1} \frac{w^{\rho(x)/\gamma(x)} - 1}{\gamma(x)\rho(x)} dw \\ \mathbb{L}_3 &= \int_{1}^{\infty} (\log w)^{r-1} \mathbb{T}_n(w; x) \frac{dw}{w}\end{aligned}$$

For $\mathbb{L}_1$, let $w^{-1/\gamma(x)} = \exp(-v)$ which implies that $\log w = \gamma(x)v$ and $w^{-1/\gamma(x)-1}dw = \gamma(x)\exp(-v)dv$, then

$$\begin{aligned}
\mathbb{L}_1 &= \int_0^\infty (\gamma(x)v)^{r-1}\gamma(x)\exp(-v)dv \\
&= \gamma^r(x)\int_0^\infty v^{r-1}\exp(-v)dv \\
&= \gamma^r(x)\Gamma(r)
\end{aligned}$$

For $\mathbb{L}_2$, let $w^{-\frac{1-\rho(x)}{\gamma(x)}} = \exp(-v)$ which implies that $\log w = \frac{\gamma(x)}{1-\rho(x)}v$ and $w^{-\frac{1-\rho(x)}{\gamma(x)}-1}dw = \frac{\gamma(x)}{1-\rho(x)}\exp(-v)dv$, then

$$\begin{aligned}
\mathbb{L}_2 &= \frac{\gamma(x)}{\rho(x)}\left[\left(\frac{\gamma(x)}{1-\rho(x)}\right)^r\int_0^\infty v^{r-1}\exp(-v)dv - \gamma^r(x)\int_0^\infty v^{r-1}\exp(-v)dv\right] \\
&= \frac{\gamma(x)}{\rho(x)}\left[\left(\frac{\gamma(x)}{1-\rho(x)}\right)^r\Gamma(r) - \gamma^r(x)\Gamma(r)\right] \\
&= \frac{\gamma^{r+1}(x)\Gamma(r)}{\rho(x)}\left[\left(\frac{1}{1-\rho(x)}\right)^r - 1)\right]
\end{aligned}$$

Now, from Theorem B 2.18 in Drees (1998), we have for a function $B_0(\cdot|x)$, possibly different from the function $B(\cdot|x)$ with $B_0(y_n|x) \approx B(y_n|x)$, $y_n \to \infty$ and for each $\epsilon, \delta > 0$, such $|\mathbb{T}_n(w)| \leqslant \epsilon w^{\rho(x)/\gamma(x)+\delta}$. Then,

$$\begin{aligned}
|\mathbb{L}_3| &\leqslant \epsilon\int_1^\infty (\log w)^{r-1}w^{\frac{1-\rho(x)}{\gamma(x)}+\delta-1}dw \\
&= \epsilon\int_0^\infty v^{r-1}\exp\left(-\left(\frac{1-\rho(x)}{\gamma(x)}-\delta\right)v\right)dv \\
&= \epsilon\Gamma(r)\left(\frac{\gamma(x)}{1-\rho(x)-\gamma(x)\delta}\right)^r
\end{aligned}$$

provided $0 < \delta < (1-\rho(x))/\gamma(x)$, thus $\mathbb{L}_3 = o(1)$ for $y_n \to \infty$. Therefore,

$$\mathbb{S}_n(y_n, r; x) = \gamma^r(x)\Gamma(r+1)\bar{H}(y_n|x)\left(1 + \frac{B(y_n|x)}{\rho(x)}\left[\frac{1}{(1-\rho(x))^r} - 1\right](1+o(1))\right).$$

The calculation of

$$\mathbb{E}\left[(\log Z - \log y_n)^r \mathbf{1}_{\{Z>y_n,\delta=1\}}|X=x\right]$$

is similar to

$$\mathbb{E}\left[(\log Z - \log y_n)^r \mathbf{1}_{\{Z>y_n\}}|X=x\right],$$

then is omitted.

Then,

$$\begin{aligned}
\widetilde{\mathbb{S}}_n(y_n,r;x) &= \mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\mathbb{S}_n(y_n,r;X)\right)\\
\widetilde{\mathbb{S}}_n(y_n,r;x)-\mu_x^{(2)}(h)\mathbb{S}_n(y_n,r;x) &= \mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\left[\mathbb{S}_n(y_n,r;X)-\mathbb{S}_n(y_n,r;x)\right]\right)\\
\widetilde{\mathbb{S}}_n(y_n,r;x)-\mu_x^{(2)}(h)\mathbb{S}_n(y_n,r;x) &= \mathbb{S}_n(y_n,r;x)\mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\left[\frac{\mathbb{S}_n(y_n,r;X)}{\mathbb{S}_n(y_n,r;x)}-1\right]\right)\\
\left|\widetilde{\mathbb{S}}_n(y_n,r;x)-\mu_x^{(2)}(h)\mathbb{S}_n(y_n,r;x)\right| &\le \mathbb{S}_n(y_n,r;x)\mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\left[\left|\frac{\mathbb{S}_n(y_n,r;X)}{\mathbb{S}_n(y_n,r;x)}-1\right|\mathbf{1}_{\{d(X,x)<h\}}\right]\right).
\end{aligned}$$

Thus, Assumption **(A5)** yields,

$$\widetilde{\mathbb{S}}_n(y_n,r;x)=\mu_x^{(2)}(h)\mathbb{S}_n(y_n,r;x)\left(1+\mathcal{O}(\Phi_n(y_n,h;x))\right).$$

By considering that $B(y_n|x)\to 0$ and $y_n\to\infty$ as $n\to\infty$, we have

$$\begin{aligned}
\mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\log(Z/y_n)\mathbf{1}_{\{Z>y_n\}}\right) &= \gamma(x)\overline{H}(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(\Phi_n(y_n,h;x))\right)\\
\mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)\log(Z/y_n)\mathbf{1}_{\{Z>y_n,\delta=1\}}\right) &= \gamma(x)\overline{H}^1(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(\Phi_n(y_n,h;x))\right)\\
\mathbb{E}\left(K^2\left(h^{-1}d(x,X)\right)(\log(Z/y_n))^2\mathbf{1}_{\{Z>y_n\}}\right) &= 2\gamma^2(x)\overline{H}(y_n|x)\mu_x^{(2)}(h)\left(1+\mathcal{O}(\Phi_n(y_n,h;x))\right).
\end{aligned}$$

The variance of $e^T\Xi_n(x)$ is given by,

$$Var(e^T\Xi_n(x))=\frac{1}{n\overline{H}(y_n|x)\mu_x^{(2)}(h)}nVar(\chi_{i,n}(x)),$$

where

$$\begin{aligned}
Var(\chi_{i,n}(x)) &= e_1^2\mu_x^{(2)}(h)\overline{H}^1(y_n|x)+2e_2^2\gamma^2(x)\mu_x^{(2)}(h)\overline{H}(y_n|x)\\
&+ 2e_1e_2\gamma(x)\mu_x^{(2)}(h)\overline{H}^1(y_n|x)+o(1)\\
&= \mu_x^{(2)}(h)\overline{H}(y_n|x)\left[e_1^2p(x)+2e_2^2\gamma^2(x)+2e_1e_2p(x)\gamma(x)\right]+o(1).
\end{aligned}$$

Therefore the variance of $e^T\Xi_n(x)$ is given by

$$\begin{aligned}
Var(e^T\Xi_n(x)) &= \left[e_1^2p(x)+2e_2^2\gamma^2(x)+2e_1e_2p(x)\gamma(x)\right]+o(1)\\
&= e^T\Pi e+o(1)
\end{aligned}$$

with

$$\Pi := \begin{pmatrix} p(x) & p(x)\gamma(x) \\ p(x)\gamma(x) & 2\gamma^2(x) \end{pmatrix}.$$

The next step is to establish the asymptotic normality of $e^T\Xi_(x)$, we check that the Lyapunov's criterion for triangular arrays of random variables holds:

$$\lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}\left|\left(\frac{1}{n\overline{H}(y_n|x)\mu_x^{(2)}(h)}\right)^{1/2}\chi_{i,n}(x)\right|^3 = 0.$$

The calculation is the same as in proof of Lemma 1, therefore it is omitted.

According to Cramér-Wold's device, for $e = (e_1, e_2) \in \mathbb{R}^2$, with $e \neq 0$ and $e^T\Xi_n(x)$ converges in distribution to

$$\mathcal{N}(0, e^T\Pi e)$$

which implies that $\Xi_n(x)$ converges in distribution to $\mathcal{N}(0, \Pi)$.
An application of Delta method shows that

$$\sigma_n^{-1}(x)\left(\hat{\gamma}_{y_n}^{c,H}(x) - \gamma_1(x)\right) = \begin{pmatrix} -\frac{\gamma_1(x)}{p(x)} \\ \frac{1}{p(x)} \end{pmatrix}^T \Xi_n(x)$$

converges in distribution to

$$\mathcal{N}\left(0, \frac{\gamma_1^3(x)}{\gamma(x)}\right). \quad \blacksquare$$